Lectures on Stochastic Programming:
Modeling and Theory

Alexander Shapiro
Darinka Dentcheva
Andrzej Ruszczyński

Darinka Dentcheva  
Department of Mathematical Sciences  
Stevens Institute of Technology  
Hoboken, NJ 07030, USA

Andrzej Ruszczyński  
Department of Management Science and Information Systems  
Rutgers University  
Piscataway, NJ 08854, USA

Alexander Shapiro  
School of Industrial and Systems Engineering  
Georgia Institute of Technology  
Atlanta, GA 30332, USA
Preface

The main topic of this book are optimization problems involving uncertain parameters, for which stochastic models are available. Although many ways have been proposed to model uncertain quantities, stochastic models have proved their flexibility and usefulness in diverse areas of science. This is mainly due to solid mathematical foundations and theoretical richness of the theory of probability and stochastic processes, and to sound statistical techniques of using real data.

Optimization problems involving stochastic models occur in almost all areas of science and engineering, so diverse as telecommunication, medicine, or finance, to name just a few. This stimulates interest in rigorous ways of formulating, analyzing, and solving such problems. Due to the presence of random parameters in the model, the theory combines concepts of the optimization theory, the theory of probability and statistics, and functional analysis. Moreover, in recent years the theory and methods of stochastic programming have undergone major advances. All these factors motivated us to present in an accessible and rigorous form contemporary models and ideas of stochastic programming. We hope that the book will encourage other researchers to apply stochastic programming models and to undertake further studies of this fascinating and rapidly developing area.

We do not try to provide a comprehensive presentation of all aspects of stochastic programming, but we rather concentrate on theoretical foundations and recent advances in selected areas. The book is organized in 7 chapters. The first chapter addresses modeling issues. The basic concepts, such as recourse actions, chance (probabilistic) constraints and the nonanticipativity principle, are introduced in the context of specific models. The discussion is aimed at providing motivation for the theoretical developments in the book, rather than practical recommendations.

Chapters 2 and 3 present detailed development of the theory of two- and multistage stochastic programming problems. We analyze properties of the models, develop optimality conditions and duality theory in a rather general setting. Our analysis covers general distributions of uncertain parameters, and also provides special results for discrete distributions, which are relevant for numerical methods. Due to specific properties of two- and multi-stage stochastic programming problems, we were able to derive many of these results without resorting to methods of functional analysis.

The basic assumption in the modeling and technical developments is that the probability distribution of the random data is not influenced by our actions (decisions). In some applications this assumption could be unjustified. However, dependence of probability distribution on decisions typically destroys the convex structure of the optimization problems considered, and our analysis exploits convexity in a significant way.
Chapter 4 deals with chance (probabilistic) constraints, which appear naturally in many applications. The chapter presents the current state of the theory, focusing on the structure of the problems, optimality theory, and duality. We present generalized convexity of functions and measures, differentiability, and approximations of probability functions. Much attention is devoted to problems with separable chance constraints and problems with discrete distributions. We also analyze problems with first order stochastic dominance constraints, which can be viewed as problems with continuum of probabilistic constraints. Many of the presented results are relatively new and were not previously available in standard textbooks.

Chapter 5 is devoted to statistical inference in stochastic programming. The starting point of the analysis is that the probability distribution of the random data vector is approximated by an empirical probability measure. Consequently the “true” (expected value) optimization problem is replaced by its sample average approximation (SAA). Origins of this statistical inference are going back to the classical theory of the maximum likelihood method routinely used in statistics. Our motivation and applications are somewhat different, because we aim at solving stochastic programming problems by Monte Carlo sampling techniques. That is, the sample is generated in the computer and its size is only constrained by the computational resources needed to solve the constructed SAA problem. One of the byproducts of this theory is the complexity analysis of two and multistage stochastic programming. Already in the case of two stage stochastic programming the number of scenarios (discretization points) grows exponentially with the increase of the number of random parameters. Furthermore, for multistage problems, the computational complexity also grows exponentially with the increase of the number of stages.

In Chapter 6 we outline the modern theory of risk averse approaches to stochastic programming. We focus on the analysis of the models, optimality theory, and duality. Static and two-stage risk-averse models are analyzed in much detail. We also outline a risk-averse approach to multistage problems, using conditional risk mappings and the principle of “time consistency.”

Chapter 7 contains formulations of technical results used in the other parts of the book. For some of these, less known, results we give proofs, while others are referred to the literature. The subject index can help the reader to find quickly a required definition or formulation of a needed technical result.

Several important aspects of stochastic programming have been left out. We do not discuss numerical methods for solving stochastic programming problems, with exception of section 5.9 where the Stochastic Approximation method, and its relation to complexity estimates, is considered. Of course, numerical methods is an important topic which deserves careful analysis. This, however, is a vast and separate area which should be considered in a more general framework of modern optimization methods and to large extent would lead outside the scope of this book.

We also decided not to include a thorough discussion of stochastic integer programming. The theory and methods of solving stochastic integer programming problems draw heavily from the theory of general integer programming. Their comprehensive presentation would entail discussion of many concepts and methods of this vast field, which would have little connection with the rest of the book.

At the beginning of each chapter we indicate the authors who were primarily responsible for writing the material, but the book is the creation of all three of us, and we share
Preface

equal responsibility for errors and inaccuracies that escaped our attention.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński
Contents

Preface iii

1 Stochastic Programming Models 1
1.1 Introduction ....................... 1
1.2 Inventory ................................ 1
1.2.1 The Newsvendor Problem .......... 1
1.2.2 Chance Constraints ................. 5
1.2.3 Multistage Models .................. 6
1.3 Multi-Product Assembly ............... 9
1.3.1 Two Stage Model ................... 9
1.3.2 Chance Constrained Model .......... 11
1.3.3 Multistage Model ................... 12
1.4 Portfolio Selection ................... 13
1.4.1 Static Model ......................... 13
1.4.2 Multistage Portfolio Selection ....... 17
1.4.3 Decision Rules ....................... 21
1.5 Supply Chain Network Design ......... 22
Exercises ................................ 25

2 Two Stage Problems 27
2.1 Linear Two Stage Problems .......... 27
2.1.1 Basic Properties .................... 27
2.1.2 The Expected Recourse Cost for Discrete Distributions 30
2.1.3 The Expected Recourse Cost for General Distributions 33
2.1.4 Optimality Conditions ............... 38
2.2 Polyhedral Two-Stage Problems ...... 42
2.2.1 General Properties ................. 42
2.2.2 Expected Recourse Cost ............. 44
2.2.3 Optimality Conditions ............... 47
2.3 General Two-Stage Problems .......... 48
2.3.1 Problem Formulation, Interchangeability 48
2.3.2 Convex Two-Stage Problems .......... 50
2.4 Nonanticipativity ..................... 53
2.4.1 Scenario Formulation ............... 53
viii Contents

2.4.2 Dualization of Nonanticipativity Constraints 55
2.4.3 Nonanticipativity Duality for General Distributions 56
2.4.4 Value of Perfect Information 60
Exercises 61

3 Multistage Problems 63
3.1 Problem Formulation 63
3.1.1 The General Setting 63
3.1.2 The Linear Case 65
3.1.3 Scenario Trees 69
3.1.4 Algebraic Formulation of Nonanticipativity Constraints 72
3.2 Duality 76
3.2.1 Convex Multistage Problems 76
3.2.2 Optimality Conditions 77
3.2.3 Dualization of Feasibility Constraints 81
3.2.4 Dualization of Nonanticipativity Constraints 82
Exercises 84

4 Optimization Models with Probabilistic Constraints 87
4.1 Introduction 87
4.2 Convexity in probabilistic optimization 94
4.2.1 Generalized concavity of functions and measures 94
4.2.2 Convexity of probabilistically constrained sets 107
4.2.3 Connectedness of probabilistically constrained sets 114
4.3 Separable probabilistic constraints 115
4.3.1 Continuity and differentiability properties of distribution functions 115
4.3.2 p-Efficient points 117
4.3.3 Optimality conditions and duality theory 123
4.4 Optimization problems with non-separable probabilistic constraints 133
4.4.1 Differentiability of probability functions and optimality conditions 134
4.4.2 Approximations of non-separable probabilistic constraints 137
4.5 Semi-infinite probabilistic problems 145
Exercises 152

5 Statistical Inference 157
5.1 Statistical Properties of SAA Estimators 157
5.1.1 Consistency of SAA Estimators 159
5.1.2 Asymptotics of the SAA Optimal Value 164
5.1.3 Second Order Asymptotics 167
5.1.4 Minimax Stochastic Programs 171
5.2 Stochastic Generalized Equations 175
5.2.1 Consistency of Solutions of the SAA Generalized Equations 177
5.2.2 Asymptotics of SAA Generalized Equations Estimators 179
Contents ix

5.3 Monte Carlo Sampling Methods ............................................. 181
  5.3.1 Exponential Rates of Convergence and Sample Size Estimates in Case of Finite Feasible Set .................. 182
  5.3.2 Sample Size Estimates in General Case ............................. 187
  5.3.3 Finite Exponential Convergence ....................................... 193
5.4 Quasi-Monte Carlo Methods ................................................. 194
5.5 Variance Reduction Techniques ............................................ 200
  5.5.1 Latin Hypercube Sampling .............................................. 200
  5.5.2 Linear Control Random Variables Method .......................... 201
  5.5.3 Importance Sampling and Likelihood Ratio Methods ............ 202
5.6 Validation Analysis ......................................................... 204
  5.6.1 Estimation of the Optimality Gap .................................... 204
  5.6.2 Statistical Testing of Optimality Conditions ....................... 209
5.7 Chance Constrained Problems .............................................. 212
  5.7.1 Monte Carlo Sampling Approach ..................................... 212
  5.7.2 Validation of an Optimal Solution ................................... 219
5.8 SAA Method Applied to Multistage Stochastic Programming ........ 222
  5.8.1 Statistical Properties of Multistage SAA Estimators ............. 223
  5.8.2 Complexity Estimates of Multistage Programs ...................... 228
5.9 Stochastic Approximation Method ......................................... 231
  5.9.1 Classical Approach ................................................... 232
  5.9.2 Robust SA Approach ................................................ 235
  5.9.3 Mirror Descent SA Method ........................................... 238
  5.9.4 Accuracy Certificates for Mirror Descent SA Solutions ......... 245
Exercises ........................................................................ 250

6 Risk Averse Optimization .................................................... 255
  6.1 Introduction ............................................................... 255
  6.2 Mean–risk models ......................................................... 256
    6.2.1 Main ideas of mean–risk analysis ................................. 256
    6.2.2 Semideviations ..................................................... 257
    6.2.3 Weighted mean deviations from quantiles ....................... 258
    6.2.4 Average Value-at-Risk ............................................. 259
  6.3 Coherent Risk Measures .................................................. 263
    6.3.1 Differentiability Properties of Risk Measures ................. 267
    6.3.2 Examples of Risk Measures ....................................... 271
    6.3.3 Law Invariant Risk Measures and Stochastic Orders .......... 281
    6.3.4 Relation to Ambiguous Chance Constraints .................... 287
  6.4 Optimization of Risk Measures ........................................... 290
    6.4.1 Dualization of Nonanticipativity Constraints ................. 293
    6.4.2 Examples .......................................................... 297
  6.5 Statistical Properties of Risk Measures ................................ 302
    6.5.1 Average Value-at-Risk .............................................. 302
    6.5.2 Absolute Semideviation Risk Measure ............................ 303
    6.5.3 Von Mises Statistical Functionals ................................. 306
  6.6 The Problem of Moments .................................................. 308
6.7 Multistage Risk Averse Optimization ........................................... 310
  6.7.1 Scenario Tree Formulation ............................................... 310
  6.7.2 Conditional Risk Mappings .............................................. 317
  6.7.3 Risk Averse Multistage Stochastic Programming .................... 320

Exercises .................................................................................. 330

7 Background Material ................................................................. 335
  7.1 Optimization and Convex Analysis .......................................... 336
    7.1.1 Directional Differentiability ........................................... 336
    7.1.2 Elements of Convex Analysis .......................................... 339
    7.1.3 Optimization and Duality .............................................. 342
    7.1.4 Optimality Conditions .................................................. 348
    7.1.5 Perturbation Analysis .................................................. 354
    7.1.6 Epiconvergence .......................................................... 360

  7.2 Probability ......................................................................... 361
    7.2.1 Probability Spaces and Random Variables ......................... 361
    7.2.2 Conditional Probability and Conditional Expectation .......... 365
    7.2.3 Measurable Multifunctions and Random Functions ............. 367
    7.2.4 Expectation Functions ................................................... 370
    7.2.5 Uniform Laws of Large Numbers .................................... 376
    7.2.6 Law of Large Numbers for Random Sets and Subdifferentials .................................................................................. 382
    7.2.7 Delta Method .................................................................. 385
    7.2.8 Exponential Bounds of the Large Deviations Theory .......... 391
    7.2.9 Uniform Exponential Bounds .......................................... 396

  7.3 Elements of Functional Analysis ........................................... 402
    7.3.1 Conjugate Duality and Differentiability ........................... 404
    7.3.2 Lattice Structure ......................................................... 407

Exercises .................................................................................. 409

8 Bibliographical Remarks .......................................................... 411

Bibliography .............................................................................. 419

Index .......................................................................................... 434
Chapter 1
Stochastic Programming Models

Andrzej Ruszczyński and Alexander Shapiro

1.1 Introduction
The readers familiar with the area of optimization can easily name several classes of optimization problems, for which advanced theoretical results exist and efficient numerical methods have been found. In that respect we can mention linear programming, quadratic programming, convex optimization, nonlinear optimization, etc. Stochastic programming sounds similar, but no specific formulation plays the role of the generic stochastic programming problem. The presence of random quantities in the model under consideration opens the door to wealth of different problem settings, reflecting different aspects of the applied problem at hand. The main purpose of this chapter is to illustrate the main approaches that can be followed when developing a suitable stochastic optimization model. For the purpose of presentation, these are very simplified versions of problems encountered in practice, but we hope that they will still help us to convey our main message.

1.2 Inventory
1.2.1 The Newsvendor Problem
Suppose that a company has to decide about order quantity \( x \) of a certain product to satisfy demand \( d \). The cost of ordering is \( c > 0 \) per unit. If the demand \( d \) is larger than \( x \), then the company makes an additional order for the unit price \( b \geq 0 \). The cost of this is equal to \( b(d - x) \) if \( d > x \), and is zero otherwise. On the other hand, if \( d < x \), then holding cost of
$h(x - d) \geq 0$ is incurred. The total cost is then equal to\(^1\)

\[ F(x, d) = cx + b[d - x]_+ + h[x - d]_+. \] (1.1)

We assume that $b > c$, i.e., the back order penalty cost is larger than the ordering cost.

The objective is to minimize the total cost $F(x, d)$. Here $x$ is the decision variable and the demand $d$ is a parameter. Therefore, if the demand is known, the corresponding optimization problem can be formulated in the form

\[ \min_{x \geq 0} F(x, d). \] (1.2)

The objective function $F(x, d)$ can be rewritten as

\[ F(x, d) = \max \left\{ (c - b)x + bd, (c + h)x - hd \right\}, \] (1.3)

which is a piecewise linear function with a minimum attained at $x = d$. That is, if the demand $d$ is known, then (as expected) the best decision is to order exactly the demand quantity $d$.

Consider now the case when the ordering decision should be made before a realization of the demand becomes known. One possible way to proceed in such situation is to view the demand $D$ as a random variable. By capital $D$ we denote the demand when viewed as a random variable in order to distinguish it from its particular realization $d$. We assume, further, that the probability distribution of $D$ is known. This makes sense in situations where the ordering procedure repeats itself and the distribution of $D$ can be estimated from historical data. Then it makes sense to talk about the expected value, denoted $\mathbb{E}[F(x, D)]$, of the total cost viewed as a function of the order quantity $x$. Consequently, we can write the corresponding optimization problem

\[ \min_{x \geq 0} \left\{ f(x) := \mathbb{E}[F(x, D)] \right\}. \] (1.4)

The above formulation approaches the problem by optimizing (minimizing) the total cost on average. What would be a possible justification of such approach? If the process repeats itself, then by the Law of Large Numbers, for a given (fixed) $x$, the average of the total cost, over many repetitions, will converge (with probability one) to the expectation $\mathbb{E}[F(x, D)]$, and, indeed, in that case the solution of problem (1.4) will be optimal on average.

The above problem gives a very simple example of a two-stage problem or a problem with a recourse action. At the first stage, before a realization of the demand $D$ is known, one has to make a decision about the ordering quantity $x$. At the second stage, after a realization $d$ of demand $D$ becomes known, it may happen that $d > x$. In that case the company takes the recourse action of ordering the required quantity $d - x$ at the higher cost of $b > c$.

The next question is how to solve the expected value problem (1.4). In the present case it can be solved in a closed form. Consider the cumulative distribution function (cdf) $H(x) := \Pr(D \leq x)$ of the random variable $D$. Note that $H(x) = 0$ for all $x < 0$.

---

\(^1\)For a number $a \in \mathbb{R}$, $[a]_+$ denotes the maximum $\max\{a, 0\}$. 
because the demand cannot be negative. The expectation $\mathbb{E}[F(x, D)]$ can be written in the following form

$$\mathbb{E}[F(x, D)] = b\mathbb{E}[D] + (c - b)x + (b + h) \int_0^x H(z)dz. \quad (1.5)$$

Indeed, the expectation function $f(x) = \mathbb{E}[F(x, D)]$ is a convex function. Moreover, since it is assumed that $f(x)$ is well defined and finite values, it is continuous. Consequently for $x \geq 0$ we have

$$f(x) = f(0) + \int_0^x f'(z)dz,$$

where at nondifferentiable points the derivative $f'(z)$ is understood as the right side derivative. Since $D \geq 0$ we have that $f(0) = b\mathbb{E}[D]$. Also we have that

$$f'(z) = c + \mathbb{E} \left[ \frac{\partial}{\partial z} (b[D - z]_+ + h[z - D]_+) \right]$$

$$= c - b \Pr(D \geq z) + h \Pr(D \leq z)$$

$$= c - b(1 - H(z)) + hH(z)$$

$$= c - b + (b + h)H(z).$$

Formula (1.5) then follows.

We have that $\frac{d}{dx} \int_0^x H(z)dz = H(x)$, provided that $H(\cdot)$ is continuous at $x$. In this case, we can take the derivative of the right hand side of (1.5) with respect to $x$ and equate it to zero. We conclude that the optimal solutions of problem (1.4) are defined by the equation

$$(b + h)H(x) + c - b = 0,$$

and hence an optimal solution of problem (1.4) is equal to the quantile

$$\bar{x} = H^{-1}(\kappa), \quad \text{with} \quad \kappa = \frac{b - c}{b + h}. \quad (1.6)$$

**Remark 1.** Recall that for $\kappa \in (0, 1)$ the left side $\kappa$-quantile of the cdf $H(\cdot)$ is defined as $H^{-1}(\kappa) := \inf \{ t : H(t) \geq \kappa \}$. In a similar way, the right side $\kappa$-quantile is defined as $\sup \{ t : H(t) \leq \kappa \}$. If the left and right $\kappa$-quantiles are the same, then problem (1.4) has unique optimal solution $\bar{x} = H^{-1}(\kappa)$. Otherwise the set of optimal solutions of problem (1.4) is given by the whole interval of $\kappa$-quantiles.

Suppose for the moment that the random variable $D$ has a finitely supported distribution, i.e., it takes values $d_1, \ldots, d_K$ (called scenarios) with respective probabilities $p_1, \ldots, p_K$. In that case its cdf $H(\cdot)$ is a step function with jumps of size $p_k$ at each $d_k$, $k = 1, \ldots, K$. Formula (1.6) for an optimal solution still holds with the corresponding left side (right side) $\kappa$-quantile, coinciding with one of the points $d_k$, $k = 1, \ldots, K$. For example, the scenarios may represent historical data collected over a period of time. In such case the corresponding cdf is viewed as the empirical cdf giving an approximation (estimation) of the true cdf, and the associated $\kappa$-quantile is viewed as the sample estimate of the $\kappa$-quantile associated with the true distribution.
Chapter 1. Stochastic Programming Models

It is instructive to compare the quantile solution \( \bar{x} \) with a solution corresponding to one specific demand value \( d := \bar{d} \), where \( d \) is, say, the mean (expected value) of \( D \). As it was mentioned earlier, the optimal solution of such (deterministic) problem is \( \bar{d} \). The mean \( \bar{d} \) can be very different from the \( \kappa \)-quantile \( \bar{x} = H^{-1}(\kappa) \). It is also worthwhile to mention that sample quantiles are typically much less sensitive than sample mean to random perturbations of the empirical data.

In applications, closed form solutions for stochastic programming problems such as (1.4) are rarely available. In the case of finitely many scenarios it is possible to model the stochastic program as a deterministic optimization problem, by writing the expected value \( \mathbb{E}[F(x, D)] \) as the weighted sum:

\[
\mathbb{E}[F(x, D)] = \sum_{k=1}^{K} p_k F(x, d_k).
\]

The deterministic formulation (1.2) corresponds to one scenario \( d \) taken with probability one. By using the representation (1.3), we can write problem (1.2) as the linear programming problem

\[
\begin{align*}
\text{Min} & \quad v \\
\text{s.t.} & \quad v \geq (c - b)x + bd, \\
& \quad v \geq (c + h)x - hd.
\end{align*}
\]  

Indeed, for fixed \( x \), the optimal value of (1.7) is equal to \( \max \{ (c - b)x + bd, (c + h)x - hd \} \), which is equal to \( F(x, d) \). Similarly, the expected value problem (1.4), with scenarios \( d_1, ..., d_K \), can be written as the linear programming problem:

\[
\begin{align*}
\text{Min} & \quad \sum_{k=1}^{K} p_k v_k \\
\text{subject to} & \quad v_k \geq (c - b)x + bd_k, \quad k = 1, ..., K, \\
& \quad v_k \geq (c + h)x - hd_k, \quad k = 1, ..., K.
\end{align*}
\]  

It is worth noting here the almost separable structure of problem (1.8). For a fixed \( x \), problem (1.8) separates into the sum of optimal values of problems of the form (1.7) with \( d = d_k \). As we shall see later such decomposable structure is typical for two-stage stochastic programming problems.

**Worst Case Approach**

One can also consider the worst case approach. That is, suppose that there are known lower and upper bounds for the demand, i.e., it is unknown that \( d \in [l, u] \), where \( l \leq u \) are given (nonnegative) numbers. Then the worst case formulation is

\[
\begin{align*}
\text{Min} & \quad \max_{x \geq 0, d \in [l, u]} F(x, d).
\end{align*}
\]  

That is, while making decision \( x \) one is prepared for the worst possible outcome of the maximal cost. By (1.3) we have that

\[
\max_{d \in [l, u]} F(x, d) = \max \{ F(x, l), F(x, u) \}.
\]
1.2. Inventory

Clearly we should look at the optimal solution in the interval \([l, u]\), and hence problem (1.9) can be written as

\[
\min_{x \in [l,u]} \max \left\{ cx + h[x-l]_+, cx + b[u-x]_+ \right\}.
\]

Consequently, if \(b > h\), then the optimal solution of problem (1.9) is \(x^* = u\), and if \(b < h\), then the optimal solution of problem (1.9) is \(x^* = l\).

These “worst case” solutions can be compared with the solution \(x\) which is optimal on average, and which is given in (1.6). That is suppose that the probability distribution of the random demand \(D\) is supported on the interval \([l, u]\). The solution \(x\), which is optimal on average, depends on values of parameters \(c, h, b\) and probability distribution of \(D\). Suppose, for example, that \(D\) has uniform distribution on the interval \([l, u]\). Then \(x = l + \kappa(u-l)\). That is, depending on \(c, h, b\), the solution \(x\) will be a point inside the interval \([l, u]\). On the other hand, the extreme solutions \(x^* = l\) or \(x^* = u\) could be too conservative in that case.

Suppose now that in addition to the lower and upper bounds of the demand we know its mean (expected value) \(\bar{d} = \mathbb{E}[D]\). Of course, we have that \(d \in [l, u]\). Then we can consider the following worst case formulation

\[
\min_{x \geq 0} \sup_{H \in \mathcal{M}} \mathbb{E}_H[F(x, D)],
\]

where \(\mathcal{M}\) denotes the set of probability measures supported on the interval \([l, u]\) and having mean \(\bar{d}\), and the notation \(\mathbb{E}_H[F(x, D)]\) emphasizes that the expectation is taken with respect to the cumulative distribution function (probability measure) \(H(\cdot)\) of \(D\). It is possible to show that the minimax problem (1.10) attains its optimal solution at one of the extreme points \(l\) or \(u\) (see problem 6.8 on page 332). We study minimax problems of the form (1.10) in section 6.6.

1.2.2 Chance Constraints

We have already observed that for a particular realization of the demand \(D\), the cost \(F(x, D)\) can be quite different from the optimal-on-average cost \(\mathbb{E}[F(x, D)]\). Therefore a natural question is whether we can control the risk of the cost \(F(x, D)\) to be not “too high.” For example, for a chosen value (threshold) \(\tau > 0\), we may add to problem (1.4) the constraint \(F(x, D) \leq \tau\) to be satisfied for all possible realizations of the demand \(D\). That is, we want to make sure that the total cost will not be larger than \(\tau\) in all possible circumstances. Assuming that the demand can vary in a specified uncertainty set \(\mathcal{D} \subset \mathbb{R}\), this means that the inequalities \((c-b)x + bd \leq \tau\) and \((c+h)x - hd \leq \tau\) should hold for all possible realizations \(d \in \mathcal{D}\) of the demand. That is, the ordering quantity \(x\) should satisfy the following inequalities

\[
\frac{bd-\tau}{b-c} \leq x \leq \frac{hd+\tau}{c+h}, \quad \forall d \in \mathcal{D}.
\]

This could be quite restrictive if the uncertainty set \(\mathcal{D}\) is large. In particular, if there is at least one realization \(d \in \mathcal{D}\) greater than \(\tau/c\), then the system (1.11) is inconsistent, i.e., the corresponding problem has no feasible solution.
In such situations it makes sense to introduce the constraint that the probability of \( F(x, D) \) being larger than \( \tau \) is less than a specified value (significance level) \( \alpha \in (0, 1) \). This leads to a chance (also called probabilistic) constraint which can be written in the form
\[
\Pr\{F(x, D) > \tau\} < \alpha,
\]
or equivalently
\[
\Pr\{F(x, D) \leq \tau\} \geq 1 - \alpha.
\]
By adding the chance constraint (1.13) to the optimization problem (1.4) we want to minimize the total cost on average, while making sure that the risk of the cost to be excessive (i.e., the probability that the cost is larger than \( \tau \)) is small (i.e., less than \( \alpha \)). We have that
\[
\Pr\{F(x, D) \leq \tau\} = \Pr\left\{ \frac{(c+h)x - \tau}{h} \leq D \leq \frac{(b-c)x + \tau}{c} \right\}.
\]
For \( x \leq \tau/c \) the inequalities on the right hand side of (1.14) are consistent and hence for such \( x \),
\[
\Pr\{F(x, D) \leq \tau\} = H\left(\frac{(b-c)x + \tau}{c}\right) - H\left(\frac{(c+h)x - \tau}{h}\right).
\]
The chance constraint (1.13) becomes:
\[
H\left(\frac{(b-c)x + \tau}{c}\right) - H\left(\frac{(c+h)x - \tau}{h}\right) \geq 1 - \alpha.
\]
Even for small (but positive) values of \( \alpha \), it can be a significant relaxation of the corresponding worst case constraints (1.11).

### 1.2.3 Multistage Models

Suppose now that the company has a planning horizon of \( T \) periods. We model the demand as a random process \( D_t \) indexed by the time \( t = 1, \ldots, T \). At the beginning, at \( t = 1 \), there is (known) inventory level \( y_1 \). At each period \( t = 1, \ldots, T \) the company first observes the current inventory level \( y_t \) and then places an order to replenish the inventory level to \( x_t \). This results in order quantity \( x_t - y_t \) which clearly should be nonnegative, i.e., \( x_t \geq y_t \).

After the inventory is replenished, demand \( d_t \) is realized and hence the next inventory level, at the beginning of period \( t + 1 \), becomes \( y_{t+1} = x_t - d_t \). We allow backlogging and the inventory level \( y_t \) may become negative. The total cost incurred in period \( t \) is
\[
c_t(x_t - y_t) + b_t [d_t - x_t]_+ + h_t [x_t - d_t]_+,
\]
where \( c_t, b_t, h_t \) are the ordering, backorder penalty, and holding costs per unit, respectively, at time \( t \). We assume that \( b_t > c_t > 0 \) and \( h_t \geq 0 \), \( t = 1, \ldots, T \). The objective is to minimize the expected value of the total cost over the planning horizon. This can be written as the following optimization problem
\[
\begin{align*}
\min_{x_t \geq y_t} & \sum_{t=1}^{T} E\{c_t(x_t - y_t) + b_t [D_t - x_t]_+ + h_t [x_t - D_t]_+\} \\
\text{s.t.} & y_{t+1} = x_t - D_t, \ t = 1, \ldots, T - 1.
\end{align*}
\]

\footnote{As before, we denote by \( D_t \) a particular realization of the random variable \( D_t \).}
1.2. Inventory

For $T = 1$ problem (1.17) is essentially the same as the (static) problem (1.4) (the only difference is the assumption here of the initial inventory level $y_1$). However, for $T > 1$ the situation is more subtle. It is not even clear what is the exact meaning of the formulation (1.17). There are several equivalent ways to give precise meaning to the above problem. One possible way is to write equations describing the dynamics of the corresponding optimization process. That is what we discuss next.

Consider the demand process $D_t$, $t = 1, \ldots, T$. We denote by $D_{[t]} := (D_1, \ldots, D_t)$ the history of the demand process up to time $t$, and by $d_{[t]} := (d_1, \ldots, d_t)$ its particular realization. At each period (stage) $t$, our decision about the inventory level $x_t$ should depend only on information available at the time of the decision, i.e., on an observed realization $d_{[t-1]}$ of the demand process, and not on future observations. This principle is called the nonanticipativity constraint. We assume, however, that the probability distribution of the demand process is known. That is, the conditional probability distribution of $D_t$, given $D_{[t-1]} = d_{[t-1]}$, is assumed to be known.

At the last stage $t = T$, for observed inventory level $y_T$, we need to solve the problem:

$$
\text{Min} \ c_T(x_T - y_T) + \mathbb{E}\{b_T[D_T - x_T]_+ + h_T[x_T - D_T]_+ | D_{[T-1]} = d_{[T-1]}\}. \quad (1.18)
$$

The expectation in (1.18) is conditional on the realization $d_{[T-1]}$ of the demand process prior to the considered time $T$. The optimal value (and the set of optimal solutions) of problem (1.18) depends on $y_T$ and $d_{[T-1]}$, and is denoted $Q_T(y_T,d_{[T-1]})$. At stage $t = T - 1$ we solve the problem

$$
\text{Min} \ c_{T-1}(x_{T-1} - y_{T-1})
+ \mathbb{E}\{b_{T-1}[D_{T-1} - x_{T-1}]_+ + h_{T-1}[x_{T-1} - D_{T-1}]_+ + Q_T(x_{T-1} - D_{T-1}, D_{[T-2]} | D_{[T-2]} = d_{[T-2]}\}. \quad (1.19)
$$

Its optimal value is denoted $Q_{T-1}(y_{T-1},d_{[T-2]})$. Proceeding in this way backwards in time we write the following dynamic programming equations

$$
Q_t(y_t, d_{[t-1]}) = \min_{x_t \geq y_t} c_t(x_t - y_t) + \mathbb{E}\{b_t[D_t - x_t]_+ \nonumber 
+ h_t[x_t - D_t]_+ + Q_{t+1}(x_t - D_t, D_{[t]} | D_{[t]} = d_{[t-1]}\}, \quad (1.20)
$$

t = T - 1, \ldots, 2. Finally, at the first stage we need to solve problem

$$
\text{Min} \ c_1(x_1 - y_1) + \mathbb{E}\{b_1[D_1 - x_1]_+ + h_1[x_1 - D_1]_+ + Q_2(x_1 - D_1, D_1) \}. \quad (1.21)
$$

Let us take a closer look at the above decision process. We need to understand how the dynamic programming equations (1.19)–(1.21) could be solved and what is the meaning of the solutions. Starting with the last stage $t = T$, we need to calculate the value functions $Q_t(y_t, d_{[t-1]})$ going backwards in time. In the present case the value functions cannot be calculated in a closed form and should be approximated numerically. For a generally
distributed demand process this could be very difficult or even impossible. The situation simplifies dramatically if we assume that the random process \( D_t \) is stagewise independent, that is, if \( D_t \) is independent of \( D_{t-1} \), \( t = 2, \ldots, T \). Then the conditional expectations in equations (1.18)–(1.19) become the corresponding unconditional expectations. Consequently, the value functions \( Q_t(y_t) \) do not depend on demand realizations and become functions of the respective univariate variables \( y_t \) only. In that case by discretization of \( y_t \) and the (one-dimensional) distribution of \( D_t \), these value functions can be calculated in a recursive way.

Suppose now that somehow we can solve the dynamic programming equations (1.19)–(1.21). Let \( \bar{x}_t \) be a corresponding optimal solution, i.e., \( \bar{x}_T \) is an optimal solution of (1.18), \( \bar{x}_t \) is an optimal solution of the right hand side of (1.20) for \( t = T-1, \ldots, 2 \), and \( \bar{x}_1 \) is an optimal solution of (1.21). We see that \( \bar{x}_t \) is a function of \( y_t \) and \( d_{t-1} \), for \( t = 2, \ldots, T \), while the first stage (optimal) decision \( \bar{x}_1 \) is independent of the data. Under the assumption of stagewise independence, \( \bar{x}_t = \bar{x}_t(y_t) \) becomes a function of \( y_t \) alone. Note that \( y_t \), in itself, is a function of \( d_{t-1} = (d_1, \ldots, d_{t-1}) \) and decisions \( (x_1, \ldots, x_{t-1}) \). Therefore we may think about a sequence of possible decisions \( x_t = x_t(d_{t-1}), t = 1, \ldots, T \), as functions of realizations of the demand process available at the time of the decision (with the convention that \( x_1 \) is independent of the data). Such a sequence of decisions \( x_t(d_{t-1}) \) is called an implementable policy, or simply a policy. That is, an implementable policy is a rule which specifies our decisions, based on information available at the current stage, for any possible realization of the demand process. By definition, an implementable policy \( x_t = x_t(d_{t-1}) \) satisfies the nonanticipativity constraint. A policy is said to be feasible if it satisfies other constraints with probability one (w.p.1). In the present case a policy is feasible if \( x_t \geq y_t, t = 1, \ldots, T \), for almost every realization of the demand process.

We can now formulate the optimization problem (1.17) as the problem of minimizing of the expectation in (1.17) with respect to all implementable feasible policies. An optimal solution of such problem will give us an optimal policy. We have that a policy \( \bar{x}_t \) is optimal if it is given by optimal solutions of the respective dynamic programming equations. Note again that under the assumption of stagewise independence, an optimal policy \( \bar{x}_t = \bar{x}_t(y_t) \) is a function of \( y_t \) alone. Moreover, in that case it is possible to give the following characterization of the optimal policy. Let \( x^*_t \) be an (unconstrained) minimizer of

\[
 c_t x_t + E\{ b_t[D_t - x_t]^+ + h_t[x_t - D_t]^+ + Q_{t+1}(x_t - D_t) \}, \ t = T, \ldots, 1, \tag{1.22}
\]

with the convention that \( Q_{T+1}(\cdot) = 0 \). Since \( Q_{t+1}(\cdot) \) is nonnegative valued and \( c_t + h_t \geq 0 \), we have that the function in (1.22) tends to \(+\infty\) if \( x_t \to -\infty \). Similarly, as \( b_t > c_t \), it also tends to \(+\infty\) if \( x_t \to -\infty \). Moreover, this function is convex and continuous (as long as it is real valued) and hence attains its minimal value. Then by using convexity of the value functions it is not difficult to show that \( \bar{x}_t = \max\{y_t, x^*_t\} \) is an optimal policy. Such policy is called the basestock policy. A similar result holds without the assumption of stagewise independence, but then the critical values \( x^*_t \) depend on realizations of the demand process up to time \( t - 1 \).

As it was mentioned above, if the stagewise independence condition is satisfied, then each value function \( Q_t(y_t) \) is a function of the variable \( y_t \). In that case we can accurately represent \( Q_t(\cdot) \) by discretization, i.e., by specifying its values at a finite number of points on the real line. Consequently, the corresponding dynamic programming equations can
be accurately solved recursively going backwards in time. The situation starts to change dramatically with increase of the number of variables on which the value functions depend, like in the example discussed in the next section. The discretization approach may still work with several state variables, but it quickly becomes impractical when the dimension of the state vector increases. This is called the “curse of dimensionality.” As we shall see it later, stochastic programming approaches the problem in a different way, by exploring convexity of the underlying problem, and thus attempting to solve problems with a state vector of high dimension. This is achieved by means of discretization of the random process $D_t$ in a form of a scenario tree, which may also become prohibitively large.

1.3 Multi-Product Assembly

1.3.1 Two Stage Model

Consider a situation where a manufacturer produces $n$ products. There are in total $m$ different parts (or subassemblies) which have to be ordered from third party suppliers. A unit of product $i$ requires $a_{ij}$ units of part $j$, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Of course, $a_{ij}$ may be zero for some combinations of $i$ and $j$. The demand for the products is modeled as a random vector $D = (D_1, \ldots, D_n)$. Before the demand is known, the manufacturer may pre-order the parts from outside suppliers, at the cost of $c_j$ per unit of part $j$. After the demand $D$ is observed, the manufacturer may decide which portion of the demand is to be satisfied, so that the available numbers of parts are not exceeded. It costs additionally $l_i$ to satisfy a unit of demand for product $i$, and the unit selling price of this product is $q_i$. The unsatisfied demand is lost.

Suppose the numbers of parts ordered are equal to $x_j$, $j = 1, \ldots, m$. After the demand $D$ becomes known, we need to determine how much of each product to make. Let us denote the numbers of units produced by $z_i$, $i = 1, \ldots, n$, and the numbers of parts left in inventory by $y_j$, $j = 1, \ldots, m$. For an observed value (a realization) $d = (d_1, \ldots, d_n)$ of the random demand vector $D$, we can find the best production plan by solving the following linear programming problem

$$\begin{align*}
\text{Minimize} & \quad (l - q) z - s^T y \\
\text{subject to} & \quad y = x - A^T z, \\
& \quad 0 \leq z_i \leq d_i, \quad i = 1, \ldots, n, \\
& \quad y_j \geq 0, \quad j = 1, \ldots, m.
\end{align*}$$

Introducing the matrix $A$ with entries $a_{ij}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, m$, we can write this problem compactly as follows:

$$\begin{align*}
\text{Minimize} & \quad (l - q)^T z - s^T y \\
\text{subject to} & \quad y = x - A^T z, \\
& \quad 0 \leq z \leq d, \quad y \geq 0.
\end{align*}$$

(1.23)
Observe that the solution of this problem, that is, the vectors $z$ and $y$, depend on realization $d$ of the demand vector $D$ as well as on $x$. Let $Q(x, d)$ denote the optimal value of problem (1.23). The quantities $x_j$ of parts to be ordered can be determined from the following optimization problem

$$
\min_{x \geq 0} c^T x + E[Q(x, D)],
$$

(1.24)

where the expectation is taken with respect to the probability distribution of the random demand vector $D$. The first part of the objective function represents the ordering cost, while the second part represents the expected cost of the optimal production plan, given ordered quantities $x$. Clearly, for realistic data with $q_i > l_i$, the second part will be negative, so that some profit will be expected.

Problem (1.23)–(1.24) is an example of a two stage stochastic programming problem, with (1.23) called the second stage problem, and (1.24) called the first stage problem. As the second stage problem contains random data (random demand $D$), its optimal value $Q(x, D)$ is a random variable. The distribution of this random variable depends on the first stage decisions $x$, and therefore the first stage problem cannot be solved without understanding of the properties of the second stage problem.

In the special case of finitely many demand scenarios $d_1, \ldots, d_K$ occurring with positive probabilities $p_1, \ldots, p_K$, with $\sum_{k=1}^K p_k = 1$, the two stage problem (1.23)–(1.24) can be written as one large scale linear programming problem:

$$
\begin{align*}
\min & \quad c^T x + \sum_{k=1}^K p_k \left[ (l - q)^T z^k - s^T y^k \right] \\
\text{s.t.} & \quad y^k = x - A^T z^k, \quad k = 1, \ldots, K, \\
& \quad 0 \leq z^k \leq d^k, \quad y^k \geq 0, \quad k = 1, \ldots, K, \\
& \quad x \geq 0,
\end{align*}
$$

(1.25)

where the minimization is performed over vector variables $x$ and $z^k, y^k, k = 1, \ldots, K$. We have integrated the second stage problem (1.23) into this formulation, but we had to allow for its solution $(z^k, y^k)$ to depend on the scenario $k$, because the demand realization $d^k$ is different in each scenario. Because of that, problem (1.25) has the numbers of variables and constraints roughly proportional to the number of scenarios $K$.

It is worthwhile to notice the following. There are three types of decision variables here. Namely, the numbers of ordered parts (vector $x$), the numbers of produced units (vector $z$) and the numbers of parts left in the inventory (vector $y$). These decision variables are naturally classified as the first and the second stage decision variables. That is, the first stage decisions $x$ should be made before a realization of the random data becomes available and hence should be independent of the random data, while the second stage decision variables $z$ and $y$ are made after observing the random data and are functions of the data. The first stage decision variables are often referred to as “here-and-now” decisions (solution), and second stage decisions are referred to as “wait-and-see” decisions (solution). It can also be noticed that the second stage problem (1.23) is feasible for every possible realization of the random data; for example, take $z = 0$ and $y = x$. In such a situation we say that the problem has relatively complete recourse.
1.3.2 Chance Constrained Model

Suppose now that the manufacturer is concerned with the possibility of losing demand. The manufacturer would like the probability that all demand be satisfied to be larger than some fixed service level \(1 - \alpha\), where \(\alpha \in (0, 1)\) is small. In this case the problem changes in a significant way.

Observe that if we want to satisfy demand \(D = (D_1, \ldots, D_n)\), we need to have \(x \geq A^T D\). If we have the parts needed, there is no need for the production planning stage, as in problem (1.23). We simply produce \(z_i = D_i, \ i = 1, \ldots, n\), whenever it is feasible.

Also, the production costs and salvage values do not affect our problem. Consequently, the requirement of satisfying the demand with probability at least \(1 - \alpha\) leads to the following formulation of the corresponding problem

\[
\min \begin{cases}
\mathbf{c}^T \mathbf{x} \\
\mathbf{x} \geq 0
\end{cases}
\text{s.t.} \quad \Pr\{\mathbf{A}^T \mathbf{D} \leq \mathbf{x}\} \geq 1 - \alpha. \tag{1.26}
\]

The chance (also called probabilistic) constraint in the above model is more difficult than in the case of the newsvendor model considered in section 1.2.2, because it involves a random vector \(W = \mathbf{A}^T \mathbf{D}\) rather than a univariate random variable.

Owing to the separable nature of the chance constraint in (1.26), we can rewrite this constraint in the following way:

\[
H_W(\mathbf{x}) \geq 1 - \alpha, \tag{1.27}
\]

where \(H_W(x) := \Pr(W \leq x)\) is the cumulative distribution function of the \(n\)-dimensional random vector \(W = \mathbf{A}^T \mathbf{D}\). Observe that if \(n = 1\) and \(c > 0\), then an optimal solution \(\bar{x}\) of (1.27) is given by the left side \((1 - \alpha)\)-quantile of \(W\), that is, \(\bar{x} = H_W^{-1}(1 - \alpha)\). On the other hand, in the case of multidimensional vector \(W\) its distribution has many “smallest (left side) \((1 - \alpha)\)-quantiles,” and the choice of \(\bar{x}\) will depend on the relative proportions of the cost coefficients \(c_j\). It is also worth mentioning that even when the coordinates of the demand vector \(D\) are independent, the coordinates of the vector \(W\) can be dependent, and thus the chance constraint of (1.27) cannot be replaced by a simpler expression featuring one dimensional marginal distributions.

The feasible set

\[
\{x \in \mathbb{R}_+^n : \Pr(\mathbf{A}^T \mathbf{D} \leq x) \geq 1 - \alpha\}
\]

of problem (1.26) can be written in the following equivalent form

\[
\{x \in \mathbb{R}_+^n : \mathbf{A}^T d \leq x, \ d \in \mathcal{D}, \ \Pr(\mathcal{D}) \geq 1 - \alpha\}. \tag{1.28}
\]

In the above formulation (1.28) the set \(\mathcal{D}\) can be any measurable subset of \(\mathbb{R}^n\) such that probability of \(D \in \mathcal{D}\) is at least \(1 - \alpha\). A considerable simplification can be achieved by choosing a fixed set \(\mathcal{D}_\alpha\) in such a way that \(\Pr(\mathcal{D}_\alpha) \geq 1 - \alpha\). In that way we obtain a simplified version of problem (1.26):

\[
\min \begin{cases}
\mathbf{c}^T \mathbf{x} \\
\mathbf{x} \geq 0
\end{cases}
\text{s.t.} \quad \mathbf{A}^T d \leq x, \ \forall \ d \in \mathcal{D}_\alpha. \tag{1.29}
\]
Chapter 1. Stochastic Programming Models

The set $\mathcal{D}_a$ in this formulation is sometimes referred to as the uncertainty set, and the whole formulation as the robust optimization problem. Observe that in our case we can solve this problem in the following way. For each part type $j$ we determine $x_j$ to be the minimum number of units necessary to satisfy every demand $d \in \mathcal{D}_a$, that is

$$x_j = \max_{d \in \mathcal{D}_a} \sum_{i=1}^{n} a_{ij}d_i, \quad j = 1, \ldots, n.$$ 

In this case the solution is completely determined by the uncertainty set $\mathcal{D}_a$ and it does not depend on the cost coefficients $c_j$.

The choice of the uncertainty set, satisfying the corresponding chance constraint, is not unique and often is governed by computational convenience. In this book we shall be mainly concerned with stochastic models, and we shall not discuss models and methods of robust optimization.

1.3.3 Multistage Model

Consider now the situation when the manufacturer has a planning horizon of $T$ periods. The demand is modeled as a stochastic process $D_t$, $t = 1, \ldots, T$, where each $D_t = (D_{t1}, \ldots, D_{tn})$ is a random vector of demands for the products. The unused parts can be stored from one period to the next, and holding one unit of part $j$ in inventory costs $h_j$. For simplicity, we assume that all costs and prices are the same in all periods.

It would not be reasonable to plan specific order quantities for the entire planning horizon $T$. Instead, one has to make orders and production decisions at successive stages, depending on the information available at the current stage. We use symbol $D_{[t]} := (D_{1}, \ldots, D_{t})$ to denote the history of the demand process in periods $1, \ldots, t$. In every multistage decision problem it is very important to specify which of the decision variables may depend on which part of the past information.

Let us denote by $x_{t-1} = (x_{t-1,1}, \ldots, x_{t-1,n})$ the vector of quantities ordered at the beginning of stage $t$, before the demand vector $D_t$ becomes known. The numbers of units produced in stage $t$, will be denoted by $z_t$, and the inventory level of parts at the end of stage $t$ by $y_t$, for $t = 1, \ldots, T$. We use the subscript $t-1$ for the order quantity to stress that it may depend on the past demand realizations $D_{[t-1]}$, but not on $D_t$, while the production and storage variables at stage $t$ may depend on $D_{[t]}$, which includes $D_t$. In the special case of $T = 1$, we have the two-stage problem discussed in section 1.3.1; the variable $x_0$ corresponds to the first stage decision vector $x$, while $z_1$ and $y_1$ correspond to the second stage decision vectors $z$ and $y$, respectively.

Suppose $T > 1$ and consider the last stage $t = T$, after the demand $D_T$ has been observed. At this time all inventory levels $y_{T-1}$ of the parts, as well as the last order quantities $x_{T-1}$ are known. The problem at stage $T$ is therefore identical to the second stage problem (1.23) of the two stage formulation:

$$\min_{z_T, y_T} (l - q)^T z_T - s^T y_T$$

s.t. $y_T = y_{T-1} + x_{T-1} - A^T z_T, \quad 0 \leq z_T \leq d_T, \quad y_T \geq 0,$

(1.30)
1.4. Portfolio Selection

1.4.1 Static Model

Suppose that we want to invest capital $W_0$ in $n$ assets, by investing an amount $x_i$ in asset $i$, for $i = 1, \ldots, n$. Suppose, further, that each asset has a respective return rate $R_i$ (per one period of time), which is unknown (uncertain) at the time we need to make our decision. We address now a question of how to distribute our wealth $W_0$ in an optimal way. The total wealth resulting from our investment after one period of time equals

$$ W_1 = \sum_{i=1}^{n} \xi_i x_i, $$
where $\xi_i := 1 + R_i$. We have here the balance constraint $\sum_{i=1}^n x_i \leq W_0$. Suppose, further, that one of possible investments is cash, so that we can write this balance condition as the equation $\sum_{i=1}^n x_i = W_0$. Viewing returns $R_i$ as random variables, one can try to maximize the expected return on an investment. This leads to the following optimization problem

$$\text{Max } x \geq 0 \quad E[W_1] \quad \text{subject to} \quad \sum_{i=1}^n x_i = W_0. \quad (1.34)$$

We have here that

$$E[W_1] = \sum_{i=1}^n E[\xi_i] x_i = \sum_{i=1}^n \mu_i x_i,$$

where $\mu_i := E[\xi_i] = 1 + E[R_i]$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Therefore, problem (1.34) has a simple optimal solution of investing everything into an asset with the largest expected return rate and has the optimal value of $\mu^* W_0$, where $\mu^* := \max_{1 \leq i \leq n} \mu_i$. Of course, from the practical point of view such solution is not very appealing. Putting everything into one asset can be very dangerous, because if its realized return rate will be bad, then one can lose much money.

An alternative approach is to maximize expected utility of the wealth represented by a concave nondecreasing function $U(W_1)$. This leads to the following optimization problem

$$\text{Max } x \geq 0 \quad E[U(W_1)] \quad \text{subject to} \quad \sum_{i=1}^n x_i = W_0. \quad (1.35)$$

This approach requires specification of the utility function. For instance, let $U(W)$ be defined as

$$U(W) := \begin{cases} 
(1 + q)(W - a), & \text{if } W \geq a, \\
(1 + r)(W - a), & \text{if } W \leq a,
\end{cases} \quad (1.36)$$

with $r > q > 0$ and $a > 0$. We can view the involved parameters as follows: $a$ is the amount that we have to pay after return on the investment, $q$ is the interest rate at which we can invest the additional wealth $W - a$, provided that $W > a$, and $r$ is the interest rate at which we will have to borrow if $W$ is less than $a$. For the above utility function, problem (1.35) can be formulated as the following two stage stochastic linear program

$$\text{Max } x \geq 0 \quad E[Q(x, \xi)] \quad \text{subject to} \quad \sum_{i=1}^n x_i = W_0, \quad (1.37)$$

where $Q(x, \xi)$ is the optimal value of the problem

$$\text{Max } y, z \in \mathbb{R}_+ \quad (1 + q)y - (1 + r)z \quad \text{subject to} \quad \sum_{i=1}^n \xi_i x_i = a + y - z. \quad (1.38)$$

We can view the above problem (1.38) as the second stage program. Given a realization $\xi = (\xi_1, \ldots, \xi_n)$ of random data, we make an optimal decision by solving the corresponding optimization problem. Of course, in the present case the optimal value $Q(x, \xi)$ is a function of $W_1 = \sum_{i=1}^n \xi_i x_i$ and can be written explicitly as $U(W_1)$. 

14 Chapter 1. Stochastic Programming Models
1.4. Portfolio Selection

Yet another possible approach is to maximize the expected return while controlling the involved risk of the investment. There are several ways how the concept of risk could be formalized. For instance, we can evaluate risk by variability of $W$ measured by its variance $\text{Var}[W] = \text{E}[W^2] - \left(\text{E}[W]\right)^2$. Since $W_1$ is a linear function of the random variables $\xi_i$, we have that

$$\text{Var}[W_1] = x^T \Sigma x = \sum_{i,j=1}^{n} \sigma_{ij} x_i x_j,$$

where $\Sigma = [\sigma_{ij}]$ is the covariance matrix of the random vector $\xi$ (note that the covariance matrices of the random vectors $\xi = (\xi_1, \ldots, \xi_n)$ and $R = (R_1, \ldots, R_n)$ are identical). This leads to the optimization problem of maximizing the expected return subject to the additional constraint $\text{Var}[W_1] \leq \nu$, where $\nu > 0$ is a specified constant. This problem can be written as

$$\text{Max} \quad x \geq 0 \quad \sum_{i=1}^{n} \mu_i x_i \text{ subject to } \sum_{i=1}^{n} x_i = W_0, \quad x^T \Sigma x \leq \nu. \quad (1.39)$$

Since the covariance matrix $\Sigma$ is positive semidefinite, the constraint $x^T \Sigma x \leq \nu$ is convex quadratic, and hence (1.39) is a convex problem. Note that problem (1.39) has at least one feasible solution of investing everything in cash, in which case $\text{Var}[W_1] = 0$, and since its feasible set is compact, the problem has an optimal solution. Moreover, since problem (1.39) is convex and satisfies the Slater condition, there is no duality gap between this problem and its dual:

$$\text{Min} \quad \lambda \geq 0 \quad \text{Max} \quad \sum_{i=1}^{n} \mu_i x_i \text{ subject to } \sum_{i=1}^{n} x_i = W_0, \quad x^T \Sigma x \leq \nu \quad \lambda \left( x^T \Sigma x - \nu \right). \quad (1.40)$$

Consequently, there exists Lagrange multiplier $\bar{\lambda} \geq 0$ such that problem (1.39) is equivalent to the problem

$$\text{Max} \quad x \geq 0 \quad \sum_{i=1}^{n} \mu_i x_i - \bar{\lambda} x^T \Sigma x \text{ subject to } \sum_{i=1}^{n} x_i = W_0. \quad (1.41)$$

The equivalence here means that the optimal value of problem (1.39) is equal to the optimal value of problem (1.41) plus the constant $\lambda \nu$, and that any optimal solution of problem (1.39) is also an optimal solution of problem (1.41). In particular, if problem (1.41) has unique optimal solution $\bar{x}$, then $\bar{x}$ is also the optimal solution of problem (1.39). The corresponding Lagrange multiplier $\bar{\lambda}$ is given by an optimal solution of the dual problem (1.40). We can view the objective function of the above problem as a compromise between the expected return and its variability measured by its variance.

Another possible formulation is to minimize $\text{Var}[W_1]$ keeping the expected return $\text{E}[W_1]$ above a specified value $\tau$. That is,

$$\text{Min} \quad x^T \Sigma x \text{ subject to } \sum_{i=1}^{n} x_i = W_0, \quad \sum_{i=1}^{n} \mu_i x_i \geq \tau. \quad (1.42)$$

For appropriately chosen constants $\nu$, $\lambda$ and $\tau$, problems (1.39)–(1.42) are equivalent to each other. Problems (1.41) and (1.42) are quadratic programming problems, while problem (1.39) can be formulated as a conic quadratic problem. These optimization problems
can be efficiently solved. Note finally that these optimization problems are based on the first and second order moments of random data $\xi$ and do not require complete knowledge of the probability distribution of $\xi$.

We can also approach risk control by imposing chance constraints. Consider the problem

$$\begin{align*}
\text{Max} & \; x_i \geq 0 \; \text{subject to} \; \sum_{i=1}^{n} x_i = W_0, \; \Pr \left\{ \sum_{i=1}^{n} \xi_i x_i \geq b \right\} \geq 1 - \alpha.
\end{align*}$$

That is, we impose the constraint that with probability at least $1 - \alpha$ our wealth $W_1 = \sum_{i=1}^{n} \xi_i x_i$ should not fall below a chosen amount $b$. Suppose the random vector $\xi$ has a multivariate normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$, written $\xi \sim N(\mu, \Sigma)$. Then $W_1$ has normal distribution with mean $\sum_{i=1}^{n} \mu_i x_i$ and variance $x^T \Sigma x$, and

$$\Pr \{ W_1 \geq b \} = \Phi \left( \frac{b - \sum_{i=1}^{n} \mu_i x_i}{\sqrt{x^T \Sigma x}} \right),$$

where $Z \sim N(0, 1)$ has the standard normal distribution and $\Phi(z) = \Pr(Z \leq z)$ is the cdf of $Z$.

Therefore we can write the chance constraint of problem (1.43) in the form

$$b - \sum_{i=1}^{n} \mu_i x_i + z_{\alpha} \sqrt{x^T \Sigma x} \leq 0,$$

where $z_{\alpha} := \Phi^{-1}(1 - \alpha)$ is the $(1 - \alpha)$-quantile of the standard normal distribution. Note that since matrix $\Sigma$ is positive semidefinite, $\sqrt{x^T \Sigma x}$ defines a seminorm on $\mathbb{R}^n$ and is a convex function. Consequently, if $0 < \alpha \leq 1/2$, then $z_{\alpha} \geq 0$ and the constraint (1.45) is convex. Therefore, provided that problem (1.43) is feasible, there exists a Lagrange multiplier $\gamma \geq 0$ such that problem (1.43) is equivalent to the problem

$$\begin{align*}
\text{Max} & \; x_i \geq 0 \; \text{subject to} \; \sum_{i=1}^{n} x_i = W_0, \\
\text{Max} & \; \sum_{i=1}^{n} \mu_i x_i - \eta \sqrt{x^T \Sigma x} \text{ subject to } \sum_{i=1}^{n} x_i = W_0,
\end{align*}$$

where $\eta = \gamma z_{\alpha}/(1 + \gamma)$.

In financial engineering the (left side) $(1 - \alpha)$-quantile of a random variable $Y$ (representing losses) is called Value-at-Risk, i.e.,

$$\text{VaR}_\alpha(Y) := H^{-1}(1 - \alpha),$$

where $H(\cdot)$ is the cdf of $Y$. The chance constraint of problem (1.43) can be written in the form of a Value at Risk constraint

$$\text{VaR}_\alpha \left( b - \sum_{i=1}^{n} \xi_i x_i \right) \leq 0.$$

Note that if $x^T \Sigma x = 0$, i.e., $\text{Var}(W_1) = 0$, then the chance constraint of problem (1.43) holds iff $\sum_{i=1}^{n} \mu_i x_i \geq b$. In that case equivalence to the constraint (1.45) obviously holds.
1.4. Portfolio Selection

It was possible to write chance (Value-at-Risk) constraint here in a closed form because of the assumption of joint normal distribution. Note that in the present case the random variables $\xi_i$ cannot be negative, which indicates that the assumption of normal distribution is not very realistic.

1.4.2 Multistage Portfolio Selection

Suppose we are allowed to rebalance our portfolio in time periods $t = 1, \ldots, T - 1$, but without injecting additional cash into it. At each period $t$ we need to make a decision about distribution of our current wealth $W_t$ among $n$ assets. Let $x_0 = (x_{10}, \ldots, x_{n0})$ be initial amounts invested in the assets. Recall that each $x_{i0}$ is nonnegative and that the balance equation $\sum_{i=1}^{n} x_{i0} = W_0$ should hold.

We assume now that respective return rates $R_{1t}, \ldots, R_{nt}$, at periods $t = 1, \ldots, T$, form a random process with a known distribution. Actually we will work with the (vector valued) random process $\xi_1, \ldots, \xi_T$, where $\xi_t = (\xi_{1t}, \ldots, \xi_{nt})$ and $\xi_{it} := 1 + R_{it}, i = 1, \ldots, n, t = 1, \ldots, T$. At time period $t = 1$ we can rebalance the portfolio by specifying the amounts $x_1 = (x_{11}, \ldots, x_{n1})$ invested in the respective assets. At that time we already know the actual returns in the first period, so it is reasonable to use this information in the rebalancing decisions. Thus, our second stage decisions, at time $t = 1$, are actually functions of realizations of the random data vector $\xi_1$, i.e., $x_1 = x_1(\xi_1)$. And so on, at time $t$ our decision $x_t = (x_{1t}, \ldots, x_{nt})$ is a function $x_t = x_t(\xi_{[t]})$ of the available information given by realization $\xi_{[t]} = (\xi_1, \ldots, \xi_t)$ of the data process up to time $t$. A sequence of specific functions $x_{it} = x_{it}(\xi_{[t]}), t = 0, 1, \ldots, T - 1, \text{ with } x_0$ being constant, defines an implementable policy of the decision process. It is said that such policy is feasible if it satisfies w.p.1 the model constraints, i.e., the nonnegativity constraints $x_{it}(\xi_{[t]}) \geq 0, i = 1, \ldots, n, t = 0, \ldots, T - 1$, and the balance of wealth constraints

$$\sum_{i=1}^{n} x_{it}(\xi_{[t]}) = W_t.$$ 

At period $t = 1, \ldots, T$, our wealth $W_t$ depends on the realization of the random data process and our decisions up to time $t$, and is equal to

$$W_t = \sum_{i=1}^{n} \xi_{it} x_{i,t-1}(\xi_{[t-1]}).$$

Suppose our objective is to maximize the expected utility of this wealth at the last period, that is, we consider the problem

$$\text{Max } \mathbb{E}[U(W_T)].$$

It is a multistage stochastic programming problem, where stages are numbered from $t = 0$ to $t = T - 1$. Optimization is performed over all implementable and feasible policies.

Of course, in order to complete the description of the problem, we need to define the probability distribution of the random process $R_1, \ldots, R_T$. This can be done in many different ways. For example, one can construct a particular scenario tree defining time
Chapter 1. Stochastic Programming Models

evolution of the process. If at every stage the random return of each asset is allowed to have just two continuations, independently of other assets, then the total number of scenarios is \(2^{nT}\). It also should be ensured that \(1 + R_{it} \geq 0\), \(i = 1, \ldots, n, t = 1, \ldots, T\), for all possible realizations of the random data.

In order to write dynamic programming equations let us consider the above multi-stage problem backwards in time. At the last stage \(t = T - 1\) a realization \(\xi[T-1] = (\xi_1, \ldots, \xi_{T-1})\) of the random process is known and \(x_{T-2}\) has been chosen. Therefore, we have to solve the problem

\[
\max_{x_{T-1} \geq 0, W_T} \mathbb{E}\{U[W_T]|\xi[T-1]\}
\]

s.t. \(W_T = \sum_{i=1}^{n} \xi_i x_{i,T-1}, \sum_{i=1}^{n} x_{i,T-1} = W_{T-1},\) \hspace{1cm} (1.50)

where \(\mathbb{E}\{U[W_T]|\xi[T-1]\}\) denotes the conditional expectation of \(U[W_T]\) given \(\xi[T-1]\). The optimal value of the above problem (1.50) depends on \(W_{T-1}\) and \(\xi[T-1]\), and is denoted \(Q_{T-1}(W_{T-1}, \xi[T-1])\).

Continuing in this way, at stage \(t = T - 2, \ldots, 1\), we consider the problem

\[
\max_{x_t \geq 0, W_{t+1}} \mathbb{E}\{Q_{t+1}(W_{t+1}, \xi[t+1])|\xi[t]\}\)

s.t. \(W_{t+1} = \sum_{i=1}^{n} \xi_{i,t} x_{i,t}, \sum_{i=1}^{n} x_{i,t} = W_t,\) \hspace{1cm} (1.51)

whose optimal value is denoted \(Q_t(W_t, \xi[t])\). Finally, at stage \(t = 0\) we solve the problem

\[
\max_{x_0 \geq 0, W_1} \mathbb{E}[Q_1(W_1, \xi_1)]
\]

s.t. \(W_1 = \sum_{i=1}^{n} \xi_{i1} x_{i0}, \sum_{i=1}^{n} x_{i0} = W_0.\) \hspace{1cm} (1.52)

For a general distribution of the data process \(\xi_t\) it may be hard to solve these dynamic programming equations. The situation simplifies dramatically if the process \(\xi_t\) is stagewise independent, i.e., \(\xi_t\) is (stochastically) independent of \(\xi_1, \ldots, \xi_{t-1}\), for \(t = 2, \ldots, T\). Of course, the assumption of stagewise independence is not very realistic in financial models, but it is instructive to see the dramatic simplifications it allows. In that case the corresponding conditional expectations become unconditional expectations and the cost-to-go (value) function \(Q_t(W_t), t = 1, \ldots, T - 1,\) does not depend on \(\xi[t]\). That is, \(Q_{T-1}(W_{T-1})\) is the optimal value of the problem

\[
\max_{x_{T-1} \geq 0, W_T} \mathbb{E}\{U[W_T]\}
\]

s.t. \(W_T = \sum_{i=1}^{n} \xi_i x_{i,T-1}, \sum_{i=1}^{n} x_{i,T-1} = W_{T-1},\)
1.4. Portfolio Selection

and $Q_t(W_t)$ is the optimal value of

$$\begin{align*}
\max_{x_i \geq 0} \quad & \mathbb{E}\{Q_{t+1}(W_{t+1})\} \\
\text{s.t.} \quad & W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1}x_{i,t}, \quad \sum_{i=1}^{n} x_{i,t} = W_t,
\end{align*}$$

for $t = T - 2, \ldots, 1$.

The other relevant question is what utility function to use. Let us consider the logarithmic utility function $U(W) := \ln W$. Note that this utility function is defined for $W > 0$. For positive numbers $a$ and $w$ and for $W_{T-1} = w$ and $W_{T-1} = aw$, there is a one-to-one correspondence $x_{T-1} \leftrightarrow ax_{T-1}$ between the feasible sets of the corresponding problem (1.50). For the logarithmic utility function this implies the following relation between the optimal values of these problems

$$Q_{T-1}(aw, \xi_{[T-1]}) = Q_{T-1}(w, \xi_{[T-1]}) + \ln a. \quad (1.53)$$

That is, at stage $t = T - 1$ we solve problem

$$\begin{align*}
\max_{x_{T-1} \geq 0} \quad & \mathbb{E}\left\{ \ln \left( \sum_{i=1}^{n} \xi_{i,T}x_{i,T-1} \right) \right\} \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_{i,T-1} = W_{T-1}.
\end{align*} \quad (1.54)$$

By (1.53) its optimal value is

$$Q_{T-1} \left( W_{T-1}, \xi_{[T-1]} \right) = \nu_{T-1} \left( \xi_{[T-1]} \right) + \ln W_{T-1},$$

where $\nu_{T-1} \left( \xi_{[T-1]} \right)$ denotes the optimal value of (1.54) for $W_{T-1} = 1$. At stage $t = T - 2$ we solve problem

$$\begin{align*}
\max_{x_{T-2} \geq 0} \quad & \mathbb{E} \left\{ \nu_{T-1} \left( \xi_{[T-1]} \right) + \ln \left( \sum_{i=1}^{n} \xi_{i,T-1}x_{i,T-2} \right) \right\} \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_{i,T-2} = W_{T-2}.
\end{align*} \quad (1.55)$$

Of course, we have that

$$\mathbb{E} \left\{ \nu_{T-1} \left( \xi_{[T-1]} \right) + \ln \left( \sum_{i=1}^{n} \xi_{i,T-1}x_{i,T-2} \right) \right\} = \mathbb{E} \left\{ \nu_{T-1} \left( \xi_{[T-2]} \right) \right\} + \mathbb{E} \left\{ \ln \left( \sum_{i=1}^{n} \xi_{i,T-1}x_{i,T-2} \right) \right\},$$

and hence by arguments similar to (1.53), the optimal value of (1.55) can be written as

$$Q_{T-2} \left( W_{T-2}, \xi_{[T-2]} \right) = \mathbb{E} \left\{ \nu_{T-1} \left( \xi_{[T-1]} \right) \right\} + \nu_{T-2} \left( \xi_{[T-2]} \right) + \ln W_{T-2},$$
where \( \nu_{T-2}(\xi_{[T-2]}) \) is the optimal value of the problem

\[
\begin{align*}
\text{Max}_{x_{T-2} \geq 0} & \quad \mathbb{E} \left\{ \ln \left( \sum_{i=1}^{n} \xi_{i,T-1}x_{i,T-2} \right) \mid \xi_{[T-2]} \right\} \\
& \text{s.t.} \quad \sum_{i=1}^{n} x_{i,T-2} = 1.
\end{align*}
\]

An identical argument applies at earlier stages. Therefore, it suffices to solve at each stage \( t = T-1, \ldots, 1, 0 \), the corresponding optimization problem

\[
\begin{align*}
\text{Max}_{x_t \geq 0} & \quad \mathbb{E} \left\{ \ln \left( \sum_{i=1}^{n} \xi_{i,t+1}x_{i,t} \right) \mid \xi_{[t]} \right\} \\
& \text{s.t.} \quad \sum_{i=1}^{n} x_{i,t} = W_t,
\end{align*}
\]

(1.56)
in a completely myopic fashion.

By definition, we set \( \xi_0 \) to be constant, so that for the first stage problem, at \( t = 0 \), the corresponding expectation is unconditional. An optimal solution \( \bar{x}_0 = \bar{x}_0(W_0, \xi_{[0]}) \) of problem (1.56) gives an optimal policy. In particular, first stage optimal solution \( \bar{x}_0 \) is given by an optimal solution of the problem:

\[
\begin{align*}
\text{Max}_{x_0 \geq 0} & \quad \mathbb{E} \left\{ \ln \left( \sum_{i=1}^{n} \xi_{1}x_{i,0} \right) \right\} \\
& \text{s.t.} \quad \sum_{i=1}^{n} x_{i,0} = W_0.
\end{align*}
\]

(1.57)

We also have here that the optimal value, denoted \( \vartheta^* \), of the optimization problem (1.49) can be written as follows

\[
\vartheta^* = \ln W_0 + \nu_0 + \sum_{t=1}^{T-1} \mathbb{E} \left[ \nu_t(\xi_{[t]}) \right],
\]

(1.58)

where \( \nu_t(\xi_{[t]}) \) is the optimal value of problem (1.56) for \( W_t = 1 \). Note that \( \nu_0 + \ln W_0 \) is the optimal value of problem (1.57) with \( \nu_0 \) being the (deterministic) optimal value of (1.57) for \( W_0 = 1 \).

If the random process \( \xi_t \) is stagewise independent, then conditional expectations in (1.56) are the same as the corresponding unconditional expectations, and hence optimal values \( \nu_t(\xi_{[t]}) = \nu_t \) do not depend on \( \xi_{[t]} \) and are given by the optimal value of the problem

\[
\begin{align*}
\text{Max}_{x_t \geq 0} & \quad \mathbb{E} \left\{ \ln \left( \sum_{i=1}^{n} \xi_{i,t+1}x_{i,t} \right) \right\} \\
& \text{s.t.} \quad \sum_{i=1}^{n} x_{i,t} = 1.
\end{align*}
\]

(1.59)

Also in the stagewise independent case the optimal policy can be described as follows. Let \( x^*_t = (x^*_{1,t}, \ldots, x^*_{n,t}) \) be the optimal solution of (1.59), \( t = 0, \ldots, T-1 \). Such optimal solution is unique by strict concavity of the logarithm function. Then

\[
\bar{x}_t(W_t) := W_t x^*_t, \quad t = 0, \ldots, T-1,
\]

defines the optimal policy.

Consider now the power utility function \( U(W) := W^\gamma \), with \( 1 \geq \gamma > 0 \), defined for \( W \geq 0 \). Suppose again that the random process \( \xi_t \) is stagewise independent. Recall that this condition implies that the cost-to-go function \( Q_t(W_t), t = 1, \ldots, T-1 \), depends only
of Portfolio Selection

1.4. Portfolio Selection

on $W_t$. By using arguments similar to the analysis for the logarithmic utility function it is not difficult to show that $Q_{T-1}(W_{T-1}) = W_{T-1}^T Q_{T-1}(1)$, and so on. The optimal policy $\bar{x}_t = \bar{x}_t(W_t)$ is obtained in a myopic way as an optimal solution of the problem

$$\max_{x_t \geq 0} \mathbb{E} \left\{ \left( \sum_{i=1}^n \xi_{i,t+1} x_{it} \right)^\gamma \right\} \text{ s.t. } \sum_{i=1}^n x_{it} = W_t. \quad (1.60)$$

That is, $\bar{x}_t(W_t) = W_t x^*_t$, where $x^*_t$ is an optimal solution of problem (1.60) for $W_t = 1$, $t = 0, ..., T - 1$. In particular, the first stage optimal solution $\bar{x}_0$ is obtained in a myopic way by solving the problem

$$\max_{x_0 \geq 0} \mathbb{E} \left\{ \left( \sum_{i=1}^n \xi_{i,1} x_{i0} \right)^\gamma \right\} \text{ s.t. } \sum_{i=1}^n x_{i0} = W_0. \quad (1.61)$$

The optimal value $\vartheta^*$ of the corresponding multistage problem (1.49) is

$$\vartheta^* = W_0^\gamma \prod_{t=0}^{T-1} \eta_t, \quad (1.61)$$

where $\eta_t$ is the optimal value of problem (1.60) for $W_t = 1$.

The above myopic behavior of multistage stochastic programs is rather exceptional. A more realistic situation occurs in the presence of transaction costs. These are costs associated with the changes in the numbers of units (stocks, bonds) held. Introduction of transaction costs will destroy such myopic behavior of optimal policies.

### 1.4.3 Decision Rules

Consider the following policy. Let $x^*_t = (x^*_1, ..., x^*_n)$, $t = 0, ..., T - 1$, be vectors such that $x^*_t \geq 0$ and $\sum_{i=1}^n x^*_it = 1$. Define the fixed mix policy

$$x_t(W_t) := W_t x^*_t, \quad t = 0, ..., T - 1. \quad (1.62)$$

As it was discussed above, under the assumption of stagewise independence, such policies are optimal for the logarithmic and power utility functions provided that $x^*_t$ are optimal solutions of the respective problems (problem (1.59) for the logarithmic utility function and problem (1.60) with $W_t = 1$ for the power utility function). In other problems, a policy of form (1.62) may be non-optimal. However, it is readily implementable, once the current wealth $W_t$ is observed. As it was mentioned before, rules for calculating decisions as functions of the observations gathered up to time $t$, similarly to (1.62), are called policies, or alternatively decision rules.

We analyze now properties of the decision rule (1.62) under the simplifying assumption of stagewise independence. We have

$$W_{t+1} = \sum_{i=1}^n \xi_{i,t+1} x_{it}(W_t) = W_t \sum_{i=1}^n \xi_{i,t+1} x^*_it. \quad (1.63)$$
Since the random process $\xi_1, \ldots, \xi_T$ is stagewise independent, by independence of $\xi_{t+1}$ and $W_t$ we have

$$E[W_{t+1}] = E[W_t]E\left(\sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}^*\right) = E[W_t] \sum_{i=1}^{n} \mu_{i,t+1} x_{i,t}^*, \quad (1.64)$$

where $\mu_t := E[\xi_t]$. Consequently, by induction,

$$E[W_t] = \prod_{\tau=1}^{t} \left(\sum_{i=1}^{n} \mu_{i,\tau} x_{i,\tau-1}^*\right) = \prod_{\tau=1}^{t} (x_{\tau-1}^{T \mu_{\tau}}).$$

In order to calculate the variance of $W_t$, we use the formula

$$\text{Var}(Y) = \frac{E[(Y - E(Y|X))^2|X] + E[(E(Y|X) - EY)^2]}{\text{Var}(E(Y|X))}, \quad (1.65)$$

where $X$ and $Y$ are random variables. Applying (1.65) to (1.63) with $Y := W_{t+1}$ and $X := W_t$ we obtain

$$\text{Var}[W_{t+1}] = E[W_t^2]\text{Var}\left(\sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}^*\right) + \text{Var}[W_t] \left(\sum_{i=1}^{n} \mu_{i,t+1} x_{i,t}^*\right)^2. \quad (1.66)$$

Recall that $E[W_t^2] = \text{Var}[W_t] + (E[W_t])^2$ and $\text{Var}\left(\sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}^*\right) = x_{t+1}^T \Sigma_{t+1} x_{t+1}^*$, where $\Sigma_{t+1}$ is the covariance matrix of $\xi_{t+1}$.

It follows from (1.64) and (1.66) that

$$\frac{\text{Var}[W_{t+1}]}{(E[W_{t+1}])^2} = \frac{x_{t+1}^{T \Sigma_{t+1} x_{t+1}^*}}{(x_{t+1}^{T \mu_{t+1}})^2} + \frac{\text{Var}[W_t]}{(E[W_t])^2}, \quad (1.67)$$

and hence

$$\frac{\text{Var}[W_t]}{(E[W_t])^2} = \sum_{\tau=1}^{t} \frac{\text{Var}\left(\sum_{i=1}^{n} \xi_{i,\tau} x_{i,\tau-1}^*\right)}{(\sum_{i=1}^{n} \mu_{i,\tau} x_{i,\tau-1}^*)^2} = \sum_{\tau=1}^{t} \frac{x_{\tau-1}^{T \Sigma_{\tau-1} x_{\tau-1}^*}}{(x_{\tau-1}^{T \mu_{\tau}})^2}, \quad t = 1, \ldots, T. \quad (1.68)$$

This shows that if the terms $x_{\tau-1}^{T \Sigma_{\tau-1} x_{\tau-1}^*}$ are of the same order for $\tau = 1, \ldots, T$, then the ratio of the standard deviation $\sqrt{\text{Var}[W_T]}$ to the expected wealth $E[W_T]$ is of order $O(\sqrt{T})$ with increase of the number of stages $T$.

### 1.5 Supply Chain Network Design

In this section we discuss a stochastic programming approach to modeling a supply chain network design. A supply chain is a network of suppliers, manufacturing plants, warehouses, and distribution channels organized to acquire raw materials, convert these raw
1.5. Supply Chain Network Design

materials to finished products, and distribute these products to customers. Let us first describe a deterministic mathematical formulation for the supply chain design problem.

Denote by $S$, $P$ and $C$ the respective (finite) sets of suppliers, processing facilities and customers. The union $N := S \cup P \cup C$ of these sets is viewed as the set of nodes of a directed graph $(N, A)$, where $A$ is a set of arcs (directed links) connecting these nodes in a way representing flow of the products. The processing facilities include manufacturing centers $M$, finishing facilities $F$ and warehouses $W$, i.e., $P = M \cup F \cup W$. Further, a manufacturing center $i \in M$ or a finishing facility $i \in F$ consists of a set of manufacturing or finishing machines $H_i$. Thus the set $P$ includes the processing centers as well as the machines in these centers. Let $K$ be the set of products flowing through the supply chain.

The supply chain configuration decisions consist of deciding which of the processing centers to build (major configuration decisions), and which processing and finishing machines to procure (minor configuration decisions). We assign a binary variable $x_i$ to denote the investment cost for building facility $i$, or procuring machine $i$, otherwise. The operational decisions consist of routing the flow of product $k \in K$ from the supplier to the customers. By $y_{ij}^k$, we denote the flow of product $k$ from a node $i$ to a node $j$ of the network where $(i,j) \in A$. A deterministic mathematical model for the supply chain design problem can be written as follows

$$
\text{Min}_{x,y} \sum_{i \in P} c_i x_i + \sum_{k \in K} \sum_{(i,j) \in A} q_{ij}^k y_{ij}^k \tag{1.69}
$$

s.t. \hspace{1cm} \sum_{i \in N} y_{ij}^k - \sum_{\ell \in N} y_{j\ell}^k = 0, \quad j \in P, \quad k \in K, \tag{1.70}

\sum_{i \in N} y_{ij}^k \geq d_j^k, \quad j \in C, \quad k \in K, \tag{1.71}

\sum_{i \in N} y_{ij}^k \leq s_j^k, \quad j \in S, \quad k \in K, \tag{1.72}

\sum_{k \in K} r_j^k \left( \sum_{i \in N} y_{ij}^k \right) \leq m_j x_j, \quad j \in P, \tag{1.73}

x \in X, \quad y \geq 0. \tag{1.74}

Here $c_i$ denotes the investment cost for building facility $i$ or procuring machine $i$, $q_{ij}^k$ denotes the per-unit cost of processing product $k$ at facility $i$ and/or transporting product $k$ on arc $(i,j) \in A$. $d_j^k$ denotes the demand of product $k$ at node $j$, $s_j^k$ denotes the supply of product $k$ at node $j$, $r_j^k$ denotes per-unit processing requirement for product $k$ at node $j$, $m_j$ denotes capacity of facility $j$, $X \subseteq \{0,1\}^{|P|}$ is a set of binary variables and $y \in \mathbb{R}^{|A| \times |K|}$ is a vector with components $y_{ij}^k$. All cost components are annualized.

The objective function (1.69) is aimed at minimizing total investment and operational costs. Of course, a similar model can be constructed for maximizing the profits. The set $X$ represents logical dependencies and restrictions, such as $x_i \leq x_j$ for all $i \in H_j$ and $j \in P$ or $j \in F$, i.e., machine $i \in H_j$ should only be procured if facility $j$ is built (since $x_i$ are binary, the constraint $x_i \leq x_j$ means that $x_i = 0$ if $x_j = 0$). Constraints (1.70) enforce the flow conservation of product $k$ across each processing node $j$. Constraints (1.71) require that the total flow of product $k$ to a customer node $j$, should exceed the
Chapter 1. Stochastic Programming Models

demand $d^k_j$ at that node. Constraints (1.72) require that the total flow of product $k$ from a supplier node $j$, should be less than the supply $s^k_j$ at that node. Constraints (1.73) enforce capacity constraints of the processing nodes. The capacity constraints then require that the total processing requirement of all products flowing into a processing node $j$ should be smaller than the capacity $m_j$ of facility $j$ if it is built ($x_j = 1$). If facility $j$ is not built ($x_j = 0$) the constraint will force all flow variables $y^k_{ij} = 0$ for all $i \in \mathcal{N}$. Finally, constraint (1.74) enforces feasibility constraint $x \in \mathcal{X}$ and the nonnegativity of the flow variables corresponding to an arc $(ij) \in \mathcal{A}$ and product $k \in \mathcal{K}$.

It will be convenient to write problem (1.69)–(1.74) in the following compact form

\[
\begin{align*}
\text{Min}_x & \quad c^T x + q^T y \\
\text{s.t.} \quad & N y = 0, \\
& C y \geq d, \\
& S y \leq s, \\
& R y \leq M x,
\end{align*}
\]

(1.75)

where vectors $c$, $q$, $d$ and $s$ correspond to investment costs, processing/transportation costs, demands, and supplies, respectively, matrices $N$, $C$ and $S$ are appropriate matrices corresponding to the summations on the left-hand-side of the respective expressions. The notation $R$ corresponds to a matrix of $r^k_j$, and the notation $M$ corresponds to a matrix with $m_j$ along the diagonal.

It is realistic to assume that at time a decision about vector $x \in \mathcal{X}$ should be made, i.e., which facilities to build and machines to procure, there is an uncertainty about parameters involved in operational decisions represented by vector $y \in \mathbb{R}^{13}$. This naturally classifies decision variables $x$ as the first stage decision variables and $y$ as the second stage decision variables. Note that problem (1.75)–(1.79) can be written in the following equivalent form as a two stage program:

\[
\begin{align*}
\text{Min}_x & \quad c^T x + Q(x, \xi) \\
\text{s.t.} \quad & N y = 0, \\
& C y \geq d, \\
& S y \leq s, \\
& R y \leq M x,
\end{align*}
\]

(1.80)

where $Q(x, \xi)$ is the optimal value of the second stage problem

\[
\begin{align*}
\text{Min}_y & \quad q^T y \\
\text{s.t.} \quad & N y = 0, \\
& C y \geq d, \\
& S y \leq s, \\
& R y \leq M x,
\end{align*}
\]

(1.81)

(1.82)

(1.83)

(1.84)

(1.85)

with $\xi = (q, d, s, R, M)$ being vector of the involved parameters. Of course, the above optimization problem depends on the data vector $\xi$. If some of the data parameters are uncertain, then the deterministic problem (1.80) does not make much sense since it depends on unknown parameters.

Suppose now that we can model uncertain components of the data vector $\xi$ as random variables with a specified joint probability distribution. Then we can formulate the
Exercises 25

following stochastic programming problem

\[
\min_{x \in X} c^T x + \mathbb{E}[Q(x, \xi)],
\]

(1.86)

where the expectation is taken with respect to the probability distribution of the random vector \( \xi \). That is, the cost of the second stage problem enters the objective of the first stage problem on average. A distinctive feature of the stochastic programming problem (1.86) is that the first stage problem here is a combinatorial problem with binary decision variables and finite feasible set \( X \). On the other hand, the second stage problem (1.81)–(1.85) is a linear programming problem and its optimal value \( Q(x, \xi) \) is convex in \( x \) (if \( x \) is viewed as a vector in \( \mathbb{R}^{|P|} \)).

It could happen that for some \( x \in X \) and some realizations of the data \( \xi \) the corresponding second stage problem (1.81)–(1.85) is infeasible, i.e., the constraints (1.82)–(1.85) define an empty set. In that case, by definition, \( Q(x, \xi) = +\infty \), i.e., we apply an infinite penalization for infeasibility of the second stage problem. For example, it could happen that demand \( d \) is not satisfied, i.e., \( Cy \leq d \) with some inequalities strict, for any \( y \geq 0 \) satisfying constraints (1.82), (1.84) and (1.85). Sometimes this can be resolved by a recourse action. That is, if demand is not satisfied, then there is a possibility of supplying the deficit \( d - Cy \) at a penalty cost. This can be modelled by writing the second stage problem in the form

\[
\begin{align*}
\min & \quad q^T y + h^T z \\
\text{s.t.} & \quad Ny = 0, \\
& \quad Cy + z \geq d, \\
& \quad Sy \leq s, \\
& \quad Ry \leq Mx,
\end{align*}
\]

(1.87)–(1.91)

where \( h \) represents the vector of (positive) recourse costs. Note that the above problem (1.87)–(1.91) is always feasible, for example \( y = 0 \) and \( z \geq d \) clearly satisfy the constraints of this problem.

Exercises

1.1. Consider the expected value function \( f(x) := \mathbb{E}[F(x, D)] \), where function \( F(x, d) \) is defined in (1.1). (i) Show that function \( F(x, d) \) is convex in \( x \), and hence \( f(x) \) is also convex. (ii) Show that \( f(\cdot) \) is differentiable at a point \( x > 0 \) if the cdf \( H(\cdot) \) of \( D \) is continuous at \( x \).

1.2. Let \( H(z) \) be the cdf of a random variable \( Z \) and \( \kappa \in (0, 1) \). Show that the minimum in the definition \( H^{-1}(\kappa) = \inf\{ t : H(t) \geq \kappa \} \) of the left side quantile is always attained.

1.3. Consider the chance constrained problem discussed in section 1.2.2. (i) Show that system (1.11) has no feasible solution if there is a realization of \( d \) greater than \( \tau/c \). (ii) Verify equation (1.15). (iii) Assume that the probability distribution of the demand \( D \) is supported on an interval \([l, u]\) with \( 0 \leq l \leq u < +\infty \). Show that if the
significance level $\alpha = 0$, then the constraint (1.16) becomes
\[
\frac{bu - \tau}{b - c} \leq x \leq \frac{hl + \tau}{c + h},
\]
and hence is equivalent to (1.11) for $D = [l, u]$.

1.4. Show that the optimal value functions $Q_t(y_t, d_{[t-1]})$, defined in (1.20), are convex in $y_t$.

1.5. Assuming the stagewise independence condition, show that the basestock policy $\bar{x}_t = \max\{y_t, x^*_t\}$, for the inventory model, is optimal (recall that $x^*_t$ denotes a minimizer of (1.22)).

1.6. Consider the assembly problem discussed in section 1.3.1 in the case when all demand has to be satisfied, by making additional orders of the missing parts. In this case, the cost of each additionally ordered part $j$ is $r_j > c_j$. Formulate the problem as a linear two stage stochastic programming problem.

1.7. Consider the assembly problem discussed in section 1.3.3 in the case when all demand has to be satisfied, by backlogging the excessive demand, if necessary. In this case, it costs $b_i$ to delay delivery of a unit of product $i$ by one period. Additional orders of the missing parts can also be made after the last demand $D_T$ becomes known. Formulate the problem as a linear multistage stochastic programming problem.

1.8. Show that for utility function $U(W)$, of the form (1.36), problems (1.35) and (1.37)--(1.38) are equivalent.

1.9. Show that variance of the random return $W_1 = \xi^T x$ is given by formula $\text{Var}[W_1] = x^T \Sigma x$, where $\Sigma = E[(\xi - \mu)(\xi - \mu)^T]$ is the covariance matrix of the random vector $\xi$ and $\mu = E[\xi]$.

1.10. Show that the optimal value function $Q_t(W_t, \xi_{[t]}$, defined in (1.51), is convex in $W_t$.

1.11. Let $D$ be a random variable with cdf $H(t) = \text{Pr}(D \leq t)$ and $D^1, ..., D^N$ be an iid random sample of $D$ with the corresponding empirical cdf $\hat{H}_N(\cdot)$. Let $a = H^{-1}(\kappa)$ and $b = \sup\{t : H(t) \leq \kappa\}$ be respective left and right-side $\kappa$-quantiles of $H(\cdot)$. Show that $\min \left\{ |\hat{H}_N^{-1}(\kappa) - a|, |\hat{H}_N^{-1}(\kappa) - b| \right\}$ tends w.p.1 to 0 as $N \to \infty$. 
Chapter 2
Two Stage Problems

Andrzej Ruszczyński and Alexander Shapiro

2.1 Linear Two Stage Problems

2.1.1 Basic Properties

In this section we discuss two-stage stochastic linear programming problems of the form

\[ \begin{array}{c}
\text{Min} & c^T x + \mathbb{E}[Q(x, \xi)] \\
\text{s.t.} & Ax = b, \quad x \geq 0,
\end{array} \]

(2.1)

where \( Q(x, \xi) \) is the optimal value of the second stage problem

\[ \begin{array}{c}
\text{Min} & q^T y \\
\text{s.t.} & Tx + Wy = h, \quad y \geq 0.
\end{array} \]

(2.2)

Here \( \xi := (q, h, T, W) \) are the data of the second stage problem. We view some or all elements of vector \( \xi \) as random and the expectation operator at the first stage problem (2.1) is taken with respect to the probability distribution of \( \xi \). Often, we use the same notation \( \xi \) to denote a random vector and its particular realization. Which one of these two meanings will be used in a particular situation will usually be clear from the context. If in doubt, then we will write \( \xi = \xi(\omega) \) to emphasize that \( \xi \) is a random vector defined on a corresponding probability space. We denote by \( \Xi \subset \mathbb{R}^d \) the support of the probability distribution of \( \xi \).

If for some \( x \) and \( \xi \in \Xi \) the second stage problem (2.2) is infeasible, then by definition \( Q(x, \xi) = +\infty \). It could also happen that the second stage problem is unbounded from below and hence \( Q(x, \xi) = -\infty \). This is somewhat pathological situation, meaning that
for some value of the first stage decision vector and a realization of the random data, the value of the second stage problem can be improved indefinitely. Models exhibiting such properties should be avoided (we discuss this later).

The second stage problem (2.2) is a linear programming problem. Its dual problem can be written in the form

\[
\begin{align*}
\max_{\pi} & \quad \pi^T (h - T x) \\
\text{s.t.} & \quad W^T \pi \leq q.
\end{align*}
\]  

(2.3)

By the theory of linear programming, the optimal values of problems (2.2) and (2.3) are equal to each other, unless both problems are infeasible. Moreover, if their common optimal value is finite, then each problem has a nonempty set of optimal solutions.

Consider the function

\[
s_q(\chi) := \inf \{ q^T y : W y = \chi, \ y \geq 0 \}.
\]  

(2.4)

Clearly, \(Q(x, \xi) = s_q(h - T x)\). By the duality theory of linear programming, if the set

\[
\Pi(q) := \{ \pi : W^T \pi \leq q \}
\]  

(2.5)

is nonempty, then

\[
s_q(\chi) = \sup_{\pi \in \Pi(q)} \pi^T \chi,
\]  

(2.6)

i.e., \(s_q(\cdot)\) is the support function of the set \(\Pi(q)\). The set \(\Pi(q)\) is convex, closed, and polyhedral. Hence, it has a finite number of extreme points. (If, moreover, \(\Pi(q)\) is bounded, then it coincides with the convex hull of its extreme points.) It follows that if \(\Pi(q)\) is nonempty, then \(s_q(\cdot)\) is a positively homogeneous polyhedral function. If the set \(\Pi(q)\) is empty, then the infimum on the right hand side of (2.4) may take only two values: \(+\infty\) or \(-\infty\). In any case it is not difficult to verify directly that the function \(s_q(\cdot)\) is convex.

**Proposition 2.1.** For any given \(\xi\), the function \(Q(\cdot, \xi)\) is convex. Moreover, if the set \(\{ \pi : W^T \pi \leq q \}\) is nonempty and problem (2.2) is feasible for at least one \(x\), then the function \(Q(\cdot, \xi)\) is polyhedral.

**Proof.** Since \(Q(x, \xi) = s_q(h - T x)\), the above properties of \(Q(\cdot, \xi)\) follow from the corresponding properties of the function \(s_q(\cdot)\).

Differentiability properties of the function \(Q(\cdot, \xi)\) can be described as follows.

**Proposition 2.2.** Suppose that for given \(x = x_0\) and \(\xi \in \Xi\), the value \(Q(x_0, \xi)\) is finite. Then \(Q(\cdot, \xi)\) is subdifferentiable at \(x_0\) and

\[
\partial Q(x_0, \xi) = -T^T \mathcal{D}(x_0, \xi),
\]  

(2.7)

where

\[
\mathcal{D}(x, \xi) := \arg \max_{\pi \in \Pi(q)} \pi^T (h - T x)
\]

is the set of optimal solutions of the dual problem (2.3).
2.1. Linear Two Stage Problems

**Proof.** Since \( Q(x_0, \xi) \) is finite, the set \( \Pi(q) \) defined in (2.5) is nonempty, and hence \( s_q(\chi) \) is its support function. It is straightforward to see from the definitions that the support function \( s_q(\cdot) \) is the conjugate function of the indicator function

\[
I_q(\pi) := \begin{cases} 
0, & \text{if } \pi \in \Pi(q), \\
+\infty, & \text{otherwise.}
\end{cases}
\]

Since the set \( \Pi(q) \) is convex and closed, the function \( I_q(\cdot) \) is convex and lower semicontinuous. It follows then by the Fenchel–Moreau Theorem (Theorem 7.5) that the conjugate of \( s_q(\cdot) \) is \( I_q(\cdot) \).

Therefore, for \( \chi_0 := h - Tx_0 \), we have (see (7.24))

\[
\partial s_q(\chi_0) = \arg \max_{\pi} \{ \pi^T \chi_0 - I_q(\pi) \} = \arg \max_{\pi \in \Pi(q)} \pi^T \chi_0. \tag{2.8}
\]

Since the set \( \Pi(q) \) is polyhedral and \( s_q(\chi_0) \) is finite, it follows that \( \partial s_q(\chi_0) \) is nonempty. Moreover, the function \( s_0(\cdot) \) is piecewise linear, and hence formula (2.7) follows from (2.8) by the chain rule of subdifferentiation. \( \Box \)

It follows that if the function \( Q(\cdot, \xi) \) has a finite value in at least one point, then it is subdifferentiable at that point, and hence is proper. Its domain can be described in a more explicit way.

The **positive hull** of a matrix \( W \) is defined as

\[
\text{pos } W := \{ \chi : \chi = Wy, \ y \geq 0 \}. \tag{2.9}
\]

It is a convex polyhedral cone generated by the columns of \( W \). Directly from the definition (2.4) we see that \( \text{dom } s_q = \text{pos } W \). Therefore,

\[
\text{dom } Q(\cdot, \xi) = \{ x : h - Tx \in \text{pos } W \}.
\]

Suppose that \( x \) is such that \( \chi = h - Tx \in \text{pos } W \) and let us analyze formula (2.7). The recession cone of \( \Pi(q) \) is equal to

\[
\Pi_0 := \Pi(0) = \{ \pi : W^T \pi \leq 0 \}. \tag{2.10}
\]

Then it follows from (2.6) that \( s_q(\chi) \) is finite iff \( \pi^T \chi \leq 0 \) for every \( \pi \in \Pi_0 \), that is, iff \( \chi \) is an element of the polar cone to \( \Pi_0 \). This polar cone is nothing else but \( \text{pos } W \), i.e.,

\[
\Pi_0^0 = \text{pos } W. \tag{2.11}
\]

If \( \chi_0 \in \text{int}(\text{pos } W) \), then the set of maximizers in (2.6) must be bounded. Indeed, if it was unbounded, there would exist an element \( \pi_0 \in \Pi_0 \) such that \( \pi_0^T \chi_0 = 0 \). By perturbing \( \chi_0 \) a little to some \( \chi \), we would be able to keep \( \chi \) within \( \text{pos } W \) and get \( \pi_0^T \chi > 0 \), which is a contradiction, because \( \text{pos } W \) is the polar of \( \Pi_0 \). Therefore the set of maximizers in (2.6) is the convex hull of the vertices \( v \) of \( \Pi(q) \) for which \( v^T \chi = s_q(\chi) \). Note that \( \Pi(q) \) must have vertices in this case, because otherwise the polar to \( \Pi_0 \) would have no interior.

If \( \chi_0 \) is a boundary point of \( \text{pos } W \), then the set of maximizers in (2.6) is unbounded. Its recession cone is the intersection of the recession cone \( \Pi_0 \) of \( \Pi(q) \) and of the subspace \( \{ \pi : \pi^T \chi_0 = 0 \} \).

This intersection is nonempty for boundary points \( \chi_0 \) and is equal to the normal cone to \( \text{pos } W \) at \( \chi_0 \). Indeed, let \( \pi_0 \) be normal to \( \text{pos } W \) at \( \chi_0 \). Since both \( \chi_0 \) and \( -\chi_0 \) are feasible directions at \( \chi_0 \), we must have \( \pi_0^T \chi_0 = 0 \). Next, for every \( \chi \in \text{pos } W \) we have \( \pi_0^T \chi = \pi_0^T (\chi - \chi_0) \leq 0 \), so \( \pi_0 \in \Pi_0 \). The converse argument is similar.
2.1.2 The Expected Recourse Cost for Discrete Distributions

Let us consider now the expected value function

$$\phi(x) := \mathbb{E}[Q(x, \xi)].$$ \hfill (2.12)

As before the expectation here is taken with respect to the probability distribution of the random vector $\xi$. Suppose that the distribution of $\xi$ has finite support. That is, $\xi$ has a finite number of realizations (called scenarios) $\xi_k = (q_k, h_k, T_k, W_k)$ with respective (positive) probabilities $p_k, k = 1, \ldots, K$, i.e., $\Xi = \{\xi_1, \ldots, \xi_K\}$. Then

$$\mathbb{E}[Q(x, \xi)] = \sum_{k=1}^{K} p_k Q(x, \xi_k).$$ \hfill (2.13)

For a given $x$, the expectation $\mathbb{E}[Q(x, \xi)]$ is equal to the optimal value of the linear programming problem

$$\begin{aligned}
\min_{y_1, \ldots, y_K} & \quad \sum_{k=1}^{K} p_k q^T_k y_k \\
\text{s.t.} & \quad T_k x + W_k y_k = h_k, \\
& \quad y_k \geq 0, \quad k = 1, \ldots, K.
\end{aligned}$$ \hfill (2.14)

If for at least one $k \in \{1, \ldots, K\}$ the system $T_k x + W_k y_k = h_k, y_k \geq 0$, has no solution, i.e., the corresponding second stage problem is infeasible, then problem (2.14) is infeasible, and hence its optimal value is $+\infty$. From that point of view the sum in the right hand side of (2.13) equals $+\infty$ if at least one of $Q(x, \xi_k) = +\infty$. That is, we assume here that $+\infty + (-\infty) = +\infty$.

The whole two stage problem is equivalent to the following large-scale linear programming problem:

$$\begin{aligned}
\min_{x, y_1, \ldots, y_K} & \quad c^T x + \sum_{k=1}^{K} p_k q^T_k y_k \\
\text{s.t.} & \quad T_k x + W_k y_k = h_k, \quad k = 1, \ldots, K, \\
& \quad A x = b, \\
& \quad x \geq 0, \quad y_k \geq 0, \quad k = 1, \ldots, K.
\end{aligned}$$ \hfill (2.15)

Properties of the expected recourse cost follow directly from properties of parametric linear programming problems.

**Proposition 2.3.** Suppose that the probability distribution of $\xi$ has finite support $\Xi = \{\xi_1, \ldots, \xi_K\}$ and that the expected recourse cost $\phi(\cdot)$ has a finite value in at least one point $\bar{x} \in \mathbb{R}^n$. Then the function $\phi(\cdot)$ is polyhedral, and for any $x_0 \in \text{dom} \ \phi$,

$$\partial \phi(x_0) = \sum_{k=1}^{K} p_k \partial Q(x_0, \xi_k).$$ \hfill (2.16)
2.1. Linear Two Stage Problems

Proof. Since \( \phi(x) \) is finite, all values \( Q(\bar{x}, \xi_k) \), \( k = 1, \ldots, K \), are finite. Consequently, by Proposition 2.2, every function \( Q(\cdot, \xi_k) \) is polyhedral. It is not difficult to see that a linear combination of polyhedral functions with positive weights is also polyhedral. Therefore, it follows that \( \phi(\cdot) \) is polyhedral. We also have that \( \text{dom } \phi = \bigcap_{k=1}^{K} \text{dom } Q_k \), where \( Q_k(\cdot) := Q(\cdot, \xi_k) \), and for any \( h \in \mathbb{R}^n \), the directional derivatives \( Q'_k(x_0, h) > -\infty \) and

\[
\phi'(x_0, h) = \sum_{k=1}^{K} p_k Q'_k(x_0, h). \tag{2.17}
\]

Formula (2.16) then follows from (2.17) by duality arguments. Note that equation (2.16) is a particular case of the Moreau–Rockafellar Theorem (Theorem 7.4). Since the functions \( Q_k \) are polyhedral, there is no need here for an additional regularity condition for (2.16) to hold true.

The subdifferential \( \partial Q(x_0, \xi_k) \) of the second stage optimal value function is described in Proposition 2.2. That is, if \( Q(x_0, \xi_k) \) is finite, then

\[
\partial Q(x_0, \xi_k) = -T_k^T \arg \max \left\{ \pi^T (h_k - T_k x_0) : W_k^T \pi \leq q_k \right\}. \tag{2.18}
\]

It follows that the expectation function \( \phi \) is differentiable at \( x_0 \) iff for every \( \xi = \xi_k \), \( k = 1, \ldots, K \), the maximum in the right hand side of (2.18) is attained at a unique point, i.e., the corresponding second stage dual problem has a unique optimal solution.

Example 2.4 (Capacity Expansion) We have a directed graph with node set \( N \) and arc set \( A \). With each arc \( a \in A \) we associate a decision variable \( x_a \) and call it the capacity of \( a \). There is a cost \( c_a \) for each unit of capacity of arc \( a \). The vector \( x \) constitutes the vector of first stage variables. They are restricted to satisfy the inequalities \( x \geq x_{min} \), where \( x_{min} \) are the existing capacities.

At each node \( n \) of the graph we have a random demand \( \xi_n \) for shipments to \( n \) (if \( \xi_n \) is negative, its absolute value represents shipments from \( n \) and we have \( \sum_{n \in N} \xi_n = 0 \)). These shipments have to be sent through the network and they can be arbitrarily split into pieces taking different paths. We denote by \( y_a \) the amount of the shipment sent through arc \( a \). There is a unit cost \( q_a \) for shipments on each arc \( a \).

Our objective is to assign the arc capacities and to organize the shipments in such a way that the expected total cost, comprising the capacity cost and the shipping cost, is minimized. The condition is that the capacities have to be assigned before the actual demands \( \xi_n \) become known, while the shipments can be arranged after that.

Let us define the second stage problem. For each node \( n \) denote by \( A_+(n) \) and \( A_-(n) \) the sets of arcs entering and leaving node \( i \). The second stage problem is the network flow problem

\[
\text{Min } \sum_{a \in A} q_a y_a \tag{2.19}
\]

s.t.

\[
\sum_{a \in A_+(n)} y_a - \sum_{a \in A_-(n)} y_a = \xi_n, \quad n \in N, \tag{2.20}
\]

\[
0 \leq y_a \leq x_a, \quad a \in A. \tag{2.21}
\]
This problem depends on the random demand vector $\xi$ and on the arc capacities, $x$. Its optimal value is denoted by $Q(x, \xi)$.

Suppose that for a given $x = x_0$ the second stage problem (2.19)–(2.21) is feasible. Denote by $\mu_n, n \in \mathcal{N}$, the optimal Lagrange multipliers (node potentials) associated with the node balance equations (2.20), and by $\pi_a, a \in \mathcal{A}$ the (nonnegative) Lagrange multipliers associated with the constraints (2.21). The dual problem has the form

$$\text{Max } - \sum_{n \in \mathcal{N}} \xi_n \mu_n - \sum_{(i,j) \in \mathcal{A}} x_{ij} \pi_{ij}$$

s.t. $- \pi_{ij} + \mu_i - \mu_j \leq q_{ij}, \quad (i,j) \in \mathcal{A}$,

$$\pi \geq 0.$$

As $\sum_{n \in \mathcal{N}} \xi_n = 0$, the values of $\mu_n$ can be translated by a constant without any change in the objective function, and thus without any loss of generality we can assume that $\mu_{n_0} = 0$ for some fixed node $n_0$. For each arc $a = (i,j)$ the multiplier $\pi_{ij}$ associated with the constraint (2.21) has the form

$$\pi_{ij} = \max\{0, \mu_i - \mu_j - q_{ij}\}.$$

Roughly, if the difference of node potentials $\mu_i - \mu_j$ is greater than $q_{ij}$, the arc is saturated and the capacity constraint $y_{ij} \leq x_{ij}$ becomes relevant. The dual problem becomes equivalent to:

$$\text{Max } - \sum_{n \in \mathcal{N}} \xi_n \mu_n - \sum_{(i,j) \in \mathcal{A}} x_{ij} \max\{0, \mu_i - \mu_j - q_{ij}\}. \quad (2.22)$$

Let us denote by $\mathcal{M}(x_0, \xi)$ the set of optimal solutions of this problem satisfying the condition $\mu_{n_0} = 0$. Since $T^T = [0 - I]$ in this case, formula (2.18) provides the description of the subdifferential of $Q(\cdot, \xi)$ at $x_0$:

$$\partial Q(x_0, \xi) = - \left\{ \{\max\{0, \mu_i - \mu_j - q_{ij}\}\}_{(i,j) \in \mathcal{A}} : \mu \in \mathcal{M}(x_0, \xi) \right\}.$$

The first stage problem has the form

$$\text{Min } \sum_{(i,j) \in \mathcal{A}} c_{ij} x_{ij} + E[Q(x, \xi)]. \quad (2.23)$$

If $\xi$ has finitely many realizations $\xi^k$ attained with probabilities $p_k, \ k = 1, \ldots, K$, the subdifferential of the overall objective can be calculated by (2.16):

$$\partial f(x_0) = c + \sum_{k=1}^{K} p_k \partial Q(x_0, \xi^k).$$
2.1. Linear Two Stage Problems

2.1.3 The Expected Recourse Cost for General Distributions

Let us discuss now the case of a general distribution of the random vector \( \xi \in \mathbb{R}^d \). The recourse cost \( Q(\cdot, \cdot) \) is the minimum value of the integrand which is a random lower semi-continuous function (see section 7.2.3). Therefore, it follows by Theorem 7.37 that \( Q(\cdot, \cdot) \) is measurable with respect to the Borel sigma algebra of \( \mathbb{R}^n \times \mathbb{R}^d \). Also for every \( \xi \) the function \( Q(\cdot, \xi) \) is lower semicontinuous. It follows that \( Q(x, \xi) \) is a random lower semicontinuous function. Recall that in order to ensure that the expectation \( \phi(x) \) is well defined we have to verify two conditions:

(i) \( Q(x, \cdot) \) is measurable (with respect to the Borel sigma algebra of \( \mathbb{R}^d \));

(ii) either \( \mathbb{E}[Q(x, \xi)_+] \) or \( \mathbb{E}[(-Q(x, \xi))_+] \) are finite.

The function \( Q(x, \cdot) \) is measurable, as the optimal value of a linear programming problem. We only need to verify condition (ii). We describe below some important particular situations where this condition is satisfied.

The two-stage problem (2.1)–(2.2) is said to have fixed recourse if the matrix \( W \) is fixed (not random). Moreover, we say that the recourse is complete if the system \( Wy = \chi \) and \( y \geq 0 \) has a solution for every \( \chi \). In other words, the positive hull of \( W \) is equal to the corresponding vector space. By duality arguments, the fixed recourse is complete if the feasible set \( \Pi(q) \) of the dual problem (2.3) is bounded (in particular, it may be empty) for every \( q \). Then its recession cone, \( \Pi_0 = \Pi(0) \), must contain only the point 0, provided that \( \Pi(q) \) is nonempty. Therefore, another equivalent condition for complete recourse is that \( \pi = 0 \) is the only solution of the system \( W^T \pi \leq 0 \).

A particular class of problems with fixed and complete recourse are simple recourse problems, in which \( W = [I; -I] \), the matrix \( T \) and the vector \( q \) are deterministic, and the components of \( q \) are positive.

It is said that the recourse is relatively complete if for every feasible first stage point \( x \) the inequality \( Q(x, \xi) < +\infty \) holds true for a.e. \( \xi \). That is, the recourse is relatively complete if for every feasible first stage point \( x \) the inequality \( Q(x, \xi) < +\infty \) holds true for a.e. \( \xi \in \Xi \), or in other words \( Q(x, \xi(\omega)) < +\infty \) w.p.1. This definition is in accordance with the general principle that an event which happens with zero probability is irrelevant for the calculation of the corresponding expected value. For example, the capacity expansion problem of Example 2.4 is not a problem with relatively complete recourse, unless \( x_{\text{min}} \) is so large that every demand \( \xi \in \Xi \) can be shipped over the network with capacities \( x_{\text{min}} \).

The following condition is sufficient for relatively complete recourse.

For every \( x \in X \) the inequality \( Q(x, \xi) < +\infty \) holds true for all \( \xi \in \Xi \). (2.24)

In general, condition (2.24) is not necessary for relatively complete recourse. It becomes necessary and sufficient in the following two cases:

(i) the random vector \( \xi \) has a finite support; or
(ii) the recourse is fixed.

Indeed, sufficiency is clear. If $\xi$ has a finite support, i.e., the set $\Xi$ is finite, then the necessity is also clear. To show the necessity in the case of fixed recourse, suppose the recourse is relatively complete. This means that if $x \in X$ then $Q(x, \xi) < +\infty$ for all $\xi$ in $\Xi$, except possibly for a subset of $\Xi$ of probability zero. We have that $Q(x, \xi) < +\infty$ iff $h - Tx \in \text{pos } W$. Let $\Xi_0(x) = \{ (h, T, q) : h - Tx \in \text{pos } W \}$. The set $\text{pos } W$ is convex and closed and thus $\Xi_0(x)$ is convex and closed as well. By assumption, $P[\Xi_0(x)] = 1$ for every $x \in X$. Thus $\bigcap_{x \in X} \Xi_0(x)$ is convex, closed, and has probability 1. The support of $\xi$ must be its subset.

Example 2.5 Consider

$$Q(x, \xi) := \inf \{ y : \xi y = x, \ y \geq 0 \},$$

with $x \in [0, 1]$ and $\xi$ being a random variable whose probability density function is $p(z) := 2z, \ 0 \leq z \leq 1$. For all $\xi > 0$ and $x \in [0, 1]$, $Q(x, \xi) = x/\xi$, and hence

$$E[Q(x, \xi)] = \int_0^1 \left( \frac{x}{z} \right) 2zdz = 2x.$$ 

That is, the recourse here is relatively complete and the expectation of $Q(x, \xi)$ is finite. On the other hand, the support of $\xi(\omega)$ is the interval $[0, 1]$, and for $\xi = 0$ and $x > 0$ the value of $Q(x, \xi)$ is $+\infty$, because the corresponding problem is infeasible. Of course, probability of the event “$\xi = 0$” is zero, and from the mathematical point of view the expected value $E[Q(x, \xi)]$ is well defined and finite for all $x \in [0, 1]$. Note, however, that arbitrary small perturbation of the probability distribution of $\xi$ may change that. Take, for example, some discretization of the distribution of $\xi$ with the first discretization point $t = 0$. Then, no matter how small the assigned (positive) probability at $t = 0$, $Q(x, \xi) = +\infty$ with positive probability. Therefore, $E[Q(x, \xi)] = +\infty$, for all $x > 0$. That is, the above problem is extremely unstable and is not well posed. As it was discussed above, such behavior cannot occur if the recourse is fixed.

Let us consider the support function $s_q(\cdot)$ of the set $\Pi(q)$. We want to find sufficient conditions for the existence of the expectation $E[Q(\omega, h - Tx)]$. By Hoffman's lemma (Theorem 7.11) there exists a constant $\kappa$, depending on $W$, such that if for some $q_0$ the set $\Pi(q_0)$ is nonempty, then for every $q$ the following inclusion is satisfied

$$\Pi(q) \subset \Pi(q_0) + \kappa \| q - q_0 \| B,$$  

(2.25)

where $B := \{ \pi : \| \pi \| \leq 1 \}$ and $\| \cdot \|$ denotes the Euclidean norm. This inclusion allows to derive an upper bound for the support function $s_q(\cdot)$. Since the support function of the unit ball $B$ is the norm $\| \cdot \|$, it follows from (2.25) that if the set $\Pi(q_0)$ is nonempty, then

$$s_q(\cdot) \leq s_{q_0}(\cdot) + \kappa \| q - q_0 \| \| \cdot \|.$$ 

(2.26)

Consider $q_0 = 0$. The support function $s_0(\cdot)$ of the cone $\Pi_0$ has the form

$$s_0(\chi) = \begin{cases} 0, & \text{if } \chi \in \text{pos } W, \\ +\infty, & \text{otherwise}. \end{cases}$$
2.1. Linear Two Stage Problems

Therefore, (2.26) with \( q_0 = 0 \) implies that if \( \Pi(q) \) is nonempty, then \( s_q(\chi) \leq \kappa \|q\| \|\chi\| \) for all \( \chi \in \text{pos } W \), and \( s_q(\chi) = +\infty \) for all \( \chi \not\in \text{pos } W \). Since \( \Pi(q) \) is polyhedral, if it is nonempty then \( s_q(\cdot) \) is piecewise linear on its domain, which coincides with \( \text{pos } W \), and

\[
|s_q(\chi_1) - s_q(\chi_2)| \leq \kappa \|q\| \|\chi_1 - \chi_2\|, \quad \forall \chi_1, \chi_2 \in \text{pos } W.
\] (2.27)

**Proposition 2.6.** Suppose that the recourse is fixed and

\[
\mathbb{E}[\|q\| \|h\|] < +\infty \quad \text{and} \quad \mathbb{E}[\|q\| \|T\|] < +\infty.
\] (2.28)

Consider a point \( x \in \mathbb{R}^n \). Then \( \mathbb{E}[Q(x, \xi)_+] \) is finite if and only if the following condition holds w.p.1:

\[
h - Tx \in \text{pos } W.
\] (2.29)

**Proof.** We have that \( Q(x, \xi) < +\infty \) iff condition (2.29) holds. Therefore, if condition (2.29) does not hold w.p.1, then \( Q(x, \xi) = +\infty \) with positive probability, and hence \( \mathbb{E}[Q(x, \xi)_+] = +\infty \).

Conversely, suppose that condition (2.29) holds w.p.1. Then \( Q(x, \xi) = s_q(h - Tx) \) with \( s_q(\cdot) \) being the support function of the set \( \Pi(q) \). By (2.26) there exists a constant \( \kappa \) such that for any \( \chi \),

\[ s_q(\chi) \leq s_0(\chi) + \kappa \|q\| \|\chi\|. \]

Also for any \( \chi \in \text{pos } W \) we have that \( s_0(\chi) = 0 \), and hence w.p.1,

\[ s_q(h - Tx) \leq \kappa \|q\| \|h - Tx\| \leq \kappa \|q\| (\|h\| + \|T\| \|x\|). \]

It follows then by (2.28) that \( \mathbb{E}[s_q(h - Tx)_+] < +\infty \). \qed

**Remark 2.** If \( q \) and \( (h, T) \) are independent and have finite first moments\(^1\), then

\[
\mathbb{E}[\|q\| \|h\|] = \mathbb{E}[\|q\|] \mathbb{E}[\|h\|] \quad \text{and} \quad \mathbb{E}[\|q\| \|T\|] = \mathbb{E}[\|q\|] \mathbb{E}[\|T\|],
\]

and hence condition (2.28) follows. Also condition (2.28) holds if \( (h, T, q) \) has finite second moments.

We obtain that, under the assumptions of Proposition 2.6, the expectation \( \phi(x) \) is well defined and \( \phi(x) < +\infty \) iff condition (2.29) holds w.p.1. If, moreover, the recourse is complete, then (2.29) holds for any \( x \) and \( \xi \), and hence \( \phi(\cdot) \) is well defined and is less than \( +\infty \). Since the function \( \phi(\cdot) \) is convex, we have that if \( \phi(\cdot) \) is less than \( +\infty \) on \( \mathbb{R}^n \) and is finite valued in at least one point, then \( \phi(\cdot) \) is finite valued on the entire space \( \mathbb{R}^n \).

**Proposition 2.7.** Suppose that: (i) the recourse is fixed, (ii) for a.e. \( q \) the set \( \Pi(q) \) is nonempty, (iii) condition (2.28) holds.

\(^1\)We say that a random variable \( Z = Z(\omega) \) has a finite \( r \)-th moment if \( \mathbb{E}[|Z|^r] < +\infty \). It is said that \( \xi(\omega) \) has finite \( r \)-th moments if each component of \( \xi(\omega) \) has a finite \( r \)-th moment.
Then the expectation function \( \phi(x) \) is well defined and \( \phi(x) > -\infty \) for all \( x \in \mathbb{R}^n \). Moreover, \( \phi \) is convex, lower semicontinuous and Lipschitz continuous on \( \text{dom} \phi \), and its domain is a convex closed subset of \( \mathbb{R}^n \) given by

\[
\text{dom} \phi = \{ x \in \mathbb{R}^n : h - Tx \in \text{pos} W \text{ w.p.1}\}.
\] (2.30)

**Proof.** By assumption (ii) the feasible set \( \Pi(q) \) of the dual problem is nonempty w.p.1. Thus \( Q(x, \xi) \) is equal to \( s_q(h - Tx) \text{ w.p.1} \) for every \( x \), where \( s_q(\cdot) \) is the support function of the set \( \Pi(q) \). Let \( \pi(q) \) be the element of the set \( \Pi(q) \) that is closest to 0. It exists, because \( \Pi(q) \) is closed. By Hoffman’s lemma (see (2.25)) there is a constant \( \kappa \) such that \( \| \pi(q) \| \leq \kappa \|q\| \). Then for every \( x \) the following holds w.p.1:

\[
s_q(h - Tx) \geq \pi(q)^T (h - Tx) \geq -\kappa\|q\| (\|h\| + \|T\| \|x\|).
\] (2.31)

Owing to condition (2.28), it follows from (2.31) that \( \phi(\cdot) \) is well defined and \( \phi(x) > -\infty \) for all \( x \in \mathbb{R}^n \). Moreover, since \( s_q(\cdot) \) is lower semicontinuous, the lower semicontinuity of \( \phi(\cdot) \) follows by Fatou’s lemma. Convexity and closedness of \( \text{dom} \phi \) follow from the convexity and lower semicontinuity of \( \phi \). We have by Proposition 2.6 that \( \phi(x) < +\infty \) iff condition (2.29) holds w.p.1. This implies (2.30).

Consider two points \( x, x' \in \text{dom} \phi \). Then by (2.30) the following holds true w.p.1:

\[
h - Tx \in \text{pos} W \text{ and } h - Tx' \in \text{pos} W.
\] (2.32)

By (2.27), if the set \( \Pi(q) \) is nonempty and (2.32) holds, then

\[
|s_q(h - Tx) - s_q(h - Tx')| \leq \kappa\|q\| \|T\| \|x - x'\|.
\]

It follows that

\[
|\phi(x) - \phi(x')| \leq \kappa \mathbb{E}[\|q\| \|T\| \|x - x'\|].
\]

Together with condition (2.28) this implies the Lipschitz continuity of \( \phi \) on its domain. \( \Box \)

Denote by \( \Sigma \) the support\(^2\) of the probability distribution (measure) of \( (h, T) \). Formula (2.30) means that a point \( x \) belongs to \( \text{dom} \phi \) iff the probability of the event \( \{h - Tx \in \text{pos} W\} \) is one. Note that the set \( \{ (h, T) : h - Tx \in \text{pos} W \} \) is convex and polyhedral, and hence is closed. Consequently \( x \) belongs to \( \text{dom} \phi \) iff for every \( (h, T) \in \Sigma \) it follows that \( h - Tx \in \text{pos} W \). Therefore, we can write formula (2.30) in the form

\[
\text{dom} \phi = \bigcap_{(h, T) \in \Sigma} \{ x : h - Tx \in \text{pos} W \}.
\] (2.33)

It should be noted that we assume that the recourse is fixed.

Let us observe that for any set \( \mathcal{H} \) of vectors \( h \), the set \( \cap_{h \in \mathcal{H}} (-h + \text{pos} W) \) is convex and polyhedral. Indeed, we have that \( \text{pos} W \) is a convex polyhedral cone, and hence can

\(^2\)Recall that the support of the probability measure is the smallest closed set such that the probability (measure) of its complement is zero.
be represented as the intersection of a finite number of half spaces \( A_i = \{ \chi : a_i^T \chi \leq 0 \}, \)
\( i = 1, \ldots, \ell. \) Since the intersection of any number of half spaces of the form \( b + A_i, \) with \( b \in B, \) is still a half space of the same form (provided that this intersection is nonempty), we have the set \( \cap_{h \in \mathcal{H}} (-h + \text{pos} W) \) can be represented as the intersection of half spaces of the form \( b_i + A_i, \) \( i = 1, \ldots, \ell, \) and hence is polyhedral. It follows that if \( T \) and \( W \) are fixed, then the set at the right hand side of (2.33) is convex and polyhedral.

Let us discuss now the differentiability properties of the expectation function \( \phi(x). \) By Theorem 7.47 and formula (2.7) of Proposition 2.2 we have the following result.

**Proposition 2.8.** Suppose that the expectation function \( \phi(\cdot) \) is proper and its domain has a nonempty interior. Then for any \( x_0 \in \text{dom} \phi, \)
\[
\partial \phi(x_0) = -\mathbb{E} \left[ T^T \mathcal{D}(x_0, \xi) \right] + \mathcal{N}_{\text{dom} \phi}(x_0),
\]
(2.34)
where
\[
\mathcal{D}(x, \xi) := \arg \max_{\pi \in \Pi(q)} \pi^T (h - Tx).
\]
Moreover, \( \phi \) is differentiable at \( x_0 \) if and only if \( x_0 \) belongs to the interior of \( \text{dom} \phi \) and the set \( \mathcal{D}(x_0, \xi) \) is a singleton w.p.1.

As we discussed earlier, when the distribution of \( \xi \) has a finite support (i.e., there is a finite number of scenarios) the expectation function \( \phi \) is piecewise linear on its domain and is differentiable everywhere only in the trivial case if it is linear.\(^3\) In the case of a continuous distribution of \( \xi \) the expectation operator smoothes the piecewise linear function \( \mathcal{Q}(\cdot; \xi) \) out.

**Proposition 2.9.** Suppose the assumptions of Proposition 2.7 are satisfied, and the conditional distribution of \( h, \) given \((T, q), \) is absolutely continuous for almost all \((T, q). \) Then \( \phi \) is continuously differentiable on the interior of its domain.

**Proof.** By Proposition 2.7, the expectation function \( \phi(\cdot) \) is well defined and greater than \(-\infty. \) Let \( x \) be a point in the interior of \( \text{dom} \phi. \) For fixed \( T \) and \( q, \) consider the multifunction
\[
\mathcal{H}(h) := \arg \max_{\pi \in \Pi(q)} \pi^T (h - Tx).
\]
Conditional on \((T, q), \) the set \( \mathcal{D}(x, \xi) \) coincides with \( \mathcal{H}(h). \) Since \( x \in \text{dom} \phi, \) relation (2.30) implies that \( h - Tx \in \text{pos} W \) w.p.1. For every \( h - Tx \in \text{pos} W, \) the set \( \mathcal{H}(h) \) is nonempty and forms a face of the polyhedral set \( \Pi(q). \) Moreover, there exists a set \( A \) given by the union of a finite number of linear subspaces of \( \mathbb{R}^m \) (where \( m \) is the dimension of \( h), \)
which are perpendicular to the faces of sets \( \Pi(q), \) such that if \( h - Tx \in (\text{pos} W) \setminus A \) then \( \mathcal{H}(h) \) is a singleton. Since an affine subspace of \( \mathbb{R}^m \) has Lebesgue measure zero, it follows that the Lebesgue measure of \( A \) is zero. As the conditional distribution of \( h, \) given \((T, q), \)
is absolutely continuous, the probability that \( \mathcal{H}(h) \) is not a singleton is zero. By integrating this probability over the marginal distribution of \((T, q), \) we obtain that probability of the

\(^3\)By linear we mean here that it is of the form \( a^T x + b. \) It is more accurate to call such a function affine.
event “$\mathcal{D}(x, \xi)$ is not a singleton” is zero. By Proposition 2.8, this implies the differentiability of $\phi(\cdot)$. Since $\phi(\cdot)$ is convex, it follows that for every $x \in \text{int}(\text{dom } \phi)$ the gradient $\nabla \phi(x)$ coincides with the (unique) subgradient of $\phi$ at $x$, and that $\nabla \phi(\cdot)$ is continuous at $x$. 

Of course, if $h$ and $(T, q)$ are independent, then the conditional distribution of $h$ given $(T, q)$ is the same as the unconditional (marginal) distribution of $h$. Therefore, if $h$ and $(T, q)$ are independent, then it suffices to assume in the above proposition that the (marginal) distribution of $h$ is absolutely continuous.

### 2.1.4 Optimality Conditions

We can now formulate optimality conditions and duality relations for linear two stage problems. Let us start from the problem with discrete distributions of the random data in (2.1)–(2.2). The problem takes on the form:

$$
\begin{aligned}
\text{Min } & c^T x + \sum_{k=1}^{K} p_k Q(x, \xi_k) \\
\text{s.t. } & Ax = b, \ x \geq 0,
\end{aligned}
$$

where $Q(x, \xi)$ is the optimal value of the second stage problem, given by (2.2).

Suppose the expectation function $\phi(\cdot) : = E[Q(\cdot, \xi)]$ has a finite value in at least one point $\bar{x} \in \mathbb{R}^n$. It follows from Propositions 2.2 and 2.3 that for every $x_0 \in \text{dom } \phi$,

$$
\partial \phi(x_0) = - \sum_{k=1}^{K} p_k T_k^T \mathcal{D}(x_0, \xi_k),
$$

where

$$
\mathcal{D}(x_0, \xi_k) : = \arg \max \left\{ \pi^T (h_k - T_k x_0) : W_k^T \pi \leq q_k \right\}.
$$

As before, we denote $X : = \{ x : Ax = b, \ x \geq 0 \}$.

**Theorem 2.10.** Let $\bar{x}$ be a feasible solution of problem (2.1)–(2.2), i.e., $\bar{x} \in X$ and $\phi(\bar{x})$ is finite. Then $\bar{x}$ is an optimal solution of problem (2.1)–(2.2) iff there exist $\pi_k \in \mathcal{D}(\bar{x}, \xi_k)$, $k = 1, \ldots, K$, and $\mu \in \mathbb{R}^m$ such that

$$
\begin{aligned}
\sum_{k=1}^{K} p_k T_k^T \pi_k + A^T \mu & \leq c, \\
\bar{x}^T \left( c - \sum_{k=1}^{K} p_k T_k^T \pi_k - A^T \mu \right) & = 0.
\end{aligned}
$$

**Proof.** Necessary and sufficient optimality conditions for minimizing $c^T x + \phi(x)$ over $x \in X$ can be written as

$$
0 \in c + \partial \phi(\bar{x}) + \mathcal{N}_X(\bar{x}),
$$

where $\mathcal{N}_X(\bar{x})$ is the normal cone to $X$ at $\bar{x}$.
2.1. Linear Two Stage Problems

where \( N_X(\bar{x}) \) is the normal cone to the feasible set \( X \). Note that condition (2.38) implies that the sets \( N_X(\bar{x}) \) and \( \partial \phi(\bar{x}) \) are nonempty and hence \( \bar{x} \in X \) and \( \phi(\bar{x}) \) is finite. Note also that there is no need here for additional regularity conditions since \( \phi(\cdot) \) and \( X \) are convex and polyhedral. Using the characterization of the subgradients of \( \phi(\cdot) \), given in (2.36), we conclude that (2.38) is equivalent to existence of \( \pi_k \in \mathcal{D}(\bar{x}, \xi_k) \) such that

\[
0 \in c - \sum_{k=1}^{K} p_k T_k^T \pi_k + N_X(\bar{x}).
\]

Observe that

\[
N_X(\bar{x}) = \{ A^T \mu : \mu \geq 0, \quad h^T \bar{x} = 0 \}. \tag{2.39}
\]

The last two relations are equivalent to conditions (2.37). \( \square \)

Conditions (2.37) can also be obtained directly from the optimality conditions for the large scale linear programming formulation

\[
\begin{align*}
\min_{x, y_1, \ldots, y_K} & \quad c^T x + \sum_{k=1}^{K} p_k y_k^T y_k \\
\text{s.t.} & \quad T_k x + W_k y_k = h_k, \quad k = 1, \ldots, K, \\
& \quad A x = b, \\
& \quad y_k \geq 0, \quad k = 1, \ldots, K. 
\end{align*} \tag{2.40}
\]

By minimizing, with respect to \( x \geq 0 \) and \( y_k \geq 0, \quad k = 1, \ldots, K \), the Lagrangian

\[
\begin{align*}
& c^T x + \sum_{k=1}^{K} p_k q_k^T y_k - \mu^T (A x - b) - \sum_{k=1}^{K} p_k \pi_k^T (T_k x + W_k y_k - h_k) = \\
& \quad \left( c - A^T \mu - \sum_{k=1}^{K} p_k T_k^T \pi_k \right)^T x + \sum_{k=1}^{K} p_k \left( q_k - W_k^T \pi_k \right)^T y_k + b^T \mu + \sum_{k=1}^{K} p_k h_k^T \pi_k,
\end{align*}
\]

we obtain the following dual of the linear programming problem (2.40):

\[
\begin{align*}
\max_{\mu, \pi_1, \ldots, \pi_K} & \quad b^T \mu + \sum_{k=1}^{K} p_k h_k^T \pi_k \\
\text{s.t.} & \quad c - A^T \mu - \sum_{k=1}^{K} p_k T_k^T \pi_k \geq 0, \\
& \quad q_k - W_k^T \pi_k \geq 0, \quad k = 1, \ldots, K.
\end{align*}
\]

Therefore optimality conditions of Theorem 2.10 can be written in the following equivalent
Chapter 2. Two Stage Problems

form

$$\sum_{k=1}^{K} p_k T_k^T \pi_k + A^T \mu \leq c,$$

$$\bar{x}^T \left( c - \sum_{k=1}^{K} p_k T_k^T \pi_k - A^T \mu \right) = 0,$$

$$q_k - W_k^T \pi_k \geq 0, \quad k = 1, \ldots, K,$$

$$\bar{y}_k^T \left( q_k - W_k^T \pi_k \right) = 0, \quad k = 1, \ldots, K.$$

The last two of the above conditions correspond to feasibility and optimality of multipliers $\pi_k$ as solutions of the dual problems.

If we deal with general distributions of problem’s data, additional conditions are needed, to ensure the subdifferentiability of the expected recourse cost, and the existence of Lagrange multipliers.

Theorem 2.11. Let $\bar{x}$ be a feasible solution of problem (2.1)–(2.2). Suppose that the expected recourse cost function $\phi(\cdot)$ is proper, $\text{int}(\text{dom } \phi) \cap X$ is nonempty and $\mathcal{N}_{\text{dom } \phi}(\bar{x}) \subseteq \mathcal{N}_X(\bar{x})$. Then $\bar{x}$ is an optimal solution of problem (2.1)–(2.2) iff there exist a measurable function $\pi(\omega) \in D(x, \xi(\omega))$, $\omega \in \Omega$, and a vector $\mu \in \mathbb{R}^m$ such that

$$E[T^T \pi] + A^T \mu \leq c,$$

$$\bar{x}^T \left( c - E[T^T \pi] - A^T \mu \right) = 0.$$

Proof. Since $\text{int}(\text{dom } \phi) \cap X$ is nonempty, we have by the Moreau-Rockafellar Theorem that

$$\partial \left( c^T \bar{x} + \phi(\bar{x}) + I_X(\bar{x}) \right) = c + \partial \phi(\bar{x}) + \partial I_X(\bar{x}).$$

Also $\partial I_X(\bar{x}) = \mathcal{N}_X(\bar{x})$. Therefore we have here that (2.38) are necessary and sufficient optimality conditions for minimizing $c^T x + \phi(x)$ over $x \in X$. Using the characterization of the subdifferential of $\phi(\cdot)$ given in (2.8), we conclude that (2.38) is equivalent to existence of a measurable function $\pi(\omega) \in \mathcal{D}(x_0, \xi(\omega))$, such that

$$0 \in c - E[T^T \pi] + \mathcal{N}_{\text{dom } \phi}(\bar{x}) + \mathcal{N}_X(\bar{x}). \quad (2.41)$$

Moreover, because of the condition $\mathcal{N}_{\text{dom } \phi}(\bar{x}) \subseteq \mathcal{N}_X(\bar{x})$, the term $\mathcal{N}_{\text{dom } \phi}(\bar{x})$ can be omitted. The proof can be completed now by using (2.41) together with formula (2.39) for the normal cone $\mathcal{N}_X(\bar{x})$.

The additional technical condition $\mathcal{N}_{\text{dom } \phi}(\bar{x}) \subseteq \mathcal{N}_X(\bar{x})$ was needed in the above derivations in order to eliminate the term $\mathcal{N}_{\text{dom } \phi}(\bar{x})$ in (2.41). In particular, this condition holds if $\bar{x} \in \text{int}(\text{dom } \phi)$, in which case $\mathcal{N}_{\text{dom } \phi}(\bar{x}) = \{0\}$, or in case of relatively complete recourse, i.e., when $X \subset \text{dom } \phi$. If the condition of relatively complete recourse is not satisfied, we may need to take into account the normal cone to the domain of $\phi(\cdot)$. In general, this requires application of techniques of functional analysis, which are beyond
2.1. Linear Two Stage Problems

the scope of this book. However, in the special case of a deterministic matrix $T$ we can carry out the analysis directly.

**Theorem 2.12.** Let $\bar{x}$ be a feasible solution of problem (2.1)–(2.2). Suppose that the assumptions of Proposition 2.7 are satisfied, $\text{int}(\phi) \cap X$ is nonempty and the matrix $T$ is deterministic. Then $\bar{x}$ is an optimal solution of problem (2.1)–(2.2) iff there exist a measurable function $\pi(\omega) \in \mathcal{D}(x, \xi(\omega)), \omega \in \Omega$, and a vector $\mu \in \mathbb{R}^m$ such that

$$
T^T \mathbb{E}[\pi] + A^T \mu \leq c,
$$

$$
\bar{x}^T (c - T^T \mathbb{E}[\pi] - A^T \mu) = 0.
$$

**Proof.** Since $T$ is deterministic, we have that $\mathbb{E}[T^T \pi] = T^T \mathbb{E}[\pi]$, and hence the optimality conditions (2.41) can be written as

$$
0 \in c - T^T \mathbb{E}[\pi] + \mathcal{N}_{\text{dom } \phi}(\bar{x}) + \mathcal{N}_X(\bar{x}).
$$

Now we need to calculate the cone $\mathcal{N}_{\text{dom } \phi}(\bar{x})$. Recall that under the assumptions of Proposition 2.7 (in particular, that the recourse is fixed and $\Pi(q)$ is nonempty w.p.1), we have that $\phi(\cdot) > -\infty$ and formula (2.30) holds true. We have here that only $q$ and $h$ are random while both matrices $W$ and $T$ are deterministic, and (2.30) simplifies to

$$
\text{dom } \phi = \left\{ x : -Tx \in \bigcap_{h \in \Sigma} (\{ -h + \text{pos } W \}) \right\},
$$

where $\Sigma$ is the support of the distribution of the random vector $h$. The tangent cone to $\text{dom } \phi$ at $\bar{x}$ has the form

$$
T_{\text{dom } \phi}(\bar{x}) = \left\{ d : -Td \in \bigcap_{h \in \Sigma} (\text{pos } W + \text{lin}(\{ -h + T\bar{x} \})) \right\}
$$

$$
= \left\{ d : -Td \in \text{pos } W + \bigcap_{h \in \Sigma} \text{lin}(\{ -h + T\bar{x} \}) \right\}.
$$

Defining the linear subspace

$$
L := \bigcap_{h \in \Sigma} \text{lin}(\{ -h + T\bar{x} \}),
$$

we can write the tangent cone as

$$
T_{\text{dom } \phi}(\bar{x}) = \{ d : -Td \in \text{pos } W + L \}.
$$

Therefore the normal cone equals

$$
\mathcal{N}_{\text{dom } \phi}(\bar{x}) = \{ -T^T v : v \in (\text{pos } W + L)^* \} = -T^T [(\text{pos } W)^* \cap L^\perp].
$$

Here we used the fact that $\text{pos } W$ is polyhedral and no interior condition is needed for calculating $(\text{pos } W + L)^*$. Recalling equation (2.11) we conclude that

$$
\mathcal{N}_{\text{dom } \phi}(\bar{x}) = -T^T (\Pi_0 \cap L^\perp).
$$
Chapter 2. Two Stage Problems

Observe that if \( \nu \in \Pi_0 \cap L^\perp \) then \( \nu \) is an element of the recession cone of the set \( \mathcal{D}(\bar{x}, \xi) \), for all \( \xi \in \Xi \). Thus \( \pi(\omega) + \nu \) is also an element of the set \( \mathcal{D}(x, \xi(\omega)) \), for almost all \( \omega \in \Omega \). Consequently

\[
- T^T \mathbb{E} \left[ \mathcal{D}(\bar{x}, \xi) \right] + \mathcal{N}_{\text{dom } \phi}(\bar{x}) = - T^T \mathbb{E} \left[ \mathcal{D}(\bar{x}, \xi) \right] - T^T (\Pi_0 \cap L^\perp) = - T^T \mathbb{E} \left[ \mathcal{D}(\bar{x}, \xi) \right],
\]

and the result follows.

**Example 2.13 (Capacity Expansion (continued))** Let us return to Example 2.13 and suppose the support \( \Xi \) of the random demand vector \( \xi \) is compact. Only the right hand side \( \xi \) in the second stage problem (2.19)–(2.21) is random and for a sufficiently large \( x \) the second stage problem is feasible for all \( \xi \in \Xi \). Thus conditions of Theorem 2.11 are satisfied. It follows from Theorem 2.11 that \( \bar{x} \) is an optimal solution of problem (2.23) iff there exist measurable functions \( \mu_n(\xi) \), \( n \in \mathcal{N} \), such that for all \( \xi \in \Xi \) we have

\[
\mu(\xi) \in \mathcal{M}(\bar{x}, \xi),
\]

and for all \( (i, j) \in \mathcal{A} \) the following conditions are satisfied:

\[
c_{ij} \geq \int \max \{0, \mu_i(\xi) - \mu_j(\xi) - q_{ij}\} \; P(d\xi),
\]

(2.42)

\[
(\bar{x}_{ij} - x_{ij}^{\min}) \left( c_{ij} - \int \max \{0, \mu_i(\xi) - \mu_j(\xi) - q_{ij}\} \; P(d\xi) \right) = 0.
\]

(2.43)

In particular, for every \( (i, j) \in \mathcal{A} \) such that \( \bar{x}_{ij} > x_{ij}^{\min} \) we have equality in equation (2.42). Each function \( \mu_n(\xi) \) can be interpreted as a random potential of node \( n \in \mathcal{N} \).

### 2.2 Polyhedral Two-Stage Problems

#### 2.2.1 General Properties

Let us consider a slightly more general formulation of a two-stage stochastic programming problem:

\[
\min_x f_1(x) + \mathbb{E}[Q(x, \omega)],
\]

(2.44)

where \( Q(x, \omega) \) is the optimal value of the second stage problem

\[
\min_y f_2(y, \omega)
\]

s.t. \( T(\omega)x + W(\omega)y = h(\omega) \).

(2.45)

We assume in this section that the above two-stage problem is *polyhedral*. That is, the following holds.

- The function \( f_1(\cdot) \) is *polyhedral* (compare with Definition 7.1). This means that there exist vectors \( c_j \) and scalars \( \alpha_j, j = 1, \ldots, J_1 \), vectors \( a_k \) and scalars \( b_k, k =

2.2. Polyhedral Two-Stage Problems

1, . . . , $K_1$, such that $f_1(x)$ can be represented as follows:

$$f_1(x) = \begin{cases} \max_{1 \leq j \leq J_1} \alpha_j + c_j^T x, & \text{if } a_k^T x \leq b_k, \quad k = 1, \ldots, K_1, \\ +\infty, & \text{otherwise,} \end{cases}$$

and its domain $\text{dom} f_1 = \{ x : a_k^T x \leq b_k, \quad k = 1, \ldots, K_1 \}$ is nonempty. (Note that any polyhedral function is convex and lower semicontinuous.)

- The function $f_2$ is random polyhedral. That is, there exist random vectors $q_j = q_j(\omega)$ and random scalars $\gamma_j = \gamma_j(\omega), \ j = 1, \ldots, J_2$, random vectors $d_k = d_k(\omega)$ and random scalars $r_k = r_k(\omega), \ k = 1, \ldots, K_2$, such that $f_2(y, \omega)$ can be represented as follows:

$$f_2(y, \omega) = \begin{cases} \max_{1 \leq j \leq J_2} \gamma_j(\omega) + q_j(\omega)^T y, & \text{if } d_k(\omega)^T y \leq r_k(\omega), \quad k = 1, \ldots, K_2, \\ +\infty, & \text{otherwise,} \end{cases}$$

and for a.e. $\omega$ the domain of $f_2(\cdot, \omega)$ is nonempty.

Note that (linear) constraints of the second stage problem which are independent of $x$, as for example $y \geq 0$, can be absorbed into the objective function $f_2(y, \omega)$. Clearly, the linear two-stage model (2.1)–(2.2) is a special case of a polyhedral two-stage problem. The converse is also true, that is, every polyhedral two-stage model can be re-formulated as a linear two-stage model. For example, the second stage problem (2.45) can be written as follows:

$$\begin{align*}
\text{Min} & \quad v \\
\text{s.t.} & \quad T(\omega)x + W(\omega)y = h(\omega), \\
& \quad \gamma_j(\omega) + q_j(\omega)^T y \leq v, \quad j = 1, \ldots, J_2, \\
& \quad d_k(\omega)^T y \leq r_k(\omega), \quad k = 1, \ldots, K_2.
\end{align*}$$

Here both $v$ and $y$ play the role of the second stage variables, and the data $(q, T, W, h)$ in (2.2) have to be re-defined in an appropriate way. In order to avoid all these manipulations and unnecessary notational complications that come together with such a conversion, we shall address polyhedral problems in a more abstract way. This will also help us to deal with multistage problems and general convex problems.

Consider the Lagrangian of the second stage problem (2.45):

$$L(y, \pi; x, \omega) := f_2(y, \omega) + \pi^T (h(\omega) - T(\omega)x - W(\omega)y).$$

We have

$$\begin{align*}
\inf_y L(y, \pi; x, \omega) &= \pi^T (h(\omega) - T(\omega)x) + \inf_y [f_2(y, \omega) - \pi^T W(\omega)y] \\
&= \pi^T (h(\omega) - T(\omega)x) - f_2^* (W(\omega)^T \pi, \omega),
\end{align*}$$
where $f^*_{2}(\cdot, \omega)$ is the conjugate\(^4\) of $f_{2}(\cdot, \omega)$. We obtain that the dual of problem (2.45) can be written as follows

$$
\text{Max}_{\pi} \left[ \pi^T (h(\omega) - T(\omega)x) - f^*_{2}(W(\omega)^T \pi, \omega) \right].
$$

(2.46)

By the duality theory of linear programming, if, for some $(x, \omega)$, the optimal value $Q(x, \omega)$ of problem (2.45) is less than $+\infty$ (i.e., problem (2.45) is feasible), then it is equal to the optimal value of the dual problem (2.46).

Let us denote, as before, by $D(x, \omega)$ the set of optimal solutions of the dual problem (2.46). We then have an analogue of Proposition 2.2.

**Proposition 2.14.** Let $\omega \in \Omega$ be given and suppose that $Q(\cdot, \omega)$ is finite in at least one point $\bar{x}$. Then the function $Q(\cdot, \omega)$ is polyhedral (and hence convex). Moreover, $Q(\cdot, \omega)$ is subdifferentiable at every $x$ at which the value $Q(x, \omega)$ is finite, and

$$
\partial Q(x, \omega) = -T(\omega)^T D(x, \omega).
$$

(2.47)

**Proof.** Let us define the function $\psi(\pi) := f^*_{2}(W^T \pi)$ (for simplicity we suppress the argument $\omega$). We have that if $Q(x, \omega)$ is finite, then it is equal to the optimal value of problem (2.46), and hence $Q(x, \omega) = \psi^*(h - Tx)$. Therefore $Q(\cdot, \omega)$ is a polyhedral function. Moreover, it follows by the Fenchel–Moreau Theorem that

$$
\partial \psi^*(h - Tx) = D(x, \omega),
$$

and the chain rule for subdifferentiation yields formula (2.47). Note that we do not need here additional regularity conditions because of the polyhedricity of the considered case.

If $Q(x, \omega)$ is finite, then the set $D(x, \omega)$ of optimal solutions of problem (2.46) is a nonempty convex closed polyhedron. If, moreover, $D(x, \omega)$ is bounded, then it is the convex hull of its finitely many vertices (extreme points), and $Q(\cdot, \omega)$ is finite in a neighborhood of $x$. If $D(x, \omega)$ is unbounded, then its recession cone (which is polyhedral) is the normal cone to the domain of $Q(\cdot, \omega)$ at the point $x$.

**2.2.2 Expected Recourse Cost**

Let us consider the expected value function $\phi(x) := E[Q(x, \omega)]$. Suppose that the probability measure $P$ has a finite support, i.e., there exists a finite number of scenarios $\omega_k$ with respective (positive) probabilities $p_k$, $k = 1, \ldots, K$. Then

$$
E[Q(x, \omega)] = \sum_{k=1}^{K} p_k Q(x, \omega_k).
$$

\(^4\)Note that since $f_{2}(\cdot, \omega)$ is polyhedral, so is $f^*_{2}(\cdot, \omega)$. 
2.2. Polyhedral Two-Stage Problems

For a given $x$, the expectation $\mathbb{E}[Q(x, \omega)]$ is equal to the optimal value of the problem

$$\min_{y_1, \ldots, y_K} \sum_{k=1}^{K} p_k f_2(y_k, \omega_k)$$

$$\text{s.t. } T_k x + W_k y_k = h_k, \ k = 1, \ldots, K,$$

where $(h_k, T_k, W_k) := (h(\omega_k), T(\omega_k), W(\omega_k))$. Similarly to the linear case, if for at least one $k \in \{1, \ldots, K\}$ the set

$$\text{dom } f_2(\cdot, \omega_k) \cap \{y : T_k x + W_k y = h_k\}$$

is empty, i.e., the corresponding second stage problem is infeasible, then problem (2.48) is infeasible, and hence its optimal value is $+\infty$.

**Proposition 2.15.** Suppose that the probability measure $P$ has a finite support and that the expectation function $\phi(\cdot) := \mathbb{E}[Q(\cdot, \omega)]$ has a finite value in at least one point $x \in \mathbb{R}^n$. Then the function $\phi(\cdot)$ is polyhedral, and for any $x_0 \in \text{dom } \phi$,

$$\partial \phi(x_0) = \sum_{k=1}^{K} p_k \partial Q(x_0, \omega_k).$$

(2.49)

The proof is identical to the proof of Proposition 2.3. Since the functions $Q(\cdot, \omega_k)$ are polyhedral, formula (2.49) follows by the Moreau–Rockafellar Theorem.

The subdifferential $\partial Q(x_0, \omega_k)$ of the second stage optimal value function is described in Proposition 2.14. That is, if $Q(x_0, \omega_k)$ is finite, then

$$\partial Q(x_0, \omega_k) = -T_k^T \arg \max \{ \pi^T (h_k - T_k x_0) - f_2^*(W_k^T \pi, \omega_k) \}.$$  

(2.50)

It follows that the expectation function $\phi$ is differentiable at $x_0$ iff for every $\omega_k, k = 1, \ldots, K$, the maximum at the right hand side of (2.50) is attained at a unique point, i.e., the corresponding second stage dual problem has a unique optimal solution.

Let us now consider the case of a general probability distribution $P$. We need to ensure that the expectation function $\phi(x) := \mathbb{E}[Q(x, \omega)]$ is well defined. General conditions are complicated, so we resort again to the case of fixed recourse.

We say that the two-stage polyhedral problem has fixed recourse if the matrix $W$ and the set $Y := \text{dom } f_2(\cdot, \omega)$ are fixed, i.e., do not depend on $\omega$. In that case,

$$f_2(y, \omega) = \begin{cases} \max_{1 \leq j \leq J_2} \gamma_j(\omega) + q_j(\omega)^T y, & \text{if } y \in Y, \\ +\infty, & \text{otherwise}. \end{cases}$$

Denote $W(Y) := \{ Wy : y \in Y \}$. Let $x$ be such that

$$h(\omega) - T(\omega)x \in W(Y) \quad \text{w.p.1.}$$

(2.51)

$^5$Note that since it is assumed that $f_2(\cdot, \omega)$ is polyhedral, it follows that the set $Y$ is nonempty and polyhedral.
This means that for a.e. \( \omega \) the system
\[
y \in \mathcal{Y}, \quad Wy = h(\omega) - T(\omega)x
\] has a solution. Let for some \( \omega_0 \in \Omega \), \( y_0 \) be a solution of the above system, i.e., \( y_0 \in \mathcal{Y} \) and \( h(\omega_0) - T(\omega_0)x = Wy_0 \). Since system (2.52) is defined by linear constraints, we have by Hoffman’s lemma that there exists a constant \( \kappa \) such that for almost all \( \omega \) we can find a solution \( \bar{y}(\omega) \) of the system (2.52) with
\[
\|\bar{y}(\omega) - y_0\| \leq \kappa\|h(\omega) - T(\omega)x - (h(\omega_0) - T(\omega_0)x)\|.
\]
Therefore the optimal value of the second stage problem can be bounded from above as follows:
\[
Q(x, \omega) \leq \max_{1 \leq j \leq J_2} \left\{ \gamma_j(\omega) + q_j(\omega)^T \bar{y}(\omega) \right\}
\leq Q(x, \omega_0) + \sum_{j=1}^{J_2} \left| \gamma_j(\omega) - \gamma_j(\omega_0) \right|
+ \kappa \sum_{j=1}^{J_2} \|q_j(\omega)\| \left( \|h(\omega) - h(\omega_0)\| + \|x\| \|T(\omega) - T(\omega_0)\| \right).
\] (2.53)

**Proposition 2.16.** Suppose that the recourse is fixed and \( E|\gamma_j| < +\infty \), \( E\|q_j\| \|h\| < +\infty \) and \( E\|q_j\| \|T\| < +\infty \), \( j = 1, \ldots, J_2 \). (2.54)

Consider a point \( x \in \mathbb{R}^n \). Then \( E|Q(x, \omega)| \) is finite if and only if condition (2.51) holds.

**Proof.** The proof uses (2.53), similarly to the proof of Proposition 2.6.

Let us now formulate conditions under which the expected recourse cost is bounded from below. Let \( \mathcal{C} \) be the recession cone of \( \mathcal{Y} \), and \( \mathcal{C}^* \) be its polar. Consider the conjugate function \( f_2^* (\cdot, \omega) \). It can be verified that
\[
\text{dom} f_2^* (\cdot, \omega) = \text{conv} \{ q_j(\omega), \ j = 1, \ldots, J_2 \} + \mathcal{C}^*.
\] (2.55)

Indeed, by the definition of the function \( f_2 (\cdot, \omega) \) and its conjugate, we have that \( f_2^* (z, \omega) \) is equal to the optimal value of the
\[
\begin{align*}
\text{Max} \ v \\
\text{s.t.} \ z^Ty - \gamma_j(\omega) - q_j(\omega)^Ty \geq v, \ j = 1, \ldots, J_2, \ y \in \mathcal{Y}.
\end{align*}
\]

Since it is assumed that the set \( \mathcal{Y} \) is nonempty, the above problem is feasible, and since \( \mathcal{Y} \) is polyhedral, it is linear. Therefore its optimal value is equal to the optimal value of its dual. In particular, its optimal value is less than \( +\infty \) iff the dual problem is feasible. Now the dual problem is feasible iff there exist \( \pi_j \geq 0, \ j = 1, \ldots, J_2 \), such that \( \sum_{j=1}^{J_2} \pi_j = 1 \) and
\[
\sup_{y \in \mathcal{Y}} y^T \left( z - \sum_{j=1}^{J_2} \pi_j q_j(\omega) \right) < +\infty.
\]
2.2. Polyhedral Two-Stage Problems

The last condition holds iff \( z - \sum_{j=1}^{J_2} \pi_j q_j(\omega) \in \mathcal{C}^* \), which completes the argument.

Let us define the set

\[
\Pi(\omega) := \{ \pi : W^T \pi \in \text{conv} \{ q_j(\omega), j = 1, \ldots, J_2 \} + \mathcal{C}^* \}.
\]

We may remark that in the case of a linear two stage problem the above set coincides with the one defined in (2.5).

**Proposition 2.17.** Suppose that: (i) the recourse is fixed, (ii) the set \( \Pi(\omega) \) is nonempty w.p.1, (iii) condition (2.54) holds.

Then the expectation function \( \phi(x) \) is well defined and \( \phi(x) > -\infty \) for all \( x \in \mathbb{R}^n \). Moreover, \( \phi \) is convex, lower semicontinuous and Lipschitz continuous on \( \text{dom} \phi \), its domain \( \text{dom} \phi \) is a convex closed subset of \( \mathbb{R}^n \) and

\[
\text{dom} \phi = \{ x \in \mathbb{R}^n : h - Tx \in W(\mathcal{Y}) \ \text{w.p.1} \}.
\]  \hspace{1cm} (2.56)

Furthermore, for any \( x_0 \in \text{dom} \phi \),

\[
\partial \phi(x_0) = -E[T^T D(x_0, \omega)] + N_{\text{dom} \phi}(x_0),
\]  \hspace{1cm} (2.57)

**Proof.** Note that the dual problem (2.46) is feasible iff \( W^T \pi \in \text{dom} f^*_2(\cdot, \omega) \). By formula (2.55) assumption (ii) means that problem (2.46) is feasible, and hence \( Q(x, \omega) \) is equal to the optimal value of (2.46), for a.e. \( \omega \). The remainder of the proof is similar to the linear case (Proposition 2.7 and Proposition 2.8). \( \square \)

2.2.3 Optimality Conditions

The optimality conditions for polyhedral two-stage problems are similar to those for linear problems. For completeness we provide the appropriate formulations. Let us start from the problem with finitely many elementary events \( \omega_k \) occurring with probabilities \( p_k, k = 1, \ldots, K \).

**Theorem 2.18.** Suppose that the probability measure \( P \) has a finite support. Then a point \( \bar{x} \) is an optimal solution of the first stage problem (2.44) iff there exist \( \pi_k \in \mathcal{D}(\bar{x}, \omega_k), \) \( k = 1, \ldots, K \), such that

\[
0 \in \partial f_1(\bar{x}) - \sum_{k=1}^{K} p_k T^T_k \pi_k.
\]  \hspace{1cm} (2.58)

**Proof.** Since \( f_1(x) \) and \( \phi(x) = E[Q(x, \omega)] \) are convex functions, a necessary and sufficient condition for a point \( \bar{x} \) to be a minimizer of \( f_1(x) + \phi(x) \) reads

\[
0 \in \partial[f_1(\bar{x}) + \phi(\bar{x})].
\]  \hspace{1cm} (2.59)

In particular, the above condition requires for \( f_1(\bar{x}) \) and \( \phi(\bar{x}) \) to be finite valued. By the Moreau–Rockafellar Theorem we have that \( \partial[f_1(\bar{x}) + \phi(\bar{x})] = \partial f_1(\bar{x}) + \partial \phi(\bar{x}) \). Note
that there is no need here for additional regularity conditions because of the polyhedricity of functions \( f_1 \) and \( \phi \). The proof can be completed now by using formula for \( \partial \phi(\bar{x}) \) given in Proposition 2.15.

In the case of general distributions, the derivation of optimality conditions requires additional assumptions.

**Theorem 2.19.** Suppose that: (i) the recourse is fixed and relatively complete, (ii) the set \( \Pi(\omega) \) is nonempty w.p.1, (iii) condition (2.54) holds.

Then a point \( \bar{x} \) is an optimal solution of problem (2.44)–(2.45) iff there exists a measurable function \( \pi(\omega) \in D(\bar{x}, \omega) \), \( \omega \in \Omega \), such that

\[
0 \in \partial f_1(\bar{x}) - \mathbb{E}[T^T \pi].
\]  

(2.60)

**Proof.** The result follows immediately from the optimality condition (2.59) and formula (2.57). Since the recourse is relatively complete, we can omit the normal cone to the domain of \( \phi(\cdot) \).

If the recourse is not relatively complete, the analysis becomes complicated. The normal cone to the domain of \( \phi(\cdot) \) enters the optimality conditions. For the domain described in (2.56) this cone is rather difficult to describe in a closed form. Some simplification can be achieved when \( T \) is deterministic. The analysis then mirrors the linear case, as in Theorem 2.12.

### 2.3 General Two-Stage Problems

#### 2.3.1 Problem Formulation, Interchangeability

In a general way two-stage stochastic programming problems can be written in the following form

\[
\min_{x \in X} \left\{ f(x) := \mathbb{E}[F(x, \omega)] \right\},
\]

(2.61)

where \( F(x, \omega) \) is the optimal value of the second stage problem

\[
\min_{y \in \mathcal{G}(x, \omega)} g(x, y, \omega).
\]

(2.62)

Here \( X \subset \mathbb{R}^n \), \( g : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \rightarrow \mathbb{R} \) and \( \mathcal{G} : \mathbb{R}^n \times \Omega \rightrightarrows \mathbb{R}^m \) is a multifunction. In particular, the linear two-stage problem (2.1)–(2.2) can be formulated in the above form with \( g(x, y, \omega) := c^T x + q(\omega)^T y \) and

\[
\mathcal{G}(x, \omega) := \{ y : T(\omega)x + W(\omega)y = h(\omega), \ y \geq 0 \}.
\]

We also use notation \( g_\omega(x, y) = g(x, y, \omega) \) and \( \mathcal{G}_\omega(x) = \mathcal{G}(x, \omega) \).

Of course, the second stage problem (2.62) can be also written in the following equivalent form

\[
\min_{y \in \mathbb{R}^m} \bar{g}(x, y, \omega),
\]

(2.63)
2.3. General Two-Stage Problems

where

\[
\bar{g}(x, y, \omega) := \begin{cases} 
g(x, y, \omega), & \text{if } y \in G(x, \omega) \\
+\infty, & \text{otherwise.} \end{cases}
\]  

(2.64)

We assume that the function \( \bar{g}(x, y, \omega) \) is random lower semicontinuous. Recall that if \( g(x, y, \cdot) \) is measurable for every \( (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \) and \( g(\cdot, \cdot, \omega) \) is continuous for a.e. \( \omega \in \Omega \), i.e., \( g(x, y, \omega) \) is a Carathéodory function, then \( g(x, y, \omega) \) is random lsc. Random lower semicontinuity of \( \bar{g}(x, y, \omega) \) implies that the optimal value function \( F(x, \cdot) \) is measurable (see Theorem 7.37). Moreover, if for a.e. \( \omega \in \Omega \) function \( F(\cdot, \omega) \) is continuous, then \( F(x, \omega) \) is a Carathéodory function, and hence is random lsc. The indicator function \( \mathbb{I}_{\bar{g}_y(x)}(y) \) is random lsc if for every \( \omega \in \Omega \) the multifunction \( \bar{g}_y(\cdot) \) is closed and \( \bar{g}(x, \omega) \) is measurable with respect to the sigma algebra of \( \mathbb{R}^n \times \Omega \) (see Theorem 7.36). Of course, if \( g(x, y, \omega) \) and \( \mathbb{I}_{\bar{g}_y(x)}(y) \) are random lsc, then their sum \( \bar{g}(x, y, \omega) \) is also random lsc.

Now let \( \mathfrak{F} \) be a linear decomposable space of measurable mappings from \( \Omega \) to \( \mathbb{R}^m \). For example, we can take \( \mathfrak{F} := L_p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m) \) with \( p \in [1, +\infty] \). Then by the interchangeability principle we have

\[
\mathbb{E} \left[ \inf_{y \in \mathfrak{F}} \bar{g}(x, y, \omega) \right] = \inf_{y \in \mathfrak{F}} \mathbb{E} \left[ \bar{g}(x, y(\omega), \omega) \right],
\]  

(2.65)

provided that the right hand side of (2.65) is less than \(+\infty\) (see Theorem 7.80). This implies the following interchangeability principle for two-stage programming.

**Theorem 2.20.** The two-stage problem (2.61)–(2.62) is equivalent to the following problem:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n, y \in \mathfrak{F}} & \quad \mathbb{E} [g(x, y(\omega), \omega)] \\
\text{s.t.} & \quad x \in X, \ y(\omega) \in G(x, \omega) \ \text{a.e.} \ \omega \in \Omega.
\end{align*}
\]  

(2.66)

The equivalence is understood in the sense that optimal values of problems (2.61) and (2.66) are equal to each other; provided that the optimal value of problem (2.66) is less than \(+\infty\). Moreover, assuming that the common optimal value of problems (2.61) and (2.66) is finite, we have at if \( (\bar{x}, \bar{y}) \) is an optimal solution of problem (2.66), then \( \bar{x} \) is an optimal solution of the first stage problem (2.61) and \( \bar{y} = y(\omega) \) is an optimal solution of the second stage problem (2.62) for \( x = \bar{x} \) and a.e. \( \omega \in \Omega \); conversely, if \( \bar{x} \) is an optimal solution of the first stage problem (2.61) and for \( x = \bar{x} \) and a.e. \( \omega \in \Omega \) the second stage problem (2.62) has an optimal solution \( \bar{y} = y(\omega) \) such that \( \bar{y} \in \mathfrak{F} \), then \( (\bar{x}, \bar{y}) \) is an optimal solution of problem (2.66).

Note that optimization in the right hand side of (2.65) and in (2.66) is performed over mappings \( y : \Omega \to \mathbb{R}^m \) belonging to the space \( \mathfrak{F} \). In particular, if \( \Omega = \{\omega_1, \ldots, \omega_K\} \) is finite, then by setting \( y_k := y(\omega_k), k = 1, \ldots, K \), every such mapping can be identified with a vector \( (y_1, \ldots, y_K) \) and the space \( \mathfrak{F} \) with the finite dimensional space \( \mathbb{R}^{mK} \). In that
case problem (2.66) takes the form (compare with (2.15)):

$$\min_{x,y_1,\ldots,y_K} \sum_{k=1}^{K} p_k g(x, y_k, \omega_k)$$

s.t. \(x \in X\), \(y_k \in G(x, \omega_k)\), \(k = 1, \ldots, K\). \hfill (2.67)

### 2.3.2 Convex Two-Stage Problems

We say that the two-stage problem (2.61)–(2.62) is convex if the set \(X\) is convex (and closed) and for every \(\omega \in \Omega\) the function \(\hat{g}(x, y, \omega)\), defined in (2.64), is convex in \((x, y) \in \mathbb{R}^n \times \mathbb{R}^m\). We leave this as an exercise to show that in such case the optimal value function \(F(\cdot, \omega)\) is convex, and hence (2.61) is a convex problem. It could be useful to understand what conditions will guarantee convexity of the function \(\hat{g}(x, y, \omega)\). We have that \(\hat{g}(x, y, \omega) \equiv g(x, y) + I_{\Omega}(\omega)\). Therefore \(\hat{g}(x, y, \omega)\) is convex if \(g(x, y)\) is convex and the indicator function \(I_{\Omega}(\omega)\) is convex in \((x, y)\). It is not difficult to see that the indicator function \(I_{\Omega}(\omega)\) is convex iff the following condition holds for any \(t \in [0, 1]\):

$$y \in \mathcal{G}_\omega(x), \quad y' \in \mathcal{G}_\omega(x') \Rightarrow ty + (1-t)y' \in \mathcal{G}_\omega(tx + (1-t)x').$$

Equivalently this condition can be written as

$$t\mathcal{G}_\omega(x) + (1-t)\mathcal{G}_\omega(x') \subseteq \mathcal{G}_\omega((tx + (1-t)x'), \quad \forall x, x' \in \mathbb{R}^n, \forall t \in [0, 1]. \hfill (2.68)$$

Multifunction \(\mathcal{G}_\omega\) satisfying the above condition (2.69) is called convex. By taking \(x = x'\) we obtain that if the multifunction \(\mathcal{G}_\omega\) is convex, then it is convex valued, i.e., the set \(\mathcal{G}_\omega(x)\) is convex for every \(x \in \mathbb{R}^n\).

In the remainder of this section we assume that the multifunction \(\mathcal{G}(x, \omega)\) is defined in the following form

$$\mathcal{G}(x, \omega) := \{y \in Y : T(x, \omega) + W(y, \omega) \in -C\}, \hfill (2.70)$$

where \(Y\) is a nonempty convex closed subset of \(\mathbb{R}^m\) and \(T = (t_1, \ldots, t_\ell) : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}^\ell\), \(W = (w_1, \ldots, w_\ell) : \mathbb{R}^m \times \Omega \rightarrow \mathbb{R}^\ell\) and \(C \subseteq \mathbb{R}^\ell\) is a closed convex cone. Cone \(C\) defines a partial order, denoted “\(\preceq\)”, on the space \(\mathbb{R}^\ell\). That is, \(a \preceq b\) iff \(b - a \in C\). In that notation the constraint \(T(x, \omega) + W(y, \omega) \in -C\) can be written as \(T(x, \omega) + W(y, \omega) \preceq 0\). For example, if \(C := \mathbb{R}_+^\ell\), then the constraint \(T(x, \omega) + W(y, \omega) \preceq 0\) means that \(t_i(x, \omega) + w_i(y, \omega) \leq 0\), \(i = 1, \ldots, \ell\). We assume that \(t_i(x, \omega)\) and \(w_i(y, \omega), i = 1, \ldots, \ell\), are Carathéodory functions and that for every \(\omega \in \Omega\), mappings \(T_\omega(\cdot) = T(\cdot, \omega)\) and \(W_\omega(\cdot) = W(\cdot, \omega)\) are convex with respect to the cone \(C\). A mapping \(G : \mathbb{R}^n \rightarrow \mathbb{R}^\ell\) is said to be convex with respect to \(C\) if the multifunction \(M(x) := G(x) + C\) is convex. Equivalently, mapping \(G\) is convex with respect to \(C\) if

$$G(tx + (1-t)x') \preceq CG(x) + (1-t)G(x') \forall x, x' \in \mathbb{R}^n, \forall t \in [0, 1].$$

For example, mapping \(G(\cdot) = (g_1(\cdot), \ldots, g_\ell(\cdot))\) is convex with respect to \(C := \mathbb{R}_+^\ell\) iff all its components \(g_i(\cdot), i = 1, \ldots, \ell\), are convex functions. Convexity of \(T_\omega\) and \(W_\omega\) implies convexity of the corresponding multifunction \(\mathcal{G}_\omega\).
2.3. General Two-Stage Problems

We assume, further, that \( g(x, y, \omega) := c(x) + q(y, \omega) \), where \( c(\cdot) \) and \( q(\cdot, \omega) \) are real valued convex functions. For \( G(x, \omega) \) of the form (2.70), and given \( x \), we can write the second stage problem, up to the constant \( c(x) \), in the form

\[
\begin{align*}
\min_{y \in Y} & \quad q_\omega(y) \\
\text{s.t.} & \quad W_\omega(y) + \chi_\omega \preceq C_0
\end{align*}
\]

(2.71)

with \( \chi_\omega := T(x, \omega) \). Let us denote by \( \vartheta(\chi, \omega) \) the optimal value of problems (2.71). Note that \( F(x, \omega) = c(x) + \vartheta(T(x, \omega), \omega) \). The (Lagrangian) dual of problem (2.71) can be written in the form

\[
\max_{\pi \succeq 0} \left\{ \pi^T \chi_\omega + \inf_{y \in Y} L_\omega(y, \pi) \right\},
\]

(2.72)

where

\[
L_\omega(y, \pi) := q_\omega(y) + \pi^T W_\omega(y)
\]

is the Lagrangian of problem (2.71). We have the following results (see Theorems 7.8 and 7.9).

**Proposition 2.21.** Let \( \omega \in \Omega \) and \( \chi_\omega \) be given and suppose that the specified above convexity assumptions are satisfied. Then the following statements hold true:

(i) The functions \( \vartheta(\cdot, \omega) \) and \( F(\cdot, \omega) \) are convex.

(ii) Suppose that problem (2.71) is subconsistent. Then there is no duality gap between problem (2.71) and its dual (2.72) if and only if the optimal value function \( \vartheta(\cdot, \omega) \) is lower semicontinuous at \( \chi_\omega \).

(iii) There is no duality gap between problems (2.71) and (2.72) and the dual problem (2.72) has a nonempty set of optimal solutions if and only if the optimal value function \( \vartheta(\cdot, \omega) \) is subdifferentiable at \( \chi_\omega \).

(iv) Suppose that the optimal value of (2.71) is finite. Then there is no duality gap between problems (2.71) and (2.72) and the dual problem (2.72) has a nonempty and bounded set of optimal solutions if and only if \( \chi_\omega \in \text{int}(\text{dom} \vartheta(\cdot, \omega)) \).

The regularity condition \( \chi_\omega \in \text{int}(\text{dom} \vartheta(\cdot, \omega)) \) means that for all small perturbations of \( \chi_\omega \) the corresponding problem (2.71) remains feasible.

We can also characterize the differentiability properties of the optimal value functions in terms of the dual problem (2.72). Let us denote by \( \mathcal{D}(\chi, \omega) \) the set of optimal solutions of the dual problem (2.72). This set may be empty, of course.

**Proposition 2.22.** Let \( \omega \in \Omega \), \( x \in \mathbb{R}^n \) and \( \chi = T(x, \omega) \) be given. Suppose that the specified convexity assumptions are satisfied and that problems (2.71) and (2.72) have finite and equal optimal values. Then

\[
\partial \vartheta(\chi, \omega) = \mathcal{D}(\chi, \omega).
\]

(2.73)
Suppose, further, that functions \( c(\cdot) \) and \( T_\omega(\cdot) \) are differentiable, and

\[
0 \in \text{int}\{T_\omega(x) + \nabla T_\omega(x)\mathbb{R}^n - \text{dom } \vartheta(\cdot, \omega)\}. \tag{2.74}
\]

Then

\[
\partial F(x, \omega) = \nabla c(x) + \nabla T_\omega(x)\mathbb{D}(\chi, \omega). \tag{2.75}
\]

**Corollary 2.23.** Let \( \omega \in \Omega, x \in \mathbb{R}^n \) and \( \chi = T(x, \omega) \) and suppose that the specified convexity assumptions are satisfied. Then \( \vartheta(\cdot, \omega) \) is differentiable at \( \chi \) if and only if \( \mathbb{D}(\chi, \omega) \) is a singleton. Suppose, further, that the functions \( f(\cdot) \) and \( T_\omega(\cdot) \) are differentiable. Then the function \( F(\cdot, \omega) \) is differentiable at every \( x \) at which \( \mathbb{D}(\chi, \omega) \) is a singleton.

**Proof.** If \( \mathbb{D}(\chi, \omega) \) is a singleton, then the set of optimal solutions of the dual problem (2.72) is nonempty and bounded, and hence there is no duality gap between problems (2.71) and (2.72). Thus formula (2.73) holds. Conversely, if \( \partial \vartheta(\chi, \omega) \) is a singleton and hence is nonempty, then again there is no duality gap between problems (2.71) and (2.72), and hence formula (2.73) holds.

Now if \( \mathbb{D}(\chi, \omega) \) is a singleton, then \( \vartheta(\cdot, \omega) \) is continuous at \( \chi \) and hence the regularity condition (2.74) holds. It follows then by formula (2.75) that \( F(\cdot, \omega) \) is differentiable at \( x \) and formula

\[
\nabla F(x, \omega) = \nabla c(x) + \nabla T_\omega(x)\mathbb{D}(\chi, \omega) \tag{2.76}
\]

holds true. \( \Box \)

Let us focus on the expectation function \( f(x) := \mathbb{E}[F(x, \omega)] \). If the set \( \Omega \) is finite, say \( \Omega = \{\omega_1, \ldots, \omega_K\} \) with corresponding probabilities \( p_k, k = 1, \ldots, K, \) then \( f(x) = \sum_{k=1}^K p_k F(x, \omega_k) \) and subdifferentiability of \( f(x) \) is described by the Moreau-Rockafellar Theorem (Theorem 7.4) together with formula (2.75). In particular, \( f(\cdot) \) is differentiable at a point \( x \) if the functions \( c(\cdot) \) and \( T_\omega(\cdot) \) are differentiable at \( x \) and for every \( \omega \in \Omega \) the corresponding dual problem (2.72) has a unique optimal solution.

Let us consider the general case, when \( \Omega \) is not assumed to be finite. By combining Proposition 2.22 and Theorem 7.47 we obtain that, under appropriate regularity conditions ensuring for a.e. \( \omega \in \Omega \) formula (2.75) and interchangeability of the subdifferential and expectation operators, it follows that \( f(\cdot) \) is subdifferentiable at a point \( \bar{x} \in \text{dom } f \) and

\[
\partial f(\bar{x}) = \nabla c(\bar{x}) + \int_\Omega \nabla T_\omega(\bar{x})\mathbb{D}(T_\omega(\bar{x}), \omega) dP(\omega) + N_{\text{dom } f}(\bar{x}). \tag{2.77}
\]

In particular, it follows from the above formula (2.77) that \( f(\cdot) \) is differentiable at \( \bar{x} \) iff \( \bar{x} \in \text{int}(\text{dom } f) \) and \( \mathbb{D}(T_\omega(\bar{x}), \omega) = \{\pi(\omega)\} \) is a singleton w.p.1, in which case

\[
\nabla f(\bar{x}) = \nabla c(\bar{x}) + \mathbb{E} [\nabla T_\omega(\bar{x})\mathbb{D}(\chi, \omega)\]. \tag{2.78}
\]

We obtain the following conditions for optimality.

**Proposition 2.24.** Let \( \bar{x} \in X \cap \text{int}(\text{dom } f) \) and assume that formula (2.77) holds. Then \( \bar{x} \) is an optimal solution of the first stage problem (2.61) iff there exists a measurable selection
2.4 Nonanticipativity

π(ω) ∈ D(T(¯x, ω), ω) such that
\[ -c(\bar{x}) - \mathbb{E}[\nabla T_\omega(\bar{x})^T \pi(\omega)] \in N_X(\bar{x}). \] (2.79)

**Proof.** Since \( \bar{x} \in X \cap \text{int}(\text{dom } f) \), we have that \( \text{int}(\text{dom } f) \neq \emptyset \) and \( \bar{x} \) is an optimal solution if \( 0 \in \partial f(\bar{x}) + N_X(\bar{x}) \). By formula (2.77) and since \( \bar{x} \in \text{int}(\text{dom } f) \), this is equivalent to condition (2.79). \[ \square \]

### 2.4 Nonanticipativity

#### 2.4.1 Scenario Formulation

An additional insight into the structure and properties of two-stage problems can be gained by introducing the concept of nonanticipativity. Consider the first stage problem (2.61). Assume that the number of scenarios is finite, i.e., \( \Omega = \{\omega_1, \ldots, \omega_K\} \) with respective (positive) probabilities \( p_1, \ldots, p_K \). Let us relax the first stage problem by replacing vector \( x \) with \( K \) vectors \( x_1, x_2, \ldots, x_K \), one for each scenario. We obtain the following relaxation of problem (2.61):

\[
\text{Min} \sum_{k=1}^{K} p_k F(x_k, \omega_k) \quad \text{subject to } x_k \in X, \quad k = 1, \ldots, K. \quad (2.80)
\]

We observe that problem (2.80) is separable in the sense that it can be split into \( K \) smaller problems, one for each scenario:

\[
\text{Min} F(x_k, \omega_k), \quad k = 1, \ldots, K, \quad (2.81)
\]

and that the optimal value of problem (2.80) is equal to the weighted sum, with weights \( p_k \), of the optimal values of problems (2.81), \( k = 1, \ldots, K \). For example, in the case of two-stage linear program (2.15), relaxation of the form (2.80) leads to solving \( K \) smaller problems

\[
\text{Min} x_k \geq 0, y_k \geq 0 \quad c^T x_k + q_k^T y_k \quad \text{s.t. } Ax_k = b, \quad T_k x_k + W_k y_k = h_k.
\]

Problem (2.80), however, is not suitable for modeling a two stage decision process. This is because the first stage decision variables \( x_k \) in (2.80) are now allowed to depend on a realization of the random data at the second stage. This can be fixed by introducing the additional constraint

\[
(x_1, \ldots, x_K) \in \mathcal{L}, \quad (2.82)
\]

where \( \mathcal{L} := \{x = (x_1, \ldots, x_K) : x_1 = \ldots = x_K\} \) is a linear subspace of the \( nK \)-dimensional vector space \( \mathcal{X} := \mathbb{R}^n \times \cdots \times \mathbb{R}^n \). Due to the constraint (2.82), all realizations \( x_k, k = 1, \ldots, K, \) of the first stage decision vector are equal to each other, that is, they do not depend on the realization of the random data. The constraint (2.82) can be written in different forms, which can be convenient in various situations, and will be referred to as the
Chapter 2. Two Stage Problems

nonanticipativity constraint. Together with the nonanticipativity constraint (2.82), problem (2.80) becomes

\[
\begin{align*}
\text{Min} & \quad \sum_{k=1}^{K} p_k F(x_k, \omega_k) \\
\text{s.t.} & \quad x_1 = \cdots = x_K, \quad x_k \in X, \quad k = 1, \ldots, K.
\end{align*}
\] (2.83)

Clearly, the above problem (2.83) is equivalent to problem (2.61). Such nonanticipativity constraints are especially important in multistage modeling which we discuss later.

A way to write the nonanticipativity constraint is to require that

\[
x_k = \sum_{i=1}^{K} p_i x_i, \quad k = 1, \ldots, K,
\] (2.84)

which is convenient for extensions to the case of a continuous distribution of problem data. Equations (2.84) can be interpreted in the following way. Consider the space \( X \) equipped with the scalar product

\[
\langle x, y \rangle := \sum_{i=1}^{K} p_i x_i^T y_i.
\] (2.85)

Define linear operator \( P : X \to X \) as

\[
P x := \left( \sum_{i=1}^{K} p_i x_i, \ldots, \sum_{i=1}^{K} p_i x_i \right).
\]

Constraint (2.84) can be compactly written as

\[
x = P x.
\]

It can be verified that \( P \) is the orthogonal projection operator of \( X \), equipped with the scalar product (2.85), onto its subspace \( L \). Indeed, \( P(P x) = P x \), and

\[
\langle P x, y \rangle = \left( \sum_{i=1}^{K} p_i x_i \right)^T \left( \sum_{k=1}^{K} p_k y_k \right) = \langle x, Py \rangle.
\] (2.86)

The range space of \( P \), which is the linear space \( L \), is called the nonanticipativity subspace of \( X \).

Another way to algebraically express nonanticipativity, which is convenient for numerical methods, is to write the system of equations

\[
\begin{align*}
x_1 &= x_2, \\
x_2 &= x_3, \\
&\vdots \\
x_{K-1} &= x_K.
\end{align*}
\] (2.87)

This system is very sparse: each equation involves only two variables, and each variable appears in at most two equations, which is convenient for many numerical solution methods.
2.4. Nonanticipativity

2.4.2 Dualization of Nonanticipativity Constraints

We discuss now a dualization of problem (2.80) with respect to the nonanticipativity constraints (2.84). Assigning to these nonanticipativity constraints Lagrange multipliers $\lambda_k \in \mathbb{R}^n$, $k = 1, \ldots, K$, we can write the Lagrangian

$$L(x, \lambda) := \sum_{k=1}^{K} p_k F(x_k, \omega_k) + \sum_{k=1}^{K} p_k \lambda_k^T \left( x_k - \sum_{i=1}^{K} p_i x_i \right).$$

Note that since $P$ is an orthogonal projection, $I - P$ is also an orthogonal projection (onto the space orthogonal to $\mathcal{L}$), and hence

$$\sum_{k=1}^{K} p_k \lambda_k^T \left( x_k - \sum_{i=1}^{K} p_i x_i \right) = \langle \lambda, (I - P)x \rangle = \langle (I - P)\lambda, x \rangle.$$

Therefore the above Lagrangian can be written in the following equivalent form

$$L(x, \lambda) = \sum_{k=1}^{K} p_k F(x_k, \omega_k) + \sum_{k=1}^{K} p_k \left( \lambda_k - \sum_{j=1}^{K} p_j \lambda_j \right)^T x_k.$$

Let us observe that shifting the multipliers $\lambda_k$, $k = 1, \ldots, K$, by a constant vector does not change the value of the Lagrangian, because the expression $\lambda_k - \sum_{j=1}^{K} p_j \lambda_j$ is invariant to such shifts. Therefore, with no loss of generality we can assume that

$$\sum_{j=1}^{K} p_j \lambda_j = 0.$$

or, equivalently, that $P \lambda = 0$. Dualization of problem (2.80) with respect to the nonanticipativity constraints takes the form of the following problem

$$\max_{\lambda} \left\{ D(\lambda) := \inf_x L(x, \lambda) \right\} \text{ subject to } P \lambda = 0. \tag{2.88}$$

By general duality theory we have that the optimal value of problem (2.61) is greater than or equal to the optimal value of problem (2.88). These optimal values are equal to each other under some regularity conditions, we will discuss a general case in the next section. In particular, if the two stage problem is linear and since the nonanticipativity constraints are linear, we have in that case that there is no duality gap between problem (2.61) and its dual problem (2.88) unless both problems are infeasible.

Let us take a closer look at the dual problem (2.88). Under the condition $P \lambda = 0$, the Lagrangian can be written simply as

$$L(x, \lambda) = \sum_{k=1}^{K} p_k \left( F(x_k, \omega_k) + \lambda_k^T x_k \right).$$

We see that the Lagrangian can be split into $K$ components:

$$L(x, \lambda) = \sum_{k=1}^{K} p_k L_k(x_k, \lambda_k),$$
where $L_k(x_k, \lambda_k) := F(x_k, \omega_k) + \lambda_k^T x_k$. It follows that

$$D(\lambda) = \sum_{j=1}^{K} p_k D_k(\lambda_k),$$

where

$$D_k(\lambda_k) := \inf_{x_k \in X} L_k(x_k, \lambda_k).$$

For example, in the case of the two-stage linear program (2.15), $D_k(\lambda_k)$ is the optimal value of the problem

$$\min_{x_k, y_k} (c + \lambda_k)^T x_k + q_k^T y_k$$

subject to

$$Ax_k = b,$$

$$T_k x_k + W_k y_k = h_k,$$

$$x_k \geq 0, \quad y_k \geq 0.$$  

We see that value of the dual function $D(\lambda)$ can be calculated by solving $K$ independent scenario subproblems.

Suppose that there is no duality gap between problem (2.61) and its dual (2.88) and their common optimal value is finite. This certainly holds true if the problem is linear, and both problems, primal and dual, are feasible. Let $\bar{\lambda} = (\bar{\lambda}_1, \ldots, \bar{\lambda}_K)$ be an optimal solution of the dual problem (2.88). Then the set of optimal solutions of problem (2.61) is contained in the set of optimal solutions of the problem

$$\min_{x_k \in X} \sum_{k=1}^{K} p_k L_k(x_k, \bar{\lambda}_k)$$

(2.89)

This inclusion can be strict, i.e., the set of optimal solutions of (2.89) can be larger than the set of optimal solutions of problem (2.61) (see an example of linear program defined in (7.32)). Of course, if problem (2.89) has unique optimal solution $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_K)$, then $\bar{x} \in X$, i.e., $\bar{x}_1 = \ldots = \bar{x}_K$, and this is also the optimal solution of problem (2.61) with $\bar{x}$ being equal to the common value of $\bar{x}_1, \ldots, \bar{x}_K$. Note also that the above problem (2.89) is separable, i.e., $\bar{x}$ is an optimal solution of (2.89) iff for every $k = 1, \ldots, K$, $\bar{x}_k$ is an optimal solution of the problem

$$\min_{x_k \in X} L_k(x_k, \bar{\lambda}_k).$$

### 2.4.3 Nonanticipativity Duality for General Distributions

In this section we discuss dualization of the first stage problem (2.61) with respect to nonanticipativity constraints in general (not necessarily finite scenarios) case. For the sake of convenience we write problem (2.61) in the form

$$\min_{x \in \mathbb{R}^n} \{ \bar{f}(x) := \mathbb{E}[F(x, \omega)] \},$$

(2.90)

where $F(x, \omega) := F(x, \omega) + \mathbb{I}_X(x)$, i.e., $\bar{F}(x, \omega) = F(x, \omega)$ if $x \in X$ and $\bar{F}(x, \omega) = +\infty$ otherwise. Let $X$ be a linear decomposable space of measurable mappings from $\Omega$ to $\mathbb{R}^n$. 

2.4. Nonanticipativity

Unless stated otherwise we use \( X := \mathcal{L}_p(\Omega, \mathcal{F}, P; \mathbb{R}^n) \) for some \( p \in [1, +\infty) \) such that for every \( x \in X \) the expectation \( \mathbb{E}[\hat{F}(x(\omega), \omega)] \) is well defined. Then we can write problem (2.90) in the following equivalent form

\[
\min_{x \in X} \mathbb{E}[\hat{F}(x(\omega), \omega)],
\]  

(2.91)

where \( \mathcal{L} \) is a linear subspace of \( X \) formed by mappings \( x : \Omega \to \mathbb{R}^n \) which are constant almost everywhere, i.e.,

\[
\mathcal{L} := \{ x \in X : x(\omega) \equiv x \text{ for some } x \in \mathbb{R}^n \},
\]

where \( x(\omega) \equiv x \) means that \( x(\omega) = x \) for a.e. \( \omega \in \Omega \).

Consider the dual\footnote{Recall that \( 1/p + 1/q = 1 \) for \( p, q \in (1, +\infty) \). If \( p = 1 \), then \( q = +\infty \). Also for \( p = +\infty \) we use \( q = 1 \). This results in a certain abuse of notation since the space \( X = \mathcal{L}_\infty(\Omega, \mathcal{F}, P; \mathbb{R}^n) \) is not reflexive and \( X^* = \mathcal{L}_1(\Omega, \mathcal{F}, P; \mathbb{R}^n) \) is smaller than its dual. Note also that if \( x \in \mathcal{L}_p(\Omega, \mathcal{F}, P; \mathbb{R}^n) \), then its expectation \( \mathbb{E}[x] = \int_\Omega x(\omega)dP(\omega) \) is well defined and is an element of vector space \( \mathbb{R}^n \).} \( X^* := \mathcal{L}_q(\Omega, \mathcal{F}, P; \mathbb{R}^n) \) of the space \( X \) and define the scalar product (bilinear form)

\[
\langle \lambda, x \rangle := \mathbb{E} \left[ \lambda^T x \right] = \int \lambda(\omega)^T x(\omega) dP(\omega), \quad \lambda \in X^*, \quad x \in X.
\]

Also, consider the projection operator \( P : X \to \mathcal{L} \) defined as \( [Px](\omega) \equiv \mathbb{E}[x] \). Clearly the space \( \mathcal{L} \) is formed by such \( x \in X \) that \( Px = x \). Note that

\[
\langle \lambda, Px \rangle = \mathbb{E} \left[ \lambda^T x \right] = \langle \lambda^*, \lambda \rangle,
\]

where \( P^* \) is a projection operator \( [P^* \lambda](\omega) \equiv \mathbb{E}[\lambda] \) from \( X^* \) onto its subspace formed by constant almost everywhere mappings. In particular, if \( p = 2 \), then \( X^* = X \) and \( P^* = P \).

With problem (2.91) is associated the following Lagrangian

\[
L(x, \lambda) := \mathbb{E}[\hat{F}(x(\omega), \omega)] + \mathbb{E} \left[ \lambda^T (x - \mathbb{E}[x]) \right].
\]

Note that

\[
\mathbb{E} \left[ \lambda^T (x - \mathbb{E}[x]) \right] = \langle \lambda, x - Px \rangle = \langle \lambda - P^* \lambda, x \rangle,
\]

and \( \lambda - P^* \lambda \) does not change by adding a constant to \( \lambda(\cdot) \). Therefore we can set \( P^* \lambda = 0 \), in which case

\[
L(x, \lambda) = \mathbb{E} \left[ \hat{F}(x(\omega), \omega) + \lambda(\omega)^T x(\omega) \right] \quad \text{for } \mathbb{E}[\lambda] = 0.
\]

(2.92)

This leads to the following dual of problem (2.90):

\[
\max_{\lambda \in X^*} \left\{ D(\lambda) := \inf_{x \in X} L(x, \lambda) \right\} \quad \text{subject to } \mathbb{E}[\lambda] = 0.
\]

(2.93)

In case of finitely many scenarios the above dual is the same as the dual problem (2.88).
By the interchangeability principle (Theorem 7.80) we have
\[ \inf_{x \in X} \mathbb{E} \left[ F(x(\omega), \omega) + \lambda(\omega)^T x(\omega) \right] = \mathbb{E} \left[ \inf_{x \in \mathbb{R}^n} (\bar{F}(x, \omega) + \lambda(\omega)^T x) \right]. \]
Consequently
\[ D(\lambda) = \mathbb{E}[D_\omega(\lambda(\omega))], \]
where \( D_\omega : \mathbb{R}^n \rightarrow \mathbb{R} \) is defined as
\[ D_\omega(\lambda) := \inf_{x \in \mathbb{R}^n} (\lambda^T x + \bar{F}_\omega(x)) = -\sup_{x \in \mathbb{R}^n} (-\lambda^T x + \bar{F}_\omega(x)) = -\bar{F}_\omega^*(-\lambda). \quad (2.94) \]
That is, in order to calculate the dual function \( D(\lambda) \) one needs to solve for every \( \omega \in \Omega \) the finite dimensional optimization problem (2.94), and then to integrate the optimal values obtained.

By the general theory we have that the optimal value of problem (2.91), which is the same as the optimal value of problem (2.90), is greater than or equal to the optimal value of its dual (2.93). We also have that there is no duality gap between problem (2.91) and its dual (2.93) and both problems have optimal solutions \( \bar{x} \) and \( \bar{\lambda} \), respectively, iff \( (\bar{x}, \bar{\lambda}) \) is a saddle point of the Lagrangian defined in (2.92). By definition a point \( (\bar{x}, \bar{\lambda}) \in X \times \mathbb{R}^n \) is a saddle point of the Lagrangian iff
\[ \bar{x} \in \arg \min_{x \in \mathbb{R}^n} L(x, \bar{\lambda}) \text{ and } \bar{\lambda} \in \arg \max_{\lambda \in \mathbb{R}^n} L(\bar{x}, \lambda). \quad (2.95) \]
By the interchangeability principle (see equation (7.253) of Theorem 7.80) we have that first condition in (2.95) can be written in the following equivalent form
\[ \bar{x}(\omega) \equiv \bar{x} \text{ and } \bar{x} \in \arg \min_{x \in \mathbb{R}^n} \{ F(x, \omega) + \bar{\lambda}(\omega)^T x \} \text{ a.e. } \omega \in \Omega. \quad (2.96) \]
Since \( \bar{x}(\omega) \equiv \bar{x} \), the second condition in (2.95) means that \( \mathbb{E}[\bar{\lambda}] = 0 \).

Let us assume now that the considered problem is convex, i.e., the set \( X \) is convex (and closed) and \( F_\omega(\cdot) \) is a convex function for a.e. \( \omega \in \Omega \). It follows that \( \bar{F}_\omega(\cdot) \) is a convex function for a.e. \( \omega \in \Omega \). Then the second condition in (2.96) holds iff \( \bar{\lambda}(\omega) \in -\partial \bar{F}_\omega(\bar{x}) \) for a.e. \( \omega \in \Omega \). Together with condition \( \mathbb{E}[\bar{\lambda}] = 0 \) this means that
\[ 0 \in \mathbb{E} \left[ \partial \bar{F}_\omega(\bar{x}) \right]. \quad (2.97) \]
It follows that the Lagrangian has a saddle point iff there exists \( \bar{x} \in \mathbb{R}^n \) satisfying condition (2.97). We obtain the following result.

**Theorem 2.25.** Suppose that the function \( F(x, \omega) \) is random lower semicontinuous, the set \( X \) is convex and closed and for a.e. \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is convex. Then there is no duality gap between problems (2.90) and (2.93) and both problems have optimal solutions iff there exists \( \bar{x} \in \mathbb{R}^n \) satisfying condition (2.97). In that case \( \bar{x} \) is an optimal solution of (2.90) and a measurable selection \( \bar{\lambda}(\omega) \in -\partial \bar{F}_\omega(\bar{x}) \) such that \( \mathbb{E}[\bar{\lambda}] = 0 \) is an optimal solution of (2.93).
Recall that the inclusion \( \mathbb{E} \left[ \partial \tilde{F}_{\omega}(\bar{x}) \right] \subset \partial \tilde{f}(\bar{x}) \) always holds (see equation (7.125) in the proof of Theorem 7.47). Therefore, condition (2.97) implies that \( 0 \in \partial \tilde{f}(\bar{x}) \), which in turn implies that \( \bar{x} \) is an optimal solution of (2.90). Conversely, if \( \bar{x} \) is an optimal solution of (2.90), then \( 0 \in \partial \tilde{f}(\bar{x}) \), and if in addition \( \mathbb{E} \left[ \partial \tilde{F}_{\omega}(\bar{x}) \right] = \partial \tilde{f}(\bar{x}) \), then (2.97) follows. Therefore, Theorems 2.25 and 7.47 imply the following result.

**Theorem 2.26.** Suppose that: (i) the function \( F(x, \omega) \) is random lower semicontinuous, (ii) the set \( X \) is convex and closed, (iii) for \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is convex, (iv) problem (2.90) possesses an optimal solution \( \bar{x} \) such that \( \bar{x} \in \text{int}(\text{dom} f) \). Then there is no duality gap between problems (2.90) and (2.93) and the dual problem (2.93) has an optimal solution \( \lambda \), and the constant mapping \( \bar{x}(\omega) \equiv \bar{x} \) is an optimal solution of the problem

\[
\min_{x \in X} \mathbb{E} \left[ \tilde{F}(x(\omega), \omega) + \lambda(\omega)^{\top} x(\omega) \right].
\]

**Proof.** Since \( \bar{x} \) is an optimal solution of problem (2.90) we have that \( \bar{x} \in X \) and \( f(\bar{x}) \) is finite. Moreover, since \( \bar{x} \in \text{int}(\text{dom} f) \) and \( f \) is convex, it follows that \( f \) is proper and \( \text{N}_{\text{dom} f}(\bar{x}) = \{ 0 \} \). Therefore, it follows by Theorem 7.47 that \( \mathbb{E} \left[ \partial \tilde{F}_{\omega}(\bar{x}) \right] = \partial \tilde{f}(\bar{x}) \). Furthermore, since \( \bar{x} \in \text{int}(\text{dom} f) \) we have that \( \partial \tilde{f}(\bar{x}) = \partial f(\bar{x}) + \text{N}_{X}(\bar{x}) \), and hence \( \mathbb{E} \left[ \partial \tilde{F}_{\omega}(\bar{x}) \right] = \partial f(\bar{x}) \). By optimality of \( \bar{x} \), we also have that \( 0 \in \partial f(\bar{x}) \). Consequently, \( 0 \in \mathbb{E} \left[ \partial \tilde{F}_{\omega}(\bar{x}) \right] \), and hence the proof can be completed by applying Theorem 2.25. \( \square \)

If \( X \) is a subset of \( \text{int}(\text{dom} f) \), then any point \( x \in X \) is an interior point of \( \text{dom} f \). In that case condition (iv) of the above theorem is reduced to existence of an optimal solution. The condition \( X \subset \text{int}(\text{dom} f) \) means that \( f(x) < +\infty \) for every \( x \) in a neighborhood of the set \( X \). This requirement is slightly stronger than the condition of relatively complete recourse.

**Example 2.27 (Capacity Expansion (continued))** Let us consider the capacity expansion problem of Examples 2.4 and 2.13. Suppose that \( \bar{x} \) is the optimal first stage decision and let \( \bar{y}_{ij}(\xi) \) be the corresponding optimal second stage decisions. The scenario problem has the form

\[
\min_{(i, j) \in \mathcal{A}} \sum_{(i, j) \in \mathcal{A}} \left[ (c_{ij} + \lambda_{ij}(\xi))x_{ij} + q_{ij}y_{ij} \right]
\]

s.t.

\[
\sum_{(i, j) \in \mathcal{A}_+(n)} y_{ij} - \sum_{(i, j) \in \mathcal{A}_-(n)} y_{ij} = \xi_n, \quad n \in \mathcal{N},
\]

\[
0 \leq y_{ij} \leq x_{ij}, \quad (i, j) \in \mathcal{A}.
\]

From Example 2.13 we know that there exist random node potentials \( \mu_n(\xi), n \in \mathcal{N} \), such that for all \( \xi \in \Xi \) we have \( \mu(\xi) \in \mathcal{M}(\bar{x}, \xi) \), and conditions (2.42)–(2.43) are satisfied. Also, the random variables \( q_{ij}(\xi) = -\max \{ 0, \mu_i(\xi) - \mu_j(\xi) - q_{ij} \} \) are the corresponding subgradients of the second stage cost. Define

\[
\lambda_{ij}(\xi) = \max \{ 0, \mu_i(\xi) - \mu_j(\xi) - q_{ij} \} - \int_{\Xi} \max \{ 0, \mu_i(\xi) - \mu_j(\xi) - q_{ij} \} P(d\xi), \quad (i, j) \in \mathcal{A}.
\]
We can easily verify that $x_{ij}(\xi) = \bar{x}_{ij}$ and $\bar{y}_{ij}(\xi), (i, j) \in \mathcal{A}$, are optimal solution of the scenario problem, because the first term of $\lambda_{ij}$ cancels with the subgradient $g_{ij}(\xi)$, while the second term satisfies the optimality conditions (2.42)–(2.43). Moreover, $\mathbb{E}[\lambda] = 0$ by construction.

### 2.4.4 Value of Perfect Information

Consider the following relaxation of the two-stage problem (2.61)–(2.62):

$$\min_{x \in \mathcal{X}} \mathbb{E}[\bar{F}(x(\omega), \omega)]. \quad (2.98)$$

This relaxation is obtained by removing the nonanticipativity constraint from the formulation (2.91) of the first stage problem. By the interchangeability principle (Theorem 7.80) we have that the optimal value of the above problem (2.98) is equal to $\mathbb{E} \left[ \inf_{x \in \mathbb{R}^n} \bar{F}(x, \omega) \right]$. The value $\inf_{x \in \mathbb{R}^n} \bar{F}(x, \omega)$ is equal to the optimal value of problem

$$\min_{x \in \mathcal{X}, y \in \mathcal{G}(x, \omega)} g(x, y, \omega). \quad (2.99)$$

That is, the optimal value of problem (2.98) is obtained by solving problems of the form (2.99), one for each $\omega \in \Omega$, and then taking the expectation of the calculated optimal values.

Solving problems of the form (2.99) makes sense if we have a perfect information about the data, i.e., the scenario $\omega \in \Omega$ is known at the time when the first stage decision should be made. The problem (2.99) is deterministic, e.g., in the case of two-stage linear program (2.1)–(2.2) it takes the form

$$\min_{x \geq 0, y \geq 0} c^T x + q^T y \text{ subject to } Ax = b, Tx + Wy = h.$$ 

An optimal solution of the second stage problem (2.99) depends on $\omega \in \Omega$ and is called the wait-and-see solution.

We have that for any $x \in \mathcal{X}$ and $\omega \in \Omega$, the inequality $F(x, \omega) \geq \inf_{x \in \mathcal{X}} F(x, \omega)$ clearly holds, and hence $\mathbb{E}[F(x, \omega)] \geq \mathbb{E} \left[ \inf_{x \in \mathcal{X}} F(x, \omega) \right]$. It follows that

$$\inf_{x \in \mathcal{X}} \mathbb{E}[F(x, \omega)] \geq \mathbb{E} \left[ \inf_{x \in \mathcal{X}} F(x, \omega) \right]. \quad (2.100)$$

Another way to view the above inequality is to observe that problem (2.98) is a relaxation of the corresponding two-stage stochastic problem, which of course implies (2.100).

Suppose that the two-stage problem has an optimal solution $\bar{x} \in \arg\min_{x \in \mathcal{X}} \mathbb{E}[F(x, \omega)]$. As $F(\bar{x}, \omega) - \inf_{x \in \mathcal{X}} F(x, \omega) \geq 0$ for all $\omega \in \Omega$, we conclude that

$$\mathbb{E}[F(\bar{x}, \omega)] = \mathbb{E} \left[ \inf_{x \in \mathcal{X}} F(x, \omega) \right] \quad (2.101)$$

iff $F(\bar{x}, \omega) = \inf_{x \in \mathcal{X}} F(x, \omega)$ w.p.1. That is, equality in (2.101) holds iff

$$F(\bar{x}, \omega) = \inf_{x \in \mathcal{X}} F(x, \omega) \text{ for a.e. } \omega \in \Omega. \quad (2.102)$$
Exercises

In particular, this happens if \( \bar{F}_\omega(x) \) has a minimizer independent of \( \omega \in \Omega \). This, of course, may happen only in rather specific situations.

The difference \( F(\bar{x}, \omega) - \inf_{x \in X} F(x, \omega) \) represents the value of perfect information of knowing \( \omega \). Consequently

\[
\text{EVPI} := \inf_{x \in X} E[F(x, \omega)] - E\left[ \inf_{x \in X} F(x, \omega) \right]
\]

is called the expected value of perfect information. It follows from \((2.100)\) that EVPI is always nonnegative and EVPI=0 iff condition \((2.102)\) holds.

Exercises

2.1. Consider the assembly problem discussed in section 1.3.1 in two cases:

(i) The demand which is not satisfied from the pre-ordered quantities of parts is lost.

(ii) All demand has to be satisfied, by making additional orders of the missing parts. In this case, the cost of each additionally ordered part \( j \) is \( r_j > c_j \).

For each of these cases describe the subdifferential of the recourse cost and of the expected recourse cost.

2.2. A transportation company has \( n \) depots among which they send cargo. The demand for transportation between depot \( i \) and depot \( j \neq i \) is modeled as a random variable \( D_{ij} \). The total capacity of vehicles currently available at depot \( i \) is denoted \( s_i \), \( i = 1, \ldots, n \). The company considers repositioning their fleet to better prepare to the uncertain demand. It costs \( c_{ij} \) to move a unit of capacity from location \( i \) to location \( j \). After repositioning, the realization of the random vector \( D \) is observed, and the demand is served, up to the limit determined by the transportation capacity available at each location. The profit from transporting a unit of cargo from location \( i \) to location \( j \) is equal \( q_{ij} \). If the total demand at location \( i \) exceeds the capacity available at location \( i \), the excessive demand is lost. It is up to the company to decide how much of each demand \( D_{ij} \) be served, and which part will remain unsatisfied. For simplicity, we consider all capacity and transportation quantities as continuous variables.

(a) Formulate the problem of maximizing the expected profit as a two stage stochastic programming problem.

(b) Describe the subdifferential of the recourse cost and of the expected recourse cost.

2.3. Show that the function \( s_q(\cdot) \), defined in \((2.4)\), is convex.

2.4. Consider the optimal value \( Q(x, \xi) \) of the second stage problem \((2.2)\). Show that \( Q(\cdot, \xi) \) is differentiable at a point \( x \) iff the dual problem \((2.3)\) has unique optimal solution \( \bar{\pi} \), in which case \( \nabla_x Q(x, \xi) = -T^T \bar{\pi} \).
2.5. Consider the two-stage problem (2.1)–(2.2) with fixed recourse. Show that the following conditions are equivalent: (i) problem (2.1)–(2.2) has complete recourse, (ii) the feasible set \( \Pi(q) \) of the dual problem is bounded for every \( q \), (iii) the system \( W^T \pi \leq 0 \) has only one solution \( \pi = 0 \).

2.6. Show that if random vector \( \xi \) has a finite support, then condition (2.24) is necessary and sufficient for relatively complete recourse.

2.7. Show that the conjugate function of a polyhedral function is also polyhedral.

2.8. Show that if \( Q(x, \omega) \) is finite, then the set \( D(x, \omega) \) of optimal solutions of problem (2.46) is a nonempty convex closed polyhedron.

2.9. Consider problem (2.63) and its optimal value \( F(x, \omega) \). Show that \( F(x, \omega) \) is convex in \( x \) if \( g(x, y, \omega) \) is convex in \( (x, y) \). Show that the indicator function \( I_G(x) \) is convex in \( (x, y) \) iff condition (2.68) holds for any \( t \in [0, 1] \).

2.10. Show that equation (2.86) implies that \( \langle x - Px, y \rangle = 0 \) for any \( x \in \mathcal{X} \) and \( y \in \mathcal{L} \), i.e., that \( P \) is the orthogonal projection of \( \mathcal{X} \) onto \( \mathcal{L} \).

2.11. Derive the form of the dual problem for the linear two stage stochastic programming problem in form (2.80) with nonanticipativity constraints (2.87).
3.1 Problem Formulation

3.1.1 The General Setting

The two-stage stochastic programming models can be naturally extended to a multistage setting. We already discussed examples of such decision processes in sections 1.2.3 and 1.4.2 for a multistage inventory model and a multistage portfolio selection problem, respectively. In the multistage setting the uncertain data $\xi_1, \ldots, \xi_T$ is revealed gradually over time, in $T$ periods, and our decisions should be adapted to this process. The decision process has the form:

$$
\text{decision }(x_1) \rightsquigarrow \text{observation }(\xi_2) \rightsquigarrow \text{decision }(x_2) \rightsquigarrow \ldots \rightsquigarrow \text{observation }(\xi_T) \rightsquigarrow \text{decision }(x_T).
$$

We view the sequence $\xi_t \in \mathbb{R}^{d_t}, t = 1, \ldots, T$, of data vectors as a stochastic process, i.e., as a sequence of random variables with a specified probability distribution. We use notation $\xi[t] := (\xi_1, \ldots, \xi_t)$ to denote the history of the process up to time $t$.

The values of the decision vector $x_t$, chosen at stage $t$, may depend on the information (data) $\xi[t]$ available up to time $t$, but not on the results of future observations. This is the basic requirement of nonanticipativity. As $x_t$ may depend on $\xi[t]$, the sequence of decisions is a stochastic process as well.

We say that the process $\{\xi_t\}$ is stagewise independent if $\xi_t$ is stochastically independent of $\xi_{[t-1]}, t = 2, \ldots, T$. It is said that the process is Markovian if for every $t = 2, \ldots, T$, the conditional distribution of $\xi_t$ given $\xi_{[t-1]}$ is the same as the conditional distribution of $\xi_t$ given $\xi_{t-1}$. Of course, if the process is stagewise independent, then it is Markovian.
As before, we often use the same notation \( \xi_t \) to denote a random vector and its particular realization. Which of these two meanings will be used in a particular situation will be clear from the context.

In a generic form a \( T \)-stage stochastic programming problem can be written in the following nested formulation

\[
\min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \mathbb{E} \left[ \inf_{x_2 \in \mathcal{X}_2(x_1, \xi_2)} f_2(x_2, \xi_2) + \mathbb{E} \left[ \cdots + \mathbb{E} \left[ \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} f_T(x_T, \xi_T) \right] \right] \right],
\]

(3.1)
driven by the random data process \( \xi_1, \xi_2, \ldots, \xi_T \). Here \( x_t \in \mathbb{R}^{n_t}, t = 1, \ldots, T \), are decision variables, \( f_t : \mathbb{R}^{n_t} \times \mathbb{R}^{d_t} \to \mathbb{R} \) are continuous functions and \( \mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \mathbb{R}^{d_t} \to \mathbb{R}^{n_t}, t = 2, \ldots, T \), are measurable closed valued multifunctions. The first stage data, i.e., the vector \( \xi_1 \), the function \( f_1 : \mathbb{R}^{n_1} \to \mathbb{R} \), and the set \( \mathcal{X}_1 \subset \mathbb{R}^{n_1} \) are deterministic. It is said that the multistage problem is linear if the objective functions and the constraint functions are linear. In a typical formulation,

\[
f_t(x_t, \xi_t) := c_t^T x_t, \quad \mathcal{X}_t := \{ x_t : A_t x_t = b_t, x_t \geq 0 \},
\]

\[
\mathcal{X}_t(x_{t-1}, \xi_t) := \{ x_t : B_t x_{t-1} + A_t x_t = b_t, x_t \geq 0 \}, \quad t = 2, \ldots, T,
\]

Here, \( \xi_1 := (c_1, A_1, b_1) \) is known at the first stage (and hence is nonrandom), and \( \xi_2 := (c_t, B_t, A_t, b_t) \in \mathbb{R}^{d_t}, t = 2, \ldots, T \), are data vectors\(^1\) some (or all) elements of which can be random.

There are several equivalent ways to make this formulation precise. One approach is to consider decision variables \( x_t = x_t(\xi_{[t]}), t = 1, \ldots, T \), as functions of the data process \( \xi_{[t]} \) up to time \( t \). Such a sequence of (measurable) mappings \( x_t : \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_t} \to \mathbb{R}^{n_t}, t = 1, \ldots, T \), is called an implementable policy (or simply a policy) (recall that \( \xi_1 \) is deterministic). An implementable policy is said to be feasible if it satisfies the feasibility constraints, i.e.,

\[
x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t), \quad t = 2, \ldots, T, \text{ w.p.1.} \tag{3.2}
\]

We can formulate the multistage problem (3.1) in the form

\[
\min_{x_{1:T}} \mathbb{E} \left[ f_1(x_1) + f_2(x_2(\xi_2)), \xi_2) + \cdots + f_T(x_T(\xi_T)), \xi_T) \right]
\]

(3.3)

\text{s.t.} \quad x_1 \in \mathcal{X}_1, \quad x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t), \quad t = 2, \ldots, T.

Note that optimization in (3.3) is performed over implementable and feasible policies and that policies \( x_2, \ldots, x_T \) are functions of the data process, and hence are elements of appropriate functional spaces, while \( x_1 \in \mathbb{R}^{n_1} \) is a deterministic vector. Therefore, unless the data process \( \xi_1, \ldots, \xi_T \) has a finite number of realizations, formulation (3.3) leads to an infinite dimensional optimization problem. This is a natural extension of the formulation (2.66) of the two-stage problem.

Another possible way is to write the corresponding dynamic programming equations. That is, consider the last stage problem

\[
\min_{x_T \in \mathcal{X}_T(x_{T-1}, \xi_T)} f_T(x_T, \xi_T).
\]

\(^1\)If data involves matrices, then their elements can be stacked columnwise to make it a vector.
The optimal value of this problem, denoted \( Q_T(x_{T-1}, \xi_T) \), depends on the decision vector \( x_{T-1} \) and data \( \xi_T \). At stage \( t = 2, \ldots, T - 1 \), we formulate the problem:

\[
\min_{x_t} f_t(x_t, \xi_t) + \mathbb{E} \{ Q_{t+1}(x_t, \xi_{t+1}) | \xi_t \}
\]

\[\text{s.t. } x_t \in \mathcal{X}_t(x_{t-1}, \xi_t),\]

where \( \mathbb{E} [ \cdot | \xi_t] \) denotes conditional expectation. Its optimal value depends on the decision \( x_{t-1} \) at the previous stage and realization of the data process \( \xi_t \), and denoted \( Q_t(x_{t-1}, \xi_t) \). The idea is to calculate the cost-to-go (or value) functions \( Q_t(x_{t-1}, \xi_t) \), recursively, going backward in time. At the first stage we finally need to solve the problem:

\[
\min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \mathbb{E} \{ Q_2(x_1, \xi_2) \}.
\]

The corresponding dynamic programming equations are

\[
Q_t(x_{t-1}, \xi_t) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \{ f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{t+1}) \},
\]

(3.4)

where

\[
Q_{t+1}(x_t, \xi_t) := \mathbb{E} \{ Q_{t+1}(x_t, \xi_{t+1}) | \xi_t \}.
\]

An implementable policy \( x_t(\xi_t) \) is optimal iff for \( t = 1, \ldots, T \),

\[
x_t(\xi_t) \in \arg \min_{x_t \in \mathcal{X}_t(x_{t-1}(\xi_{t-1}), \xi_t)} \{ f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{t+1}) \}, \quad \text{w.p.1},
\]

(3.5)

where for \( t = T \) the term \( Q_{T+1} \) is omitted and for \( t = 1 \) the set \( \mathcal{X}_1 \) depends only on \( \xi_1 \). In the dynamic programming formulation the problem is reduced to solving a family of finite dimensional problems, indexed by \( t \) and by \( \xi_t \). It can be viewed as an extension of the formulation (2.61)–(2.62) of the two-stage problem.

If the process \( \xi_1, \ldots, \xi_T \) is Markovian, then conditional distributions in the above equations, given \( \xi_{t-1} \), are the same as the respective conditional distributions given \( \xi_t \). In that case each cost-to-go function \( Q_t \) depends on \( \xi_t \) rather than the whole \( \xi_{t-1} \) and we can write it as \( Q_t(x_{t-1}, \xi_t) \). If, moreover, the stagewise independence condition holds, then each expectation function \( Q_t \) does not depend on realizations of the random process, and we can write it simply as \( Q_t(x_{t-1}) \).

### 3.1.2 The Linear Case

We discuss linear multistage problems in more detail. Let \( x_1, \ldots, x_T \) be decision vectors corresponding to time periods (stages) \( 1, \ldots, T \). Consider the following linear programming problem

\[
\begin{align*}
\min \quad & c_1^T x_1 + c_2^T x_2 + c_3^T x_3 + \ldots + c_T^T x_T \\
\text{s.t.} \quad & A_1 x_1 + B_2 x_1 + A_2 x_2 + B_3 x_2 + A_3 x_3 + \ldots + B_T x_{T-1} + A_T x_T = b_1, \\
& b_1 = b_2, \\
& b_2 = b_3, \quad \ldots \quad b_T, \\
& x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad \ldots \quad x_T \geq 0.
\end{align*}
\]

(3.6)
We can view this problem as a multiperiod stochastic programming problem where \( c_1, A_1 \) and \( b_1 \) are known, but some (or all) the entries of the cost vectors \( c_t \), matrices \( B_t \) and \( A_t \), and right hand side vectors \( b_t \), \( t = 2, \ldots, T \), are random. In the multistage setting the values (realizations) of the random data become known in the respective time periods (stages), and we have the following sequence of actions:

\[
\begin{align*}
\text{decision (} x_1 \text{)} \\
\text{observation } \xi_2 := (c_2, B_2, A_2, b_2) \\
\text{decision (} x_2 \text{)} \\
\vdots \\
\text{observation } \xi_T := (c_T, B_T, A_T, b_T) \\
\text{decision (} x_T \text{)}.
\end{align*}
\]

Our objective is to design the decision process in such a way that the expected value of the total cost is minimized while optimal decisions are allowed to be made at every time period \( t = 1, \ldots, T \).

Let us denote by \( \xi_t \) the data vector, realization of which becomes known at time period \( t \). In the setting of the multiperiod problem (3.6), \( \xi_t \) is assembled from the components of \( c_t, B_t, A_t, b_t \), some (or all) of which can be random, while the data \( \xi_1 = (c_1, A_1, b_1) \) at the first stage of problem (3.6) is assumed to be known. The important condition in the above multistage decision process is that every decision vector \( x_t \) may depend on the information available at time \( t \) (that is, \( \xi_t \)), but not on the results of observations to be made at later stages. This differs multistage stochastic programming problems from deterministic multiperiod problems, in which all the information is assumed to be available at the beginning of the decision process.

As it was outlined in section 3.1.1, there are several possible ways to formulate multistage stochastic programs in a precise mathematical form. In one such formulation \( x_t = x_t(\xi_t) \), \( t = 2, \ldots, T \), is viewed as a function of \( \xi_t \), and the minimization in (3.6) is performed over appropriate functional spaces of such functions. If the number of scenarios is finite, this leads to a formulation of the linear multistage stochastic program as one large (deterministic) linear programming problem. We discuss that further in section 3.1.4.

Another possible approach is to write dynamic programming equations, which we discuss next.

Let us look at our problem from the perspective of the last stage \( T \). At that time the values of all problem data, \( \xi_T \), are already known, and the values of the earlier decision vectors, \( x_1, \ldots, x_{T-1} \), have been chosen. Our problem is, therefore, a simple linear programming problem

\[
\begin{align*}
\text{Min } c_T^T x_T \\
\text{s.t. } B_T x_{T-1} + A_T x_T = b_T, \\
x_T \geq 0.
\end{align*}
\]

The optimal value of this problem depends on the earlier decision vector \( x_{T-1} \in \mathbb{R}^n \) and data \( \xi_T = (c_T, B_T, A_T, b_T) \), and is denoted by \( Q_T(x_{T-1}, \xi_T) \).
3.1. Problem Formulation

At stage $T-1$ we know $x_{T-2}$ and $\xi_{[T-1]}$. We face, therefore, the following stochastic programming problem

$$\begin{align*}
\text{Min} & \quad c_T^T x_{T-1} + E \left[ Q_T(x_{T-1}, \xi_T) \big| \xi_{[T-1]} \right] \\
\text{s.t.} & \quad B_{T-1} x_{T-2} + A_{T-1} x_{T-1} = b_{T-1}, \\
& \quad x_{T-1} \geq 0.
\end{align*}$$

The optimal value of the above problem depends on $x_{T-2} \in \mathbb{R}^{n_{T-2}}$ and data $\xi_{[T-1]}$, and is denoted $Q_{T-1}(x_{T-2}, \xi_{[T-1]})$.

Generally, at stage $t = 2, \ldots, T - 1$, we have the problem

$$\begin{align*}
\text{Min} & \quad c_t^T x_t + E \left[ Q_{t+1}(x_t, \xi_{[t+1]}) \big| \xi_t \right] \\
\text{s.t.} & \quad B_t x_{t-1} + A_t x_t = b_t, \\
& \quad x_t \geq 0.
\end{align*}$$

(3.7)

Its optimal value, called cost-to-go function, is denoted $Q_t(x_{t-1}, \xi_t)$.

On top of all these problems is the problem to find the first decisions, $x_1 \in \mathbb{R}^{n_1}$,

$$\begin{align*}
\text{Min} & \quad c_1^T x_1 + E \left[ Q_2(x_1, \xi_2) \right] \\
\text{s.t.} & \quad A_1 x_1 = b_1, \\
& \quad x_1 \geq 0.
\end{align*}$$

(3.8)

Note that all subsequent stages $t = 2, \ldots, T$ are absorbed in the above problem into the function $Q_2(x_1, \xi_2)$ through the corresponding expected values. Note also that since $\xi_1$ is not random, the optimal value $Q_2(x_1, \xi_2)$ does not depend on $\xi_1$. In particular, if $T = 2$, then (3.8) coincides with the formulation (2.1) of a two-stage linear problem.

The dynamic programming equations here take the form (compare with (3.4)):

$$Q_t(x_{t-1}, \xi_t) = \inf_{x_t} \left\{ c_t^T x_t + Q_{t+1}(x_t, \xi_{t+1}) : B_t x_{t-1} + A_t x_t = b_t, \ x_t \geq 0 \right\},$$

where

$$Q_{t+1}(x_t, \xi_{t+1}) := E \left[ Q_{t+1}(x_t, \xi_{t+1}) \big| \xi_t \right].$$

Also an implementable policy $\bar{x}_t(\xi_t)$ is optimal if for $t = 1, \ldots, T$ the condition

$$\bar{x}_t(\xi_t) = \arg \min_{x_t} \left\{ c_t^T x_t + Q_{t+1}(x_t, \xi_{t+1}) : A_t x_t = b_t - B_t \bar{x}_{t-1}(\xi_{t-1}), \ x_t \geq 0 \right\},$$

holds for almost every (a.e.) realization of the random process (for $t = T$ the term $Q_{T+1}$ is omitted and for $t = 1$ the term $B_t \bar{x}_{t-1}$ is omitted). If the process $\xi_t$ is Markovian, then each cost-to-go function depends on $\xi_t$ rather than $\xi_{[t]}$, and we can simply write $Q_t(x_{t-1}, \xi_t)$, $t = 2, \ldots, T$. If, moreover, the stagewise independence condition holds, then each expectation function $Q_t$ does not depend on realizations of the random process, and we can write it as $Q_t(x_{t-1})$, $t = 2, \ldots, T$. 
The nested formulation of the linear multistage problem can be written as follows (compare with (3.1)):

$$
\begin{align*}
\min_{x_1 \geq 0, x_2 \geq 0} & \quad c_1^T x_1 + \mathbb{E} \left[ \min_{x_2 \geq 0} \left( c_2^T x_2 + \mathbb{E} \left[ \min_{x_3 \geq 0} \left( c_3^T x_3 + \cdots + \mathbb{E} \left[ c_T^T x_T \right] \right) \right] \right) \right] \\
\text{s.t.} & \quad A_{11} x_1 + A_{21} x_2 + A_{31} x_3 + \cdots + A_T x_T = b_1, \\
& \quad A_{21} x_1 + A_{22} x_2 = b_2, \\
& \quad A_{31} x_1 + A_{32} x_2 + A_{33} x_3 = b_3, \\
& \quad \vdots \\
& \quad A_{T1} x_1 + A_{T2} x_2 + \cdots + A_{T,T-1} x_{T-1} + A_{TT} x_T = b_T, \\
& \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad \cdots \quad x_T \geq 0.
\end{align*}
$$

(3.9)

Suppose now that we deal with an underlying model with a full lower block triangular constraint matrix:

$$
\begin{align*}
\min_{x_t} & \quad c_t^T x_t + \mathbb{E} \left[ Q_{t+1}(x_{t+1}, \xi_{t+1}) \right] \\
\text{s.t.} & \quad A_{t1} x_1 + \cdots + A_{t,t-1} x_{t-1} + A_{tt} x_t = b_t, \\
& \quad x_t \geq 0.
\end{align*}
$$

(3.11)

Its optimal value (i.e., the cost-to-go function) $Q_t(x_{t-1}, \xi_t)$ is now a function of the whole history $x_{t-1}$ of the decision process rather than its last decision vector $x_{t-1}$.

Sometimes it is convenient to convert such a lower triangular formulation into the staircase formulation from which we started our presentation. This can be accomplished by introducing additional variables $r_t$ which summarize the relevant history of our decisions. We shall call these variables the model state variables (to distinguish from information states discussed before). The relations that describe the next values of the state variables as a function of the current values of these variables, current decisions and current random parameters are called model state equations.

For the general problem (3.10) the vectors $x_t = (x_1, \ldots, x_t)$ are sufficient model state variables. They are updated at each stage according to the state equation $x_t = (x_{t-1}, x_t)$ (which is linear), and the constraint in (3.11) can be formally written as

$$
\begin{bmatrix}
A_{t1} & A_{t2} & \cdots & A_{t,t-1}
\end{bmatrix} x_{t-1} + A_{tt} x_t = b_t.
$$

Although it looks a little awkward in this general case, in many problems it is possible to
3.1. Problem Formulation

define model state variables of reasonable size. As an example let us consider the structure

\[
\begin{align*}
\text{Min} & \quad c_1^T x_1 + c_2^T x_2 + c_3^T x_3 + \ldots + c_T^T x_T \\
\text{s.t.} & \quad A_{11} x_1 + B_1 x_1 + A_{21} x_2 + B_2 x_2 + \ldots + A_{33} x_3 = b_1, \\
& \quad B_1 x_1 + B_2 x_2 + \ldots + B_{T-1} x_{T-1} + A_{TT} x_T = b_T, \\
& \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad \ldots \quad x_T \geq 0,
\end{align*}
\]

in which all blocks \(A_{it}, i = t + 1, \ldots, T\) are identical and observed at time \(t\). Then we can define the state variables \(r_t, t = 1, \ldots, T\) recursively by the state equation \(r_t = r_{t-1} + B_t x_t, t = 1, \ldots, T - 1\), where \(r_0 = 0\). Subproblem (3.11) simplifies substantially:

\[
\begin{align*}
\text{Min} & \quad c_t^T x_t + \mathbb{E} \left[ Q_{t+1} (r_t, \xi_{t+1}) \mid \xi_t \right] \\
\text{s.t.} & \quad r_{t-1} + A_{t,t} x_t = b_t, \\
& \quad r_t = r_{t-1} + B_t x_t, \\
& \quad x_t \geq 0.
\end{align*}
\]

Its optimal value depends on \(r_{t-1}\) and is denoted \(Q_t(r_{t-1}, \xi_t)\).

Let us finally remark that the simple sign constraints \(x_t \geq 0\) can be replaced in our model by a general constraint \(x_t \in X_t\), where \(X_t\) is a convex polyhedron defined by some linear equations and inequalities (local for stage \(t\)). The set \(X_t\) may be random, too, but has to become known at stage \(t\).

3.1.3 Scenario Trees

In order to proceed with numerical calculations one needs to make a discretization of the underlying random process. It is useful and instructive to discuss this in detail. That is, we consider in this section the case where the random process \(\xi_1, \ldots, \xi_T\) has a finite number of realizations. It is useful to depict the possible sequences of data in a form of scenario tree. It has nodes organized in levels which correspond to stages \(1, \ldots, T\). At level \(t = 1\) we have only one root node, and we associate with it the value of \(\xi_1\) (which is known at stage \(t = 1\)). At level \(t = 2\) we have as many nodes as many different realizations of \(\xi_2\) may occur. Each of them is connected with the root node by an arc. For each node \(t\) of level \(t = 2\) (which corresponds to a particular realization \(\xi_2^t\) of \(\xi_2\)) we create at least as many nodes at level 3 as different values of \(\xi_3\) may follow \(\xi_2^t\), and we connect them with the node \(t\), etc.

Generally, nodes at level \(t\) correspond to possible values of \(\xi_t\) that may occur. Each of them is connected to a unique node at level \(t - 1\), called the ancestor node, which corresponds to the identical first \(t - 1\) parts of the process \(\xi_{[t]}\), and is also connected to nodes at level \(t + 1\), called children nodes, which correspond to possible continuations of history \(\xi_{[t]}\). Note that, in general, realizations \(\xi_t^t\) are vectors and it may happen that some of the values \(\xi_t^t\), associated with nodes at a given level \(t\), are equal to each other. Nevertheless, such equal values may be represented by different nodes, because they may correspond to different histories of the process (see Figure 3.1 in Example 3.1 of the next section).
We denote by $\Omega_t$ the set of all nodes at stage $t = 1, \ldots, T$. In particular, $\Omega_1$ consists of unique root node, $\Omega_2$ has as many elements as many different realizations of $\xi_2$ may occur, etc. For a node $i \in \Omega_t$ we denote by $C_i \subset \Omega_{t+1}$, $t = 1, \ldots, T - 1$, the set of all children nodes of $i$, and by $a(i) \in \Omega_{t-1}$, $t = 2, \ldots, T$, the ancestor node of $i$. We have that $\Omega_{t+1} = \cup_{i \in \Omega_t} C_i$ and the sets $C_i$ are disjoint, i.e., $C_i \cap C_{i'} = \emptyset$ if $i \neq i'$. Note again that with different nodes at stage $t \geq 3$ may be associated the same numerical values (realizations) of the corresponding data process $\xi_t$. Scenario is a path from the root node at stage $t = 1$ to a node at the last stage $T$. Each scenario represents a history of the process $\xi_1, \ldots, \xi_T$. By construction, there is one-to-one correspondence between scenarios and the set $\Omega_T$, and hence the total number $K$ of scenarios is equal to the cardinality$^3$ of the set $\Omega_T$, i.e., $K = |\Omega_T|$.

Next we should define a probability distribution on a scenario tree. In order to deal with the nested structure of the decision process we need to specify the conditional distribution of $\xi_{t+1}$ given $\xi_t$, $t=1, \ldots, T-1$. That is, if we are currently at a node $i \in \Omega_t$, we need to specify probability of moving from $i$ to a node $j \in C_i$. Let us denote this probability by $p_{i\eta}$. Note that $p_{i\eta} \geq 0$ and $\sum_{\eta \in C_i} p_{i\eta} = 1$, and that probabilities $p_{i\eta}$ are in one-to-one correspondence with arcs of the scenario tree. Probabilities $p_{i\eta}$, $\eta \in C_i$, represent conditional distribution of $\xi_{t+1}$ given that the path of the process $\xi_1, \ldots, \xi_t$ ended at the node $i$.

Every scenario can be defined by its nodes $i_1, \ldots, i_T$ arranged in the chronological order, i.e., node $i_2$ (at level $t=2$) is connected to the root node, $i_3$ is connected to the node $i_2$, etc. The probability of that scenario is then given by the product $p_{i_1 \cdot i_2} \cdot p_{i_2 \cdot i_3} \cdot \ldots \cdot p_{i_T \cdot i_{T-1}}$. That is, a set of conditional probabilities defines a probability distribution on the set of scenarios. Conversely, it is possible to derive these conditional probabilities from scenario probabilities $p_k$, $k = 1, \ldots, K$, as follows. Let us denote by $S^{(i)}$ the set of scenarios passing through node $i$ (at level $t$) of the scenario tree, and let $p^{(i)} := \Pr[S^{(i)}]$, i.e., $p^{(i)}$ is the sum of probabilities of all scenarios passing through node $i$. If $i_1, i_2, \ldots, i_T$, with $i_1$ being the root node and $i_T = i$, is the history of the process up to node $i$, then the probability $p^{(i)}$ is given by the product

$$p^{(i)} = p_{i_1 \cdot i_2} \cdot p_{i_2 \cdot i_3} \cdot \ldots \cdot p_{i_T \cdot i_{T-1}}$$

of the corresponding conditional probabilities. In another way we can write this in the recursive form $p^{(i)} = p_{a(i)} p^{(a)}$, where $a = a(i)$ is the ancestor of the node $i$. This equation defines the conditional probability $p_{a(i)}$ from the probabilities $p^{(i)}$ and $p^{(a)}$. Note that if $a = a(i)$ is the ancestor of the node $i$, then $S^{(i)} \subset S^{(a)}$ and hence $p^{(i)} \leq p^{(a)}$. Consequently if $p^{(a)} > 0$, then $p_{a(i)} = p^{(i)}/p^{(a)}$. Otherwise $S^{(a)}$ is empty, i.e., no scenario is passing through the node $a$, and hence no scenario is passing through the node $i$.

If the process $\xi_1, \ldots, \xi_T$ is stagewise independent, then the conditional distribution of $\xi_{t+1}$ given $\xi_t$ is the same as the unconditional distribution of $\xi_{t+1}$, $t=1, \ldots, T-1$. In that case at every stage $t = 1, \ldots, T - 1$, with every node $i \in \Omega_t$ is associated an identical set of children, with the same set of respective conditional probabilities and with the same respective numerical values.

Recall that a stochastic process $Z_t$, $t = 1, 2, \ldots$, that can take a finite number

$^3$We denote by $|\Omega|$ the number of elements in a (finite) set $\Omega$. 
3.1. Problem Formulation

\{z_1, \ldots, z_m\} of different values, is a Markov chain if

\[ \Pr\{Z_{t+1} = z_j \mid Z_t = z_i, Z_{t-1} = z_{i_{t-1}}, \ldots, Z_1 = z_{i_1}\} = \Pr\{Z_{t+1} = z_j \mid Z_t = z_i\}, \]

for all states \(z_{i_{t-1}}, \ldots, z_i, z_j\) and all \(t = 1, 2, \ldots\). Denote

\[ p_{ij} := \Pr\{Z_{t+1} = z_j \mid Z_t = z_i\}, \ i, j = 1, \ldots, m. \]

In some situations it is natural to model the data process as a Markov chain with the corresponding state space \(\{\zeta^1, \ldots, \zeta^m\}\) and probabilities \(p_{ij}\) of moving from state \(\zeta^i\) to state \(\zeta^j, i, j = 1, \ldots, m\). We can model such a process by a scenario tree. At stage \(t = 1\) there is one root node to which is assigned one of the values from the state space, say \(\zeta^1\). At stage \(t = 2\) there are \(m\) nodes to which are assigned values \(\zeta^1, \ldots, \zeta^m\) with the corresponding probabilities \(p_{11}, \ldots, p_{1m}\). At stage \(t = 3\) there are \(m^2\) nodes, such that each node at stage \(t = 2\), associated with a state \(\zeta^a, a = 1, \ldots, m\), is the ancestor of \(m\) nodes at stage \(t = 3\) to which are assigned values \(\zeta^1, \ldots, \zeta^m\) with the corresponding conditional probabilities \(p_{a1}, \ldots, p_{am}\). At stage \(t = 4\) there are \(m^3\) nodes, etc. At each stage \(t\) of such \(T\)-stage Markov chain process there are \(m^{T-1}\) nodes, the corresponding random vector (variable) \(\xi_t\) can take values \(\zeta^1, \ldots, \zeta^m\) with respective probabilities which can be calculated from the history of the process up to time \(t\), and the total number of scenarios is \(m^{T-1}\). We have here that the random vectors (variables) \(\xi_1, \ldots, \xi_T\) are independently distributed iff \(p_{ij} = p_{ij'}, j = 1, \ldots, m\), i.e., the conditional probability \(p_{ij}\) of moving from state \(\zeta^i\) to state \(\zeta^j\) does not depend on \(i\).

In the above formulation of the Markov chain the corresponding scenario tree represents the total history of the process with the number of scenarios growing exponentially with the number of stages. Now if we approach the problem by writing the cost-to-go functions \(Q_t(x_{t-1}, \xi_t)\), going backwards, then we do not need to keep track of the history of the process. That is, at every stage \(t\) the cost-to-go function \(Q_t(\cdot, \xi_t)\) depends only on the current state (realization) \(\xi_t = \zeta^i, i = 1, \ldots, m\), of the process. On the other hand, if we want to write the corresponding optimization problem (in the case of a finite number of scenarios) as one large linear programming problem, we still need the scenario tree formulation. This is the basic difference between the stochastic and dynamic programming approaches to the problem. That is, the stochastic programming approach does not necessarily rely on the Markovian structure of the process considered. This makes it more general at the price of considering a possibly very large number of scenarios.

An important concept associated with the data process is the corresponding filtration. We associate with the set \(\Omega_T\) the sigma algebra \(\mathcal{F}_T\) of all its subsets. The set \(\Omega_T\) can be represented as the union of disjoint sets \(C_i, i \in \Omega_{T-1}\). Let \(\mathcal{F}_{T-1}\) be the subalgebra of \(\mathcal{F}_T\) generated by the sets \(C_i, i \in \Omega_{T-1}\). As they are disjoint, they are the elementary events of \(\mathcal{F}_{T-1}\). By this construction, there is one-to-one correspondence between elementary events of \(\mathcal{F}_{T-1}\) and the set \(\Omega_{T-1}\) of nodes at stage \(T - 1\). By continuing in this way we construct a sequence of sigma algebras \(\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T\), called filtration. In this construction elementary events of sigma algebra \(\mathcal{F}_t\) are subsets of \(\Omega_T\) which are in one-to-one correspondence with the nodes \(i \in \Omega_T\). Of course, the cardinality \(|\mathcal{F}_t| = 2^{\Omega_t}|\). In particular, \(\mathcal{F}_1\) corresponds to the unique root at stage \(t = 1\) and hence \(\mathcal{F}_1 = \{\emptyset, \Omega_T\}\).

\(^1\)In our model, values \(\zeta^1, \ldots, \zeta^m\) can be numbers or vectors.
3.1.4 Algebraic Formulation of Nonanticipativity Constraints

Suppose that in our basic problem (3.6) there are only finitely many, say $K$, different scenarios the problem data can take. Recall that each scenario can be considered as a path of the respective scenario tree. With each scenario, numbered $k$, is associated probability $p_k$ and the corresponding sequence of decisions $x^k = (x^k_1, x^k_2, \ldots, x^k_T)$. That is, with each possible scenario $k = 1, \ldots, K$ (i.e., a realization of the data process) we associate a sequence of decisions $x^k$. Of course, it would not be appropriate to try to find the optimal values of these decisions by solving the relaxed version of (3.6):

$$\text{Min} \sum_{k=1}^{K} p_k \left[ c^T_1 x^k_1 + (c^k_2)^T x^k_2 + (c^k_3)^T x^k_3 + \ldots + (c^k_T)^T x^k_T \right]$$

s.t.

$$A_1 x^k_1 + B_2^k x^k_1 = b_1^k,$$
$$B_3^k x^k_2 + A_3 x^k_3 = b_3^k,$$
$$\ldots,$$

$$x^k_1 \geq 0, \quad x^k_2 \geq 0, \quad x^k_3 \geq 0, \quad \ldots, \quad x^k_T \geq 0,$$

(3.12)

The reason is the same as in the two-stage case. That is, in problem (3.12) all parts of the decision vector are allowed to depend on all parts of the random data, while each part $x_t$ should be allowed to depend only on the data known up to stage $t$. In particular, problem (3.12) may suggest different values of $x_1$, one for each scenario $k$, while our first stage decision should be independent of possible realizations of the data process.

In order to correct this problem we enforce the constraints

$$x^k_1 = x^\ell_1 \text{ for all } k, \ell \in \{1, \ldots, K\},$$

(3.13)

similarly to the two stage case (2.83). But this is not sufficient, in general. Consider the second part of the decision vector, $x_2$. It should be allowed to depend only on $\xi^{[2]} = (\xi_1, \xi_2)$, so it has to have the same value for all scenarios $k$ for which $\xi^{[2]}_k$ are identical. We must, therefore, enforce the constraints

$$x^k_2 = x^\ell_2 \text{ for all } k, \ell \text{ for which } \xi^{[2]}_k = \xi^{[2]}_\ell,$$

Generally, at stage $t = 1, \ldots, T$, the scenarios that have the same history $\xi^{[t]}$ cannot be distinguished, so we need to enforce the nonanticipativity constraints:

$$x^k_t = x^\ell_t \text{ for all } k, \ell \text{ for which } \xi^{[t]}_k = \xi^{[t]}_\ell, \quad t = 1, \ldots, T.$$  

(3.14)

Problem (3.12) together with the nonanticipativity constraints (3.14) becomes equivalent to our original formulation (3.6).

---

\(^4\)To avoid ugly collisions of subscripts we change our notation a little and we put the index of the scenario, $k$, as a superscript.
Remark 3. Let us observe that if in problem (3.12) only the constraints (3.13) are enforced, then from the mathematical point of view the problem obtained becomes a two-stage stochastic linear program with $K$ scenarios. In this two-stage program the first stage decision vector is $x_1$, the second stage decision vector is $(x_2, \ldots, x_K)$, the technology matrix is $B_2$ and the recourse matrix is the block matrix

$$
\begin{bmatrix}
A_2 & 0 & \cdots & 0 & 0 \\
B_3 & A_3 & \cdots & 0 & 0 \\
0 & 0 & \cdots & B_T & A_T
\end{bmatrix}.
$$

Since the two-stage program obtained is a relaxation of the multistage problem (3.6), its optimal value gives a lower bound for the optimal value of problem (3.6) and in that sense it may be useful. Note, however, that this model does not make much sense in any application, because it assumes that at the end of the process, when all realizations of the random data become known, one can go back in time and make all decisions $x_2, \ldots, x_{K-1}$.

Example 3.1 (Scenario Tree) As it was discussed in section 3.1.3 it can be useful to depict the possible sequences of data $\xi_t$ in a form of a scenario tree. An example of such scenario tree is given in Figure 3.1. Numbers along the arcs represent conditional probabilities of moving from one node to the next. The associated process $\xi_t = (c_t, B_t, A_t, b_t)$, $t = 1, \ldots, T$, with $T = 4$, is defined as follows. All involved variables are assumed to be one dimensional, with $c_t, B_t, A_t$, $t = 2, 3, 4$, being fixed and only right hand side variables $b_t$ being random. The values (realizations) of the random process $b_1, \ldots, b_T$ are indicated by the bold numbers at the nodes of the tree. (The numerical values of $c_t, B_t, A_t$ are not written explicitly, although, of course, they also should be specified.) That is, at level $t = 1$, $b_1$ has the value 36. At level $t = 2$, $b_2$ has two values 15 and 50 with respective probabilities 0.4 and 0.6. At level $t = 3$ we have 5 nodes with which are associated the following numerical values (from left to right) 10, 20, 12, 20, 70. That is, $b_3$ can take 4 different values with respective probabilities $\Pr\{b_3 = 10\} = 0.4 \cdot 0.1$, $\Pr\{b_3 = 20\} = 0.4 \cdot 0.4 + 0.6 \cdot 0.4$, $\Pr\{b_3 = 12\} = 0.4 \cdot 0.5$ and $\Pr\{b_3 = 70\} = 0.6 \cdot 0.6$. At level $t = 4$, the numerical values associated with 8 nodes are defined, from left to right, as 10, 10, 30, 12, 10, 20, 40, 70. The respective probabilities can be calculated by using the corresponding conditional probabilities. For example,

$$
\Pr\{b_4 = 10\} = 0.4 \cdot 0.1 \cdot 1.0 + 0.4 \cdot 0.4 \cdot 0.5 + 0.6 \cdot 0.4 \cdot 0.4.
$$

Note that although some of the realizations of $b_3$, and hence of $\xi_3$, are equal to each other, they are represented by different nodes. This is necessary in order to identify different histories of the process corresponding to different scenarios. The same remark applies to $b_4$ and $\xi_4$. Altogether, there are eight scenarios in this tree. Figure 3.2 illustrates the way in which sequences of decisions are associated with scenarios from Figure 3.1.

The process $b_t$ (and hence the process $\xi_t$) in this example is not Markovian. For instance,

$$
\Pr\{b_4 = 10 \mid b_3 = 20, b_2 = 15, b_1 = 36\} = 0.5,
$$

which sequences of decisions are associated with scenarios from Figure 3.1.
Figure 3.1. Scenario tree. Nodes represent information states. Paths from the root to leaves represent scenarios. Numbers along the arcs represent conditional probabilities of moving to the next node. Bold numbers represent numerical values of the process.

Figure 3.2. Sequences of decisions for scenarios from Figure 3.1. Horizontal dotted lines represent the equations of nonanticipativity.

while

\[
\Pr \{b_4 = 10 \mid b_3 = 20\} = \frac{\Pr \{b_4 = 10, b_3 = 20\}}{\Pr \{b_3 = 20\}} = \frac{0.5 \cdot 0.4 \cdot 0.4 + 0.4 \cdot 0.4 \cdot 0.6}{0.4 \cdot 0.4 + 0.4 \cdot 0.6} = 0.44.
\]

That is, \( \Pr \{b_4 = 10 \mid b_3 = 20\} \neq \Pr \{b_4 = 10 \mid b_3 = 20, b_2 = 15, b_1 = 36\} \).}

Relaxing the nonanticipativity constraints means that decisions \( x_t = x_t(\omega) \) are viewed as functions of all possible realizations (scenarios) of the data process. This was the case in formulation (3.12) where the problem was separated into \( K \) different problems, one for each scenario \( \omega_k = (\xi^k_1, \ldots, \xi^k_T) \), \( k = 1, \ldots, K \). There are several ways how the
corresponding nonanticipativity constraints can be written. One possible way is to write them, similarly to (2.84) for two stage models, as
\[ x_t = \mathbb{E}[x_t | \xi(t)], \quad t = 1, \ldots, T. \] (3.15)

Another way is to use filtration associated with the data process. Let \( \mathcal{F}_t \) be the sigma algebra generated by \( \xi(t) \), \( t = 1, \ldots, T \). That is, \( \mathcal{F}_t \) is the minimal subalgebra of the sigma algebra \( \mathcal{F} \) such that \( \xi_1(\omega), \ldots, \xi_t(\omega) \) are \( \mathcal{F}_t \)-measurable. Since \( \xi_1 \) is not random, \( \mathcal{F}_1 \) contains only two sets: \( \emptyset \) and \( \Omega \). We have that \( \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_T \subset \mathcal{F} \). In the case of finitely many scenarios, we discussed construction of such a filtration at the end of section 3.1.3. We can write (3.15) in the following equivalent form
\[ x_t = \mathbb{E}[x_t | \mathcal{F}_t], \quad t = 1, \ldots, T, \] (3.16)

(see section 7.2.2 for a definition of conditional expectation with respect to a sigma subalgebra). Condition (3.16) holds iff \( x_t(\omega) \) is measurable with respect \( \mathcal{F}_t \), \( t = 1, \ldots, T \). One can use this measurability requirement as a definition of the nonanticipativity constraints.

Suppose, for the sake of simplicity, that there is a finite number \( K \) of scenarios. To each scenario corresponds a sequence \( (x_{k1}, \ldots, x_{kT}) \) of decision vectors which can be considered as an element of a vector space of dimension \( n_1 + \cdots + n_T \). The space of all such sequences \( (x_{11}, \ldots, x_{K1}), k = 1, \ldots, K, \) is a vector space, denoted \( \mathcal{X} \), of dimension \( (n_1 + \cdots + n_T)K \). The nonanticipativity constraints (3.14) define a linear subspace of \( \mathcal{X} \), denoted \( \mathcal{L} \). Define the following scalar product on the space \( \mathcal{X} \),
\[ \langle x, y \rangle := \sum_{k=1}^{K} \sum_{t=1}^{T} p_k (x_{kt})^T y_{kt}, \] (3.17)

and let \( P \) be the orthogonal projection of \( \mathcal{X} \) onto \( \mathcal{L} \) with respect to this scalar product. Then
\[ x = Px \]
is yet another way to write the nonanticipativity constraints.

A computationally convenient way of writing the nonanticipativity constraints (3.14) can be derived by using the following construction, which extends to the multistage case the system (2.87).

Let \( \Omega_t \) be the set of nodes at level \( t \). For a node \( \iota \in \Omega_t \) we denote by \( S^{(i)} \) the set of scenarios that pass through node \( \iota \) and are, therefore, indistinguishable on the basis of the information available up to time \( t \). As explained before, the sets \( S^{(i)} \) for all \( i \in \Omega_t \) are the atoms of the sigma-subalgebra \( \mathcal{F}_t \) associated with the time stage \( t \). We order them and denote them by \( S_t^1, \ldots, S_t^{\gamma_t} \).

Let us assume that all scenarios \( 1, \ldots, K \) are ordered in such a way that each set \( S_t^{\nu} \) is a set of consecutive numbers \( l_t^{\nu}, l_t^{\nu} + 1, \ldots, r_t^{\nu} \). Then nonanticipativity can be expressed by the system of equations
\[ x_s - x_{s+1} = 0, \quad s = l_t^{\nu}, \ldots, r_t^{\nu} - 1, \quad t = 1, \ldots, T - 1, \quad \nu = 1, \ldots, \gamma_t. \] (3.18)

In other words, each decision is related to its neighbors from the left and from the right, if they correspond to the same node of the scenario tree.
The coefficients of constraints (3.18) define a giant matrix
\[ M = [M^1 \ldots M^K], \]
whose rows have two nonzeros each: 1 and -1. Thus, we obtain an algebraic description of the nonanticipativity constraints:
\[ M^1 x^1 + \cdots + M^K x^K = 0. \quad (3.19) \]
Owing to the sparsity of the matrix \( M \), this formulation is very convenient for various numerical methods for solving linear multistage stochastic programming problems: the simplex method, interior point methods and decomposition methods.

**Example 3.2** Consider the scenario tree depicted in Figure 3.1. Let us assume that the scenarios are numbered from the left to the right. Our nonanticipativity constraints take on the form:
\[
\begin{align*}
    x_1^1 - x_2^1 &= 0, \quad x_2^1 - x_3^1 = 0, \quad \ldots, \quad x_7^1 - x_8^1 = 0, \\
    x_2^2 - x_3^2 &= 0, \quad x_2^3 - x_3^3 = 0, \quad x_3^3 - x_4^3 = 0, \\
    x_5^2 - x_6^2 &= 0, \quad x_5^3 - x_6^3 = 0, \quad x_6^3 - x_7^3 = 0, \\
    x_2^3 - x_3^3 &= 0, \quad x_3^3 - x_4^3 = 0, \quad x_4^3 - x_5^3 = 0.
\end{align*}
\]
Using \( I \) to denote the identity matrix of an appropriate dimension, we may write the constraint matrix \( M \) as shown in Figure 3.3.

\( M \) is always a very sparse matrix: each row of it has only two nonzeros, each column at most two nonzeros. Moreover, all nonzeros are either 1 or -1, which is also convenient for numerical methods.

### 3.2 Duality

#### 3.2.1 Convex Multistage Problems

In this section we consider multistage problems of the form (3.1) with
\[
\mathcal{X}_t(x_{t-1}, \xi_t) := \{ x_t : B_t x_{t-1} + A_t x_t = b_t \}, \quad t = 2, \ldots, T, \quad (3.20)
\]
\[
\mathcal{X}_1 := \{ x_1 : A_1 x_1 = b_1 \} \text{ and } f_t(x_t, \xi_t), \quad t = 1, \ldots, T, \]
being random lower semicontinuous functions. We assume that functions \( f_t(\cdot, \xi_t) \) are convex for a.e. \( \xi_t \). In particular, if
\[
f_t(x_t, \xi_t) := \begin{cases} c_t^T x_t, & \text{if } x_t \geq 0, \\ +\infty, & \text{otherwise,} \end{cases} \quad (3.21)
\]
then the problem becomes the linear multistage problem given in the nested formulation (3.9). All constraints involving only variables and quantities associated with stage \( t \) are absorbed in the definition of the functions \( f_t \). It is be implicitly assumed that the data \( (A_t, B_t, b_t) = (A_t(\xi_t), B_t(\xi_t), b_t(\xi_t)) \), \( t = 1, \ldots, T \), form a random process.

Dynamic programming equations take here the form:
\[
Q_t(x_{t-1}, \xi_{[t]} = \inf_{x_t} \{ f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{[t]} : B_t x_{t-1} + A_t x_t = b_t) \}, \quad (3.22)
\]
3.2. Duality

![Matrix](image)

**Figure 3.3.** The nonanticipativity constraint matrix \( M \) corresponding to the scenario tree from Figure 3.1. The subdivision corresponds to the scenario submatrices \( M_1, \ldots, M_8 \).

where

\[
Q_t(x_t, \xi_t) := \mathbb{E} \{ Q_{t+1}(x_t, \xi_{t+1}) | \xi_t \}.
\]

For every \( t = 1, \ldots, T \) and \( \xi_t \), the function \( Q_T(\cdot, \xi) \) is convex. Indeed,

\[
Q_T(x_{T-1}, \xi_T) = \inf_{x_T} \phi(x_T, x_{T-1}, \xi_T),
\]

where

\[
\phi(x_T, x_{T-1}, \xi_T) := \begin{cases} 
  f_T(x_T, \xi_T), & \text{if } B_T x_{T-1} + A_T x_T = b_T, \\
  +\infty, & \text{otherwise}.
\end{cases}
\]

It follows from the convexity of \( f_T(\cdot, \xi) \), that \( \phi(\cdot, \xi) \) is convex, and hence the optimal value function \( Q_T(\cdot, \xi_T) \) is also convex. Convexity of functions \( Q_t(\cdot, \xi_t) \) can be shown in the same way by induction in \( t = T, \ldots, 1 \). Moreover, if the number of scenarios is finite and functions \( f_t(x_t, \xi_t) \) are random polyhedral, then the cost-to-go functions \( Q_t(x_{t-1}, \xi_t) \) are also random polyhedral.

### 3.2.2 Optimality Conditions

Consider the cost-to-go functions \( Q_t(x_{t-1}, \xi_t) \) defined by the dynamic programming equations (3.22). With the optimization problem on the right hand side of (3.22) is associated the following Lagrangian,

\[
L_t(x_t, \pi_t) := f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{t+1}) + \pi_t^T(b_t - B_t x_{t-1} - A_t x_t).
\]

This Lagrangian also depends on \( \xi_t \) and \( x_{t-1} \) which we omit for brevity of the notation. Denote

\[
\psi_t(x_t, \xi_t) := f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_{t+1}).
\]
The dual functional is
\[ D_t(\pi_t) := \inf_{\pi_t} L_t(x_t, \pi_t) \]
\[ = -\sup_{x_t} \{ \pi_t^T A_t x_t - \psi_t(x_t, \xi_{[t]}) \} + \pi_t^T (b_t - B_t x_{t-1}) \]
\[ = -\psi_t^* (A_t^T \pi_t, \xi_{[t]}) + \pi_t^T (b_t - B_t x_{t-1}), \]
where \( \psi_t^*(\cdot, \xi_{[t]} \) is the conjugate function of \( \psi_t(\cdot, \xi_{[t]} \) ). Therefore we can write the Lagrangian dual of the optimization problem on the right hand side of (3.22) as follows
\[ \operatorname{Max}_{\pi_t} \left\{ -\psi_t^* (A_t^T \pi_t, \xi_{[t]}) + \pi_t^T (b_t - B_t x_{t-1}) \right\}. \]  
(3.23)

Both optimization problems, (3.22) and its dual (3.23), are convex. Under various regularity conditions there is no duality gap between problems (3.22) and (3.23). In particular, we can formulate the following two conditions.

\textbf{(D1)} The functions \( f_t(x_t, \xi_t), t = 1, \ldots, T \), are random polyhedral and the number of scenarios is finite.

\textbf{(D2)} For all sufficiently small perturbations of the vector \( b_t \) the corresponding optimal value \( Q_t(x_{t-1}, \xi_{[t]} \) is finite, i.e., there is a neighborhood of \( b_t \) such that for any \( b'_t \) in that neighborhood the optimal value of the right hand side of (3.22) with \( b_t \) replaced by \( b'_t \) is finite.

We denote by \( \mathcal{D}_t (x_{t-1}, \xi_{[t]} \) the set of optimal solutions of the dual problem (3.23). All subdifferentials in the subsequent formulas are taken with respect to \( x_t \) for an appropriate \( t = 1, \ldots, T \).

\textbf{Proposition 3.3.} Suppose that either condition \textbf{(D1)} holds and \( Q_t (x_{t-1}, \xi_{[t]} \) is finite, or condition \textbf{(D2)} holds. Then:

(i) there is no duality gap between problems (3.22) and (3.23), i.e.,
\[ Q_t (x_{t-1}, \xi_{[t]} \) = \sup_{\pi_t} \left\{ -\psi_t^* (A_t^T \pi_t, \xi_{[t]}) + \pi_t^T (b_t - B_t x_{t-1}) \right\}, \]  
(3.24)

(ii) \( \bar{x}_t \) is an optimal solution of (3.22) if there exists \( \bar{\pi}_t = \bar{\pi}_t (\xi_{[t]} \) such that \( \bar{\pi}_t \in \mathcal{D}_t (x_{t-1}, \xi_{[t]} \) and
\[ 0 \in \partial L_t (\bar{x}_t, \bar{\pi}_t), \]  
(3.25)

(iii) the function \( Q_t (\cdot, \xi_{[t]} \) is subdifferentiable at \( x_{t-1} \) and
\[ \partial Q_t (x_{t-1}, \xi_{[t]} \) = -B_t^T \mathcal{D}_t (x_{t-1}, \xi_{[t]}). \]  
(3.26)

\textbf{Proof.} Consider the optimal value function
\[ \vartheta(y) := \inf_{x_t} \{ \psi_t(x_t, \xi_{[t]} : A_t x_t = y \}. \]
Since \( \psi_t(\cdot, \xi_{[t]} \) is convex, the function \( \vartheta(\cdot) \) is also convex. Condition \textbf{(D2)} means that \( \vartheta(y) \) is finite valued for all \( y \) in a neighborhood of \( \bar{y} := b_t - B_t x_{t-1} \). It follows that \( \vartheta(\cdot) \) is
3.2. Duality

continuous and subdifferentiable at \( \bar{y} \). By conjugate duality (see Theorem 7.8) this implies assertion (i). Moreover, the set of optimal solutions of the corresponding dual problem coincides with the subdifferential of \( \vartheta(\cdot) \) at \( \bar{y} \). Formula (3.26) then follows by the chain rule. Condition (3.25) means that \( \bar{x}_t \) is a minimizer of \( L(\cdot, \pi_t) \), and hence the assertion (ii) follows by (i).

If condition (D1) holds, then the functions \( Q_t(\cdot, \xi_{[t]}^t) \) are polyhedral, and hence \( \vartheta(\cdot) \) is polyhedral. It follows that \( \vartheta(\cdot) \) is lower semicontinuous and subdifferentiable at any point where it is finite valued. Again, the proof can be completed by applying the conjugate duality theory. \( \square \)

Note that condition (D2) actually implies that the set \( D_t(x_{t-1}, \xi_{[t]}^t) \) of optimal solutions of the dual problem is nonempty and bounded, while condition (D1) only implies that \( D_t(x_{t-1}, \xi_{[t]}^t) \) is nonempty.

Now let us look at the optimality conditions (3.5), which in the present case can be written as follows:

\[
\bar{x}_t(\xi_{[t]}^t) \in \arg\min_{x_t} \left\{ f_t(x_t, \xi_{[t]}^t) + Q_{t+1} \left( x_t, \xi_{[t]}^t \right) : A_t x_t = b_t - B_t x_{t-1}(\xi_{[t-1]}^{t-1}) \right\}, \quad (3.27)
\]

Since the optimization problem on the right hand side of (3.27) is convex, subject to linear constraints, we have that a feasible policy is optimal iff it satisfies the following optimality conditions: for \( t = 1, \ldots, T \) and a.e. \( \xi_{[t]} \) there exists \( \pi_t(\xi_{[t]}^t) \) such that the following condition holds

\[
0 \in \partial \left[ f_t(\bar{x}_t(\xi_{[t]}^t), \xi_{[t]}^t) + Q_{t+1} \left( x_t(\xi_{[t]}^t), \xi_{[t]}^t \right) \right] - A_t^T \pi_t(\xi_{[t]}^t). \quad (3.28)
\]

Recall that all subdifferentials are taken with respect to \( x_t \), and for \( t = T \) the term \( Q_{T+1} \) is omitted.

We shall use the following regularity condition.

(D3) For \( t = 2, \ldots, T \) and a.e. \( \xi_{[t]} \) the function \( Q_t(\cdot, \xi_{[t-1]}^{t-1}) \) is finite valued.

The above condition implies, of course, that \( Q_t(\cdot, \xi_{[t]}^t) \) is finite valued for a.e. \( \xi_{[t]} \) conditional on \( \xi_{[t-1]}^{t-1} \), which in turn implies relatively complete recourse. Note also that condition (D3) does not necessarily imply condition (D2), because in the latter the function \( Q_t(\cdot, \xi_{[t]}^t) \) is required to be finite for all small perturbations of \( b_t \).

Proposition 3.4. Suppose that either conditions (D2) and (D3) or condition (D1) are satisfied. Then a feasible policy \( \bar{x}_t(\xi_{[t]}^t) \) is optimal iff there exist mappings \( \pi_t(\xi_{[t]}^t) \), \( t = 1, \ldots, T \), such that the following condition holds true:

\[
0 \in \partial f_t(\bar{x}_t(\xi_{[t]}^t), \xi_{[t]}^t) - A_t^T \pi_t(\xi_{[t]}^t) + \mathbb{E} \left[ \partial Q_{t+1} \left( \bar{x}_t(\xi_{[t]}^t), \xi_{[t+1]}^{t+1} \right) \right| \xi_{[t]}^t] \quad (3.29)
\]

for a.e. \( \xi_{[t]}^t \) and \( t = 1, \ldots, T \). Moreover, multipliers \( \pi_t(\xi_{[t]}^t) \) satisfy (3.29) iff for a.e. \( \xi_{[t]}^t \) it holds that

\[
\pi_t(\xi_{[t]}^t) \in D(\bar{x}_{t-1}(\xi_{[t-1]}^{t-1}), \xi_{[t]}^t). \quad (3.30)
\]

Proof. Suppose that condition (D3) holds. Then by the Moreau–Rockafellar Theorem (Theorem 7.4) we have that at \( \bar{x}_t = \bar{x}_t(\xi_{[t]}^t) \),

\[
\partial \left[ f_t(\bar{x}_t, \xi_{[t]}^t) + Q_{t+1} \left( \bar{x}_t, \xi_{[t]}^t \right) \right] = \partial f_t(\bar{x}_t, \xi_{[t]}^t) + \partial Q_{t+1} \left( \bar{x}_t, \xi_{[t]}^t \right).
\]
Also by Theorem 7.47 the subdifferential of \( Q_{t+1} (\bar{x}_t, \xi_{[t]} ) \) can be taken inside the expectation to obtain the last term in the right hand side of (3.29). Note that conditional on \( \xi_{[t]} \) the term \( \bar{x}_t = \bar{x}_t(\xi_{[t]} ) \) is fixed. Optimality conditions (3.29) then follow from (3.28). Suppose, further, that condition (D2) holds. Then there is no duality gap between problems (3.22) and (3.23), and the second assertion follows by (3.27) and Proposition 3.3(ii).

If condition (D1) holds, then functions \( f_t(\bar{x}_t, \xi_{[t]} ) \) and \( Q_{t+1} (\bar{x}_t, \xi_{[t]} ) \) are random polyhedral, and hence the same arguments can be applied without additional regularity conditions. \( \square \)

Formula (3.26) makes it possible to write optimality conditions (3.29) in the following form.

**Theorem 3.5.** Suppose that either conditions (D2) and (D3) or condition (D1) are satisfied. Then a feasible policy \( x_t(\xi_{[t]} ) \) is optimal iff there exist measurable \( \pi_t(\xi_{[t]} ) \), \( t = 1, \ldots, T \), such that

\[
0 \in \partial f_1(\bar{x}_1(\xi_{[1]}), \xi_{[1]} ) - A_1^T \pi_1(\xi_{[1]} ) - E \left[ B_{t+1}^T \pi_{t+1}(\xi_{[t+1]} ) \right| \xi_{[t]} ]
\]

(3.31)

for a.e. \( \xi_{[t]} \) and \( t = 1, \ldots, T \), where for \( t = T \) the corresponding term \( T + 1 \) is omitted.

**Proof.** By Proposition 3.4 we have that a feasible policy \( x_t(\xi_{[t]} ) \) is optimal iff conditions (3.29) and (3.30) hold true. For \( t = 1 \) this means the existence of \( \bar{x}_1 \in D_1 \) such that

\[
0 \in \partial f_1(\bar{x}_1) - A_1^T \bar{\pi}_1 + E \left[ \partial Q_2(\bar{x}_1, \xi_{[2]} ) \right].
\]

(3.32)

Recall that \( \xi_1 \) is known, and hence the set \( D_1 \) is fixed. By (3.26) we have

\[
\partial Q_2(\bar{x}_1, \xi_{[2]} ) = -B_1^T D_2(\bar{x}_1, \xi_{[2]} )
\]

(3.33)

Formulas (3.32) and (3.33) mean that there exists a measurable selection

\[
\bar{\pi}_2(\xi_{[2]} ) \in D_2(\bar{x}_1, \xi_{[2]} )
\]

such that (3.31) holds for \( t = 1 \). By the second assertion of Proposition 3.4, the same selection \( \bar{\pi}_2(\xi_{[2]} ) \) can be used in (3.29) for \( t = 2 \). Proceeding in that way we obtain existence of measurable selections

\[
\pi_1(\xi_{[1]} ) \in D_t\left( \bar{x}_{t-1}(\xi_{[t-1]} ), \xi_{[t]} \right)
\]

satisfying (3.31). \( \square \)

In particular, consider the multistage linear problem given in the nested formulation (3.9). That is, functions \( f_t(\bar{x}_t, \xi_{[t]} ) \) are defined in the form (3.21) which can be written as

\[
f_t(\bar{x}_t, \xi_{[t]} ) = c_t^T \bar{x}_t + I_{\mathbb{R}^n_+}(\bar{x}_t)
\]

Then \( \partial f_t(\bar{x}_t, \xi_{[t]} ) = \{ c_t + N_{\mathbb{R}^n_+}(\bar{x}_t) \} \) at every point \( \bar{x}_t \geq 0 \), and hence optimality conditions (3.31) take the form

\[
0 \in N_{\mathbb{R}^n_+}(\bar{x}_t(\xi_{[t]} )) + c_t - A^T \pi_t(\xi_{[t]} ) - E \left[ B_{t+1}^T \pi_{t+1}(\xi_{[t+1]} ) \right| \xi_{[t]} ].
\]
3.2. Duality

3.2.3 Dualization of Feasibility Constraints

Consider the linear multistage program given in the nested formulation (3.9). In this section we discuss dualization of that problem with respect to the feasibility constraints. As it was discussed before, we can formulate that problem as an optimization problem with respect to decision variables \( x_t = x_t(\xi_{[t]}) \) viewed as functions of the history of the data process. Recall that the vector \( \xi_t \) of the data process of that problem is formed from some (or all) elements of \((c_t, B_t, A_t, b_t), t = 1, \ldots, T\). As before, we use the same symbols \( c_t, B_t, A_t, b_t \) to denote random variables and their particular realization. It will be clear from the context which of these meanings is used in a particular situation.

With problem (3.9) we associate the following Lagrangian:

\[
L(x, \pi) := \mathbb{E} \left\{ \sum_{t=1}^{T} \left[ c_t^T x_t + \pi_t^T (b_t - B_t x_{t-1} - A_t x_t) \right] \right\}
\]

\[
= \mathbb{E} \left\{ \sum_{t=1}^{T} \left[ c_t^T x_t + \pi_t^T b_t - \pi_t^T A_t x_t - \pi_{t+1}^T B_{t+1} x_t \right] \right\}
\]

\[
= \mathbb{E} \left\{ \sum_{t=1}^{T} \left[ b_t^T \pi_t + (c_t - A_t^T \pi_t - B_{t+1}^T \pi_{t+1})^T x_t \right] \right\},
\]

with the convention that \( x_0 = 0 \) and \( B_{T+1} = 0 \). Here the multipliers \( \pi_t = \pi_t(\xi_{[t]}) \), as well as decisions \( x_t = x_t(\xi_{[t]}) \), are functions of the data process up to time \( t \).

The dual functional is defined as

\[
D(\pi) := \inf_{x \geq 0} L(x, \pi),
\]

where the minimization is performed over variables \( x_t = x_t(\xi_{[t]}), t = 1, \ldots, T \), in an appropriate functional space subject to the nonnegativity constraints. The Lagrangian dual of (3.9) is the problem

\[
\max_{\pi} D(\pi), \tag{3.34}
\]

where \( \pi \) lives in an appropriate functional space. Since, for a given \( \pi \), the Lagrangian \( L(\cdot, \pi) \) is separable in \( x_t = x_t(\cdot) \), by the interchangeability principle (Theorem 7.80) we can move the operation of minimization with respect to \( x_t \) inside the conditional expectation \( \mathbb{E}[\cdot|\xi_{[t]}] \). Therefore, we obtain

\[
D(\pi) = \mathbb{E} \left\{ \sum_{t=1}^{T} \left[ b_t^T \pi_t + \inf_{x_t \in \mathbb{R}^n_+} \left( c_t - A_t^T \pi_t - \mathbb{E} [B_{t+1}^T \pi_{t+1}|\xi_{[t]}] \right)^T x_t \right] \right\}.
\]

Clearly we have that \( \inf_{x_t \in \mathbb{R}^n_+} \left( c_t - A_t^T \pi_t - \mathbb{E} [B_{t+1}^T \pi_{t+1}|\xi_{[t]}] \right)^T x_t \) is equal to zero if \( A_t^T \pi_t - \mathbb{E} [B_{t+1}^T \pi_{t+1}|\xi_{[t]}] \leq c_t \), and to \(-\infty\) otherwise. It follows that in the present case the dual problem (3.34) can be written as

\[
\max_{\pi} \mathbb{E} \left\{ \sum_{t=1}^{T} b_t^T \pi_t \right\}
\]

\[
\text{s.t. } A_t^T \pi_t + \mathbb{E} [B_{t+1}^T \pi_{t+1}|\xi_{[t]}] \leq c_t, \ t = 1, \ldots, T, \tag{3.35}
\]
where for the uniformity of notation we set all “$T + 1$” terms equal zero. Each multiplier vector $\pi_t = \pi_t(\xi_t)$, $t = 1, \ldots, T$, of problem (3.35) is a function of $\xi_t$. In that sense these multipliers form a dual implementable policy. Optimization in (3.35) is performed over all implementable and feasible dual policies.

If the data process has a finite number of scenarios, then implementable policies $x_t(\cdot)$ and $\pi_t(\cdot)$, $t = 1, \ldots, T$, can be identified with finite dimensional vectors. In that case the primal and dual problems form a pair of mutually dual linear programming problems. Therefore, the following duality result is a consequence of the general duality theory of linear programming.

**Theorem 3.6.** Suppose that the data process has a finite number of possible realizations (scenarios). Then the optimal values of problems (3.9) and (3.35) are equal unless both problems are infeasible. If the (common) optimal value of these problems is finite, then both problems have optimal solutions.

If the data process has a general distribution with an infinite number of possible realizations, then some regularity conditions are needed to ensure zero duality gap between problems (3.9) and (3.35).

### 3.2.4 Dualization of Nonanticipativity Constraints

In this section we deal with a problem which is slightly more general than linear problem (3.12). Let $f_t(x, \xi_t)$, $t = 1, \ldots, T$, be random polyhedral functions, and consider the problem

\[ \min \sum_{k=1}^{K} p_k \left[ f_1(x_1^k) + f_2^k(x_2^k) + f_3^k(x_3^k) + \ldots + f_T^k(x_T^k) \right] \]

subject to

\[
\begin{align*}
A_1 x_1^k & = b_1, \\
B_2 x_1^k & = b_2, \\
A_3 x_2^k & = b_3, \\
& \vdots \\
B_T x_{T-1}^k & = b_T,
\end{align*}
\]

where $x_1^k \geq 0, x_2^k \geq 0, x_3^k \geq 0, \ldots, x_T^k \geq 0, k = 1, \ldots, K.

Here $\xi_t^k$, $k = 1, \ldots, K$, is a particular realization (scenario) of the corresponding data process, $f_t^k(x_T^k) := f_t(x_T^k, \xi_t^k)$ and $(B_t^k, A_t^k, b_t^k) := (B_t(\xi_t^k), A_t(\xi_t^k), b_t(\xi_t^k))$, $t = 2, \ldots, T$. This problem can be formulated as a multistage stochastic programming problem by enforcing the corresponding nonanticipativity constraints.

As it was discussed in section 3.1.4, there are many ways to write nonanticipativity constraints. For example, let $\mathcal{X}$ be the linear space of all sequences $(x_1^k, \ldots, x_T^k)$, $k = 1, \ldots, K$, and $\mathcal{L}$ be the linear subspace of $\mathcal{X}$ defined by the nonanticipativity constraints (these spaces were defined above equation (3.17)). We can write the corresponding
3.2. Duality

multistage problem in the following lucid form

\[
\min_{\mathbf{x} \in \mathbb{X}} f(\mathbf{x}) := K \sum_{k=1}^{K} T \sum_{t=1}^{T} p_k f^k_{t}(x^k_t) \quad \text{subject to} \quad \mathbf{x} \in \mathbb{L}.
\] (3.36)

Clearly, \(f(\cdot)\) is a polyhedral function, so if problem (3.36) has a finite optimal value, then it has an optimal solution and the optimality conditions and duality relations hold true. Let us introduce the Lagrangian associated with (3.36):

\[
L(\mathbf{x}, \lambda) := f(\mathbf{x}) + \langle \lambda, \mathbf{x} \rangle,
\]

with the scalar product \(\langle \cdot, \cdot \rangle\) defined in (3.17). By the definition of the subspace \(\mathbb{L}\), every point \(\mathbf{x} \in \mathbb{L}\) can be viewed as an implementable policy. By \(L^\perp := \{\mathbf{y} \in \mathbb{X} : \langle \mathbf{y}, \mathbf{x} \rangle = 0, \forall \mathbf{x} \in \mathbb{L}\}\) we denote the orthogonal subspace to the subspace \(\mathbb{L}\).

**Theorem 3.7.** A policy \(\bar{x} \in \mathbb{L}\) is an optimal solution of (3.36) if and only if there exist a multiplier vector \(\bar{\lambda} \in L^\perp\) such that

\[
\bar{x} \in \arg \min_{\mathbf{x} \in \mathbb{X}} L(\mathbf{x}, \bar{\lambda}).
\] (3.37)

**Proof.** Let \(\bar{\lambda} \in \mathbb{L}^\perp\) and \(\bar{x} \in \mathbb{L}\) be a minimizer of \(L(\cdot, \bar{\lambda})\) over \(\mathbb{X}\). Then by the first order optimality conditions we have that \(0 \in \partial_x L(\bar{x}, \bar{\lambda})\). Note that there is no need here for a constraint qualification since the problem is polyhedral. Now \(\partial_x L(\bar{x}, \bar{\lambda}) = \partial f(\bar{x}) + \bar{\lambda}\). Since \(N_{\mathbb{L}}(\bar{x}) = \mathbb{L}^\perp\), it follows that \(0 \in \partial f(\bar{x}) + N_{\mathbb{L}}(\bar{x})\), which is a sufficient condition for \(\bar{x}\) to be an optimal solution of (3.36). Conversely, if \(\bar{x}\) is an optimal solution of (3.36), then necessarily \(0 \in \partial f(\bar{x}) + N_{\mathbb{L}}(\bar{x})\). This implies existence of \(\bar{\lambda} \in \mathbb{L}^\perp\) such that \(0 \in \partial f(\bar{x}) + N_{\mathbb{L}}(\bar{x})\). This, in turn, implies that \(\bar{x} \in \mathbb{L}\) is a minimizer of \(L(\cdot, \bar{\lambda})\) over \(\mathbb{X}\). \(\Box\)

Also, we can define the dual function

\[
D(\lambda) := \inf_{\mathbf{x} \in \mathbb{X}} L(\mathbf{x}, \lambda),
\]

and the dual problem

\[
\max_{\lambda \in \mathbb{L}^\perp} D(\lambda).
\] (3.38)

Since the problem considered is polyhedral, we have by the standard theory of linear programming the following results.

**Theorem 3.8.** The optimal values of problems (3.36) and (3.38) are equal unless both problems are infeasible. If their (common) optimal value is finite, then both problems have optimal solutions.

The crucial role in our approach is played by the requirement that \(\lambda \in \mathbb{L}^\perp\). Let us decipher this condition. For \(\lambda = (\lambda^k_t)_{t=1,...,T, k=1,...,K}\), the condition \(\lambda \in \mathbb{L}^\perp\) is equivalent to

\[
\sum_{t=1}^{T} \sum_{k=1}^{K} p_k \langle \lambda^k_t, x^k_t \rangle = 0, \quad \text{for all } \mathbf{x} \in \mathbb{L}.
\]
84 Chapter 3. Multistage Problems

We can write this in a more abstract form as

\[ \mathbb{E} \left[ \sum_{t=1}^{T} \langle \lambda_t, x_t \rangle \right] = 0, \text{ for all } x \in \mathcal{L}. \]  

(3.39)

Since \( \mathbb{E}_{\xi_t} x_t = x_t \) for all \( x \in \mathcal{L} \), and \( \langle \lambda_t, \mathbb{E}_{\xi_t} x_t \rangle = \langle \mathbb{E}_{\xi_t} \lambda_t, x_t \rangle \), we obtain from (3.39) that

\[ \mathbb{E} \left[ \sum_{t=1}^{T} \langle \mathbb{E}_{\xi_t} \lambda_t, x_t \rangle \right] = 0, \text{ for all } x \in \mathcal{L}, \]

which is equivalent to

\[ \mathbb{E}_{\xi_t} [\lambda_t] = 0, \text{ } t = 1, \ldots, T. \]

(3.40)

We can now rewrite our necessary conditions of optimality and duality relations in a more explicit form. We can write the dual problem in the form

\[
\max_{\lambda \in \mathcal{X}} D(\lambda) \text{ subject to } \mathbb{E}_{\xi_t} [\lambda_t] = 0, \text{ } t = 1, \ldots, T. 
\]

(3.41)

Corollary 3.9. A policy \( \bar{x} \in \mathcal{L} \) is an optimal solution of (3.36) if and only if there exist multipliers vector \( \bar{\lambda} \) satisfying (3.40) such that

\[ \bar{x} \in \arg \min_{x \in \mathcal{X}} L(x, \bar{\lambda}). \]

Moreover, problem (3.36) has an optimal solution if and only if problem (3.41) has an optimal solution. The optimal values of these problems are equal unless both are infeasible.

There are many different ways to express the nonanticipativity constraints, and thus there are many equivalent ways to formulate the Lagrangian and the dual problem. In particular, a dual formulation based on (3.18) is quite convenient for dual decomposition methods. We leave it to the reader to develop the particular form of the dual problem in this case.

Exercises

3.1. Consider the inventory model of section 1.2.3.

(a) Specify for this problem the variables, the data process, the functions, and the sets in the general formulation (3.1). Describe the sets \( \mathcal{X}_t(x_{t-1}, \xi_t) \) as in formula (3.20).

(b) Transform the problem to an equivalent linear multistage stochastic programming problem.

---

\footnote{In order to simplify notation we denote in the remainder of this section by \( \mathbb{E}_{\xi_t} \) the conditional expectation, conditional on \( \xi_{[t]} \).}
3.2. Consider the cost-to-go function $Q_t(x_{t-1}, \xi_{[t]}), t = 2, \ldots, T$, of the linear multistage problem defined as the optimal value of problem (3.7). Show that $Q_t(x_{t-1}, \xi_{[t]})$ is convex in $x_{t-1}$.

3.3. Consider the assembly problem discussed in section 1.3.3 in the case when all demand has to be satisfied, by backlogging the orders. It costs $b_i$ to delay delivery of a unit of product $i$ by one period. Additional orders of the missing parts can also be made after the last demand $D(T)$ is known. Write the dynamic programming equations of the problem. How they can be simplified, if the demand is stage-wise independent?

3.4. A transportation company has $n$ depots among which they move cargo. They are planning their operation in the next $T$ days. The demand for transportation between depot $i$ and depot $j \neq i$ on day $t$, where $t = 1, 2, \ldots, T$, is modeled as a random variable $D_{ij}(t)$. The total capacity of vehicles currently available at depot $i$ is denoted $s_i, i = 1, \ldots, n$. Before each day $t$, the company considers repositioning their fleet to better prepare to the uncertain demand on the coming day. It costs $c_{ij}$ to move a unit of capacity from location $i$ to location $j$. After repositioning, the realization of the random variables $D_{ij}(t)$ is observed, and the demand is served, up to the limit determined by the transportation capacity available at each location. The profit from transporting a unit of cargo from location $i$ to location $j$ is equal $q_{ij}$. If the total demand at location $i$ exceeds the capacity available at this location, the excessive demand is lost. It is up to the company to decide how much of each demand $D_{ij}$ will be served, and which part will remain unsatisfied. For simplicity, we consider all capacity and transportation quantities as continuous variables. After the demand is served, the transportation capacity of the vehicles at each location changes, as a result of the arrivals of vehicles with cargo from other locations. Before the next day, the company may choose to reposition some of the vehicles to prepare for the next demand. On the last day, the vehicles are repositioned so that initial quantities $s_i, i = 1, \ldots, n$, are restored.

(a) Formulate the problem of maximizing the expected profit as a multistage stochastic programming problem.

(b) Write the dynamic programming equations for this problem. Assuming that the demand is stage-wise independent, identify the state variables and simplify the dynamic programming equations.

(c) Develop a scenario tree based formulation of the problem.

3.5. Derive the dual problem to the linear multistage stochastic programming problem (3.12) with nonanticipativity constraints in the form (3.18).

3.6. You have initial capital $C_0$ which you may invest in a stock or keep in cash. You plan your investments for the next $T$ periods. The return rate on cash is deterministic and equals $r$ per each period. The price of the stock is random and equals $S_t$ in period $t = 1, \ldots, T$. The current price $S_0$ is known to you and you have a model of the price process $S_t$ in the form of a scenario tree. At the beginning, several American options on the stock prize are available. There are $n$ put options with strike prices $p_1, \ldots, p_n$ and corresponding costs $c_1, \ldots, c_n$. For example, if you buy one put...
option i, at any time $t = 1, \ldots, T$ you have the right to exercise the option and cash $p_i - S_t$ (this makes sense only when $p_i > S_t$). Also, $m$ call options are available, with strike prices $\pi_1, \ldots, \pi_m$ and corresponding costs $q_1, \ldots, q_m$. For example, if you buy one call option $j$, at any time $t = 1, \ldots, T$ you may exercise it and cash $S_t - \pi_j$ (this makes sense only when $\pi_j < S_t$). The options are available only at $t = 0$. At any time period $t$ you may buy or sell the underlying stock. Borrowing cash and short selling, that is, selling shares which are not actually owned (with the hope of repurchasing them later with profit), are not allowed. At the end of period $T$ all options expire. There are no transaction costs, and shares and options can be bought, sold (in the case of shares), or realized (in the case of options) in any quantities (not necessarily whole numbers). The amounts gained by exercising options are immediately available for purchasing shares.

Consider two objective functions.

(i) The expected value of your holdings at the end of period $T$;

(ii) The expected value of a piecewise linear utility function evaluated at the value of your final holdings. Its form is

$$u(C_T) = \begin{cases} C_T, & \text{if } C_T \geq 0, \\ (1 + R)C_T, & \text{if } C_T < 0, \end{cases}$$

where $R > 0$ is some known constant.

For both objective functions:

(a) Develop a linear multistage stochastic programming model.

(b) Derive the dual problem by dualizing with respect to feasibility constraints.
Chapter 4
Optimization Models with Probabilistic Constraints

Darinka Dentcheva

4.1 Introduction

In this chapter, we discuss stochastic optimization problems with probabilistic (also called chance) constraints of the form

$$\begin{align*}
\text{Min} & \quad c(x) \\
\text{s.t.} & \quad \Pr\{g_j(x, Z) \leq 0, \; j \in J\} \geq p, \\
& \quad x \in X.
\end{align*}$$

(4.1)

Here $X \subset \mathbb{R}^n$ is a nonempty set, $c : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$, $j \in J$, with $J$ being an index set, $Z$ is an $s$-dimensional random vector, and $p$ is a modeling parameter. We denote by $P_Z$ the probability measure (probability distribution) induced by the random vector $Z$ on $\mathbb{R}^s$. The event $A(x) = \{g_j(x, Z) \leq 0, \; j \in J\}$ in (4.1) depends on the decision vector $x$, and its probability $\Pr\{A(x)\}$ is calculated with respect to the probability distribution $P_Z$.

This model reflects the point of view that for a given decision $x$ we do not reject the statistical hypothesis that the constraints $g_j(x, Z) \leq 0, \; j \in J$, are satisfied. We discussed examples and a motivation for such problems in Chapter 1 in the contexts of inventory, multi-product and portfolio selection models. We emphasize that imposing constraints on probability of events is particularly appropriate whenever high uncertainty is involved and reliability is a central issue. In such cases, constraints on the expected value may not be sufficient to reflect our attitude to undesirable outcomes.

We also note that the objective function $c(x)$ can represent an expected value function, i.e., $c(x) = \mathbb{E}[f(x, Z)]$; however, we focus on the analysis of the probabilistic constraints at the moment.
Chapter 4. Optimization Models with Probabilistic Constraints

Figure 4.1. Vehicle routing network

We can write the probability $\Pr\{A(x)\}$ as the expected value of the characteristic function of the event $A(x)$, i.e., $\Pr\{A(x)\} = E[1_{A(x)}]$. The discontinuity of the characteristic function and the complexity of the event $A(x)$ make such problems qualitatively different from the expectation models. Let us consider two examples.

Example 4.1 (Vehicle Routing Problem) Consider a network with $m$ arcs on which a random transportation demand arises. A set of $n$ routes in the network is described by the incidence matrix $T$. More precisely, $T$ is an $m \times n$ dimensional matrix such that

$$t_{ij} = \begin{cases} 1 & \text{if route } j \text{ contains arc } i, \\ 0 & \text{otherwise.} \end{cases}$$

We have to allocate vehicles to the routes so as satisfy transportation demand. Figure 4.1 depicts a small network, and the table provides the incidence information for 19 routes on this network. For example, route 5 consists of the arcs AB, BC, and CA.

Our aim is to satisfy the demand with high prescribed probability $p \in (0, 1)$. Let $x_j$ be the number of vehicles assigned to route $j$, $j = 1, \ldots, n$. The demand for transportation on each arc is given by the random variables $Z_i$, $i = 1, \ldots, m$. We set $Z = (Z_1, \ldots, Z_m)^T$. A cost $c_j$ is associated with operating a vehicle on route $j$. Setting $c = (c_1, \ldots, c_n)^T$, the model can be formulated as follows:

$$\begin{align*}
\text{Min} & \quad c^T x \\
\text{s.t.} & \quad \Pr\{Tx \geq Z\} \geq p,
\end{align*}$$

In practical applications, we may have heterogeneous fleet of vehicles with different capacities; we may consider imposing constraints on transportation time or other requirements.

In the context of portfolio optimization, probabilistic constraints arise in a natural way, as discussed in Chapter 1.

---

1The notation $\mathbb{Z}_+$ used to denote the set of nonnegative integer numbers.
4.1. Introduction

<table>
<thead>
<tr>
<th>Arc</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CB</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DC</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.2. Vehicle routing incidence matrix

Example 4.2 (Portfolio Optimization with Value-at-Risk Constraint) We consider $n$ investment opportunities, with random return rates $R_1, \ldots, R_n$ in the next year. We have certain initial capital and our aim is to invest it in such a way that the expected value of our investment after a year is maximized, under the condition that the chance of losing no more than a given fraction of this amount is at least $p$, where $p \in (0, 1)$. Such a requirement is called a Value-at-Risk ($V@R$) constraint, discussed in Chapter 1.

Let $x_1, \ldots, x_n$ be the fractions of our capital invested in the $n$ assets. After a year, our investment changes in value according to a rate that can be expressed as

$$g(x, R) = \sum_{i=1}^{n} R_i x_i.$$ 

We formulate the following stochastic optimization problem with a probabilistic constraint:

$$\text{Max } \sum_{i=1}^{n} E[R_i] x_i$$

s.t. $\Pr\left\{ \sum_{i=1}^{n} R_i x_i \geq \eta \right\} \geq p,$

$$\sum_{i=1}^{n} x_i = 1,$$

$$x \geq 0.$$ 

(4.5)

For example, $\eta = -0.1$ may be chosen, if we aim at protecting against losses larger than 10%.
The constraint
\[ \Pr\{g_j(x, Z) \leq 0, \ j \in J\} \geq p \]
is called a \textit{joint probabilistic constraint}, while the constraints
\[ \Pr\{g_j(x, Z) \leq 0\} \geq p_j, \ j \in J, \text{ where } p_j \in [0, 1], \]
are called \textit{individual probabilistic constraints}.

In the vehicle routing example, we have a joint probabilistic constraint. If we were to
cover the demand on each arc separately with high probability, then the constraints would
be formulated as follows:
\[ \Pr\{T_i x \geq Z_i\} \geq p_i, \ i = 1, \ldots, m, \]
where \(T_i\) denotes the \(i\)-th row of the matrix \(T\). However, the latter formulation would not
ensure reliability of the network as a whole.

Infinitely many individual probabilistic constraints appear naturally in the context
of stochastic orders. For an integrable random variable \(X\), we consider its distribution
function \(F_X(\cdot)\).

**Definition 4.3.** A random variable \(X\) dominates in the first order a random variable \(Y\)
(denoted \(X \succeq_{(1)} Y\)) if
\[ F_X(\eta) \leq F_Y(\eta), \ \forall \eta \in \mathbb{R}. \]

The left-continuous inverse \(F_X^{(-1)}(\cdot)\) of the cumulative distribution function of a random
variable \(X\) is defined as follows:
\[ F_X^{(-1)}(p) = \inf \{ \eta : F_1(X; \eta) \geq p \}, \ p \in (0, 1). \]

Given \(p \in (0, 1)\), the number \(q = q(X; p)\) is called a \textit{p-quantile} of the random variable \(X\) if
\[ \Pr\{X < q\} \leq p \leq \Pr\{X \leq q\}. \]

For \(p \in (0, 1)\) the set of \(p\)-quantiles is a closed interval and \(F_X^{(-1)}(p)\) represents its left end.
Directly from the definition of the first order dominance we see that
\[ X \succeq_{(1)} Y \iff F_X^{(-1)}(p) \geq F_Y^{(-1)}(p) \text{ for all } 0 < p < 1. \] (4.6)

The first order dominance constraint can be interpreted as a continuum of probabilistic
(chance) constraints.

Denoting \(F_X^{(1)}(\eta) = F_X(\eta)\), we define higher order distribution functions of a random
variable \(X \in \mathcal{L}_{k-1}(\Omega, \mathcal{F}, P)\) as follows:
\[ F_X^{(k)}(\eta) = \int_{-\infty}^{\eta} F_X^{(k-1)}(t) \, dt \text{ for } k = 2, 3, 4, \ldots. \]
4.1. Introduction

We can express the integrated distribution function $F^{(2)}_X$ as the expected shortfall function. Integrating by parts, for each value $\eta$, we have the following formula:

$$F^{(2)}_X(\eta) = \int_{-\infty}^{\eta} F_X(\alpha) \, d\alpha = E[(\eta - X)_+] .$$

(4.7)

The function $F^{(2)}_X(\cdot)$ is well defined and finite for every integrable random variable. It is continuous, nonnegative and nondecreasing. The function $F^{(2)}_X(\cdot)$ is also convex because its derivative is nondecreasing as it is a cumulative distribution function. By the same arguments, the higher order distribution functions are continuous, nonnegative, nondecreasing, and convex as well.

Due to (4.7) second order dominance relation can be expressed in an equivalent way as follows:

$$X \succeq^{(2)} Y \text{ iff } E\{[\eta - X]_+\} \leq E\{[\eta - Y]_+\} \text{ for all } \eta \in \mathbb{R} .$$

(4.8)

The stochastic dominance relation generalizes to higher orders as follows:

**Definition 4.4.** Given two random variables $X$ and $Y$ in $L_{k-1}(\Omega, \mathcal{F}, P)$ we say that $X$ dominates $Y$ in the $k$-th order if

$$F^{(k)}_X(\eta) \leq F^{(k)}_Y(\eta), \quad \forall \eta \in \mathbb{R} .$$

We denote this relation by $X \succeq^{(k)} Y$.

We call the following semi-infinite (probabilistic) problem a stochastic optimization problem with a stochastic ordering constraint:

$$\begin{align*}
\text{Min} \quad & E[f(x, Z)] \\
\text{s.t.} \quad & \Pr\{g(x, Z) \leq \eta\} \leq F_Y(\eta), \quad \eta \in [a, b], \\
& x \in X .
\end{align*}$$

(4.9)

Here the dominance relation is restricted to an interval $[a, b] \subset \mathbb{R}$. There are technical reasons for this restriction, which will become apparent later. In the case of discrete distributions with finitely many realizations, we can assume that the interval $[a, b]$ contains the entire support of the probability measures.

In general, we formulate the following semi-infinite probabilistic problem, which we refer to as a stochastic optimization problem with a stochastic dominance constraint of order $k \geq 2$:

$$\begin{align*}
\text{Min} \quad & c(x) \\
\text{s.t.} \quad & F^{(k)}_{g(x, Z)}(\eta) \leq F^{(k)}_Y(\eta), \quad \eta \in [a, b], \\
& x \in X .
\end{align*}$$

(4.10)

\footnote{Recall that $[a]_+ = \max\{a, 0\}$.}
Example 4.5 (Portfolio Selection Problem with Stochastic Ordering Constraints)
Returning to Example 4.2, we can require that the net profit on our investment dominates certain benchmark outcome $Y$, which may be the return rate of our current portfolio or the return rate of some index. Then the Value-at-Risk constraint has to be satisfied at a continuum of points $\eta \in \mathbb{R}$. Setting $\Pr\{ Y \leq \eta \} = p_\eta$, we formulate the following model:

$$\begin{align*}
\text{Max} & \quad \sum_{i=1}^{n} \mathbb{E}[R_i] x_i \\
\text{s.t.} & \quad \Pr\left\{ \sum_{i=1}^{n} R_i x_i \leq \eta \right\} \leq p_\eta, \quad \forall \eta \in \mathbb{R}, \\
& \quad \sum_{i=1}^{n} x_i = 1, \\
& \quad x \geq 0.
\end{align*}$$

(4.11)

Using higher order stochastic dominance relations, we formulate a portfolio optimization model of form:

$$\begin{align*}
\text{Max} & \quad \sum_{i=1}^{n} \mathbb{E}[R_i] x_i \\
\text{s.t.} & \quad \sum_{i=1}^{n} R_i x_i \succeq_{(k)} Y, \\
& \quad \sum_{i=1}^{n} x_i = 1, \\
& \quad x \geq 0.
\end{align*}$$

(4.12)

A second order dominance constraint on the portfolio return rate represents a constraint on the shortfall function:

$$\begin{align*}
\sum_{i=1}^{n} R_i x_i \succeq_{(2)} Y & \iff \mathbb{E}\left[ \left( \eta - \sum_{i=1}^{n} R_i x_i \right)_{+} \right] \leq \mathbb{E}\left[ (\eta - Y)_{+} \right], \forall \eta \in \mathbb{R}.
\end{align*}$$

The second order dominance constraint can also be viewed as a continuum of Average Value-at-Risk\(^3\) (AV@R) constraints. For more information on this connection we refer to Dentcheva and Ruszczyński [56]. □

We stress that if $a = b$, then the semi-infinite model (4.9) reduces to a problem with a single probabilistic constraint, and problem (4.10) for $k = 2$ becomes a problem with a single shortfall constraint.

We shall pay special attention to problems with separable functions $g_i$, $i = 1, \ldots, m$, that is, functions of form $g_i(x, z) = \hat{g}_i(x) + h_i(z)$. The probabilistic constraint becomes

$$\Pr\{ \hat{g}_i(x) \geq -h_i(Z), \quad i = 1, \ldots, m \} \geq p.$$\\
\(^3\)Average Value-at-Risk is also called Conditional Value-at-Risk.
4.1. Introduction

We can view the inequalities under the probability as a deterministic vector function $\hat{g} : \mathbb{R}^n \to \mathbb{R}^m$, $\hat{g} = [\hat{g}_1, \ldots, \hat{g}_m]^T$ constrained from below by a random vector $Y$ with $Y_i = -h_i(Z)$, $i = 1, \ldots, m$. The problem can be formulated as follows:

$$\begin{align*}
\text{Min} & \quad c(x) \\
\text{s.t.} & \quad \Pr\{\hat{g}(x) \geq Y\} \geq p, \quad x \in X, \\
\end{align*}$$

(4.13)

where the inequality $a \leq b$ for two vectors $a, b \in \mathbb{R}^n$ is understood componentwise.

We note again that the objective function can have a more specific form:

$$c(x) = \mathbb{E}[f(x, Z)].$$

By virtue of Theorem 7.43, if the function $f(\cdot, Z)$ is continuous at $x_0$ and there exists an integrable random variable $\hat{Z}$ such that $|f(x, Z(\omega))| \leq \hat{Z}(\omega)$ for $P$-almost every $\omega \in \Omega$ and for all $x$ in a neighborhood of $x_0$, then for all $x$ in a neighborhood of $x_0$ the value function $c(x)$ is well defined and continuous at $x_0$. Furthermore, convexity of $f(\cdot, z)$ implies convexity of the expectation function $c(x)$. Therefore, we can carry out the analysis of probabilistically constrained problems using a general objective function $c(x)$ with the understanding that in some cases it may be defined as an expectation function.

Problems with separable probabilistic constraints arise frequently in the context of serving certain demand, as in the vehicle routing Example 4.1. Another type of examples is an inventory problem, as the following one.

**Example 4.6 (Cash Matching with Probabilistic Liquidity Constraint)**

We have random liabilities $L_t$ in periods $t = 1, \ldots, T$. We consider an investment in a bond portfolio from a basket of $n$ bonds. The payment of bond $i$ in period $t$ is denoted by $a_{it}$. It is zero for the time periods $t$ before purchasing of the bond is possible, as well as for $t$ greater than the maturity time of the bond. At the time period of purchase, $a_{it}$ is the negative of the price of the bond. At the following periods, $a_{it}$ is equal to the coupon payment, and at the time of maturity it is equal to the face value plus the coupon payment. All prices of bonds and coupon payments are deterministic and no default is assumed. Our initial capital equals $c_0$.

The objective is to design a bond portfolio such that the probability of covering the liabilities over the entire period $1, \ldots, T$ is at least $p$. Subject to this condition, we want to maximize the final cash on hand, guaranteed with probability $p$.

Let us introduce the cumulative liabilities

$$Z_t = \sum_{\tau=1}^{t} L_{\tau}, \quad t = 1, \ldots, T.$$  

Denoting by $x_i$ the amount invested in bond $i$, we observe that the cumulative cash flows up to time $t$, denoted $c_t$, can be expressed as follows:

$$c_t = c_{t-1} + \sum_{i=1}^{n} a_{it}x_i, \quad t = 1, \ldots, T.$$
Chapter 4. Optimization Models with Probabilistic Constraints

Using cumulative cash flows and cumulative liabilities permits the carry-over of capital from one stage to the next one, while keeping the random quantities at the right hand side of the constraints. We represent the cumulative cash flow during the entire period by the vector \( c = (c_1, \ldots, c_T)^T \). Let us assume that we quantify our preferences by using concave utility function \( U : \mathbb{R} \to \mathbb{R} \). We would like to maximize the final capital at hand in a risk-averse manner. The problem takes on the form

\[
\max_{x, c} \mathbb{E}[U(c_T - Z_T)]
\]

s.t. \( \Pr\{c_t \geq Z_t, t = 1, \ldots, T\} \geq p \),

\[
c_t = c_{t-1} + \sum_{i=1}^{n} a_{it} x_i, \quad t = 1, \ldots, T,
\]

\( x \geq 0 \).

This optimization problem has the structure of model (4.13). The first constraint can be called a probabilistic liquidity constraint.

4.2 Convexity in probabilistic optimization

Fundamental questions for every optimization model concern convexity of the feasible set, as well as continuity and differentiability of the constraint functions. The analysis of models with probability functions is based on specific properties of the underlying probability distributions. In particular, the generalized concavity theory plays a central role in probabilistic optimization as it facilitates the application of powerful tools of convex analysis.

4.2.1 Generalized concavity of functions and measures

We consider various nonlinear transformations of functions \( f(x) : \Omega \to \mathbb{R}^+ \) defined on a convex set \( \Omega \subset \mathbb{R}^n \).

Definition 4.7. A nonnegative function \( f(x) \) defined on a convex set \( \Omega \subset \mathbb{R}^n \) is said to be \( \alpha \)-concave, where \( \alpha \in [-\infty, +\infty] \), if for all \( x, y \in \Omega \) and all \( \lambda \in [0, 1] \) the following inequality holds true

\[
f(\lambda x + (1 - \lambda) y) \geq m_\alpha(f(x), f(y), \lambda),
\]

where \( m_\alpha : \mathbb{R}^+ \times \mathbb{R}^+ \times [0, 1] \to \mathbb{R} \) is defined as follows:

\[
m_\alpha(a, b, \lambda) = 0, \quad \text{if} \quad ab = 0,
\]

and if \( a > 0, b > 0, 0 \leq \lambda \leq 1 \), then

\[
m_\alpha(a, b, \lambda) = \begin{cases} 
    a^{\lambda}b^{1-\lambda}, & \text{if } \alpha = 0, \\
    \max\{a, b\}, & \text{if } \alpha = \infty, \\
    \min\{a, b\}, & \text{if } \alpha = -\infty, \\
    (\lambda a^\alpha + (1 - \lambda)b^\alpha)^{1/\alpha}, & \text{otherwise}.
\end{cases}
\]
4.2. Convexity in probabilistic optimization

In the case of \( \alpha = 0 \) the function \( f \) is called logarithmically concave or log-concave because \( \ln f(\cdot) \) is a concave function. In the case of \( \alpha = 1 \), the function \( f \) is simply concave.

It is important to note that if \( f \) and \( g \) are two measurable functions, then the function \( m_\alpha(f(\cdot), g(\cdot), \lambda) \) is a measurable function for all \( \alpha \) and all \( \lambda \in (0, 1) \). Furthermore, \( m_\alpha(a, b, \lambda) \) has the following important property:

**Lemma 4.8.** The mapping \( \alpha \mapsto m_\alpha(a, b, \lambda) \) is nondecreasing and continuous.

**Proof.** First we show the continuity of the mapping at \( \alpha = 0 \). We have the following chain of equations:

\[
\ln m_\alpha(a, b, \lambda) = \ln(\lambda a^\alpha + (1 - \lambda)b^\alpha)^{1/\alpha} = \frac{1}{\alpha} \ln(\lambda e^{\alpha \ln a} + (1 - \lambda)e^{\alpha \ln b}) = \frac{1}{\alpha} \ln \left[1 + \alpha(\lambda \ln a + (1 - \lambda) \ln b) + o(\alpha^2)\right].
\]

Applying the l'Hospital rule to the right hand side in order to calculate its limit when \( \alpha \to 0 \), we obtain

\[
\lim_{\alpha \to 0} \ln m_\alpha(a, b, \lambda) = \lim_{\alpha \to 0} \frac{\lambda \ln a + (1 - \lambda) \ln b + o(\alpha)}{1 + \alpha(\lambda \ln a + (1 - \lambda) \ln b) + o(\alpha^2)} = \lim_{\alpha \to 0} \frac{\ln(\lambda^\alpha b^{(1-\lambda)}) + o(\alpha)}{1 + \alpha \ln(\lambda^\beta b^{(1-\beta)}) + o(\alpha^2)} = \ln(\lambda^\beta b^{(1-\beta)}).
\]

Now, we turn to the monotonicity of the mapping. First, let us consider the case of \( 0 < \alpha < \beta \). We set

\[ h(\alpha) = m_\alpha(a, b, \lambda) = \exp\left(\frac{1}{\alpha} \ln \left[\lambda a^\alpha + (1 - \lambda)b^\alpha\right]\right). \]

Calculating its derivative, we obtain:

\[ h'(\alpha) = h(\alpha) \left(\frac{1}{\alpha} \cdot \frac{\lambda a^\alpha \ln a + (1 - \lambda)b^\alpha \ln b}{\lambda a^\alpha + (1 - \lambda)b^\alpha} - \frac{1}{\alpha^2} \ln \left[\lambda a^\alpha + (1 - \lambda)b^\alpha\right]\right). \]

We have to demonstrate that the expression on the right hand side is nonnegative. Substituting \( x = a^\alpha \) and \( y = b^\alpha \), we obtain

\[ h'(\alpha) = \frac{1}{\alpha^2} h(\alpha) \left(\frac{\lambda x \ln x + (1 - \lambda)y \ln y}{\lambda x + (1 - \lambda)y} - \ln \left[\lambda x + (1 - \lambda)y\right]\right). \]

Using the fact that the function \( z \mapsto z \ln z \) is convex for \( z > 0 \) and that both \( x, y > 0 \), we have that

\[ \frac{\lambda x \ln x + (1 - \lambda)y \ln y}{\lambda x + (1 - \lambda)y} - \ln \left[\lambda x + (1 - \lambda)y\right] \geq 0. \]

As \( h'(\alpha) > 0 \), we conclude that \( h(\cdot) \) is nondecreasing in this case. If \( \alpha < \beta < 0 \), we have the following chain of relations:

\[ m_\alpha(a, b, \lambda) = \left[m_{-\alpha} \left(\frac{1}{a}, \frac{1}{b}, \lambda\right)\right]^{-1} \leq \left[m_{-\beta} \left(\frac{1}{a}, \frac{1}{b}, \lambda\right)\right]^{-1} = m_{-\beta}(a, b, \lambda). \]
In the case of $0 = \alpha < \beta$, we can select a sequence $\{\alpha_k\}$, such that $\alpha_k > 0$ and $\lim_{k \to \infty} \alpha_k = 0$. We use the monotonicity of $h(\cdot)$ for positive arguments and the continuity at 0 to obtain the desired assertion. In the case $\alpha < \beta = 0$, we proceed in the same way, choosing appropriate sequence approaching 0.

If $\alpha < 0 < \beta$, then the inequality

$$m_\alpha(a, b, \lambda) \leq m_0(a, b, \lambda) \leq m_\beta(a, b, \lambda)$$

follows from the previous two cases. It remains to investigate how the mapping behaves when $\alpha \to \infty$ or $\alpha \to -\infty$. We observe that

$$\max\{\lambda^{1/\alpha}a, (1 - \lambda)^{1/\alpha}b\} \leq m_\alpha(a, b, \lambda) \leq \max\{a, b\}.$$ 

Passing to the limit, we obtain that

$$\lim_{\alpha \to \infty} m_\alpha(a, b, \lambda) = \max\{a, b\}.$$ 

We also conclude that

$$\lim_{\alpha \to -\infty} m_\alpha(a, b, \lambda) = \lim_{\alpha \to -\infty} [m_\alpha(1/a, 1/b, \lambda)]^{-1} = \max\{1/a, 1/b\}^{-1} = \min\{a, b\}.$$ 

This completes the proof. \(\square\)

This statement has the very important implication that $\alpha$-concavity entails $\beta$-concavity for all $\beta \leq \alpha$. Therefore, all $\alpha$-concave functions are $(-\infty)$-concave, that is, quasi-concave.

**Example 4.9** Consider the density of a nondegenerate multivariate normal distribution on $\mathbb{R}^s$:

$$\theta(x) = \frac{1}{\sqrt{(2\pi)^s \det(\Sigma)}} \exp\left\{-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right\},$$

where $\Sigma$ is a positive definite symmetric matrix of dimension $s \times s$, $\det(\Sigma)$ denotes the determinant of the matrix $\Sigma$, and $\mu \in \mathbb{R}^s$. We observe that

$$\ln \theta(x) = -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) - \ln \left(\sqrt{(2\pi)^s \det(\Sigma)}\right)$$

is a concave function. Therefore, we conclude that $\theta$ is 0-concave, or log-concave. \(\blacksquare\)

**Example 4.10** Consider a convex body (a convex compact set with nonempty interior) $\Omega \subset \mathbb{R}^s$. The uniform distribution on this set has density defined as follows:

$$\theta(x) = \begin{cases} \frac{1}{V_s(\Omega)}, & x \in \Omega, \\ 0, & x \notin \Omega, \end{cases}$$

where $V_s(\Omega)$ denotes the Lebesgue measure of $\Omega$. The function $\theta(x)$ is quasi-concave on $\mathbb{R}^s$ and $+\infty$-concave on $\Omega$. \(\blacksquare\)
4.2. Convexity in probabilistic optimization

We point out that for two Borel measurable sets $A, B$ in $\mathbb{R}^s$, the Minkowski sum $A + B = \{ x + y : x \in A, y \in B \}$ is Lebesgue measurable in $\mathbb{R}^s$.

**Definition 4.11.** A probability measure $P$ defined on the Lebesgue subsets of a convex set $\Omega \subset \mathbb{R}^s$ is said to be $\alpha$-concave if for any Borel measurable sets $A, B \subset \Omega$ and for all $\lambda \in [0, 1]$ we have the inequality

$$ P(\lambda A + (1 - \lambda)B) \geq m_\alpha(P(A), P(B), \lambda), $$

where $\lambda A + (1 - \lambda)B = \{ \lambda x + (1 - \lambda)y : x \in A, y \in B \}$.

We say that a random vector $Z$ with values in $\mathbb{R}^n$ has an $\alpha$-concave distribution if the probability measure $P_Z$ induced by $Z$ on $\mathbb{R}^n$ is $\alpha$-concave.

**Lemma 4.12.** If a random vector $Z$ induces an $\alpha$-concave probability measure on $\mathbb{R}^s$, then its cumulative distribution function $F_Z$ is an $\alpha$-concave function.

**Proof.** Indeed, for given points $z^1, z^2 \in \mathbb{R}^s$ and $\lambda \in [0, 1]$, we define

$$ A = \{ z \in \mathbb{R}^s : z_i \leq z_i^1, i = 1, \ldots, s \} \quad \text{and} \quad B = \{ z \in \mathbb{R}^s : z_i \leq z_i^2, i = 1, \ldots, s \}. $$

Then the inequality for $F_Z$ follows from the inequality in Definition 4.11. □

**Lemma 4.13.** If a random vector $Z$ has independent components with log-concave marginal distributions, then $Z$ has a log-concave distribution.

**Proof.** For two Borel sets $A, B \subset \mathbb{R}^s$ and $\lambda \in (0, 1)$, we define the set $C = \lambda A + (1 - \lambda)B$. Denote the projections of $A, B$ and $C$ on the coordinate axis by $A_i, B_i$ and $C_i, i = 1, \ldots, s$, respectively. For any number $r \in C_i$ there is $c \in C$ such that $c_i = r$, which implies that we have $a \in A$ and $b \in B$ with $\lambda a + (1 - \lambda)b = c$ and $r = \lambda a_i + (1 - \lambda)b_i$. In other words, $r \in \lambda A_i + (1 - \lambda)B_i$, and we conclude that $C_i \subset \lambda A_i + (1 - \lambda)B_i$. On the other hand, if $r \in \lambda A_i + (1 - \lambda)B_i$, then we have $a \in A$ and $b \in B$ such that $r = \lambda a_i + (1 - \lambda)b_i$. Setting $c = \lambda a + (1 - \lambda)b$, we conclude that $r \in C_i$. We obtain

$$ \ln[P_Z(C)] = \sum_{i=1}^s \ln[P_{Z_i}(C_i)] = \sum_{i=1}^s \ln[P_{Z_i}(\lambda A_i + (1 - \lambda)B_i)] $$

$$ \geq \sum_{i=1}^s (\lambda \ln[P_{Z_i}(A_i)] + (1 - \lambda) \ln[P_{Z_i}(B_i)]) $$

$$ = \lambda \ln[P_Z(A)] + (1 - \lambda) \ln[P_Z(B)]. $$

□

As usually, concavity properties of a function imply certain continuity of the function. We formulate without proof two theorems addressing this issue.
Chapter 4. Optimization Models with Probabilistic Constraints

Theorem 4.14 (Borell [24]). If $P$ is a quasi-concave measure on $\mathbb{R}^s$ and the dimension of its support is $s$, then $P$ has a density with respect to the Lebesgue measure.

We can relate the $\alpha$-concavity property of a measure to generalized concavity of its density (see Brascamp and Lieb [26], Prékopa [159], Rinott[168] and the references therein.)

Theorem 4.15. Let $\Omega$ be a convex subset of $\mathbb{R}^s$ and let $m > 0$ be the dimension of the smallest affine subspace $L$ containing $\Omega$. The probability measure $P$ on $\Omega$ is $\gamma$-concave with $\gamma \in [-\infty, 1/m]$, if and only if its probability density function with respect to the Lebesgue measure on $L$ is $\alpha$-concave with $\alpha = \begin{cases} \gamma/(1 - m\gamma), & \text{if } \gamma \in (-\infty, 1/m), \\ -1/m, & \text{if } \gamma = -\infty, \\ +\infty, & \text{if } \gamma = 1/m. \end{cases}$

Corollary 4.16. Let an integrable function $\theta(x)$ be defined and positive on a non-degenerated convex set $\Omega \subset \mathbb{R}^s$. Denote $c = \int_\Omega \theta(x) \, dx$. If $\theta(x)$ is $\alpha$-concave with $\alpha \in [-1/s, \infty]$ and positive on the interior of $\Omega$, then the measure $P$ on $\Omega$ defined by setting $$P(A) = \frac{1}{c} \int_A \theta(x) \, dx, \quad A \subset \Omega,$$
is $\gamma$-concave with
$$\gamma = \begin{cases} \alpha/(1 + \alpha s), & \text{if } \alpha \in (-1/s, \infty), \\ 1/s, & \text{if } \alpha = \infty, \\ -\infty, & \text{if } \alpha = -1/s. \end{cases}$$

In particular, if a measure $P$ on $\mathbb{R}^s$ has a density function $\theta(x)$ such that $\theta^{-1/s}$ is convex, then $P$ is quasi-concave.

Example 4.17 We have observed in Example 4.10 that the density of the uniform distribution on a convex body $\Omega$ is a $\alpha$-concave function. Hence, it generates a $1/s$-concave measure on $\Omega$. On the other hand, the density of the normal distribution (Example 4.9) is log-concave, and, therefore, it generates a log-concave probability measure.

Example 4.18 Consider positive numbers $\alpha_1, \ldots, \alpha_s$ and the simplex
$$S = \left\{ x \in \mathbb{R}^s : \sum_{i=1}^s x_i \leq 1, \ x_i \geq 0, \ i = 1, \ldots, s \right\}.$$The density function of the Dirichlet distribution with parameters $\alpha_1, \ldots, \alpha_s$ is defined as follows:
$$\theta(x) = \begin{cases} \frac{\Gamma(\alpha_1 + \cdots + \alpha_s)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_s)} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_s^{\alpha_s-1}, & \text{if } x \in \text{int } S, \\ 0, & \text{otherwise}. \end{cases}$$
4.2. Convexity in probabilistic optimization

Here $\Gamma(\cdot)$ stands for the Gamma function $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$.
Assuming that $x \in \text{int } S$, we consider

$$
\ln \theta(x) = \sum_{i=1}^s (\alpha_i - 1) \ln x_i + \ln \Gamma(\alpha_1 + \cdots + \alpha_s) - \sum_{i=1}^s \ln \Gamma(\alpha_i).
$$

If $\alpha_i \geq 1$ for all $i = 1, \ldots, s$, then $\ln \theta(\cdot)$ is a concave function on the interior of $S$ and, therefore, $\theta(x)$ is log-concave on $\text{cl } S$. If all parameters satisfy $\alpha_i \leq 1$, then $\theta(x)$ is log-convex on $\text{cl } (S)$. For other sets of parameters, this density function does not have any generalized concavity properties.

The next results provide calculus rules for $\alpha$-concave functions.

**Theorem 4.19.** If the function $f : \mathbb{R}^n \to \mathbb{R}_+$ is $\alpha$-concave and the function $g : \mathbb{R}^n \to \mathbb{R}_+$ is $\beta$-concave, where $\alpha, \beta \geq 1$, then the function $h : \mathbb{R}^n \to \mathbb{R}$, defined as $h(x) = f(x) + g(x)$ is $\gamma$-concave with $\gamma = \min\{\alpha, \beta\}$.

**Proof.** Given points $x_1, x_2 \in \mathbb{R}^n$ and a scalar $\lambda \in (0, 1)$, we set $x_\lambda = \lambda x_1 + (1 - \lambda)x_2$. Both functions $f$ and $g$ are $\gamma$-concave by virtue of Lemma 4.8. Using Minkowski inequality, which holds true for $\gamma \geq 1$, we obtain:

$$
\begin{align*}
\lambda f(x_1) + (1-\lambda)g(x_2) &\geq [\lambda f(x_1)]^{\gamma} + (1-\lambda)[g(x_2)]^{\gamma} \left[ \frac{\lambda}{\gamma} + \left(1 - \lambda \right) \frac{1}{\gamma} \right]^{\frac{1}{\gamma}} \\
&\geq \left[ \lambda f(x_1) + g(x_2) \right]^{\gamma} \left[ \frac{\lambda}{\gamma} + \left(1 - \lambda \right) \frac{1}{\gamma} \right]^{\frac{1}{\gamma}}.
\end{align*}
$$

This completes the proof.

**Theorem 4.20.** Let $f$ be a concave function defined on a convex set $C \subset \mathbb{R}^n$ and $g : \mathbb{R} \to \mathbb{R}$ be a nonnegative nondecreasing $\alpha$-concave function, $\alpha \in [-\infty, \infty]$. Then the function $g \circ f$ is $\alpha$-concave.

**Proof.** Given $x, y \in \mathbb{R}^n$ and a scalar $\lambda \in (0, 1)$, we consider $z = \lambda x + (1 - \lambda)y$. We have $f(z) \geq f(x) + (1 - \lambda)f(y)$. By monotonicity and $\alpha$-concavity of $g$, we obtain the following chain of inequalities:

$$
|g \circ f|(z) \geq g(\lambda f(x) + (1 - \lambda)f(y)) \geq m_\alpha(g(f(x)), g(f(y)), \lambda).
$$

This proves the assertion.

**Theorem 4.21.** Let the function $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}_+$ be such that for all $y \in Y \subset \mathbb{R}^n$ the function $f(\cdot, y)$ is $\alpha$-concave ($\alpha \in [-\infty, \infty]$) on the convex set $X \subset \mathbb{R}^m$. Then the function $\varphi(x) = \inf_{y \in Y} f(x, y)$ is $\alpha$-concave on $X$. 
Proof. Let \( x_1, x_2 \in X \) and a scalar \( \lambda \in (0, 1) \) be given. We set \( z = \lambda x_1 + (1 - \lambda) x_2 \). We can find a sequence of points \( y_k \in Y \) such that

\[
\varphi(z) = \inf_{y \in Y} f(z, y) = \lim_{k \to \infty} f(z, y_k).
\]

Using the \( \alpha \)-concavity of the function \( f(\cdot, y) \), we conclude that

\[
f(z, y_k) \geq m_\alpha(\{f(x_1, y_k), f(x_2, y_k), \lambda\}).
\]

The mapping \((a, b) \mapsto m_\alpha(a, b, \lambda)\) is monotone for nonnegative \( a \) and \( b \) and \( \lambda \in (0, 1) \). Therefore, we have that

\[
f(z, y_k) \geq m_\alpha(\varphi(x_1), \varphi(x_2), \lambda).
\]

Passing to the limit, we obtain the assertion. \( \square \)

**Lemma 4.22.** If \( \alpha_i > 0 \), \( i = 1, \ldots, m \), and \( \sum_{i=1}^m \alpha_i = 1 \), then the function \( f : \mathbb{R}_+^m \to \mathbb{R} \), defined as \( f(x) = \prod_{i=1}^m x_i^{\alpha_i} \), is concave.

**Proof.** We shall show the statement for the case of \( m = 2 \). For points \( x, y \in \mathbb{R}^2_+ \) and a scalar \( \lambda \in (0, 1) \), we consider \( \lambda x + (1 - \lambda) y \). Define the quantities:

\[
a_1 = (\lambda x_1)^{\alpha_1}, \quad a_2 = ((1 - \lambda) y_1)^{\alpha_1}, \quad b_1 = (\lambda x_2)^{\alpha_2}, \quad b_2 = ((1 - \lambda) y_2)^{\alpha_2}.
\]

Using Hölder’s inequality, we obtain the following:

\[
f(\lambda x + (1 - \lambda) y) = \left( \frac{a_1}{a_2} + \frac{b_1}{b_2} \right)^{\alpha_1} \left( \frac{b_1}{b_2} + \frac{a_2}{a_1} \right)^{\alpha_2}
\]

\[
\geq a_1 b_1 + a_2 b_2 = \lambda x_1^{\alpha_1} x_2^{\alpha_2} + (1 - \lambda) y_1^{\alpha_1} y_2^{\alpha_2}.
\]

The assertion in the general case follows by induction. \( \square \)

**Theorem 4.23.** If the functions \( f_i : \mathbb{R}^n \to \mathbb{R}_+ \), \( i = 1, \ldots, m \) are \( \alpha_i \)-concave and \( \alpha_i \) are such that \( \sum_{i=1}^m \alpha_i^{-1} > 0 \), then the function \( g : \mathbb{R}^{mn} \to \mathbb{R}_+ \), defined as \( g(x) = \prod_{i=1}^m f_i(x_i) \) is \( \gamma \)-concave with \( \gamma = \left( \sum_{i=1}^m \alpha_i^{-1} \right)^{-1} \).

**Proof.** Fix points \( x_1, x_2 \in \mathbb{R}^n_+ \), a scalar \( \lambda \in (0, 1) \) and set \( x_\lambda = \lambda x_1 + (1 - \lambda) x_2 \). By the generalized concavity of the functions \( f_i \), \( i = 1, \ldots, m \), we have the following inequality:

\[
\prod_{i=1}^m f_i(x_\lambda) \geq \prod_{i=1}^m \left( \lambda f_i(x_1)^{\alpha_i} + (1 - \lambda) f_i(x_2)^{\alpha_i} \right)^{1/\alpha_i}.
\]
4.2. Convexity in probabilistic optimization

We denote $y_{ij} = f_i(x_j)^{\alpha_i}$, $j = 1, 2$. Substituting into the last displayed inequality and raising both sides to power $\gamma$, we obtain

$$\left( \prod_{i=1}^{m} f_i(x_\lambda) \right)^\gamma \geq \prod_{i=1}^{m} \left( \lambda y_{i1} + (1 - \lambda) y_{i2} \right)^{\gamma/\alpha_i}. $$

We continue the chain of inequalities using Lemma 4.22:

$$\prod_{i=1}^{m} \left( \lambda y_{i1} + (1 - \lambda) y_{i2} \right)^{\gamma/\alpha_i} \geq \lambda \prod_{i=1}^{m} \left[ f_i(x_1) \right]^{\gamma} + (1 - \lambda) \prod_{i=1}^{m} \left[ f_i(x_2) \right]^{\gamma},$$

Putting the inequalities together and using the substitutions at the right hand side of the last inequality, we conclude that

$$\prod_{i=1}^{m} \left[ f_i(x_\lambda) \right]^{\gamma} \geq \lambda \prod_{i=1}^{m} \left[ f_i(x_1) \right]^{\gamma} + (1 - \lambda) \prod_{i=1}^{m} \left[ f_i(x_2) \right]^{\gamma},$$

as required.

In the special case, when the functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, k$ are concave, we can apply Theorem 4.23 consecutively, to conclude that $f_1f_2$ is $\frac{1}{2}$-concave and $f_1 \ldots f_k$ is $\frac{1}{k}$-concave.

**Lemma 4.24.** If $A$ is a symmetric positive definite matrix of size $n \times n$, then the function $A \mapsto \det(A)$ is $\frac{1}{n}$-concave.

**Proof.** Consider two $n \times n$ symmetric positive definite matrices $A, B$ and $\gamma \in (0, 1)$. We note that for every eigenvalue $\lambda$ of $A$, $\gamma \lambda$ is an eigenvalue of $\gamma A$, and, hence, $\det(\gamma A) = \gamma^n \det(A)$. We could apply the following Minkowski inequality for matrices:

$$[\det(A + B)]^{\frac{1}{n}} \geq [\det(A)]^{\frac{1}{n}} + [\det(B)]^{\frac{1}{n}}, \quad (4.14)$$

which implies the $\frac{1}{n}$-concavity of the function. As inequality (4.14) is not well-known, we provide a proof of it. First, we consider the case of diagonal matrices. In this case the determinants of $A$ and $B$ are products of their diagonal elements and inequality (4.14) follows from Lemma 4.22.

In the general case, let $A^{1/2}$ stand for the symmetric positive definite square-root of $A$ and let $A^{-1/2}$ be its inverse. We have

$$\det(A + B) = \det \left( A^{1/2}A^{-1/2}(A + B)A^{-1/2}A^{1/2} \right) = \det \left( A^{-1/2}(A + B)A^{-1/2} \right) \det(A) = \det \left( I + A^{-1/2}BA^{-1/2} \right) \det(A) \quad (4.15)$$

Notice that $A^{-1/2}BA^{-1/2}$ is symmetric positive definite and, therefore, we can choose an
Chapter 4. Optimization Models with Probabilistic Constraints

$n \times n$ orthogonal matrix $R$, which diagonalizes it. We obtain
\[
\det \left( I + A^{-1/2}BA^{-1/2} \right) = \det \left( R^T (I + A^{-1/2}BA^{-1/2}) R \right) = \det \left( I + R^T A^{-1/2}BA^{-1/2}R \right).
\]

At the right hand side of the equation, we have a sum of two diagonal matrices and we can apply inequality (4.14) for this case. We conclude that
\[
\left[ \det \left( I + A^{-1/2}BA^{-1/2} \right) \right]^{\frac{1}{n}} \leq \left[ \det \left( I + R^T A^{-1/2}BA^{-1/2}R \right) \right]^{\frac{1}{n}}
\geq 1 + \left[ \det \left( R^T A^{-1/2}BA^{-1/2}R \right) \right]^{\frac{1}{n}} = 1 + \left[ \det(B) \right]^{\frac{1}{n}} \left[ \det(A) \right]^{-\frac{1}{n}}.
\]

Combining this inequality with equation (4.15), we obtain (4.14) in the general case. □

Example 4.25 (Dirichlet distribution continued) We return to Example 4.18. We see that the functions $x_i \mapsto x_i^{\beta_i}$ are $1/\beta_i$-concave, provided that $\beta_i > 0$. Therefore, the density function of the Dirichlet distribution is a product of $1/\alpha_i$-concave functions, given that all parameters $\alpha_i > 1$. By virtue of Theorem 4.23, we obtain that this density is $\gamma$-concave with $\gamma = (\alpha_1 + \cdots + \alpha_m - s)^{-1}$ provided that $\alpha_i > 1, i = 1, \ldots, m$. Due to Corollary 4.16, the Dirichlet distribution is a $(\alpha_1 + \cdots + \alpha_m)^{-1}$-concave probability measure. □

Theorem 4.26. If the $s$-dimensional random vector $Z$ has an $\alpha$-concave probability distribution, $\alpha \in [-\infty, +\infty]$, and $T$ is a constant $m \times s$ matrix, then the $m$-dimensional random vector $Y = TZ$ has an $\alpha$-concave probability distribution.

Proof. Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^m$ be two Borel sets. We define
\[
A_1 = \{ z \in \mathbb{R}^s : Tz \in A \} \quad \text{and} \quad B_1 = \{ z \in \mathbb{R}^s : Tz \in B \}.
\]

The sets $A_1$ and $A_2$ are Borel sets as well due to the continuity of the linear mapping $z \mapsto Tz$. Furthermore, for $\lambda \in [0, 1]$ we have the relation:
\[
\lambda A_1 + (1 - \lambda) B_1 \subset \{ z \in \mathbb{R}^s : Tz \in \lambda A + (1 - \lambda) B \}.
\]

Denoting $P_Z$ and $P_Y$ the probability measure of $Z$ and $Y$ respectively, we obtain,
\[
P_Y \{ \lambda A + (1 - \lambda) B \} \geq P_Z \{ \lambda A_1 + (1 - \lambda) B_1 \} \\
\geq m_\alpha (P_Z \{ A_1 \}, P_Z \{ B_1 \}, \lambda) \\
= m_\alpha (P_Y \{ A \}, P_Y \{ B \}, \lambda).
\]

This completes the proof. □
4.2. Convexity in probabilistic optimization

Example 4.27 A univariate Gamma distribution is given by the following probability density function

\[
f(z) = \begin{cases} \\
\lambda \vartheta z^{\vartheta - 1} e^{-\lambda z} / \Gamma(\vartheta), & \text{for } z > 0, \\
0, & \text{otherwise},
\end{cases}
\]

where \(\lambda > 0\) and \(\vartheta > 0\) are constants. For \(\lambda = 1\) the distribution is the standard gamma distribution. If a random variable \(Y\) has the gamma distribution, then \(\vartheta Y\) has the standard gamma distribution. It is not difficult to check that this density function is log-concave, provided \(\vartheta \geq 1\).

A multivariate gamma distribution can be defined by a certain linear transformation of \(m\) independent random variables \(Z_1, \ldots, Z_m\) \((1 \leq m \leq 2^s - 1)\), that have the standard gamma distribution. Let an \(s \times m\) matrix \(A\) with 0–1 elements be given. Setting \(Z = (Z_1, \ldots, Z_2s - 1)\), we define \(Y = AZ\).

The random vector \(Y\) has a multivariate standard gamma distribution.

We observe that the distribution of the vector \(Z\) is log-concave by virtue of Lemma 4.13. Hence, the \(s\)-variate standard gamma distribution is log-concave by virtue of Theorem 4.26.

Example 4.28 The Wishart distribution arises in estimation of covariance matrices and can be considered as a multidimensional version of the \(\chi^2\) distribution. More precisely, let us assume that \(Z\) is an \(s\)-dimensional random vector having multivariate normal distribution with a nonsingular covariance matrix \(\Sigma\) and expectation \(\mu\). Given a sample with observed values \(z^1, \ldots, z^N\) from this distribution, we consider the matrix:

\[
\sum_{i=1}^{N} (z^i - \bar{z})(z^i - \bar{z})^T,
\]

where \(\bar{z}\) is the sample mean. This matrix has the Wishart distribution with \(N - 1\) degrees of freedom. We denote the trace of a matrix \(A\) by \(\text{tr}(A)\).

If \(N > s\), the Wishart distribution is a continuous distribution on the space of symmetric square matrices with probability density function defined by

\[
f(A) = \begin{cases} \\
\frac{\det(A)^{N - 2 - s - 2} \exp\left(-\frac{1}{2} \text{tr}(\Sigma^{-1} A)\right)}{2^{N - 1} \pi^{s(s - 1)/2} \det(\Sigma)^{N - 2} \prod_{i=1}^{s} \Gamma\left(\frac{N - 1}{2}\right)}, & \text{for } A \text{ positive definite,} \\
0, & \text{otherwise.}
\end{cases}
\]

If \(s = 1\) and \(\Sigma = 1\) this density becomes the \(\chi^2\) distribution density with \(N - 1\) degrees of freedom.

If \(A_1\) and \(A_2\) are two positive definite matrices and \(\lambda \in (0, 1)\), then the matrix \(\lambda A_1 + (1 - \lambda)A_2\) is positive definite as well. Using Lemma 4.24 and Lemma 4.8 we conclude that function \(A \mapsto \ln \det(A)\), defined on the set of positive definite Hermitian matrices, is concave. This implies that, if \(N \geq s + 2\), then \(f\) is a log-concave function on the set of symmetric positive definite matrices. If \(N = s + 1\), then \(f\) is a log-convex on the convex set of symmetric positive definite matrices.
Recall that a function \( f : \mathbb{R}^n \to \mathbb{R} \) is called 
regular in the sense of Clarke
 or Clarke-
regular if
\[
 f'(x; d) = \lim_{t \to 0} \frac{1}{t} [f(x + td) - f(x)] = \lim_{y \to x, t \to 0} \frac{1}{t} [f(y + td) - f(y)].
\]
It is known that convex functions are regular in this sense. We call a concave function 
\( f \) regular with the understanding that the regularity requirement applies to \(-f\). In this
 case, we have \( \partial^\circ (-f)(x) = -\partial^\circ f(x) \), where \( \partial^\circ f(x) \) refers to the collection of the Clarke 
generalized gradients of \( f \) at the point \( x \). For convex functions 
\( \partial^\circ f(x) = \partial f(x) \).

**Theorem 4.29.** If \( f : \mathbb{R}^n \to \mathbb{R} \) is \( \alpha \)-concave (\( \alpha \in \mathbb{R} \)) on some open set \( U \subset \mathbb{R}^n \) and 
\( f(x) > 0 \) for all \( x \in U \), then \( f(x) \) is locally Lipschitz continuous, directionally differentiable, and Clarke-regular. Its Clarke generalized gradients are given by the formula:
\[
 \partial^\circ f(x) = \begin{cases} 
 \frac{1}{\alpha} [f(x)]^{1/\alpha} \partial [f(x)]^\alpha, & \text{if } \alpha \neq 0, \\
 \frac{\partial f}{\partial \ln f(x)}, & \text{if } \alpha = 0.
\end{cases}
\]

**Proof.** If \( f \) is an \( \alpha \)-concave function, then an appropriate transformation of \( f \), is a concave 
function on \( U \). We define
\[
 \tilde{f}(x) = \begin{cases} 
 (f(x))^\alpha, & \text{if } \alpha \neq 0 \\
 \ln f(x), & \text{if } \alpha = 0.
\end{cases}
\]
If \( \alpha < 0 \), then \( f^\alpha(\cdot) \) is convex. This transformation is well-defined on the open subset \( U \) 
since \( f(x) > 0 \) for \( x \in U \), and, thus, \( \tilde{f}(x) \) is subdifferentiable at any \( x \in U \). Further, we 
represent \( f \) as follows:
\[
 f(x) = \begin{cases} 
 [(\tilde{f}(x))^{1/\alpha}, & \text{if } \alpha \neq 0, \\
 \exp(\tilde{f}(x)), & \text{if } \alpha = 0.
\end{cases}
\]
In this representation, \( f \) is a composition of a continuously differentiable function and a 
concave function. By virtue of Clarke [38, Theorem 2.3.9(3)], the function \( f \) is locally Lipschitz continuous, directionally differentiable, and Clarke-regular. Its Clarke general-
ized gradient set is given by the formula:
\[
 \partial^\circ f(x) = \begin{cases} 
 \frac{1}{\alpha} [\tilde{f}(x)]^{1/\alpha - 1} \partial \tilde{f}(x), & \text{if } \alpha \neq 0, \\
 \exp(\tilde{f}(x)) \partial \tilde{f}(x), & \text{if } \alpha = 0.
\end{cases}
\]
Substituting the definition of \( \tilde{f} \) yields the result. \( \Box \)

For a function \( f : \mathbb{R}^n \to \mathbb{R} \), we consider the set of points at which it takes positive 
values. It is denoted by \( \text{dompos} f \), i.e.,
\[
 \text{dompos} f = \{ x \in \mathbb{R}^n : f(x) > 0 \}.
\]
Recall that \( N_X(x) \) denotes the normal cone to the set \( X \) at \( x \in X \).
4.2. Convexity in probabilistic optimization

Definition 4.30. We call a point $\hat{x} \in \mathbb{R}^n$ a stationary point of an $\alpha$-concave function $f$, if there is a neighborhood $U$ of $\hat{x}$ such that $f$ is Lipschitz continuous on $U$, and $0 \in \partial^0 f(\hat{x})$. Furthermore, for a convex set $X \subset \text{dompos} f$, we call $\hat{x} \in X$ a stationary point of $f$ on $X$, if there is a neighborhood $U$ of $\hat{x}$ such that $f$ is Lipschitz continuous on $U$ and $0 \in \partial^0 f_X(\hat{x}) + N_X(\hat{x})$.

We observe that certain properties of the maxima of concave functions extend to generalized concave functions.

Theorem 4.31. Let $f$ be an $\alpha$-concave function $f$ and the set $X \subset \text{dompos} f$ be convex. Then all the stationary points of $f$ on $X$ are global maxima and the set of global maxima of $f$ on $X$ is convex.

Proof. First, assume that $\alpha = 0$. Let $\hat{x}$ be a stationary point of $f$ on $X$. This implies that $0 \in f(\hat{x}) \partial (\ln f(\hat{x})) + N_X(\hat{x})$, \hspace{1cm} (4.16)

Using that $f(\hat{x}) > 0$, we obtain $0 \in \partial (\ln f(\hat{x})) + N_X(\hat{x})$, \hspace{1cm} (4.17)

As the function $\overline{f}(x) = \ln f(x)$ is concave, this inclusion implies that $\hat{x}$ is a global maximal point of $\overline{f}$ on $X$. By the monotonicity of $\ln(\cdot)$, we conclude that $\hat{x}$ is a global maximal point of $f$ on $X$. If a point $\tilde{x} \in X$ is a maximal point of $\overline{f}$ on $X$, then inclusion (4.16) as $X \subset \text{dompos} f$, and, therefore, $\tilde{x}$ is a stationary point of $f$ on $X$. Therefore, the set of maximal points of $f$ on $X$ is convex because this is the set of maximal points of the concave function $\overline{f}$.

In the case of $\alpha \neq 0$, the statement follows by the same line of argument using the function $f(x) = \left[ f(x) \right]^\alpha$. \hspace{1cm} \square

Another important property of $\alpha$-concave measures is the existence of so-called floating body for all probability levels $p \in (\frac{1}{2}, 1)$.

Definition 4.32. A measure $P$ on $\mathbb{R}^s$ has a floating body at level $p > 0$ if there exists a convex body $C_p \subset \mathbb{R}^s$ such for all vectors $z \in \mathbb{R}^s$:

$P \{ x \in \mathbb{R}^s : z^T x \geq s_{C_p}(z) \} = 1 - p,$

where $s_{C_p}(\cdot)$ is the support function of the set $C_p$. The set $C_p$ is called the floating body of $P$ at level $p$.

Symmetric log-concave measures have floating bodies. We formulate this result of Meyer and Reisner [128] without proof.

Theorem 4.33. Any non-degenerate probability measure with symmetric log-concave density function has a floating body $C_p$ at all levels $p \in (\frac{1}{2}, 1)$.

We see that $\alpha$-concavity as introduced so far implies continuity of the distribution function. As empirical distributions are very important in practical applications, we would
like to find a suitable generalization of this notion applicable to discrete distributions. For this purpose, we introduce the following notion.

**Definition 4.34.** A distribution function $F$ is called $\alpha$-concave on the set $A \subset \mathbb{R}^s$ with $\alpha \in [-\infty, \infty]$, if

$$F(z) \geq m_\alpha(F(x), F(y), \lambda)$$

for all $z, x, y \in A$ and $\lambda \in (0, 1)$ such that $z \geq \lambda x + (1 - \lambda)y$.

Observe that if $A = \mathbb{R}^s$, then this definition coincides with the usual definition of $\alpha$-concavity of a distribution function.

To illustrate the relation between Definition 4.7 and Definition 4.34 let us consider the case of integer random vectors which are roundups of continuously distributed random vectors.

**Remark 4.** If the distribution function of a random vector $Z$ is $\alpha$-concave on $\mathbb{R}^s$ then the distribution function of $Y = \lfloor Z \rfloor$ is $\alpha$-concave on $\mathbb{Z}^s$.

This property follows from the observation that at integer points both distribution functions coincide.

**Example 4.35** Every distribution function of an $s$-dimensional binary random vector is $\alpha$-concave on $\mathbb{Z}^s$ for all $\alpha \in [-\infty, \infty]$.

Indeed, let $x$ and $y$ be binary vectors, $\lambda \in (0, 1)$, and $z \geq \lambda x + (1 - \lambda)y$. As $z$ is integer and $x$ and $y$ binary, then $z \geq x$ and $z \geq y$. Hence, $F(z) \geq \max\{F(x), F(y)\}$ by the monotonicity of the cumulative distribution function. Consequently, $F$ is $\infty$-concave. Using Lemma 4.8 we conclude that $F_Z$ is $\alpha$-concave for all $\alpha \in [-\infty, \infty]$. 

For a random vector with independent components, we can relate concavity of the marginal distribution functions to the concavity of the joint distribution function. Note that the statement applies not only to discrete distributions, as we can always assume that the set $A$ is the whole space or some convex subset of it.

**Theorem 4.36.** Consider the $s$-dimensional random vector $Z = (Z^1, \ldots, Z^L)$, where the sub-vectors $Z^l$, $l = 1, \ldots, L$, are $s_l$-dimensional and $\sum_{l=1}^{L} s_l = s$. Assume that $Z^l$, $l = 1, \ldots, L$ are independent and that their marginal distribution functions $F_{Z^l} : \mathbb{R}^{s_l} \to [0, 1]$ are $\alpha_l$-concave on the sets $A_l \subset \mathbb{R}^{s_l}$. Then the following statements hold true:

1. If $\sum_{l=1}^{L} \alpha_l^{-1} > 0$, $l = 1, \ldots, L$, then $F_Z$ is $\alpha$-concave on $A = A_1 \times \cdots \times A_L$ with $\alpha = (\sum_{l=1}^{L} \alpha_l^{-1})^{-1}$.
2. If $\alpha_l = 0$, $l = 1, \ldots, L$, then $F_Z$ is log-concave on $A = A_1 \times \cdots \times A_L$.

**Proof.** The proof of the first statement follows by virtue of Theorem 4.23 using the monotonicity of the cumulative distribution function.

For the second statement consider $\lambda \in (0, 1)$ and points $x = (x^1, \ldots, x^L) \in A$, $y = (y^1, \ldots, y^L) \in A$, and $z = (z^1, \ldots, z^L) \in A$, such that $z \geq \lambda x + (1 - \lambda)y$. Using
4.2. Convexity in probabilistic optimization

the monotonicity of the function \( \ln(\cdot) \) and of \( F_Z(\cdot) \), along with the log-concavity of the marginal distribution functions, we obtain the following chain of inequalities:

\[
\ln[F_Z(z)] \geq \ln[F_Z(\lambda x + (1 - \lambda)y)] = \sum_{l=1}^{L} \ln [F_{Z_l}(\lambda x^l + (1 - \lambda)y^l)] \\
\geq \sum_{l=1}^{L} [\lambda \ln[F_{Z_l}(x^l)] + (1 - \lambda) \ln[F_{Z_l}(y^l)]] \\
\geq \lambda \sum_{l=1}^{L} \ln[F_{Z_l}(x^l)] + (1 - \lambda) \sum_{l=1}^{L} \ln[F_{Z_l}(y^l)] \\
= \lambda [F_Z(x)] + (1 - \lambda)[F_Z(y)].
\]

This concludes the proof. \( \Box \)

For integer random variables our definition of \( \alpha \)-concavity is related to log-concavity of sequences.

**Definition 4.37.** A sequence \( p_k, k \in \mathbb{Z} \), is called log-concave, if

\[
p_k^2 \geq p_{k-1} p_{k+1} \quad \text{for all} \quad k \in \mathbb{Z}.
\]

We have the following property (see Prékopa [159, Thm 4.7.2]).

**Theorem 4.38.** Suppose that for an integer random variable \( Y \) the probabilities \( p_k = \Pr\{Y = k\}, k \in \mathbb{Z} \) form a log-concave sequence. Then the distribution function of \( Y \) is \( \alpha \)-concave on \( \mathbb{Z} \) for every \( \alpha \in [-\infty, 0] \).

### 4.2.2 Convexity of probabilistically constrained sets

One of the most general results in the convexity theory of probabilistic optimization is the following theorem.

**Theorem 4.39.** Let the functions \( g_j : \mathbb{R}^n \times \mathbb{R}^s, j \in J \), be quasi-concave. If \( Z \in \mathbb{R}^s \) is a random vector that has an \( \alpha \)-concave probability distribution, then the function

\[
G(x) = \Pr\{g_j(x, Z) \geq 0, j \in J\} \tag{4.18}
\]

is \( \alpha \)-concave on the set

\[
D = \{x \in \mathbb{R}^n : \exists z \in \mathbb{R}^s \text{ such that } g_j(x, z) \geq 0, j \in J\}.
\]

**Proof.** Given the points \( x_1, x_2 \in D \) and \( \lambda \in (0, 1) \), we define the sets

\[
A_i = \{z \in \mathbb{R}^s : g_j(x_i, z) \geq 0, j \in J\}, \quad i = 1, 2,
\]
and \( B = \lambda A_1 + (1 - \lambda)A_2 \). We consider
\[
G(\lambda x_1 + (1 - \lambda)x_2) = P\{g_j(\lambda x_1 + (1 - \lambda)x_2, Z) \geq 0, j \in \mathcal{J}\}.
\]
If \( z \in B \), then there exist points \( z_i \in A_i \) such that \( z = \lambda z_1 + (1 - \lambda)z_2 \). By virtue of the quasi-concavity of \( g_j \) we obtain that
\[
g_j(\lambda x_1 + (1 - \lambda)x_2, \lambda z_1 + (1 - \lambda)z_2) \geq \min\{g_j(x_1, z_1), g_j(x_2, z_2)\} \geq 0 \quad \text{for all } j \in \mathcal{J}.
\]
This implies that \( z \in \{z \in \mathbb{R}^s : g_j(\lambda x_1 + (1 - \lambda)x_2, z) \geq 0, j \in \mathcal{J}\} \), which entails that \( \lambda x_1 + (1 - \lambda)x_2 \in D \) and that
\[
G(\lambda x_1 + (1 - \lambda)x_2) \geq P\{B\}.
\]
Using the \( \alpha \)-concavity of the measure, we conclude that
\[
G(\lambda x_1 + (1 - \lambda)x_2) \geq P\{B\} \geq m_\alpha\{P\{A_1\}, P\{A_2\}, \lambda\} = m_\alpha\{G(x_1), G(x_2), \lambda\},
\]
as desired.

**Example 4.40 (The log-normal distribution)** The probability density function of the one-dimensional log-normal distribution with parameters \( \mu \) and \( \sigma \) is given by
\[
f(x) = \begin{cases} \frac{1}{\sqrt{2\pi} \sigma x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}
\]
This density is neither log-concave, nor log-convex. However, we can show that the cumulative distribution function is log-concave. We demonstrate it for the multidimensional case.

The \( m \)-dimensional random vector \( Z \) has the log-normal distribution, if the vector \( y = (\ln Z_1, \ldots, \ln Z_m)^T \) has a multivariate normal distribution. Recall that the normal distribution is log-concave. The distribution function of \( Z \) at a point \( z \in \mathbb{R}^m \), \( z > 0 \), can be written as
\[
F_Z(z) = \Pr\{Z_1 \leq z_1, \ldots, Z_m \leq z_m\} = \Pr\left\{ z_1 - e_1^Y \geq 0, \ldots, z_m - e_m^Y \geq 0 \right\}.
\]
We observe that the assumptions of Theorem 4.39 are satisfied for the probability function on the right hand side. Thus, \( F_Z \) is a log-concave function.

As a consequence, under the assumptions of Theorem 4.39, we obtain convexity statements for sets described by probabilistic constraints.

**Corollary 4.41.** Assume that the functions \( g_j(\cdot, \cdot), j \in \mathcal{J} \), are quasi-concave jointly in both arguments, and that \( Z \in \mathbb{R}^s \) is a random variable that has an \( \alpha \)-concave probability distribution. Then the following set is convex and closed:
\[
X_0 = \{x \in \mathbb{R}^n : \Pr\{g_i(x, Z) \geq 0, i = 1, \ldots, m\} \geq p\}.
\]
4.2. Convexity in probabilistic optimization

Proof. Let \( G(x) \) be defined as in (4.18), and let \( x_1, x_2 \in X_0, \lambda \in [0, 1] \). We have

\[
G(\lambda x_1 + (1 - \lambda) x_2) \geq m_\alpha \{ G(x_1), G(x_2), \lambda \} \geq \min \{ G(x_1), G(x_2) \} \geq p.
\]

The closedness of the set follows from the continuity of \( \alpha \)-concave functions.

We consider the case of a separable mapping \( g \), when the random quantities appear only on the right hand side of the inequalities.

Theorem 4.42. Let the mapping \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) be such that each component \( g_i \) is a concave function. Furthermore, assume that the random vector \( Z \) has independent components and the one-dimensional marginal distribution functions \( F_{Z_i}, i = 1, \ldots, m \) are \( \alpha_i \)-concave. Furthermore, let \( \sum_{i=1}^k \alpha_i^{-1} > 0 \). Then the set

\[
X_0 = \left\{ x \in \mathbb{R}^n : \Pr \{ g(x) \geq Z \} \geq p \right\}
\]

is convex.

Proof. Indeed, the probability function appearing in the definition of the set \( X_0 \) can be described as follows:

\[
G(x) = P \{ g_i(x) \geq Z_i, i = 1, \ldots, m \} = \prod_{i=1}^m F_{Z_i}(g_i(x_i))
\]

Due to Theorem 4.20, the functions \( F_{Z_i} \circ g_i \) are \( \alpha_i \)-concave. Using Theorem 4.23, we conclude that \( G(\cdot) \) is \( \gamma \)-concave with \( \gamma = \left( \sum_{i=1}^k \alpha_i^{-1} \right)^{-1} \). The convexity of \( X_0 \) follows the same way as in Corollary 4.41.

Under the same assumptions the set determined by the first order stochastic dominance constraint with respect to any random variable \( Y \) is convex and closed.

Theorem 4.43. Assume that \( g(\cdot, \cdot) \) is a quasi-concave function jointly in both arguments, and that \( Z \) has an \( \alpha \)-concave distribution. Then the following sets are convex and closed:

\[
X_d = \left\{ x \in \mathbb{R}^n : g(x, Z) \succeq_1 Y \right\},
\]

\[
X_c = \left\{ x \in \mathbb{R}^n : \Pr \{ g(x, Z) \geq \eta \} \geq \Pr \{ Y \geq \eta \}, \forall \eta \in [a, b] \right\}.
\]

Proof. Let us fix \( \eta \in \mathbb{R} \) and observe that the relation \( g(x, Z) \succeq_1 Y \) can be formulated in the following equivalent way:

\[
\Pr \{ g(x, Z) \geq \eta \} \geq \Pr \{ Y \geq \eta \} \quad \text{for all} \quad \eta \in \mathbb{R}.
\]

Therefore, the first set can be defined as follows:

\[
X_d = \left\{ x \in \mathbb{R}^n : \Pr \{ g(x, Z) - \eta \geq 0 \} \geq \Pr \{ Y \geq \eta \} \quad \forall \eta \in \mathbb{R} \right\}.
\]
For any \( \eta \in \mathbb{R} \), we define the set

\[
X(\eta) = \{ x \in \mathbb{R}^n : \Pr\{ g(x, Z) - \eta \geq 0 \} \geq \Pr\{ Y \geq \eta \} \}.
\]

This set is convex and closed by virtue of Corollary 4.41. The set \( X_\eta \) is the intersection of the sets \( X(\eta) \) for all \( \eta \in \mathbb{R} \), and, therefore, it is convex and closed as well. Analogously, the set \( X_\varepsilon \) is convex and closed as \( X_\varepsilon = \bigcap_{\eta \in [a,b]} X(\eta) \).

Let us observe that affine in each argument functions \( g_i(x, z) = z^T x + b_i \) are not necessarily quasi-concave in both arguments \((x, z)\). We can apply Theorem 4.39 to conclude that the set

\[
X_l = \{ x \in \mathbb{R}^n : \Pr\{ x^T a_i \leq b_i(Z), i = 1, \ldots, m \} \geq p \}
\]

is convex if \( a_i, i = 1, \ldots, m \) are deterministic vectors. We have the following.

**Corollary 4.44.** The set \( X_l \) is convex whenever \( b_i(\cdot) \) are quasi-concave functions and \( Z \) has a quasi-concave probability distribution function.

**Example 4.45 (Vehicle Routing continued)** We return to Example 4.1. The probabilistic constraint (4.3) has the form:

\[
\Pr\{ Tx \geq Z \} \leq p_\eta.
\]

If the vector \( Z \) of a random demand has an \( \alpha \)-concave distribution, then this constraint defines a convex set. For example, this is the case if each component \( Z_i \) has a Poisson distribution and the components (the demand on each arc) are independent of each other.

If the functions \( g_i \) are not separable, we can invoke Theorem 4.33.

**Theorem 4.46.** Let \( p_i \in (\frac{1}{2}, 1) \) for all \( i = 1, \ldots, n \). The following set is convex

\[
X_p = \{ x \in \mathbb{R}^n : P_{Z_i}\{ x^T Z_i \leq b_i \} \geq p_i, i = 1, \ldots, m \},
\]

whenever \( Z_i \) has a non-degenerated log-concave probability distribution, which is symmetric around some point \( \mu_i \in \mathbb{R}^n \).

**Proof.** If the random vector \( Z_i \) has a non-degenerated log-concave probability distribution, which is symmetric around some point \( \mu_i \in \mathbb{R}^n \), then the vector \( Y_i = Z_i - \mu_i \) has a symmetric and non degenerate log-concave distribution.

Given points \( x_1, x_2 \in X_p \) and a number \( \lambda \in [0, 1] \), we define

\[
K_i(x) = \{ a \in \mathbb{R}^n : a^T x \leq b_i \}, \quad i = 1, \ldots, n.
\]

Let us fix an index \( i \). The probability distribution of \( Y_i \) satisfies the assumptions of Theorem 4.33. Thus, there is a convex set \( C_{p_i} \) such that any supporting plane defines a half-plane containing probability \( p_i \):

\[
P_{Y_i}\{ y \in \mathbb{R}^n : y^T x \leq s_{c_{p_i}}(x) \} = p_i, \quad \forall x \in \mathbb{R}^n.
\]
4.2. Convexity in probabilistic optimization

Thus,
\[
P_z \{ z \in \mathbb{R}^n : z^T x \leq s_{c_{p_i}}(x) + \mu_i^T x \} = p_i, \quad \forall x \in \mathbb{R}^n. \tag{4.22}
\]

Since \( P_z \{ K_i(x_1) \} \geq p_i \) and \( P_z \{ K_i(x_2) \} \geq p_i \) by assumption, then
\[
K_i(x_j) \subset \{ z \in \mathbb{R}^n : z^T x \leq s_{c_{p_i}}(x) + \langle \mu_i, x \rangle \}, \quad j = 1, 2,
\]

The properties of the support function entail that
\[
b_i \geq s_{c_{p_i}}(x_1) + \mu_i^T x_j, \quad j = 1, 2.
\]

Consequently, the set \( K_i(x_\lambda) \) with \( x_\lambda = \lambda x_1 + (1 - \lambda) x_2 \) contains the set
\[
\{ z \in \mathbb{R}^n : z^T x_\lambda \leq s_{c_{p_i}}(x_\lambda) + \mu_i^T x_\lambda \}
\]
and, therefore, using equation (4.22) we obtain that
\[
P_z \{ K_i(\lambda x_1 + (1 - \lambda) x_2) \} \geq p_i.
\]

Since \( i \) was arbitrary, we obtain that \( \lambda x_1 + (1 - \lambda) x_2 \in X_p \). \( \square \)

**Example 4.47 (Portfolio optimization continued)** Let us consider the Portfolio Example 4.2. The probabilistic constraint has the form:
\[
\Pr \left\{ \sum_{i=1}^{n} R_i x_i \leq \eta \right\} \leq p_\eta.
\]

If the random vector \( R = (R_1, \ldots, R_n)^T \) has a multidimensional normal distribution or a uniform distribution then the feasible set in this example is convex by virtue of the last corollary since both distributions are symmetric and log-concave.

There is an important relation between the sets constrained by first and second order stochastic dominance relation to a benchmark random variable (see Dentcheva and Ruszczyński [53]). We denote the space of integrable random variables by \( L_1(\Omega, \mathcal{F}, P) \) and set:
\[
A_1(Y) = \{ X \in L_1(\Omega, \mathcal{F}, P) : X \succeq_{(1)} Y \},
\]
\[
A_2(Y) = \{ X \in L_1(\Omega, \mathcal{F}, P) : X \succeq_{(2)} Y \}.
\]

**Proposition 4.48.** For every \( Y \in L_1(\Omega, \mathcal{F}, P) \) the set \( A_2(Y) \) is convex and closed.

**Proof.** By changing the order of integration in the definition of the second order function \( F^{(2)} \), we obtain
\[
F_X^{(2)}(\eta) = \mathbb{E}[(\eta - X)_+]. \tag{4.23}
\]
Thereby, an equivalent representation of the second order stochastic dominance relation is given by the relation
\[ \mathbb{E}[(\eta - X)_+] \leq \mathbb{E}[(\eta - Y)_+] \quad \text{for all } \eta \in \mathbb{R}. \] (4.24)
For every \( \eta \in \mathbb{R} \) the functional \( X \to \mathbb{E}[(\eta - X)_+] \) is convex and continuous in \( L^1(\Omega, \mathcal{F}, P) \), as a composition of a linear function, the "max" function and the expectation operator. Consequently, the set \( A_2(Y) \) is convex and closed. \( \Box \)

The set \( A_1(Y) \) is closed, because convergence in \( L^1 \) implies convergence in probability, but it is not convex in general.

**Example 4.49** Suppose that \( \Omega = \{\omega_1, \omega_2\} \), \( P\{\omega_1\} = P\{\omega_2\} = 1/2 \) and \( Y(\omega_1) = -1, Y(\omega_2) = 1 \). Then \( X_1 = Y \) and \( X_2 = -Y \) both dominate \( Y \) in the first order. However, \( X = (X_1 + X_2)/2 = 0 \) is not an element of \( A_1(Y) \) and, thus, the set \( A_1(Y) \) is not convex. We notice that \( X \) dominates \( Y \) in the second order. \( \blacksquare \)

Directly from the definition we see that first order dominance relation implies the second order dominance. Hence, \( A_1(Y) \subset A_2(Y) \). We have demonstrated that the set \( A_2(Y) \) is convex; therefore, we also have
\[ \text{conv}(A_1(Y)) \subset A_2(Y). \] (4.25)
We find sufficient conditions for the opposite inclusion.

**Theorem 4.50.** Assume that \( \Omega = \{\omega_1, \ldots, \omega_N\} \), \( \mathcal{F} \) contains all subsets of \( \Omega \), and \( P\{\omega_k\} = 1/N, k = 1, \ldots, N \). If \( Y : (\Omega, \mathcal{F}, P) \to \mathbb{R} \) is a random variable, then
\[ \text{conv}(A_1(Y)) = A_2(Y). \]

**Proof.** To prove the inverse inclusion to (4.25), suppose that \( X \in A_2(Y) \). Under the assumptions of the theorem, we can identify \( X \) and \( Y \) with vectors \( x = (x_1, \ldots, x_N) \) and \( y = (y_1, \ldots, y_N) \) such that \( x_i = X(i) \) and \( y_i = Y(i) \), \( i = 1, \ldots, N \). As the probabilities of all elementary events are equal, the second order stochastic dominance relation coincides with the concept of weak majorization, which is characterized by the following system of inequalities:
\[ \sum_{k=1}^{l} x[k] \geq \sum_{k=1}^{l} y[k], \quad l = 1, \ldots, N, \]
where \( x[k] \) denotes the \( k \)-th smallest component of \( x \).

As established by Hardy, Littlewood and Polya [73], weak majorization is equivalent to the existence of a doubly stochastic matrix \( \Pi \) such that
\[ x \geq \Pi y. \]

By Birkhoff’s Theorem [20], we can find permutation matrices \( Q^1, \ldots, Q^M \) and nonnegative reals \( \alpha_1, \ldots, \alpha_M \) totaling 1, such that
\[ \Pi = \sum_{j=1}^{M} \alpha_j Q^j. \]
4.2. Convexity in probabilistic optimization

Setting $z^j = Q^j y$, we conclude that

$$x \geq \sum_{j=1}^{M} \alpha_j z^j.$$  

Identifying random variables $Z^j$ on $(\Omega, \mathcal{F}, P)$ with the vectors $z^j$, we also see that

$$X(\omega) \geq \sum_{j=1}^{M} \alpha_j Z^j(\omega),$$

for all $\omega \in \Omega$. Since each vector $z^j$ is a permutation of $y$ and the probabilities are equal, the distribution of $Z^j$ is identical with the distribution of $Y$. Thus

$$Z^j \succeq_{(1)} Y, \quad j = 1, \ldots, M.$$  

Let us define

$$\hat{Z}^j(\omega) = Z^j(\omega) + (X(\omega) - \sum_{k=1}^{M} \alpha_k Z^k(\omega)), \quad \omega \in \Omega, \quad j = 1, \ldots, M.$$  

Then the last two inequalities render $\hat{Z}^j \in A_1(Y), j = 1, \ldots, M$, and

$$X(\omega) = \sum_{j=1}^{M} \alpha_j \hat{Z}^j(\omega),$$

as required.

This result does not extend to general probability spaces, as the following example illustrates.

**Example 4.51** We consider the probability space $\Omega = \{\omega_1, \omega_2\}$, $P\{\omega_1\} = 1/3$, $P\{\omega_2\} = 2/3$. The benchmark variable $Y$ is defined as $Y(\omega_1) = -1$, $Y(\omega_2) = 1$. It is easy to see that $X \succeq_{(1)} Y$ if and only if $X(\omega_1) \geq -1$ and $X(\omega_2) \geq 1$. Thus, $A_1(Y)$ is a convex set.

Now, consider the random variable $Z = \mathbb{E}[Y] = 1/3$. It dominates $Y$ in the second order, but it does not belong to $\text{conv} A_1(Y) = A_1(Y)$.

It follows from this example that the probability space must be sufficiently rich to observe our phenomenon. If we could define a new probability space $\Omega' = \{\omega_1, \omega_2, \omega_2\}$, in which the event $\omega_2$ is split in two equally likely events $\omega_2, \omega_2'$, then we could use Theorem 4.50 to obtain the equality $\text{conv} A_1(Y) = A_2(Y)$. In the context of optimization, however, the probability space has to be fixed at the outset and we are interested in sets of random variables as elements of $L_p(\Omega, \mathcal{F}, P; \mathbb{R}^n)$, rather than in sets of their distributions.

**Theorem 4.52.** Assume that the probability space $(\Omega, \mathcal{F}, P)$ is nonatomic. Then

$$A_2(Y) = \text{cl}\{\text{conv}(A_1(Y))\}.$$
Proof. If the space \((\Omega, \mathcal{F}, P)\) is nonatomic, we can partition \(\Omega\) into \(N\) disjoint subsets, each of the same \(P\)-measure \(1/N\), and we verify the postulated equation for random variables which are piecewise constant on such partitions. This reduces the problem to the case considered in Theorem 4.50. Passing to the limit with \(N \to \infty\), we obtain the desired result. We refer the interested reader to Dentcheva and Ruszczyński [55] for technical details of the proof. \(\square\)

4.2.3 Connectedness of probabilistically constrained sets

Let \(\mathcal{X} \subset \mathbb{R}^n\) be a closed convex set. In this section we focus on the following set:

\[
X = \{ x \in \mathcal{X} : \Pr\{ g_j(x, Z) \geq 0, j \in \mathcal{J} \} \geq p \},
\]

where \(\mathcal{J}\) is an arbitrary index set. The functions \(g_j : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}\) are continuous, \(Z\) is an \(s\)-dimensional random vector, and \(p \in (0, 1)\) is a prescribed probability. It will be demonstrated later (Lemma 4.61) that the probabilistically constrained set \(X\) with separable functions \(g_j\) is a union of cones intersected by \(\mathcal{X}\). Thus, \(\mathcal{X}\) could be disconnected. The following result provides a sufficient condition for \(\mathcal{X}\) to be topologically connected. A more general version of this result is proved in Henrion [84].

**Theorem 4.53.** Assume that the functions \(g_j(\cdot, Z)\), \(j \in \mathcal{J}\) are quasi-concave and that they satisfy the following condition: for all \(x^1, x^2 \in \mathbb{R}^n\) there exists a point \(x^* \in \mathcal{X}\) such that

\[
g_j(x^*, z) \geq \min \{ g_j(x^1, z), g_j(x^2, z) \} \quad \forall z \in \mathbb{R}^s, \forall j \in \mathcal{J}.
\]

Then the set \(\mathcal{X}\) is connected.

**Proof.** Let \(x^1, x^2 \in \mathcal{X}\) be arbitrary points. We construct a path joining the two points, which is contained entirely in \(\mathcal{X}\). Let \(x^* \in \mathcal{X}\) be the point that exists according to the assumption. We set

\[
\pi(t) = \begin{cases} 
(1 - 2t)x^1 + 2tx^*, & \text{for } 0 \leq t \leq 1/2, \\
2(1 - t)x^* + (2t - 1)x^2, & \text{for } 1/2 < t \leq 1.
\end{cases}
\]

First, we observe that \(\pi(t) \in \mathcal{X}\) for every \(t \in [0, 1]\) since \(x^1, x^2, x^* \in \mathcal{X}\) and the set \(\mathcal{X}\) is convex. Furthermore, the quasi-concavity of \(g_j, j \in \mathcal{J}\), and the assumptions of the theorem imply for every \(j\) and for \(0 \leq t \leq 1/2\) the following inequality:

\[
g_j((1 - 2t)x^1 + 2tx^*, z) \geq \min \{ g_j(x^1, z), g_j(x^*, z) \} = g_j(x^1, z).
\]

Therefore,

\[
\Pr\{ g_j(\pi(t), Z) \geq 0, j \in \mathcal{J}\} \geq \Pr\{ g(x^1) \geq 0, j \in \mathcal{J}\} \geq p \quad \text{for} \quad 0 \leq t \leq 1/2.
\]

Similar argument applies for \(1/2 < t \leq 1\). Consequently, \(\pi(t) \in \mathcal{X}\), and this proves the assertion. \(\square\)
4.3 Separable probabilistic constraints

We focus our attention on problems with separable probabilistic constraints. The problem that we analyze in this section, is the following:

\[
\begin{align*}
\min_{x} & \quad c(x) \\
\text{s.t.} & \quad \Pr\{g(x) \geq Z\} \geq p, \quad \forall x \in \mathcal{X}.
\end{align*}
\] (4.26)

We assume that \(c : \mathbb{R}^n \to \mathbb{R}\) is a convex function and \(g : \mathbb{R}^n \to \mathbb{R}^m\) is such that each component \(g_i : \mathbb{R}^n \to \mathbb{R}\) is a concave function. We assume that the deterministic constraints are expressed by a closed convex set \(\mathcal{X} \subset \mathbb{R}^n\). The vector \(Z\) is an \(m\)-dimensional random vector.

4.3.1 Continuity and differentiability properties of distribution functions

When the probabilistic constraint involves inequalities with random variables on the right hand side only as in problem (4.26), we can express it as a constraint on a distribution function:

\[
\Pr\{g(x) \geq Z\} \geq p \iff F_Z(g(x)) \geq p.
\]

Therefore, it is important to analyze the continuity and differentiability properties of distribution functions. These properties are relevant to the numerical solution of probabilistic optimization problems.

Suppose that \(Z\) has an \(\alpha\)-concave distribution function with \(\alpha \in \mathbb{R}\) and that the support of it, \(\text{supp } P_Z\), has nonempty interior in \(\mathbb{R}^s\). Then \(F_Z(\cdot)\) is locally Lipschitz continuous on \(\text{int supp } P_Z\) by virtue of Theorem 4.29.

**Example 4.54** We consider the following density function

\[
\theta(z) = \begin{cases} 
\frac{1}{2\sqrt{\pi}}, & \text{for } z \in (0, 1), \\
0, & \text{otherwise}.
\end{cases}
\]

The corresponding cumulative distribution function is

\[
F(z) = \begin{cases} 
0, & \text{for } z \leq 0, \\
\sqrt{z}, & \text{for } z \in (0, 1), \\
1, & \text{for } z \geq 1.
\end{cases}
\]

The density \(\theta\) is unbounded. We observe that \(F\) is continuous but it is not Lipschitz continuous at \(z = 0\). The density \(\theta\) is also not \((-1)\)-concave and that means that the corresponding probability distribution is not quasi-concave.

**Theorem 4.55.** Suppose that all one-dimensional marginal distribution functions of an \(s\)-dimensional random vector \(Z\) are locally Lipschitz continuous. Then \(F_Z\) is locally Lipschitz continuous as well.
The statement can be proved by straightforward estimation of the distribution function by its marginals for \( s = 2 \) and induction on the dimension of the space.

It should be noted that even if the multivariate probability measure \( P_Z \) has a continuous and bounded density, then the distribution function \( F_Z \) is not necessarily Lipschitz continuous.

**Theorem 4.56.** Assume that \( P_Z \) has a continuous density \( \theta(\cdot) \) and that all one-dimensional marginal distribution functions are continuous as well. Then the distribution function \( F_Z \) is continuously differentiable.

**Proof.** In order to simplify the notation, we demonstrate the statement for \( s = 2 \). It will be clear how to extend the proof for \( s > 2 \). We have that

\[
F_Z(z_1, z_2) = \Pr(Z_1 \leq z_1, Z_2 \leq z_2) = \int_{-\infty}^{z_1} \int_{-\infty}^{z_2} \theta(t_1, t_2)dt_2dt_1 = \int_{-\infty}^{z_1} \psi(t_1, z_2)dt_1,
\]

where \( \psi(t_1, z_2) = \int_{-\infty}^{z_2} \theta(t_1, t_2)dt_2 \). Since \( \psi(\cdot, z_2) \) is continuous, by the Newton-Leibnitz Theorem we have that

\[
\frac{\partial F_Z}{\partial z_1}(z_1, z_2) = \psi(z_1, z_2) = \int_{-\infty}^{z_2} \theta(z_1, t_2)dt_2.
\]

In a similar way

\[
\frac{\partial F_Z}{\partial z_2}(z_1, z_2) = \int_{-\infty}^{z_1} \theta(t_1, z_2)dt_1.
\]

Let us show continuity of \( \frac{\partial F_Z}{\partial z_1}(z_1, z_2) \). Given the points \( z \in \mathbb{R}^2 \) and \( y^k \in \mathbb{R}^2 \), such that \( \lim_{k \to \infty} y^k = z \), we have:

\[
\left| \frac{\partial F_Z}{\partial z_1}(z) - \frac{\partial F_Z}{\partial z_1}(y^k) \right| = \left| \int_{-\infty}^{z_2} \theta(z_1, t)dt - \int_{-\infty}^{y^k_2} \theta(y_1^k, t)dt \right| \\
\leq \left| \int_{z_2}^{y^k_2} \theta(y_1^k, t)dt \right| + \left| \int_{-\infty}^{z_2} \left[ \theta(z_1, t) - \theta(y_1^k, t) \right] dt \right|.
\]

First, we observe that the mapping \( (z_1, z_2) \mapsto \int_{-\infty}^{z_2} \theta(z_1, t)dt \) is continuous for every \( a \in \mathbb{R} \) by the uniform continuity of \( \theta(\cdot) \) on compact sets in \( \mathbb{R}^2 \). Therefore, \( \left| \int_{-\infty}^{y^k_2} \theta(y_1^k, t)dt \right| \to 0 \) whenever \( k \to \infty \). Furthermore, \( \left| \int_{-\infty}^{z_2} \left[ \theta(z_1, t) - \theta(y_1^k, t) \right] dt \right| \to 0 \) as well, due to the continuity of the one-dimensional marginal function \( F_Z \). Moreover, by the same reason, the convergence is uniform about \( z_1 \). This proves that \( \frac{\partial F_Z}{\partial z_1}(z) \) is continuous.

The continuity of the second partial derivative follows by the same line of argument. As both partial derivatives exist and are continuous, the function \( F_Z \) is continuously differentiable.  

4.3. Separable probabilistic constraints

4.3.2 \textit{p}-Efficient points

We concentrate on deriving an equivalent algebraic description for the feasible set of problem (4.26).

The \textit{p}-level set of the distribution function $F_Z(z) = \Pr\{Z \leq z\}$ of $Z$ is defined as follows:

$$Z_p = \{ z \in \mathbb{R}^m : F_Z(z) \geq p \}.$$  \hfill (4.27)

Clearly, problem (4.26) can be compactly rewritten as

$$\begin{aligned}
\min_x & \ c(x) \\
n\text{s.t.} & \ g(x) \in Z_p, \\
& \ x \in \mathcal{X}.
\end{aligned}$$ \hfill (4.28)

\textbf{Lemma 4.57.} For every $p \in (0, 1)$ the level set $Z_p$ is nonempty and closed.

\textbf{Proof.} The statement follows from the monotonicity and the right continuity of the distribution function.

We introduce the key concept of a \textit{p}-efficient point.

\textbf{Definition 4.58.} Let $p \in (0, 1)$. A point $v \in \mathbb{R}^m$ is called a \textit{p}-efficient point of the probability distribution function $F$, if $F(v) \geq p$ and there is no $z \leq v$, $z \neq v$ such that $F(z) \geq p$.

The \textit{p}-efficient points are minimal points of the level set $Z_p$ with respect to the partial order in $\mathbb{R}^m$ generated by the nonnegative cone $\mathbb{R}_+^m$.

Clearly, for a scalar random variable $Z$ and for every $p \in (0, 1)$ there is exactly one \textit{p}-efficient point, which is the smallest $v$ such that $F_Z(v) \geq p$, i.e., $F_Z^{(-1)}(p)$.

\textbf{Lemma 4.59.} Let $p \in (0, 1)$ and let

$$l = (F_{Z_1}^{(-1)}(p), \ldots, F_{Z_m}^{(-1)}(p)).$$ \hfill (4.29)

Then every $v \in \mathbb{R}^m$ such that $F_Z(v) \geq p$ must satisfy the inequality $v \geq l$.

\textbf{Proof.} Let $v_i = F_{Z_i}^{(-1)}(p)$ be the \textit{p}-efficient point of the \textit{i}th marginal distribution function. We observe that $F_Z(v) \leq F_{Z_i}(v_i)$ for every $v \in \mathbb{R}^m$ and $i = 1, \ldots, m$, and, therefore, we obtain that the set of \textit{p}-efficient points is bounded from below.

Let $p \in (0, 1)$ and let $v^j$, $j \in \mathcal{E}$, be all \textit{p}-efficient points of $Z$. Here $\mathcal{E}$ is an arbitrary index set. We define the cones

$$K_j = v^j + \mathbb{R}_+^m, \quad j \in \mathcal{E}.$$ 

The following result can be derived from Phelps theorem [150, Lemma 3.12] about the existence of conical support points, but we can easily prove it directly.
Theorem 4.60. It holds that $Z_p = \bigcup_{j \in E} K_j$.

Proof. If $y \in Z_p$ then either $y$ is $p$-efficient or there exists a vector $w$ such that $w \leq y$, $w \neq y$, $w \in Z_p$. By Lemma 4.59, one must have $l \leq w \leq y$. The set $Z_1 = \{z \in Z_p : l \leq z \leq y\}$ is compact because the set $Z_p$ is closed by virtue of Lemma 4.57. Thus, there exists $w^1 \in Z_1$ with the minimal first coordinate. If $w^1$ is a $p$-efficient point, then $y \in w^1 + R^m_+$, what had to be shown. Otherwise, we define $Z_2 = \{z \in Z_p : l \leq z \leq w^1\}$, and choose a point $w^2 \in Z_2$ with the minimal second coordinate. Proceeding in the same way, we shall find the minimal element $w^m$ in the set $Z_p$ with $w^m \leq w^{m-1} \leq \cdots \leq y$. Therefore, $y \in w^m + R^m_+$, and this completes the proof. 

By virtue of Theorem 4.60 we obtain (for $0 < p < 1$) the following disjunctive semi-infinite formulation of problem (4.28):

$$\begin{align*}
\text{Min } & c(x) \\
\text{s.t. } & g(x) \in \bigcup_{j \in E} K_j, \\
& x \in X.
\end{align*}$$

(4.30)

This formulation provides an insight into the structure of the feasible set and the nature of its non-convexity. The main difficulty here is the implicit character of the disjunctive constraint.

Let $S$ stand for the simplex in $\mathbb{R}^{m+1}$,

$$S = \left\{ \alpha \in \mathbb{R}^{m+1} : \sum_{i=1}^{m+1} \alpha_i = 1, \alpha_i \geq 0 \right\}.$$ 

Denote the convex hull of the $p$-efficient points by $E$, i.e., $E = \text{conv}\{v^j, j \in E\}$. We obtain a semi-infinite disjunctive representation of the convex hull of $Z_p$.

Lemma 4.61. It holds that

$$\text{conv}(Z_p) = E + R^m_+.$$ 

Proof. By Theorem 4.60 every point $y \in \text{conv}Z$ can be represented as a convex combination of points in the cones $K_j$. By the theorem of Caratheodory the number of these points is no more than $m + 1$. Thus, we can write $y = \sum_{i=1}^{m+1} \alpha_i (v^j + w^i)$, where $w^i \in R^m_+$, $\alpha \in S$ and $j_i \in E$. The vector $w = \sum_{i=1}^{m+1} \alpha_i w^i$ belongs to $R^m_+$. Therefore, $y \in \sum_{i=1}^{m+1} \alpha_i v^j + R^m_+$. 

We also have the representation $E = \{ \sum_{i=1}^{m+1} \alpha_i v^j : \alpha \in S, j_i \in E\}$.

Theorem 4.62. For every $p \in (0, 1)$ the set $\text{conv}Z_p$ is closed.
4.3. Separable probabilistic constraints

Proof. Consider a sequence \( \{z^k\} \) of points of \( \text{conv} \mathcal{Z}_p \) which is convergent to a point \( \bar{z} \). Using Caratheodory’s theorem again, we have

\[
z^k = \sum_{i=1}^{m+1} \alpha^k_i y^k_i,
\]

with \( y^k_i \in \mathcal{Z}_p, \alpha^k_i \geq 0 \), and \( \sum_{i=1}^{m+1} \alpha^k_i = 1 \). By passing to a subsequence, if necessary, we can assume that the limits

\[
\bar{\alpha}_i = \lim_{k \to \infty} \alpha^k_i
\]

exist for all \( i = 1, \ldots, m+1 \). By Lemma 4.59 all points \( y^k_i \) are bounded below by some vector \( l \). For simplicity of notation we may assume that \( l = 0 \).

Let \( I = \{i : \bar{\alpha}_i > 0\} \). Clearly, \( \sum_{i \in I} \bar{\alpha}_i = 1 \). We obtain

\[
z^k \geq \sum_{i \in I} \alpha^k_i y^k_i.
\] (4.31)

We observe that \( 0 \leq \alpha^k_i y^k_i \leq \bar{z}^k \) for all \( i \in I \) and all \( k \). Since \( \{z^k\} \) is convergent and \( \alpha^k_i \to \bar{\alpha}_i > 0 \), each sequence \( \{y^k_i\}, i \in I \), is bounded. Therefore, we can assume that each of them is convergent to some limit \( \bar{y}_i, i \in I \). By virtue of Lemma 4.57 \( \bar{y}_i \in \mathcal{Z}_p \). Passing to the limit in inequality (4.31), we obtain

\[
\bar{z} \geq \sum_{i \in I} \bar{\alpha}_i \bar{y}_i \in \text{conv} \mathcal{Z}.
\]

Due to Lemma 4.61, we conclude that \( \bar{z} \in \text{conv} \mathcal{Z}_p \).

For a general random vector the set of \( p \)-efficient points may be unbounded and not closed, as illustrated on Figure 4.3.

We encounter also a relation between the \( p \)-efficient points and the extreme points of the convex hull of \( \mathcal{Z}_p \).

Theorem 4.63. For every \( p \in (0, 1) \) the set of extreme points of \( \text{conv} \mathcal{Z}_p \) is nonempty and it is contained in the set of \( p \)-efficient points.

Proof. Consider the lower bound \( l \) defined in (4.29). The set \( \text{conv} \mathcal{Z}_p \) is included in \( l + \mathbb{R}^m_+ \), by virtue of Lemma 4.59 and Lemma 4.61. Therefore, it does not contain any line. Since \( \text{conv} \mathcal{Z}_p \) is closed by Theorem 4.62, it has at least one extreme point.

Let \( w \) be an extreme point of \( \text{conv} \mathcal{Z}_p \). Suppose that \( w \) is not a \( p \)-efficient point. Then Theorem 4.60 implies that there exists a \( p \)-efficient point \( v \leq w, v \neq w \). Since \( w + \mathbb{R}^m_+ \subset \text{conv} \mathcal{Z}_p \), the point \( w \) is a convex combination of \( v \) and \( w + (w - v) \). Consequently, \( w \) cannot be extreme.

The representation becomes very handy when the vector \( Z \) has a discrete distribution on \( \mathbb{Z}^m \), in particular, if the problem is of form (4.57). We shall discuss this special case in more detail. Let us emphasize that our investigations extend to the case when the random
vector \( Z \) has a discrete distribution with values on a grid. Our further study can be adapted to the case of distributions on non-uniform grids for which a uniform lower bound on the distance of grid points in each coordinate exists. In this presentation, we assume that \( Z \in \mathbb{Z}^m \). In this case, we can establish that the distribution function \( F_Z \) has finitely many \( p \)-efficient points.

**Theorem 4.64.** For each \( p \in (0, 1) \) the set of \( p \)-efficient points of an integer random vector is nonempty and finite.

**Proof.** First we shall show that at least one \( p \)-efficient point exists. Since \( p < 1 \), there exists a point \( y \) such that \( F_Z(y) \geq p \). By Lemma 4.59, the level set \( Z \) is bounded from below by the vector \( l \) of \( p \)-efficient points of one-dimensional marginals. Therefore, if \( y \) is not \( p \)-efficient, one of finitely many integer points \( v \) such that \( l \leq v \leq y \) must be \( p \)-efficient.

Now we prove the finiteness of the set of \( p \)-efficient points. Suppose that there exists an infinite sequence of different \( p \)-efficient points \( v^j, j = 1, 2, \ldots \). Since they are integer, and the first coordinate \( v^j_1 \) is bounded from below by \( l_1 \), with no loss of generality we may select a subsequence which is non-decreasing in the first coordinate. By a similar token, we can select further subsequences which are non-decreasing in the first \( k \) coordinates \((k = 1, \ldots, m)\). Since the dimension \( m \) is finite, we obtain a subsequence of different \( p \)-efficient points which is non-decreasing in all coordinates. This contradicts the definition of a \( p \)-efficient point.

Note the crucial role of Lemma 4.59 in this proof. In conclusion, we have obtained that the disjunctive formulation (4.30) of problem (4.28) has a finite index set \( \mathcal{E} \).

Figure 4.4 illustrates the structure of the probabilistically constrained set for a discrete random variable.
4.3. Separable probabilistic constraints

The concept of $\alpha$-concavity on a set can be used at this moment to find an equivalent representation of the set $Z_p$ for a random vector with a discrete distribution.

**Theorem 4.65.** Let $Z_p$ be the set of all possible values of an integer random vector $Z$. If the distribution function $F_Z$ of $Z$ is $\alpha$-concave on $Z_p + Z_m^+$, for some $\alpha \in [-\infty, \infty]$, then for every $p \in (0, 1)$ one has

$$Z_p = \left\{ y \in \mathbb{R}^m : y \geq z \geq \sum_{j \in \mathcal{E}} \lambda_j v^j, \sum_{j \in \mathcal{E}} \lambda_j = 1, \lambda_j \geq 0, z \in \mathbb{Z}^m \right\},$$

where $v^j, j \in \mathcal{E}$, are the $p$-efficient points of $F$.

**Proof.** The representation (4.30) implies that

$$Z_p \subset \left\{ y \in \mathbb{R}^m : y \geq z \geq \sum_{j \in \mathcal{E}} \lambda_j v^j, \sum_{j \in \mathcal{E}} \lambda_j = 1, \lambda_j \geq 0, z \in \mathbb{Z}^m \right\},$$

We have to show that every point $y$ from the set at the right hand side belongs to $Z$. By the monotonicity of the distribution function $F_Z$, we have $F_Z(y) \geq F_Z(z)$ whenever $y \geq z$. Therefore, it is sufficient to show that $Pr\{Z \leq z\} \geq p$ for all $z \in \mathbb{Z}^m$ such that $z \geq \sum_{j \in \mathcal{E}} \lambda_j v^j$ with $\lambda_j \geq 0, \sum_{j \in \mathcal{E}} \lambda_j = 1$. We consider five cases with respect to $\alpha$.

**Case 1:** $\alpha = \infty$. It follows from the definition of $\alpha$-concavity that

$$F_Z(z) \geq \max\{F_Z(v^j), j \in \mathcal{E} : \lambda_j \neq 0\} \geq p.$$

**Case 2:** $\alpha = -\infty$. Since $F_Z(v^j) \geq p$ for each index $j \in \mathcal{E}$ such that $\lambda_j \neq 0$, the assertion follows as in Case 1.
Case 3: $\alpha = 0$. By the definition of $\alpha$-concavity, we have the following inequalities:

$$F_Z(z) \geq \prod_{j \in E} [F_Z(v^j)]^{\lambda_j} \geq \prod_{j \in E} p^{\lambda_j} = p.$$ 

Case 4: $\alpha \in (-\infty, 0)$. By the definition of $\alpha$-concavity,

$$[F_Z(z)]^\alpha \leq \sum_{j \in E} \lambda_j [F_Z(v^j)]^\alpha \leq \sum_{j \in E} \lambda_j p^\alpha = p^\alpha.$$ 

Since $\alpha < 0$, we obtain $F_Z(z) \geq p$.

Case 5: $\alpha \in (0, \infty)$. By the definition of $\alpha$-concavity,

$$[F_Z(z)]^\alpha \geq \sum_{j \in E} \lambda_j [F_Z(v^j)]^\alpha \geq \sum_{j \in E} \lambda_j p^\alpha = p^\alpha,$$

concluding that $z \in Z$, as desired. 

The consequence of this theorem is that under the $\alpha$-concavity assumption all integer point contained in $\text{conv} Z_p = E + R^n_m$ satisfy the probabilistic constraint. This demonstrates the importance of the notion of $\alpha$-concavity for discrete distribution functions as introduced in Definition 4.34. For example, the set $Z_p$ illustrated in Figure 4.4 does not correspond to any $\alpha$-concave distribution function, because its convex hull contains integer points which do not belong to $Z_p$. These are the points $(3,6), (4,5)$ and $(6,2)$.

Under the conditions of Theorem 4.65, problem (4.28) can be formulated in the following equivalent way:

$$\text{Min}_{x,z,\lambda} \ c(x) \quad (4.32)$$

s.t. $g(x) \geq z, \quad (4.33)$

$$z \geq \sum_{j \in E} \lambda_j v^j, \quad (4.34)$$

$$z \in Z^n, \quad (4.35)$$

$$\sum_{j \in E} \lambda_j = 1, \quad (4.36)$$

$$\lambda_j \geq 0, \ j \in E, \quad (4.37)$$

$$x \in X. \quad (4.38)$$

In this way, we have replaced the probabilistic constraint by algebraic equations and inequalities, together with the integrality requirement (4.35). This condition cannot be dropped, in general. However, if other conditions of the problem imply that $g(x)$ is integer, then we may remove $z$ entirely form the problem formulation. In this case, we replace constraints (4.33)–(4.35) with

$$g(x) \geq \sum_{j \in E} \lambda_j v^j.$$ 

For example, if the definition of $X$ contains the constraint $x \in Z^n$, and, in addition, $g(x) = T^T x$, where $T$ is a matrix with integer elements, then we can dispose of the variable $z$. 

4.3. Separable probabilistic constraints

If $Z$ takes values on a non-uniform grid, condition (4.35) should be replaced by the requirement that $z$ is a grid point.

4.3.3 Optimality conditions and duality theory

In this section, we return to problem formulation (4.28). We assume that $c : \mathbb{R}^n \to \mathbb{R}$ is a convex function. The mapping $g : \mathbb{R}^n \to \mathbb{R}^m$ has concave components $g_i : \mathbb{R}^n \to \mathbb{R}$. The set $\mathcal{X} \subset \mathbb{R}^n$ is closed and convex; the random vector $Z$ takes values in $\mathbb{R}^m$. The set $Z_p$ is defined as in (4.27). We split variables and consider the following formulation of the problem:

$$
\begin{align*}
\text{Min} \quad & c(x) \\
\text{subject to} \quad & g(x) \geq z, \\
& x \in \mathcal{X}, \\
& z \in Z_p.
\end{align*}
$$

(4.39)

Associating a Lagrange multiplier $u \in \mathbb{R}_+^m$ with the constraint $g(x) \geq z$, we obtain the Lagrangian function:

$$
L(x, z, u) = c(x) + u^T(z - g(x)).
$$

The dual functional has the form

$$
\Psi(u) = \inf_{(x,z) \in \mathcal{X} \times Z_p} L(x, z, u) = h(u) + d(u),
$$

where

$$
\begin{align*}
h(u) &= \inf \{ c(x) - u^T g(x) : x \in \mathcal{X} \}, \\
d(u) &= \inf \{ u^T z : z \in Z_p \}.
\end{align*}
$$

(4.40)

(4.41)

For any $u \in \mathbb{R}_+^m$ the value of $\Psi(u)$ is a lower bound on the optimal value $c^*$ of the original problem. The best Lagrangian lower bound will be given by the optimal value $\Psi^*$ of the problem:

$$
\sup_{u \geq 0} \Psi(u).
$$

(4.42)

We call (4.42) the dual problem to problem (4.39). For $u \not\geq 0$ one has $d(u) = -\infty$, because the set $Z_p$ contains a translation of $\mathbb{R}_+^m$. The function $d(\cdot)$ is concave. Note that $d(u) = -s_{Z_p}(-u)$, where $s_{Z_p}(\cdot)$ is the support function of the set $Z_p$. By virtue of Theorem 4.62 and Hiriart-Urruty and Lemaréchal [89, Chapter V, Prop.2.2.1], we have

$$
d(u) = \inf \{ u^T z : z \in \text{conv} Z_p \}.
$$

(4.43)

Let us consider the convex hull problem:

$$
\begin{align*}
\text{Min} \quad & c(x) \\
\text{subject to} \quad & g(x) \geq z, \\
& x \in \mathcal{X}, \\
& z \in \text{conv} Z_p.
\end{align*}
$$

(4.44)
We impose the following constraint qualification condition:

There exist points \( x^0 \in X \) and \( z^0 \in \text{conv} Z_p \) such that \( g(x^0) > z^0 \). \hspace{1cm} (4.45)

If this constraint qualification condition is satisfied, then the duality theory in convex programming Rockafellar [174, Corollary 28.2.1] implies that there exists \( \tilde{u} \geq 0 \) at which the minimum in (4.42) is attained, and \( \Psi^* = \Psi(\tilde{u}) \) is the optimal value of the convex hull problem (4.44).

We now study in detail the structure of the dual functional \( \Psi \). We shall characterize the solution sets of the two subproblems (4.40) and (4.41), which provide the values of the dual functional. Observe that the normal cone to the positive orthant at a point \( u \geq 0 \) is the following:

\[
N_{\text{conv} Z_p}(u) = \{ d \in \mathbb{R}^m : d_i = 0 \text{ if } u_i > 0, \ i = 1, \ldots, m \}. \hspace{1cm} (4.46)
\]

We define the set

\[
V(u) = \{ v \in \mathbb{R}^m : u^T v = d(u) \text{ and } v \text{ is } p\text{-efficient} \}. \hspace{1cm} (4.47)
\]

**Lemma 4.66.** For every \( u > 0 \) the solution set of (4.41) is nonempty. For every \( u \geq 0 \) it has the following form: \( \hat{Z}(u) = V(u) - N_{\text{conv} Z_p}(u) \).

**Proof.** First we consider the case \( u > 0 \). Then every recession direction \( q \) of \( Z_p \) satisfies \( u^T q > 0 \). Since \( Z_p \) is closed, a solution to (4.41) must exist. Suppose that a solution \( z \) to (4.41) is not a \( p\)-efficient point. By virtue of Theorem 4.60, there is a \( p\)-efficient \( v \in Z_p \) such that \( v \leq z \), and \( v \neq z \). Thus, \( u^T v < u^T z \), which is a contradiction. Therefore, we conclude that there is a \( p\)-efficient point \( v \), which solves problem (4.41).

Consider the general case \( u \geq 0 \) and assume that the solution set of problem (4.41) is nonempty. In this case, the solution set always contains a \( p\)-efficient point. Indeed, if a solution \( z \) is not \( p\)-efficient, we must have a \( p\)-efficient point \( v \) dominated by \( z \), and \( u^T v \leq u^T z \) holds by the nonnegativity of \( u \). Consequently, \( u^T v = u^T z \) for all \( p\)-efficient \( v \leq z \), which is equivalent to \( z \in \{ v \} - N_{\text{conv} Z_p}(u) \), as required.

If the solution set of (4.41) is empty then \( V(u) = \emptyset \) by definition and the assertion is true as well. \( \Box \)

The last result allows us to calculate the subdifferential of the function \( d \) in a closed form.

**Lemma 4.67.** For every \( u \geq 0 \) one has \( \partial d(u) = \text{conv}(V(u)) - N_{\text{conv} Z_p}(u) \). If \( u > 0 \) then \( \partial d(u) \) is nonempty.

**Proof.** From (4.41) we obtain \( d(u) = -s_{z_p}(u) \), where \( s_{z_p}(\cdot) \) is the support function of \( Z_p \) and, consequently, of \( \text{conv} Z_p \). Consider the indicator function \( I_{\text{conv} Z_p}(\cdot) \) of the set \( \text{conv} Z_p \). By virtue of Corollary 16.5.1 in Rockafellar [174], we have

\[
s_{z_p}(u) = I_{\text{conv} Z_p}^*(u),
\]

where the latter function is the conjugate of the indicator function \( I_{\text{conv} Z_p}(\cdot) \). Thus,

\[
\partial d(u) = -\partial I_{\text{conv} Z_p}^*(u) \hspace{1cm} (4.48)
\]
4.3. Separable probabilistic constraints

Recall that \( \text{conv} \ Z_p \) is closed, by Theorem 4.62. Using Rockafellar [174, Thm 23.5], we observe that \( y \in \partial I^*_\text{conv} \ Z_p (u) \) if and only if \( I^*_\text{conv} \ Z_p (u) + I^*_\text{conv} \ Z_p (y) = -y^T u \). It follows that \( y \in \text{conv} \ Z_p \) and \( I^*_\text{conv} \ Z_p (u) = -y^T u \). Consequently,

\[
y^T u = d(u). \tag{4.48}
\]

Since \( y \in \text{conv} \ Z_p \) we can represent it as follows:

\[
y = \sum_{j=1}^{m+1} \alpha_j e^j + w,
\]

where \( e^j, j = 1, \ldots, m + 1 \), are extreme points of \( \text{conv} \ Z_p \) and \( w \geq 0 \). Using Theorem 4.63 we conclude that \( e^j \) are \( p \)-efficient points. Moreover, applying \( u \), we obtain

\[
y^T u = \sum_{j=1}^{m+1} \alpha_j u^T e^j + u^T w \geq d(u), \tag{4.49}
\]

because \( u^T e^j \geq d(u) \) and \( u^T w \geq 0 \). Combining (4.48) and (4.49) we conclude that \( u^T e^j = d(u) \) for all \( j \), and \( u^T w = 0 \). Thus \( y \in \text{conv} V(u) - N_{\mathbb{R}^m}(u) \).

Conversely, if \( y \in \text{conv} V(u) - N_{\mathbb{R}^m}(u) \) then (4.48) holds true by the definitions of the set \( V(u) \) and the normal cone. This implies that \( y \in \partial d(u) \), as required.

Furthermore, the set \( \partial d(u) \) is nonempty for \( u > 0 \) due to Lemma 4.66.

Now, we analyze the function \( h(\cdot) \). Define the set of minimizers in (4.40),

\[
X(u) = \{ x \in \mathcal{X} : c(x) - u^T g(x) = h(u) \}.
\]

Since the set \( \mathcal{X} \) is convex and the objective function of problem (4.40) is convex for all \( u \geq 0 \), we conclude that the solution set \( X(u) \) is convex for all \( u \geq 0 \).

**Lemma 4.68.** Assume that the set \( \mathcal{X} \) is compact. For every \( u \in \mathbb{R}^m \), the subdifferential of the function \( h \) is described as follows:

\[
\partial h(u) = \text{conv} \{ -g(x) : x \in X(u) \}.
\]

**Proof.** The function \( h \) is concave on \( \mathbb{R}^m \). Since the set \( \mathcal{X} \) is compact, \( c \) is convex, and \( g_i, i = 1, \ldots, m \) are concave, the set \( X(u) \) is compact. Therefore, the subdifferential of \( h(u) \) for every \( u \in \mathbb{R}^m \) is the closure of \( \text{conv} \{ -g(x) : x \in X(u) \} \) (see Hiriart-Urruty and C. Lemaréchal [89, Chapter VI, Lemma 4.4.2]). By the compactness of \( X(u) \) and concavity of \( g \), the set \( \{ -g(x) : x \in X(u) \} \) is closed. Therefore, we can omit taking the closure in the description of the subdifferential of \( h(u) \).

This analysis provides the basis for the following necessary and sufficient optimality conditions for problem (4.42).
Theorem 4.69. Assume that the constraint qualification condition (4.45) is satisfied and that the set $X$ is compact. A vector $u \geq 0$ is an optimal solution of (4.42) if and only if there exists a point $x \in X(u)$, points $v^1, \ldots, v^{m+1} \in V(u)$ and scalars $\beta_1, \ldots, \beta_{m+1} \geq 0$ with $\sum_{j=1}^{m+1} \beta_j = 1$, such that

$$\sum_{j=1}^{m+1} \beta_j v^j - g(x) \in N_{\mathbb{R}^m}(u). \quad (4.50)$$

Proof. Using Rockafellar [174, Thm. 27.4], the necessary and sufficient optimality condition for (4.42) has the form

$$0 \in -\partial \Psi(u) + N_{\mathbb{R}^m}(u). \quad (4.51)$$

Since $\text{int dom } d \neq \emptyset$ and $\text{dom } h = \mathbb{R}^m$ we have $\partial \Psi(u) = \partial h(u) + \partial d(u)$. Using Lemma 4.67 and Lemma 4.68, we conclude that there exist $p$-efficient points $v^j \in V(u)$, $j = 1, \ldots, m + 1$, $\beta^j \geq 0$, $j = 1, \ldots, m + 1$, $\sum_{j=1}^{m+1} \beta_j = 1$, $x^j \in X(u)$, $j = 1, \ldots, m + 1$, $\alpha^j \geq 0$, $j = 1, \ldots, m + 1$, $\sum_{j=1}^{m+1} \alpha_j = 1$, $\sum_{j=1}^{m+1} \beta_j v^j \in -N_{\mathbb{R}^m}(u). \quad (4.52)$

such that

$$\sum_{j=1}^{m+1} \alpha_j g(x^j) - \sum_{j=1}^{m+1} \beta_j v^j \in -N_{\mathbb{R}^m}(u). \quad (4.53)$$

If the function $c$ was strictly convex, or $g$ was strictly concave, then the set $X(u)$ would be a singleton. In this case, all $x^j$ would be identical and the above relation would immediately imply (4.50). Otherwise, let us define

$$x = \sum_{j=1}^{m+1} \alpha_j x^j.$$ 

By the convexity of $X(u)$ we have $x \in X(u)$. Consequently,

$$c(x) - \sum_{i=1}^{m} u_ig_i(x) = h(u) = c(x^j) - \sum_{i=1}^{m} u_ig_i(x^j), \quad j = 1, \ldots, m + 1. \quad (4.54)$$

Multiplying the last equation by $\alpha_j$ and adding, we obtain

$$c(x) - \sum_{i=1}^{m} u_ig_i(x) = \sum_{j=1}^{m+1} \alpha_j \left[ c(x^j) - \sum_{i=1}^{m} u_ig_i(x^j) \right] \geq c(x) - \sum_{i=1}^{m} u_1 \sum_{j=1}^{m+1} \alpha_j g_i(x^j).$$
4.3. Separable probabilistic constraints

The last inequality follows from the convexity of \( c \). We have the following inequality:

\[
\sum_{i=1}^{m} u_i \left[ g_i(x) - \sum_{j=1}^{m+1} \alpha_j g_i(x^j) \right] \leq 0.
\]

Since the functions \( g_i \) are concave, we have \( g_i(x) \geq \sum_{j=1}^{m+1} \alpha_j g_i(x^j) \). Therefore, we conclude that \( u_i = 0 \) whenever \( g_i(x) > \sum_{j=1}^{m+1} \alpha_j g_i(x^j) \). This implies that

\[
g(x) - \sum_{j=1}^{m+1} \alpha_j g(x^j) \in -\mathcal{N}_{\mathbb{R}^{m+1}}(u).
\]

Since \( \mathcal{N}_{\mathbb{R}^{m+1}}(u) \) is a convex cone, we can combine the last relation with (4.53) and obtain (4.50), as required.

Now, we prove the converse implication. Assume that we have \( x \in X(u) \), points \( v^1, \ldots, v^{m+1} \in V(u) \) and scalars \( \beta_1, \ldots, \beta_{m+1} \geq 0 \) with \( \sum_{j=1}^{m+1} \beta_j = 1 \), such that (4.50) holds true. By Lemma 4.67 and Lemma 4.68 we have

\[
-g(x) + \sum_{j=1}^{m+1} \beta_j v^j \in \partial \Psi(u).
\]

Thus (4.50) implies (4.51), which is a necessary and sufficient optimality condition for problem (4.42).

Using these optimality conditions we obtain the following duality result.

**Theorem 4.70.** Assume that the constraint qualification condition (4.45) for problem (4.39) is satisfied, the probability distribution of the vector \( Z \) is \( \alpha \)-concave for some \( \alpha \in [-\infty, \infty] \), and the set \( X \) is compact. If a point \( (\hat{x}, \hat{z}) \) is an optimal solution of (4.39), then there exists a vector \( \hat{u} \geq 0 \), which is an optimal solution of (4.42) and the optimal values of both problems are equal. If \( \hat{u} \) is an optimal solution of problem (4.42), then there exist a point \( \hat{x} \), such that \( (\hat{x}, g(\hat{x})) \) is a solution of problem (4.39), and the optimal values of both problems are equal.

**Proof.** The \( \alpha \)-concavity assumption implies that problems (4.39) and (4.44) are the same. If \( \hat{u} \) is optimal solution of problem (4.42), we obtain the existence of points \( \hat{x} \in X(\hat{u}) \), \( v^1, \ldots, v^{m+1} \in V(u) \) and scalars \( \beta_1, \ldots, \beta_{m+1} \geq 0 \) with \( \sum_{j=1}^{m+1} \beta_j = 1 \), such that the optimality conditions in Theorem 4.69 are satisfied. Setting \( \bar{z} = g(\hat{x}) \) we have to show that \( (\hat{x}, \bar{z}) \) is an optimal solution of problem (4.39) and that the optimal values of both problems are equal. First we observe that this point is feasible. We choose \( y \in -\mathcal{N}_{\mathbb{R}^{m+1}}(\hat{u}) \) such that \( y = g(\hat{x}) - \sum_{j=1}^{m+1} \beta_j v^j \). From the definitions of \( X(\hat{u}), V(\hat{u}) \), and the normal cone, we
obtain
\[ h(\hat{u}) = c(\hat{x}) - \hat{u}^T g(\hat{x}) = c(\hat{x}) - \hat{u}^T \left( \sum_{j=1}^{m+1} \beta_j v^j + y \right) \]
\[ = c(\hat{x}) - \sum_{j=1}^{m+1} \beta_j d(\hat{u}) - \hat{u}^T y = c(\hat{x}) - d(\hat{u}). \]
Thus,
\[ c(\hat{x}) = h(\hat{u}) + d(\hat{u}) = \Psi^* \geq c^*, \]
which proves that \((\hat{x}, \hat{z})\) is an optimal solution of problem (4.39) and \(\Psi^* = c^*\).

If \((\hat{x}, \hat{z})\) is a solution of (4.39), then by Rockafellar [174, Thm. 28.4] there is a vector \(\hat{u} \geq 0\) such that
\[ \hat{u}_i (\hat{z}_i - g_i(\hat{x})) = 0 \]
and
\[ 0 \in \partial c(\hat{x}) + \partial \hat{u}^T g(\hat{x}) - \hat{z} + N_{X \times Z}(\hat{x}, \hat{z}). \]
This means that
\[ 0 \in \partial c(\hat{x}) - \partial u^T g(\hat{x}) + N_X(\hat{x}) \quad (4.55) \]
and
\[ 0 \in \hat{u} + N_Z(\hat{z}). \quad (4.56) \]
The first inclusion (4.55) is optimality condition for problem (4.40), and thus \(x \in X(\hat{u})\). By virtue of Rockafellar [174, Thm. 23.5] the inclusion (4.56) is equivalent to \(\hat{z} \in \partial I_{Z^p}^*(\hat{u})\). Using Lemma 4.67 we obtain that
\[ \hat{z} \in \partial d(\hat{u}) = \text{conv} V(\hat{u}) - N_{R^m}(\hat{u}). \]
Thus, there exists points \(v^1, \ldots, v^{m+1} \in V(u)\) and scalars \(\beta_1, \ldots, \beta_{m+1} \geq 0\) with \(\sum_{j=1}^{m+1} \beta_j = 1\), such that
\[ \hat{z} - \sum_{j=1}^{m+1} \beta_j v^j \in -N_{R^m}(\hat{u}). \]
Using the complementarity condition \(\hat{u}_i (\hat{z}_i - g_i(\hat{x})) = 0\) we conclude that the optimality conditions of Theorem 4.69 are satisfied. Thus, \(\hat{u}\) is an optimal solution of problem (4.42).

For the special case of discrete distribution and linear constraints we can obtain more specific necessary and sufficient optimality conditions.

In the linear probabilistic optimization problem, we have \(g(x) = T^T x\), where \(T\) is an \(m \times n\) matrix, and \(c(x) = c^T x\) with \(c \in \mathbb{R}^n\). Furthermore, we assume that \(\mathcal{X}\) is a closed convex polyhedral set, defined by a system of linear inequalities. The problem reads as follows:
\[
\begin{align*}
\text{Min} & \quad c^T x \\
\text{s.t.} & \quad \Pr\{T^T x \geq Z\} \geq p, \\
& \quad Ax \geq b, \\
& \quad x \geq 0.
\end{align*}
\quad (4.57)
\]
4.3. Separable probabilistic constraints

Here $A$ is an $s \times n$ matrix and $b \in \mathbb{R}^s$.

**Definition 4.71.** Problem (4.57) satisfies the dual feasibility condition if

$$\Lambda = \{(u, w) \in \mathbb{R}^{m+s}_+ : A^Tw + T^Tu \leq c\} \neq \emptyset.$$  

**Theorem 4.72.** Assume that the feasible set of (4.57) is nonempty and that $Z$ has a discrete distribution on $\mathbb{Z}^m$. Then (4.57) has an optimal solution if and only if it satisfies the LQ condition, defined in (4.71).

**Proof.** If (4.57) has an optimal solution, then for some $j \in \mathcal{E}$ the linear optimization problem

$$\begin{align*}
\min_x c^T x \\
\text{s.t. } Tx \geq v^j, \\
Ax \geq b, \\
x \geq 0,
\end{align*}$$

(4.58)

has an optimal solution. By duality in linear programming, its dual problem

$$\begin{align*}
\max_{u, w} & \quad u^Tv^j + b^Tw \\
\text{s.t. } & \quad T^Tu + A^Tw \leq c, \\
& \quad u, w \geq 0,
\end{align*}$$

(4.59)

has an optimal solution and the optimal values of both programs are equal. Thus, the dual feasibility condition (4.71) must be satisfied. On the other hand, if the dual feasibility condition is satisfied, all dual programs (4.59) for $j \in \mathcal{E}$ have nonempty feasible sets, so the objective values of all primal problems (4.58) are bounded from below. Since at least one of them has a nonempty feasible set by assumption, an optimal solution must exist. □

**Example 4.73 (Vehicle Routing continued)** We return to the vehicle routing Example 4.1, introduced at the beginning of the chapter. The convex hull problem reads:

$$\begin{align*}
\min_{x, \lambda} & \quad c^T x \\
\text{s.t. } & \quad \sum_{i=1}^n t_{ii}x_i \geq \sum_{j \in \mathcal{E}} \lambda_j v^j, \\
& \quad \sum_{j \in \mathcal{E}} \lambda_j = 1, \\
& \quad x \geq 0, \lambda \geq 0.
\end{align*}$$

(4.60)

We assign a Lagrange multiplier $u$ to constraint (4.60) and a multiplier $\mu$ to constraint...
Chapter 4. Optimization Models with Probabilistic Constraints

(4.61). The dual problem has the form:

$$\text{Max } \mu, \mu$$

s.t. \( \sum_{i=1}^{m} \lambda_{i} u_{i} \leq c_{i}, \ i = 1, 2, \ldots, n, \)

\( \mu \leq u^{\top} v^{j}, \ j \in E, \)

\( u \geq 0. \)

We see that \( u_{i} \) provides the increase of routing cost if the demand on arc \( l \) increases by 1 unit, \( \mu \) is the minimum cost for covering the demand with probability \( p \), and the \( p \)-efficient points \( v^{j} \) correspond to critical demand levels, that have to be covered. The auxiliary problem \( \text{Min } u^{\top} z \) identifies \( p \)-efficient points, which represent critical demand levels. The optimal value of this problem provides the minimum total cost of a critical demand.

Our duality theory finds interesting interpretation in the context of the cash matching problem in Example 4.6.

Example 4.74 (Cash Matching continued) Recall the problem formulation

$$\text{Max } \mathbb{E}[U(c_{T} - Z_{T})]$$

s.t. \( \text{Pr}\{c_{t} \geq Z_{t}, \ t = 1, \ldots, T\} \geq p, \)

\( c_{t} = c_{t-1} + \sum_{i=1}^{n} a_{it} x_{i}, \ t = 1, \ldots, T, \)

\( x \geq 0. \)

If the vector \( Z \) has a quasi-concave distribution (e.g., joint normal distribution), the resulting problem is convex.

The convex hull problem (4.44) can be written as follows:

$$\text{Max } \mathbb{E}[U(c_{T} - Z_{T})]$$

s.t. \( c_{t} = c_{t-1} + \sum_{i=1}^{n} a_{it} x_{i}, \ t = 1, \ldots, T, \) (4.63)

\( c_{t} \geq \sum_{j=1}^{T+1} \lambda_{j} v_{t}^{j}, \ t = 1, \ldots, T, \) (4.64)

\( \sum_{j=1}^{T+1} \lambda_{j} = 1, \) (4.65)

\( \lambda \geq 0, \ x \geq 0. \) (4.66)

In constraint (4.64) the vectors \( v^{j} = (v_{1}^{j}, \ldots, v_{n}^{j}) \), for \( j = 1, \ldots, T + 1 \), are \( p \)-efficient trajectories of the cumulative liabilities \( Z = (Z_{1}, \ldots, Z_{T}) \). Constraints (4.64)–(4.66)
4.3. Separable probabilistic constraints

require that the cumulative cash flows are greater than or equal to some convex combination of \( p \)-efficient trajectories. Recall that by Lemma 4.61, no more than \( T + 1 \) \( p \)-efficient trajectories are needed. Unfortunately, we do not know the optimal collection of these trajectories.

Let us assign non-negative Lagrange multipliers \( u = (u_1, \ldots, u_T) \) to the constraint (4.64), multipliers \( w = (w_1, \ldots, w_T) \) to the constraints (4.63), and a multiplier \( \rho \in \mathbb{R} \) to the constraint (4.65). To simplify notation, we define the function \( \bar{U} : \mathbb{R} \to \mathbb{R} \) as follows:

\[
\bar{U}(y) = E[U(y - Z_T)].
\]

It is a concave nondecreasing function of \( y \) due to the properties of \( U(\cdot) \). We make the convention that its conjugate is defined as follows:

\[
\bar{U}^*(u) = \inf_y \{uy - \bar{U}(y)\}.
\]

Consider the dual function of the convex hull problem:

\[
D(w, u, \rho) = \min_{x \geq 0, \lambda \geq 0, c} \left\{ -\bar{U}(c_T) + \sum_{t=1}^T w_t \left( c_t - c_{t-1} - \sum_{i=1}^n a_{it} x_i \right) + \sum_{t=1}^T u_t \left( \sum_{j=1}^{T+1} \lambda_j v_{jt}^2 - c_t \right) + \rho \left( 1 - \sum_{j=1}^{T+1} \lambda_j \right) \right\}
\]

\[
= -\max_{x \geq 0} \sum_{i=1}^n \sum_{t=1}^T a_{it} w_t x_i + \min_{\lambda \geq 0} \sum_{j=1}^{T+1} \left( \sum_{t=1}^T v_{jt}^2 u_t - \rho \right) \lambda_j + \rho
\]

\[
+ \min_c \left\{ \sum_{t=1}^{T-1} c_t (w_t - u_t - w_{t+1}) - w_1 c_0 + c_T (w_T - u_T) - \bar{U}(c_T) \right\}
\]

\[
= \rho - w_1 c_0 + \bar{U}^*(w_T - u_T).
\]

The dual problem becomes:

\[
\begin{align*}
\text{Min}_{u, w, \rho} & \quad -\bar{U}^*(w_T - u_T) + w_1 c_0 - \rho \\
\text{s.t.} \quad & \quad w_t = w_{t+1} + u_t, \quad t = T - 1, \ldots, 1, \\
& \quad \sum_{t=1}^T w_t a_{it} \leq 0, \quad i = 1, \ldots, n, \\
& \quad \rho \leq \sum_{t=1}^T u_t v_{jt}^2, \quad j = 1, \ldots, T + 1. \\
& \quad u \geq 0.
\end{align*}
\]

We can observe that each dual variable \( u_t \) is the cost of borrowing a unit of cash for one time period, \( t \). The amount \( u_t \) is to be paid at the end of the planning horizon. It follows from (4.68) that each multiplier \( w_t \) is the amount that has to be returned at the end of the planning horizon if a unit of cash is borrowed at \( t \) and held until time \( T \).
Chapter 4. Optimization Models with Probabilistic Constraints

The constraints (4.69) represent the non-arbitrage condition. For each bond \( i \) we can consider the following operation: borrow money to buy the bond and lend away its coupon payments, according to the rates implied by \( w_t \)'s. At the end of the planning horizon, we collect all loans and pay off the debt. The profit from this operation should be non-positive for each bond in order to comply with the “no free lunch” condition, which is expressed via (4.69).

Let us observe that each product \( u_t v_{jt} \) is the amount that has to be paid at the end, for having a debt in the amount \( v_{jt} \) in period \( t \). Recall that \( v_{jt} \) is the \( p \)-efficient cumulative liability up to time \( t \). Denote the implied one-period liabilities by

\[
L^j_t = v_{jt} - v_{j,t-1}, \quad t = 2, \ldots, T,
\]

\[
L^j_1 = v_{j1}.
\]

Changing the order of summation, we obtain

\[
\sum_{t=1}^T u_t v_{jt} = \sum_{t=1}^T u_t \sum_{r=1}^t L^j_r = \sum_{t=1}^T \sum_{r=t}^T u_t = \sum_{t=1}^T L^j_t (w_T + u_T - w_T).
\]

It follows that the sum appearing on the right hand side of (4.70) can be viewed as the extra cost of covering the \( j \)th \( p \)-efficient liability sequence by borrowed money, that is, the difference between the amount that has to be returned at the end of the planning horizon, and the total liability discounted by \( w_T - u_T \).

If we consider the special case of a linear expected utility:

\[
\hat{U}(c_T) = c_T - \mathbb{E}[Z_T],
\]

Then we can skip the constant \( \mathbb{E}[Z_T] \) in the formulation of the optimization problem. The dual function of the convexified cash matching problem becomes:

\[
D(w, u, \rho) = - \max_{x \geq 0} \sum_{i=1}^n \sum_{t=1}^T a_{it} w_t x_i + \min_{\lambda \geq 0} \left( \sum_{t=1}^T v^j_t u_t - \rho \right) \lambda_j + \rho
\]

\[
+ \min_{c} \left\{ \sum_{t=1}^{T-1} c_t (w_t - u_t - w_{t+1}) - w_1 c_0 + c_T (w_T - u_T - 1) \right\}
\]

\[
= \rho - w_1 c_0.
\]

The objective function of the dual problem takes on the form:

\[
\min_{u, w, \rho} w_1 c_0 - \rho
\]

and the constraints (4.68) extends to all time periods:

\[
w_t = w_{t+1} + u_t, \quad t = T, T - 1, \ldots, 1
\]

with the convention \( w_{T+1} = 1 \).

In this case, the sum on the right hand side of (4.70) is the difference between the cost of covering the \( j \)th \( p \)-efficient liability sequence by borrowed money and the total liability.
The variable \( \rho \) represents the minimal cost of this form, for all \( p \)-efficient trajectories. This allows us to interpret the dual objective function in this special case as the amount obtained at \( T \) for lending away our capital \( c_0 \) decreased by the extra cost of covering a \( p \)-efficient liability sequence by borrowed money. By duality this quantity is the same as \( c_T \), which implies that both ways of covering the liabilities are equally profitable. In the case of a general utility function, the dual objective function contains an additional adjustment term.

### 4.4 Optimization problems with non-separable probabilistic constraints

In this section, we concentrate on the following problem:

\[
\begin{align*}
\min_{x} & \quad c(x) \\
\text{s.t.} & \quad \Pr \{ g(x, Z) \geq 0 \} \geq p, \\
& \quad x \in X.
\end{align*}
\]

The parameter \( p \in (0, 1) \) denotes some probability level. We assume that the functions \( c : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R} \) and \( g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m \) are continuous and the set \( X \subset \mathbb{R}^n \) is a closed convex set. We define the constraint function as follows:

\[
G(x) = \Pr \{ g(x, Z) \geq 0 \}.
\]

Recall that if \( G(\cdot) \) is \( \alpha \)-concave function, \( \alpha \in \mathbb{R} \), then a transformation of it is a concave function. In this case, we define

\[
\tilde{G}(x) = \begin{cases} 
\ln p - \ln [G(x)], & \text{if } \alpha = 0, \\
p^\alpha - [G(x)]^\alpha, & \text{if } \alpha > 0, \\
[G(x)]^\alpha - p^\alpha, & \text{if } \alpha < 0.
\end{cases}
\]

We obtain the following equivalent formulation of problem (4.72):

\[
\begin{align*}
\min_{x} & \quad c(x) \\
\text{s.t.} & \quad \tilde{G}(x) \leq 0, \\
& \quad x \in X.
\end{align*}
\]

Assuming that \( c(\cdot) \) is convex, we have a convex problem.

Recall that Slater’s condition is satisfied for problem (4.72) if there is a point \( x^* \in \text{int} X \) such that \( \tilde{G}(x^*) > 0 \). Using optimality conditions for convex optimization problems, we can infer the following conditions for problem (4.72):

**Theorem 4.75.** Assume that \( c(\cdot) \) is a continuous convex function, the functions \( g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^m \) are quasi-concave, \( Z \) has an \( \alpha \)-concave distribution, and the set \( X \subset \mathbb{R}^n \) is closed and convex. Furthermore, let Slater’s condition be satisfied and \( \text{int} \, \text{dom} \, G \neq \emptyset \).
A point $\hat{x} \in X$ is an optimal solution of problem (4.72) if and only if there is a number $\lambda \in \mathbb{R}_+$ such that

$$0 \in \partial c(\hat{x}) + \lambda \frac{1}{\alpha} G(\hat{x})^{1-\alpha} \partial G(\hat{x})^{\alpha} + N_X(\hat{x}) \quad \text{if } \alpha \neq 0,$$

or

$$0 \in \partial c(\hat{x}) + \lambda G(\hat{x}) \partial \ln G(\hat{x}) + N_X(\hat{x}) \quad \text{if } \alpha = 0.$$

Proof. Under the assumptions of the theorem problem (4.72) can be reformulated in form (4.74), which is a convex optimization problem. The optimality conditions follow from the optimality conditions for convex optimization problems using Theorem 4.29. Due to Slater’s condition, we have that $G(x) > 0$ on a set with non-empty interior, and therefore, the assumptions of Theorem 4.29 are satisfied.

4.4.1 Differentiability of probability functions and optimality conditions

We can avoid concavity assumptions and replace them by differentiability requirements. Under certain assumptions, we can differentiate the probability function and obtain optimality conditions in a differential form. For this purpose, we assume that $Z$ has a probability density function $\theta(z)$, and that the support of $P_Z$ is a closed set with a piece-wise smooth boundary such that $\text{supp} P_Z = \text{cl}\{\text{int}(\text{supp} P_Z)\}$. For example, it can be the union of several disjoint sets, but cannot contain isolated points, or surfaces of zero Lebesgue measure.

Consider the multifunction $H : \mathbb{R}^n \Rightarrow \mathbb{R}^s$, defined as follows:

$$H(x) = \{z \in \mathbb{R}^s : g_i(x, z) \geq 0, i = 1, \ldots, m\}.$$

We denote the boundary of a set $H(x)$ by $\partial H(x)$. For an open set $U \subset \mathbb{R}^n$ containing the origin, we set:

$$H_U = \text{cl}(\cup_{x \in U} H(x)) \quad \text{and} \quad \triangle H_U = \text{cl}(\cup_{x \in U} \partial H(x)),$$

$$V_U = \text{cl}(U \times H_U) \quad \text{and} \quad \triangle V_U = \text{cl}(U \times \triangle H_U).$$

For any of these sets, we indicate with upper subscript $r$ its restriction to the $\text{supp} P_Z$, e.g., $H_U^r = H_U \cap \text{supp} P_Z$. Let

$$S_i(x) = \{z \in \text{supp} P_Z : g_i(x, z) = 0, g_j(x, z) \geq 0, j \neq i\}, \quad i = 1, \ldots, m.$$

We use the notation

$$S(x) = \cup_{i=1}^m S_i(x), \quad \triangle H_i = \text{int}\left(\cup_{x \in U} (\partial \{g_i(x, z) \geq 0\} \cap H^r(x))\right).$$

The $(m - 1)$-dimensional Lebesgue measure is denoted by $P_{m-1}$. We assume that the functions $g_i(x, z)$, $i = 1, \ldots, m$, are continuously differentiable, and such that $\partial H(x) = S(x)$.
4.4. Optimization problems with non-separable probabilistic constraints

with \( S(x) \) being the \((s-1)\)-dimensional surface of the set \( H(x) \subset \mathbb{R}^s \). The set \( H_U \) is the union of all sets \( H(x) \) when \( x \in U \), and, correspondingly, \( \Delta H_U \) contains all surfaces \( S(x) \) when \( x \in U \).

First we formulate and prove a result about the differentiability of the probability function for a single constraint function \( g(x, z) \), that is, \( m = 1 \). In this case we omit the index for the function \( g \), as well as for the set \( S(x) \).

**Theorem 4.76.** Assume that:
(i) the vector functions \( \nabla_x g(x, z) \) and \( \nabla_z g(x, z) \) are continuous on \( \Delta V_U \);
(ii) the vector functions \( \nabla_z g(x, z) > 0 \) (componentwise) on the set \( \Delta V_U \);
(iii) the function \( \| \nabla_x g(x, z) \| > 0 \) on \( \Delta V_U \).

Then the probability function \( G(x) = \text{Pr}\{g(x, Z) \geq 0\} \) has partial derivatives for almost all \( x \in U \) that can be represented as a surface integral

\[
\left( \frac{\partial G(x)}{\partial x_i} \right)^n_{i=1} = \int_{bdH(x) \cap \text{supp} P_z} \frac{\theta(z)}{\| \nabla_x g(x, z) \|} \nabla_z g(x, z) dS.
\]

**Proof.** Without loss of generality, we shall assume that \( x \in U \subset \mathbb{R} \).

For two points \( x, y \in U \), we consider the difference:

\[
G(x) - G(y) = \int_{H(x)} \theta(z) dz - \int_{H(y)} \theta(z) dz = \int_{H(x) \setminus H(y) \cap H(x)} \theta(z) dz - \int_{H(y) \setminus H(x) \cap H(x)} \theta(z) dz.
\]

By the Implicit Function Theorem, the equation \( g(x, z) = 0 \) determines a differentiable function \( x(z) \) such that

\[
g(x(z), z) = 0 \quad \text{and} \quad \nabla_z x(z) = -\frac{\nabla_x g(x, z)}{\nabla_z g(x, z)}|_{x=x(z)}.
\]

Moreover, the constraint \( g(x, z) \geq 0 \) is equivalent to \( x \geq x(z) \) for all \( (x, z) \in U \times \Delta H_U^r \), because the function \( g(\cdot, z) \) strictly increases on this set due to the assumption (iii). Thus, for all points \( x, y \in U \) such that \( x < y \), we can write:

\[
H^r(x) \setminus H^r(y) = \{ z \in \mathbb{R}^s : g(x, z) \geq 0, g(y, z) < 0 \} = \{ z \in \mathbb{R}^s : x \geq x(z) > y \} = 0,
\]

\[
H^r(y) \setminus H^r(x) = \{ z \in \mathbb{R}^s : g(y, z) \geq 0, g(x, z) < 0 \} = \{ z \in \mathbb{R}^s : y \geq x(z) > x \}.
\]

Hence, we can continue our representation of the difference (4.75) as follows:

\[
G(x) - G(y) = -\int_{\{ z \in \mathbb{R}^s : y \geq x(z) > x \}} \theta(z) dz.
\]
Now, we apply Schwarz [194, Vol. 1, Theorem 108] and obtain
\[
G(x) - G(y) = -\int_y^x \int_{\{z \in \mathbb{R}^s : x(z) = t\}} \frac{\theta(z)}{\|\nabla z g(x(z))\|} dS dt \\
= \int_x^y \int_{\text{bd}(H(x))} \frac{\|\nabla_z g(t, z)\|}{\|\nabla_z g(x, z)\|} dS dt.
\]

By Fubini's theorem [194, Vol. 1, Theorem 77], the inner integral converges almost everywhere with respect to the Lebesgue measure. Therefore, we can apply Schwarz [194, Vol. 1, Theorem 90] to conclude that the difference \( G(x) - G(y) \) is differentiable almost everywhere with respect to \( x \in U \) and we have:
\[
\frac{\partial}{\partial x} G(x) = \int_{\text{bd}(H(x))} \frac{\nabla_z g(x, z) \theta(z)}{\|\nabla_z g(x, z)\|} dS.
\]

We have used assumption (ii) to set \( |\nabla_z g(x, z)| = \nabla_z g(x, z) \). □

Obviously, the statement remains valid, if the assumption (ii) is replaced by the opposite strict inequality, so that the function \( g(x, z) \) would be strictly decreasing on \( U \times \Delta H_i^c \).

We note, that this result does not imply the differentiability of the function \( G \) at any fixed point \( x_0 \in U \). However, this type of differentiability is sufficient for many applications as it is elaborated in Ermoliev [65] and Uryasev [216].

The conditions of this theorem can be slightly modified, so that the result and the formula for the derivative are valid for piece-wise smooth function.

**Theorem 4.77 (Raik [166]).** Given a bounded open set \( U \subset \mathbb{R}^n \), we assume that:

- (i) the density function \( \theta(\cdot) \) is continuous and bounded on the set \( \Delta H_i \) for each \( i = 1, \ldots, m \);
- (ii) the vector functions \( \nabla_z g_i(x, z) \) and \( \nabla_x g_i(x, z) \) are continuous and bounded on the set \( U \times \Delta H_i \) for each \( i = 1, \ldots, m \);
- (iii) the function \( \|\nabla_x g_i(x, z)\| \geq \delta > 0 \) on the set \( U \times \Delta H_i \) for each \( i = 1, \ldots, m \);
- (iv) the following conditions are satisfied for all \( i = 1, \ldots, m \) and all \( x \in U \):
  \[
  P_{m-1}\{ S_i(x) \cap S_j(x) \} = 0, \quad i \neq j, \quad P_{m-1}\{ \text{bd}(\text{supp}P_Z \cap S_i(x)) \} = 0.
  \]

Then the probability function \( G(x) \) is differentiable on \( U \) and
\[
\nabla G(x) = \sum_{i=1}^m \int_{S_i(x)} \frac{\theta(z)}{\|\nabla_z g_i(x, z)\|} \nabla_z g_i(x, z) dS. \tag{4.76}
\]

The precise proof of this theorem is omitted. We refer to Kibzun and Tretyakov [104], and Kibzun and Uryasev [105] for more information on this topic.
4.4. Optimization problems with non-separable probabilistic constraints

For example, if \( g(x, Z) = x^T Z, m = 1 \), and \( Z \) has a nondegenerate multivariate normal distribution \( \mathcal{N}(\bar{z}, \Sigma) \), then \( g(x, Z) \sim \mathcal{N}(x^T \bar{z}, x^T \Sigma x) \), and hence the probability function \( G(x) = \Pr\{g(x, Z) \geq 0\} \) can be written in the form

\[
G(x) = \Phi\left( \frac{x^T \bar{z}}{\sqrt{x^T \Sigma x}} \right),
\]

where \( \Phi(\cdot) \) is the cdf of the standard normal distribution. In this case \( G(x) \) is continuously differentiable at every \( x \neq 0 \).

For problem (4.72), we impose the following constraint qualification at a point \( \hat{x} \in \mathcal{X} \). There exists a point \( x^r \in \mathcal{X} \) such that:

\[
\sum_{i=1}^{m} \int_{S_i(x)} \theta(z) \left\| \nabla_z g_i(\hat{x}, z) \right\| \nabla_x g_i(\hat{x}, z) dS < 0. \tag{4.77}
\]

This condition implies Robinson’s condition. We obtain the following necessary optimality conditions.

**Theorem 4.78.** Under the assumption of Theorem 4.77, let the constraint qualification (4.77) be satisfied, the function \( c(\cdot) \) be continuously differentiable, and let \( \hat{x} \in \mathcal{X} \) be an optimal solution of problem (4.72). Then there is a multiplier \( \lambda \geq 0 \) such that:

\[
0 \in \nabla c(\hat{x}) - \lambda \sum_{i=1}^{m} \int_{S_i(x)} \frac{\theta(z)}{\left\| \nabla_z g_i(\hat{x}, z) \right\| } \nabla_x g_i(\hat{x}, z) dS + \mathcal{N}_\mathcal{X}(\hat{x}), \tag{4.78}
\]

\[
\lambda [G(\hat{x}) - p] = 0. \tag{4.79}
\]

**Proof.** The statement follows from the necessary optimality conditions for smooth optimization problems and formula (4.76).

4.4.2 Approximations of non-separable probabilistic constraints

**Smoothing approximation via Steklov transformation**

In order to apply the optimality conditions formulated in Theorem 4.75 we need to calculate the subdifferential of the probability function \( \tilde{G} \) defined by the formula (4.73). The calculation involves the subdifferential of the probability function and the characteristic function of the event

\[
\{g_i(x, z) \geq 0, i = 1, \ldots, m\}.
\]

The latter function may be discontinuous. To alleviate this difficulty, we shall approximate the function \( G(x) \) by smooth functions.

Let \( k : \mathbb{R} \to \mathbb{R} \) be a nonnegative integrable symmetric function such that

\[
\int_{-\infty}^{+\infty} k(t) dt = 1.
\]
138 Chapter 4. Optimization Models with Probabilistic Constraints

It can be used as a density function of a random variable $K$, and, thus,

$$F_K(\tau) = \int_{-\infty}^{\tau} k(t) dt.$$  

Taking the characteristic function of the interval $[0, \infty)$, we consider the Steklov-Sobolev average functions for $\varepsilon > 0$:

$$F^\varepsilon_K(\tau) = \int_{-\infty}^{+\infty} 1_{[0, \infty)}(\tau + \varepsilon t) k(t) dt = \frac{1}{\varepsilon} \int_{-\infty}^{+\infty} 1_{[0, \infty)}(t) k\left(\frac{t - \tau}{\varepsilon}\right) dt.$$  

(4.80)

We see that by the definition of $F^\varepsilon_K$ and $1_{[0, \infty)}$ and by the symmetry of $k(\cdot)$ we have:

$$F^\varepsilon_K(\tau) = \int_{-\infty}^{+\infty} 1_{[0, \infty)}(\tau + \varepsilon t) k(t) dt = \int_{-\tau/\varepsilon}^{+\infty} k(t) dt$$

$$= \int_{-\infty}^{+\infty} k(-t) dt = \int_{-\tau/\varepsilon}^{+\infty} k(t) dt$$

(4.81)

$$= F_K \left(\frac{\tau}{\varepsilon}\right).$$

Setting

$$g_M(x, z) = \min_{1 \leq i \leq m} g_i(x, z),$$

we note that $g_M$ is quasi-concave, provided that all $g_i$ are quasi-concave functions. If the functions $g_i(\cdot, z)$ are continuous, then $g_M(\cdot, z)$ is continuous as well.

Using (4.81), we can approximate the constraint function $G(\cdot)$ by the function:

$$G_\varepsilon(x) = \int_{\mathbb{R}^\varepsilon} F^\varepsilon_K(g_M(x, z) - c) dP_z$$

$$= \int_{\mathbb{R}^\varepsilon} F_K \left(\frac{g_M(x, z) - c}{\varepsilon}\right) dP_z$$

(4.82)

$$= \frac{1}{\varepsilon} \int_{\mathbb{R}^\varepsilon} \int_{-\infty}^{-c} k\left(\frac{t + g_M(x, z)}{\varepsilon}\right) dt dP_z.$$

Now, we show that the functions $G_\varepsilon(\cdot)$ uniformly converge to $G(\cdot)$ when $\varepsilon$ converges to zero.

**Theorem 4.79.** Assume that $Z$ has a continuous distribution, the functions $g_i(\cdot, z)$ are continuous for almost all $z \in \mathbb{R}^\varepsilon$ and that, for certain constant $c \in \mathbb{R}$, we have

$$\Pr\{z \in \mathbb{R}^\varepsilon : g_M(x, z) = c\} = 0.$$

Then for any compact set $C \subseteq X$ the functions $G_\varepsilon$ uniformly converge on $C$ to $G$ when $\varepsilon \to 0$, i.e.,

$$\lim_{\varepsilon \to 0} \max_{x \in C} |G_\varepsilon(x) - G(x)| = 0.$$
4.4. Optimization problems with non-separable probabilistic constraints

**Proof.** Defining \( \delta(\epsilon) = \epsilon^{1-\beta} \) with \( \beta \in (0, 1) \), we have

\[
\lim_{\epsilon \to 0} \delta(\epsilon) = 0 \quad \text{and} \quad \lim_{\epsilon \to 0} F_K \left( \frac{\delta(\epsilon)}{\epsilon} \right) = 1, \lim_{\epsilon \to 0} F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right) = 0. \tag{4.83}
\]

Define for any \( \delta > 0 \) the sets:

\[
A(x, \delta) = \{ z \in \mathbb{R}^s : g_M(x, z) - c \leq -\delta \},
\]

\[
B(x, \delta) = \{ z \in \mathbb{R}^s : g_M(x, z) - c \geq \delta \},
\]

\[
C(x, \delta) = \{ z \in \mathbb{R}^s : |g_M(x, z) - c| \leq \delta \}.
\]

On the set \( A(x, \delta(\epsilon)) \) we have \( 1_{[0, \infty)}(g_M(x, z) - c) = 0 \) and, using (4.81), we obtain

\[
F_K^*(g_M(x, z) - c) = F_K \left( \frac{g_M(x, z) - c}{\epsilon} \right) \leq F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right).
\]

On the set \( B(x, \delta(\epsilon)) \) we have \( 1_{[0, \infty)}(g_M(x, z) - c) = 1 \) and

\[
F_K^*(g_M(x, z) - c) = F_K \left( \frac{g_M(x, z) - c}{\epsilon} \right) \geq F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right).
\]

On the set \( C(\delta(\epsilon)) \) we use the fact that \( 0 \leq 1_{[0, \infty)}(t) \leq 1 \) and \( 0 \leq F_K(t) \leq 1 \). We obtain the following estimate:

\[
|G(x) - G_\epsilon(x)| \leq \int_{\mathbb{R}^s} \left| 1_{[0, \infty)}(g_M(x, z) - c) - F_K^*(g_M(x, z) - c) \right| dP_Z
\]

\[
\leq F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right) \int_{A(x, \delta(\epsilon))} dP_Z + \left( 1 - F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right) \right) \int_{B(x, \delta(\epsilon))} dP_Z + 2 \int_{C(x, \delta(\epsilon))} dP_Z
\]

\[
\leq F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right) + \left( 1 - F_K \left( \frac{-\delta(\epsilon)}{\epsilon} \right) \right) + 2P_Z(C(x, \delta(\epsilon))).
\]

The first two terms on the right hand side of the inequality converge to zero when \( \epsilon \to 0 \) by the virtue of (4.83). It remains to show that \( \lim_{\epsilon \to 0} P_Z \{ C(x, \delta(\epsilon)) \} = 0 \) uniformly with respect to \( x \in C \). The function \( (x, z, \delta) \mapsto |g_M(x, z) - c| - \delta \) is continuous in \( (x, \delta) \) and measurable in \( z \). Thus, it is uniformly continuous with respect to \( (x, \delta) \) on any compact set \( C \times [-\delta_0, \delta_0] \) with \( \delta_0 > 0 \). The probability measure \( P_Z \) is continuous, and, therefore, the function

\[
\Theta(x, \delta) = P \{ |g_M(x, z) - c| - \delta \leq 0 \} = P \{ \cap_{\beta > \delta} C(x, \beta) \}
\]

is uniformly continuous with respect to \( (x, \delta) \) on \( C \times [-\delta_0, \delta_0] \). By the assumptions of the theorem

\[
\Theta(x, 0) = P_Z \{ z \in \mathbb{R}^s : |g_M(x, z) - c| = 0 \} = 0,
\]

and, thus,

\[
\lim_{\epsilon \to 0} P_Z \{ z \in \mathbb{R}^s : |g_M(x, z) - c| \leq \delta(\epsilon) \} = \lim_{\delta \to 0} \Theta(x, \delta) = 0.
\]
As $\Theta(\cdot, \delta)$ is continuous, the convergence is uniform on compact sets with respect to the first argument.

Now, we derive a formula for the Clarke generalized gradients of the approximation $G_\varepsilon$. We define the index set:

$$I(x, z) = \{ i : g_i(x, z) = g_M(x, z), 1 \leq i \leq m \}.$$ 

**Theorem 4.80.** Assume that the density function $k(\cdot)$ is nonnegative, bounded, and continuous. Furthermore, let the functions $g_i(\cdot, z)$ be concave for every $z \in \mathbb{R}^s$, and their subgradients be uniformly bounded as follows:

$$\sup\{ s \in \partial g_i(y, z), \|y - x\| \leq \delta \} \leq l_s(x, z), \delta > 0, \text{ for all } i = 1, \ldots, m,$$

where $l_s(x, z)$ is an integrable function of $z$ for all $x \in X$. Then $G_\varepsilon(\cdot)$ is Lipschitz continuous, Clarke regular and its Clarke generalized gradient set is given by

$$\partial^o G_\varepsilon(x) = \frac{1}{\varepsilon} \int_{\mathbb{R}^s} k \left( \frac{g_M(x, z) - c}{\varepsilon} \right) \text{conv} \{ \partial g_i(x, z) : i \in I(x, z) \} \, dP_Z.$$

**Proof.** Under the assumptions of the theorem, the function $F_K(\cdot)$ is monotone and continuously differentiable. The function $g_M(\cdot, z)$ is concave for every $z \in \mathbb{R}^s$ and its subdifferential are given by the formula:

$$\partial g_M(y, z) = \text{conv} \{ s_i \in \partial g_i(y, z) : g_i(y, z) = g_M(y, z) \}.$$ 

Thus the subgradients of $g_M$ are uniformly bounded:

$$\sup\{ s \in \partial g_M(y, z), \|y - x\| \leq \delta \} \leq l_s(x, z), \delta > 0.$$ 

Therefore, the composite function $F_K \left( \frac{g_M(x, z) - c}{\varepsilon} \right)$ is subdifferentiable and its subdifferential can be calculated as:

$$\partial^o F_K \left( \frac{g_M(x, z) - c}{\varepsilon} \right) = \frac{1}{\varepsilon} k \left( \frac{g_M(x, z) - c}{\varepsilon} \right) \cdot \partial g_M(x, z).$$

The mathematical expectation function

$$G_\varepsilon(x) = \int_{\mathbb{R}^s} F_K^\varepsilon(g_M(x, z) - c) \, dP_Z = \int_{\mathbb{R}^s} F_K \left( \frac{g_M(x, z) - c}{\varepsilon} \right) \, dP_Z,$$

is regular by Clarke [38, Theorem 2.7.2] and its Clarke generalized gradient set has the form:

$$\partial^o G_\varepsilon(x) = \int_{\mathbb{R}^s} \partial^o F_K \left( \frac{g_M(x, z) - c}{\varepsilon} \right) \, dP_Z = \frac{1}{\varepsilon} \int_{\mathbb{R}^s} k \left( \frac{g_M(x, z) - c}{\varepsilon} \right) \cdot \partial g_M(x, z) \, dP_Z.$$
4.4. Optimization problems with non-separable probabilistic constraints

Using the formula for the subdifferential of \( g_M(x, z) \), we obtain the statement. □

Now we show that if we choose \( K \) to have an \( \alpha \)-concave distribution, and all assumptions of Theorem 4.39 are satisfied, the generalized concavity property of the approximated probability function is preserved.

**Theorem 4.81.** If the density function \( k \) is \( \alpha \)-concave (\( \alpha \geq 0 \)), \( Z \) has \( \gamma \)-concave distribution (\( \gamma \geq 0 \)), the functions \( g_i(\cdot, z) \), \( i = 1, \ldots, m \) are quasi-concave, then the approximate probability function \( G_\varepsilon \) has a \( \beta \)-concave distribution, where

\[
\beta = \begin{cases} 
(\gamma^{-1} + (1 + s\alpha)/\alpha)^{-1}, & \text{if } \alpha + \gamma > 0, \\
0, & \text{if } \alpha + \gamma = 0.
\end{cases}
\]

**Proof.** If the density function \( k \) is \( \alpha \)-concave (\( \alpha \geq 0 \)), then \( K \) has a \( \gamma \)-concave distribution with \( \gamma = \alpha/(1 + s\alpha) \). If \( Z \) has \( \gamma' \)-concave distribution (\( \gamma \geq 0 \)), then the random vector \( (Z, K)^T \) has a \( \beta \)-concave distribution according to Theorem 4.36, where

\[
\beta = \begin{cases} 
(\gamma^{-1} + \gamma'^{-1})^{-1}, & \text{if } \gamma + \gamma' > 0, \\
0, & \text{if } \gamma + \gamma' = 0.
\end{cases}
\]

Using the definition \( G_\varepsilon(x) \) of (4.82), we can write:

\[
G_\varepsilon(x) = \int_{\mathbb{R}^s} F^*_K \left( \frac{g_M(x, z) - c}{\varepsilon} \right) dP_Z = \int_{\mathbb{R}^s} \int_{-\infty}^{(g_M(x, z) - c)/\varepsilon} k(t) dt \, dP_Z 
\]

\[
= \int_{\mathbb{R}^s} \int_{-\infty}^{\infty} 1_{(g_M(x, z) - c)/\varepsilon > t} dP_K \, dP_Z = \int_{\mathbb{R}^s} \int_{H_\varepsilon(x)} dP_K \, dP_Z, \quad (4.84)
\]

where

\[
H_\varepsilon(x) = \{(z, t) \in \mathbb{R}^{s+1} : g_M(x, z) - \varepsilon t \geq c\}.
\]

Since \( g_M(\cdot, z) \) is quasi-concave, the set \( H_\varepsilon(x) \) is convex. Representation (4.84) of \( G_\varepsilon \) and the \( \beta \)-concavity of \((Z, K)\) imply the assumptions of Theorem 4.39, and, thus, the function \( G_\varepsilon \) is \( \beta \)-concave. □

This theorem shows that if the random vector \( Z \) has a generalized concave distribution, we can choose a suitable generalized concave density function \( k(\cdot) \) for smoothing and obtain an approximate convex optimization problem.

**Theorem 4.82.** In addition to the assumptions of Theorems 4.79, 4.80 and 4.81. Then on the set \( \{x \in \mathbb{R}^n : G(x) > 0\} \), the function \( G_\varepsilon \) is Clarke-regular and the set of Clarke generalized gradients \( \partial^G G_\varepsilon(x^\varepsilon) \) converge to the set of Clarke generalized gradients of \( G \), \( \partial^G G(x) \), in the following sense: if for any sequences \( \varepsilon \downarrow 0 \), \( x^\varepsilon \rightarrow x \) and \( s^\varepsilon \in \partial^G G_\varepsilon(x^\varepsilon) \) such that \( s^\varepsilon \rightarrow s \), then \( s \in \partial^G G(x) \).

**Proof.** Consider a point \( x \) such that \( G(x) > 0 \) and points \( x^\varepsilon \rightarrow x \) as \( \varepsilon \downarrow 0 \). All points \( x^\varepsilon \) can be included in some compact set containing \( x \) in its interior. The function \( G \) is generalized concave by virtue of Theorem 4.39. It is locally Lipschitz continuous, directionally
differentiable, and Clarke-regular due to Theorem 4.29. It follows that $G(y) > 0$ for all point $y$ in some neighborhood of $x$. By virtue of Theorem 4.79, this neighborhood can be chosen small enough, so that $G_{\varepsilon}(y) > 0$ for all $\varepsilon$ small enough, as well. The functions $G_{\varepsilon}$ are generalized concave by virtue of Theorem 4.81. It follows that $G_{\varepsilon}$ are locally Lipschitz continuous, directionally differentiable, and Clarke-regular due to Theorem 4.29. Using the uniform convergence of $G_{\varepsilon}$ on compact sets and the definition of Clarke generalized gradient, we can pass to the limit with $\varepsilon \downarrow 0$ in the inequality

$$
\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \left[ G_{\varepsilon}(y + \varepsilon d) - G_{\varepsilon}(y) \right] \geq d^T s \quad \text{for any } d \in \mathbb{R}^n.
$$

Consequently, $s \in \partial^c G(x)$. □

Using the approximate probability function we can solve the following approximation of problem (4.72):

$$
\begin{align*}
\text{Min} & \quad c(x) \\
\text{s.t.} & \quad G_{\varepsilon}(x) \geq p, \quad x \in X.
\end{align*}
$$

(4.85)

Under the conditions of Theorem 4.82 the function $G_{\varepsilon}$ is $\beta$-concave for some $\beta \geq 0$. We can specify the necessary and sufficient optimality conditions for the approximate problem.

**Theorem 4.83.** *In addition to the assumptions of Theorem 4.82, assume that $c(\cdot)$ is a convex function, the Slater condition for problem (4.85) is satisfied, and $\text{int} G_{\varepsilon} \neq \emptyset$. A point $\hat{x} \in X$ is an optimal solution of problem (4.85) if and only if a nonpositive number $\lambda$ exists such that

$$
0 \in \partial c(\hat{x}) + s \lambda \int_{\mathbb{R}^n} k\left( \frac{g_M(\hat{x}, z) - c}{\varepsilon} \right) \text{conv} \left\{ \partial g_i(\hat{x}, z) : i \in I(\hat{x}, z) \right\} dP_Z + N_X(\hat{x}),
$$

$$
\lambda [G_{\varepsilon}(\hat{x}) - p] = 0.
$$

Here

$$
s = \begin{cases} 
\alpha \varepsilon^{-1} [G_{\varepsilon}(\hat{x})]^{\alpha-1}, & \text{if } \beta \neq 0, \\
[\varepsilon G_{\varepsilon}(\hat{x})]^{-1}, & \text{if } \beta = 0.
\end{cases}
$$

**Proof.** We shall show the statement for $\beta = 0$. The proof for the other case is analogous. Setting $\tilde{G}_{\varepsilon}(x) = \ln G_{\varepsilon}(x)$, we obtain a concave function $\tilde{G}_{\varepsilon}$, and formulate the problem:

$$
\begin{align*}
\text{Min} & \quad c(x) \\
\text{s.t.} & \quad \ln p - \tilde{G}_{\varepsilon}(x) \leq 0, \quad x \in X.
\end{align*}
$$

(4.86)

Clearly, $\hat{x}$ is a solution of the problem (4.86) if and only if it is a solution of problem (4.85). Problem (4.86) is a convex problem and Slater’s condition is satisfied for it as well.
4.4. Optimization problems with non-separable probabilistic constraints

Therefore, we can write the following optimality conditions for it. The point \( \hat{x} \in \mathcal{X} \) is a solution if and only if a number \( \lambda_0 > 0 \) exists such that

\[
0 \in \partial c(x) + \lambda_0 \partial \left[ - \hat{g}_\varepsilon(\hat{x}) \right] + \mathcal{N}_{\mathcal{X}}(\hat{x}),
\]

\[
\lambda_0 [G_\varepsilon(\hat{x}) - p] = 0. \tag{4.87}
\]

We use the formula for the Clarke generalized gradients of generalized concave functions to obtain

\[
\partial^p \hat{g}_\varepsilon(\hat{x}) = \frac{1}{G_\varepsilon(\hat{x})} \partial^p G_\varepsilon(\hat{x}).
\]

Moreover, we have a representation of the Clarke generalized gradient set of \( G_\varepsilon \), which yields

\[
\partial^p \hat{g}_\varepsilon(\hat{x}) = \frac{1}{\varepsilon G_\varepsilon(\hat{x})} \int_{\mathbb{R}^s} k \left( \frac{g_M(\hat{x}, z) - c}{\varepsilon} \right) \cdot \partial g_M(\hat{x}, z) dP_z.
\]

Substituting the last expression into (4.87), we obtain the result. \( \square \)

Normal approximation

In this section we analyze approximation for problems with individual probabilistic constraints, defined by linear inequalities. In this setting it is sufficient to consider a problem with a single probabilistic constraint of form:

\[
\begin{align*}
\text{Max} & \quad c(x) \\
\text{s.t.} & \quad \Pr \{ x^T \mathbf{Z} \geq \eta \} \geq p, \\
& \quad x \in \mathcal{X}.
\end{align*} \tag{4.89}
\]

Before developing the normal approximation for this problem, let us illustrate its potential on an example. We return to our Example 4.2, in which we have formulated a portfolio optimization problem under a Value-at-Risk constraint.

\[
\begin{align*}
\text{Max} & \quad \sum_{i=1}^{n} \mathbb{E}[R_i] x_i \\
\text{s.t.} & \quad \Pr \{ \sum_{i=1}^{n} R_i x_i \geq -\eta \} \geq p, \\
& \quad \sum_{i=1}^{n} x_i \leq 1, \\
& \quad x \geq 0.
\end{align*} \tag{4.90}
\]

We denote the net increase of the value of our investment after a period of time by

\[
G(x, R) = \sum_{i=1}^{n} \mathbb{E}[R_i] x_i.
\]
Chapter 4. Optimization Models with Probabilistic Constraints

Let us assume that the random return rates \( R_1, \ldots, R_n \) have a joint normal probability distribution. Recall that the normal distribution is log-concave and the probabilistic constraint in problem (4.90) determines a convex feasible set, according to Theorem 4.39.

Another direct way to see that the last transformation of the probabilistic constraint results in a convex constraint is the following. We denote \( \bar{r}_i = \mathbb{E}[R_i] \), \( \bar{r} = (\bar{r}_1, \ldots, \bar{r}_n)^T \), and assume that \( \bar{r} \) is not the zero-vector. Further, let \( \Sigma \) be the covariance matrix of the joint distribution of the return rates. We observe that the total profit (or loss) \( G(x, R) \) is a normally distributed random variable with expected value \( \mathbb{E}[G(x, R)] = \bar{r}^T x \) and variance \( \text{Var}[G(x, R)] = x^T \Sigma x \). Assuming that \( \Sigma \) is positive definite, the probabilistic constraint

\[
\Pr\{G(x, R) \geq -\eta\} \geq p
\]

can be written in the form (see the discussion on page 16)

\[
z_p \sqrt{x^T \Sigma x} - \bar{r}^T x \leq \eta.
\]

Hence problem (4.90) can be written in the following form:

\[
\begin{align*}
\text{Max} & \ \bar{r}^T x \\
\text{s.t.} & \ z_p \sqrt{x^T \Sigma x} - \bar{r}^T x \leq \eta, \\
& \sum_{i=1}^n x_i \leq 1, \\
& x \geq 0.
\end{align*}
\]

(4.91)

Note that \( \sqrt{x^T \Sigma x} \) is a convex function, of \( x \), and \( z_p = \Phi^{-1}(p) \) is positive for \( p > 1/2 \), and hence (4.91) is a convex programming problem.

Now, we consider the general optimization problem (4.89). Assuming that the \( n \)-dimensional random vector \( Z \) has independent components and the dimension \( n \) is relatively large, we may invoke the central limit theorem. Under mild additional assumptions, we can conclude that the distribution of \( x^T Z \) is approximately normal and convert the probabilistic constraint into an algebraic constraint in a similar manner. Note, that this approach is appropriate if \( Z \) has a substantial number of components and the vector \( x \) has appropriately large number of non-zero components, so that the central limit theorem would be applicable to \( x^T Z \). Furthermore, we assume that the probability parameter \( p \) is not too close to one, such as 0.9999.

We recall several versions of the Central Limit Theorem (CLT). Let \( Z_i, i = 1, 2, \ldots \) be a sequence of independent random variables defined on the same probability space. We assume that each \( Z_i \) has finite expected value \( \mu_i = \mathbb{E}[Z_i] \) and finite variance \( \sigma_i^2 = \text{Var}[Z_i] \). Setting

\[
s_n^2 = \sum_{i=1}^n \sigma_i^2, \quad \text{and} \quad r_n^3 = \sum_{i=1}^n \mathbb{E}(|Z_i - \mu_i|^3),
\]

we assume that \( r_n^3 \) is finite for every \( n \), and that

\[
\lim_{n \to \infty} \frac{r_n}{s_n} = 0.
\]

(4.92)
4.5. Semi-infinite probabilistic problems

Then, the distribution of the random variable
\[ \frac{\sum_{i=1}^{n} (Z_i - \mu_i)}{s_n} \]
converges towards the standard normal distribution as \( n \to \infty \).

The condition (4.92) is called Lyapunov’s condition. In the same setting, we can replace the Lyapunov’s condition with the following weaker condition, proposed by Lindeberg. For every \( \varepsilon > 0 \) we define
\[ Y_i = \begin{cases} \frac{(X_i - \mu_i)^2}{s_n^2}, & \text{if } |X_i - \mu_i| > \varepsilon s_n, \\ 0, & \text{otherwise}. \end{cases} \]
The Lindeberg’s condition reads:
\[ \lim_{n \to \infty} n \sum_{i=1}^{n} E(Y_i) = 0. \]

Let us denote \( \bar{\mu} = (\mu_1, \ldots, \mu_n)^T \). Under the conditions of the Central Limit Theorem, the distribution of our random variable \( x^T Z \) is close to the normal distribution with expected value \( x^T \bar{\mu} \) and variance \( \sum_{i=1}^{n} \sigma_i^2 x_i^2 \) for problems of large dimensions. Our probabilistic constraint takes on the form:
\[ \frac{\bar{\mu}^T x - \eta}{\sqrt{\sum_{i=1}^{n} \sigma_i^2 x_i^2}} \geq z_p. \]

Define \( X = \{ x \in \mathbb{R}^n_+ : \sum_{i=1}^{n} x_i \leq 1 \} \). Denoting the matrix with diagonal elements \( \sigma_1, \ldots, \sigma_n \) by \( D \), problem (4.89) can be replaced by the following approximate problem:
\[ \begin{align*}
\text{Min } & c(x) \\
\text{s.t. } & z_p\|Dx\| \leq \frac{\bar{\mu}^T x - \eta}{\sqrt{\sum_{i=1}^{n} \sigma_i^2 x_i^2}}, \\
& x \in X.
\end{align*} \]
The probabilistic constraint in this problem is approximated by an algebraic convex constraint. Due to the independence of the components of the random vector \( Z \), the matrix \( D \) has a simple diagonal form. There are versions of the central limit theorem which treat the case of sums of dependent variables, for instance the \( n \)-dependent CLT, the martingale CLT and the CLT for mixing processes. These statements will not be presented here. One can follow the same line of argument to formulate a normal approximation of the probabilistic constraint, which is very accurate for problems with large decision space.

4.5 Semi-infinite probabilistic problems

In this section, we concentrate on the semi-infinite probabilistic problem (4.9). We recall its formulation:
\[ \begin{align*}
\text{Min } & c(x) \\
\text{s.t. } & \text{Pr}\{ g(x, Z) \geq \eta \} \geq \text{Pr}\{ Y \geq \eta \}, \quad \eta \in [a, b], \\
& x \in X.
\end{align*} \]
Our goal is to derive necessary and sufficient optimality conditions for this problem. Denote the space of regular countably additive measures on \([a, b]\) having finite variation by \(\mathcal{M}([a, b])\), and its subset of nonnegative measures by \(\mathcal{M}_+([a, b])\).

We define the constraint function \(G(x, \eta) = P\{z : g(x, z) \geq \eta\}\). As we shall develop optimality conditions in differential form, we impose additional assumptions on problem (4.9):

(i) The function \(c\) is continuously differentiable on \(X\);
(ii) The constraint function \(G(\cdot, \cdot)\) is continuous with respect to the second argument and continuously differentiable with respect to the first argument;
(iii) The reference random variable \(Y\) has a continuous distribution.

The differentiability assumption on \(G\) may be enforced taking into account the results in section 4.4.1. For example, if the vector \(Z\) has a probability density \(\theta(z)\), the function \(g(\cdot, \cdot)\) is continuously differentiable with nonzero gradient \(\nabla_z g(x, z)\) and such that the quantity \(\frac{\theta(z)}{\|\nabla_z g(x, z)\|}\) is uniformly bounded (in a neighborhood of \(x\)) by an integrable function, then the function \(G\) is differentiable. Moreover, we can express its gradient with respect to \(x\) as follows:

\[
\nabla_x G(x, \eta) = \int_{\partial H(x, \eta)} \frac{\theta(z)}{\|\nabla_z g(x, z)\|} \nabla_z g(x, z) \, dP_{m-1},
\]

where \(\partial H(x, \eta)\) is the surface of the set \(H(x, \eta) = \{z : g(x, z) \geq \eta\}\) and \(P_{m-1}\) refers to Lebesgue measure on the \(m-1\)-dimensional surface.

We define the set \(U([a, b])\) of functions \(u(\cdot)\) satisfying the following conditions:

\(u(\cdot)\) is nondecreasing and right continuous;
\(u(t) = 0\) for all \(t \leq a\);
\(u(t) = u(b)\), for all \(t \geq b\).

It is evident that \(U([a, b])\) is a convex cone.

First we derive a useful formula.

**Lemma 4.84.** For any real random variable \(Z\) and any measure \(\mu \in \mathcal{M}([a, b])\) we have

\[
\int_a^b \Pr\{Z \geq \eta\} \, d\mu(\eta) = \mathbb{E}[u(Z)],
\]

where \(u(z) = \mu([a, z])\).

**Proof.** We extend the measure \(\mu\) to the entire real line, by assigning measure 0 to sets not intersecting \([a, b]\). Using the probability measure \(P_Z\) induced by \(Z\) on \(\mathbb{R}\) and applying Fubini's theorem we obtain

\[
\int_a^b \Pr\{Z \geq \eta\} \, d\mu(\eta) = \int_a^\infty \Pr\{Z \geq \eta\} \, d\mu(\eta) = \int_a^\infty \int_\eta^\infty dP_Z(z) \, d\mu(\eta)
\]

\[
= \int_a^\infty \int_a^z d\mu(\eta) \, dP_Z(z) = \int_a^\infty \mu([a, z]) \, dP_Z(z) = \mathbb{E}[\mu([a, Z])].
\]
4.5. Semi-infinite probabilistic problems

We define \( u(z) = \mu([a, z]) \) and obtain the stated result. \( \square \)

Let us observe that if the measure \( \mu \) in the above lemma is nonnegative, then \( u \in U([a, b]) \). Indeed, \( u(\cdot) \) is nondecreasing since for \( z_1 > z_2 \) we have

\[
u(z_1) = \mu([a, z_1]) = \mu([a, z_2]) + \mu([z_1, z_2]) \geq \mu([a, z_2]) = u(z_2).
\]

Furthermore, \( u(z) = \mu([a, z]) = u(b) \) for \( z \geq b \).

We introduce the functional \( L : \mathbb{R}^n \times U \rightarrow \mathbb{R} \) associated with problem (4.9):

\[
L(x, u) = c(x) + \mathbb{E} \left[ u(g(x, Z)) - u(Y) \right].
\]

We shall see that the functional \( L \) plays the role of a Lagrangian of the problem.

We also set \( v(\eta) = \Pr\{Y \geq \eta\} \).

**Definition 4.85.** Problem (4.9) satisfies the differential uniform dominance condition at the point \( \hat{x} \in \mathcal{X} \) if there exists \( x^0 \in \mathcal{X} \) such that

\[
\min_{a \leq \eta \leq b} \left[ G(\hat{x}, \eta) + \nabla_x G(\hat{x}, \eta)(x^0 - \hat{x}) - v(\eta) \right] > 0.
\]

**Theorem 4.86.** Assume that \( \hat{x} \) is an optimal solution of problem (4.9) and that the differential uniform dominance condition is satisfied at the point \( \hat{x} \). Then there exists a function \( \hat{u} \in U \), such that

\[
-\nabla_x L(\hat{x}, \hat{u}) \in \mathcal{N}_{\mathcal{X}}(\hat{x}),
\]

\[
\mathbb{E}\left[ \hat{u}(g(\hat{x}, Z)) \right] = \mathbb{E}\left[ \hat{u}(Y) \right].
\]

**Proof.** We consider the mapping \( \Gamma : \mathcal{X} \rightarrow C([a, b]) \) defined as follows:

\[
\Gamma(x)(\eta) = \Pr\{g(x, Z) \geq \eta\} - v(\eta), \quad \eta \in [a, b].
\]

We define \( K \) as the cone of nonnegative functions in \( C([a, b]) \). Problem (4.9) can be formulated as follows:

\[
\begin{align*}
\min_x & \quad c(x) \\
\text{s.t.} & \quad \Gamma(x) \in K, \\
& \quad x \in \mathcal{X}.
\end{align*}
\]

At first we observe that the functions \( c(\cdot) \) and \( \Gamma(\cdot) \) are continuously differentiable by the assumptions made at the beginning of this section. Secondly, the differential uniform dominance condition is equivalent to Robinson’s constraint qualification condition:

\[
0 \in \text{int} \left\{ \Gamma(\hat{x}) + \nabla_x \Gamma(\hat{x})(\mathcal{X} - \hat{x}) - K \right\}.
\]

Indeed, it is easy to see that the uniform dominance condition implies Robinson’s condition. On the other hand, if Robinson’s condition holds true, then there exists \( \varepsilon > 0 \) such that the
function identically equal to \( \varepsilon \) is an element of the set at the right hand side of (4.99). Then we can find \( x^0 \) such that

\[
\Gamma(\hat{x})(\eta) + \left[ \nabla_x \Gamma(\hat{x})(\eta) \right] (x^0 - \hat{x}) \geq \varepsilon \quad \text{for all } \eta \in [a, b].
\]

Consequently, the uniform dominance condition is satisfied.

By the Riesz representation theorem, the space dual to \( C([a, b]) \) is the space \( M([a, b]) \) of regular countably additive measures on \([a, b]\) having finite variation. The Lagrangian \( \Lambda : \mathbb{R}^n \times M([a, b]) \to \mathbb{R} \), for problem (4.98) is defined as follows:

\[
\Lambda(x, \mu) = c(x) + \int_a^b \Gamma(x)(\eta) \, d\mu(\eta). \tag{4.100}
\]

The necessary optimality conditions for problem (4.98) have the form: there exists a measure \( \hat{\mu} \in M_+([a, b]) \), such that

\[
-\nabla_x \Lambda(\hat{x}, \hat{\mu}) \in \mathcal{N}_X(\hat{x}), \tag{4.101}
\]

\[
\int_a^b \Gamma(\hat{x})(\eta) \, d\hat{\mu}(\eta) = 0. \tag{4.102}
\]

Using Lemma 4.84, we obtain the equation for all \( x \):

\[
\int_a^b \Gamma(x)(\eta) \, d\hat{\mu}(\eta) = \int_a^b \left( \Pr\{g(x, z) \geq \eta \} - \Pr\{Y \geq \eta \} \right) \, d\hat{\mu}(\eta)
= E[\hat{u}(g(x, Z))] - E[\hat{u}(Y)],
\]

where \( \hat{u}(\eta) = \hat{\mu}([a, \eta]) \). Since \( \hat{\mu} \) is nonnegative, the corresponding utility function \( \hat{u} \) is an element of \( \mathcal{U}([a, b]) \). The correspondence between nonnegative measures \( \mu \in M([a, b]) \) and utility functions \( u \in \mathcal{U} \) and the last equation imply that (4.102) is equivalent to (4.96). Moreover, \( \Lambda(x, \mu) = L(x, u) \)

and, therefore, (4.101) is equivalent to (4.95). \( \square \)

We note that the functions \( u \in \mathcal{U}([a, b]) \) can be interpreted as von Neumann-Morgenstern utility functions of rational decision makers. The theorem demonstrates that one can view the maximization of expected utility as a dual model to the model with stochastic dominance constraints. Utility functions of decision makers are very difficult to elicit. This task becomes even more complicated when there is a group of decision-makers who have to come to a consensus. The model (4.9) avoids these difficulties, by requiring that a benchmark random outcome, considered reasonable, be specified. Our analysis, departing from the benchmark outcome, generates the utility function of the decision maker. It is implicitly defined by the benchmark used, and by the problem under consideration.

We will demonstrate that it is sufficient to consider only the subset of \( \mathcal{U}([a, b]) \) containing piecewise constant utility functions.

**Theorem 4.87.** Under the assumptions of Theorem 4.86 there exist piecewise constant utility function \( w(\cdot) \in \mathcal{U} \) satisfying the necessary optimality conditions (4.95)–(4.96). Moreover, the function \( w(\cdot) \) has at most \( n + 2 \) jump points: there exist numbers \( \eta_i \in [a, b], \quad i = 0, \ldots, n + 1 \)
4.5. Semi-infinite probabilistic problems

1, . . . , k, such that the function \( w(\cdot) \) is constant on the intervals \( (-\infty, \eta_1], (\eta_1, \eta_2], \ldots, (\eta_k, \infty) \), and \( 0 \leq k \leq n + 2 \).

**Proof.** Consider the mapping \( \Gamma \) defined by (4.97). As already noted in the proof of the previous theorem, it is continuously differentiable due to the assumptions about the probability function. Therefore, the derivative of the Lagrangian has the form

\[
\nabla_x \Lambda(\hat{x}, \hat{\mu}) = \nabla_x c(\hat{x}) + \int_a^b \nabla_x \Gamma(\hat{x})(\eta) \, d\hat{\mu}(\eta).
\]

The necessary condition of optimality (4.101) can be rewritten as follows

\[
-\nabla_x c(\hat{x}) - \int_a^b \nabla_x \Gamma(\hat{x})(\eta) \, d\hat{\mu}(\eta) \in N_{\mathcal{X}}(\hat{x}).
\]

Considering the vector

\[
g = \nabla_x c(\hat{x}) - \nabla_x \Lambda(\hat{x}, \hat{\mu}),
\]

we observe that the optimal values of multipliers \( \hat{\mu} \) have to satisfy the equation

\[
\int_a^b \nabla_x \Gamma(\hat{x})(\eta) \, d\mu(\eta) = g. \tag{4.103}
\]

At the optimal solution \( \hat{x} \) we have \( \Gamma(\hat{x})(\cdot) \leq 0 \) and \( \hat{\mu} \geq 0 \). Therefore, the complementarity condition (4.102) can be equivalently expressed as the equation

\[
\int_a^b \Gamma(\hat{x})(\eta) \, d\mu(\eta) = 0. \tag{4.104}
\]

Every nonnegative solution \( \mu \) of equations (4.103)–(4.104) can be used as the Lagrange multiplier satisfying conditions (4.101)–(4.102) at \( \hat{x} \). Define

\[
a = \int_a^b d\hat{\mu}(\eta).
\]

We can add to equations (4.103)–(4.104) the condition

\[
\int_a^b d\mu(\eta) = a. \tag{4.105}
\]

The system of three equations (4.103)–(4.105) still has at least one nonnegative solution, namely \( \hat{\mu} \). If \( \hat{\mu} \equiv 0 \), then the dominance constraint is not active. In this case, we can set \( w(\eta) \equiv 0 \) and the statement of the theorem follows from the fact that conditions (4.103)–(4.104) are equivalent to (4.101)–(4.102).

Now, consider the case of \( \hat{\mu} \neq 0 \). In this case, we have \( a > 0 \). Normalizing by \( a \), we notice that equations (4.103)–(4.105) are equivalent to the following inclusion:

\[
\left[ \frac{g}{a} \ 0 \right] \in \text{conv} \left\{ \begin{bmatrix} \nabla_x \Gamma(\hat{x})(\eta) \\ \Gamma(\hat{x})(\eta) \end{bmatrix} : \eta \in [a, b] \right\} \subset \mathbb{R}^{n+1}.
\]
By Carathéodory’s Theorem, there exist numbers \( \eta_i \in [a, b] \), and \( \alpha_i \geq 0 \), \( i = 1, \ldots, k \) such that
\[
\begin{bmatrix}
g/a \\
0
\end{bmatrix} = \sum_{i=1}^{k} \alpha_i \begin{bmatrix}
\nabla_x \Gamma(x)(\eta_i) \\
\Gamma(x)(\eta_i)
\end{bmatrix},
\]
and
\[
\sum_{i=1}^{k} \alpha_i = 1,
\]
and
\[
1 \leq k \leq n + 2.
\]
We define atomic measure \( \nu \) having atoms of mass \( c\alpha_i \) at points \( \eta_i \), \( i = 1, \ldots, k \). It satisfies equations (4.103)–(4.104):
\[
\int_{a}^{b} \nabla_x \Gamma(x)(\eta) \, d\nu(\eta) = \sum_{i=1}^{k} \nabla_x \Gamma(x)(\eta_i) c\alpha_i = g,
\]
\[
\int_{a}^{b} \Gamma(x)(\eta) \, d\nu(\eta) = \sum_{i=1}^{k} \Gamma(x)(\eta_i) c\alpha_i = 0.
\]
Recall that equations (4.103)–(4.104) are equivalent to (4.101)–(4.102). Now, applying Lemma 4.84, we obtain the utility functions
\[
w(\eta) = \nu[a, \eta], \quad \eta \in \mathbb{R}.
\]
It is straightforward to check that \( w \in U([a, b]) \) and the assertion of the theorem holds true.

It follows from Theorem 4.87 that if the dominance constraint is active, then there exist at least one and at most \( n + 2 \) target values \( \eta_i \) and target probabilities \( v_i = \Pr\{Y \geq \eta_i\} \), \( i = 1, \ldots, k \) which are critical for problem (4.9). They define a relaxation of (4.9) involving finitely many probabilistic constraints:
\[
\begin{aligned}
\text{Min} & \quad c(x) \\
\text{s.t.} & \quad \Pr\{g(x, Z) \geq \eta_i\} \geq v_i, \quad i = 1, \ldots, k, \\
& \quad x \in X.
\end{aligned}
\]
The necessary conditions of optimality for this relaxation yield a solution of the optimality conditions of the original problem (4.9). Unfortunately, the target values and the target probabilities are not known in advance.

A particular situation, in which the target values and the target probabilities can be specified in advance, occurs when \( Y \) has a discrete distribution with finite support. Denote the realizations of \( Y \) by
\[
\eta_1 < \eta_2 < \cdots < \eta_k,
\]
4.5. Semi-infinite probabilistic problems

and the corresponding probabilities by \( p_i, \ i = 1, \ldots, k \). Then the dominance constraint is equivalent to

\[
\Pr\{g(x, Z) \geq \eta_i\} \geq \sum_{j=i}^{k} p_j, \quad i = 1, \ldots, k.
\]

Here, we use the fact that the probability distribution function of \( g(x, Z) \) is continuous and nondecreasing.

Now, we shall derive sufficient conditions of optimality for problem (4.9). We assume additionally that the function \( g \) is jointly quasi-concave in both arguments, and \( Z \) has an \( \alpha \)-concave probability distribution.

**Theorem 4.88.** Assume that a point \( \hat{x} \) is feasible for problem (4.9). Suppose that there exists a function \( \hat{u} \in U, \ \hat{u} \neq 0 \), such that conditions (4.95)–(4.96) are satisfied. If the functions \( c \) is convex, the function \( g \) satisfies the concavity assumptions above and the variable \( Z \) has an \( \alpha \)-concave probability distribution, then \( \hat{x} \) is an optimal solution of problem (4.9).

**Proof.** By virtue of Theorem 4.43, the feasible set of problem (4.98) is convex and closed.

Let the operator \( \Gamma \) and the cone \( K \) be defined as in the proof of Theorem 4.86. Using Lemma 4.84, we observe that optimality conditions (4.101)–(4.102) for problem (4.98) are satisfied. Consider a feasible direction \( d \) at the point \( \hat{x} \). As the feasible set is convex, we conclude that

\[
\Gamma(\hat{x} + \tau d) \in K,
\]

for all sufficiently small \( \tau > 0 \). Since \( \Gamma \) is differentiable, we have

\[
\frac{1}{\tau} \left[ \Gamma(\hat{x} + \tau d) - \Gamma(\hat{x}) \right] \rightarrow \nabla_x \Gamma(\hat{x})(d) \quad \text{whenever} \quad \tau \downarrow 0.
\]

This implies, that

\[
\nabla_x \Gamma(\hat{x})(d) \in T_K(\Gamma(\hat{x})),
\]

where \( T_K(\gamma) \) denotes the tangent cone to \( K \) at \( \gamma \). Since

\[
T_K(\gamma) = K + \{ t\gamma : t \in \mathbb{R} \},
\]

there exists \( t \in \mathbb{R} \) such that

\[
\nabla_x \Gamma(\hat{x})(d) + t \Gamma(\hat{x}) \in K.
\] (4.106)

Condition (4.101) implies that there exists \( q \in \mathcal{N}_X(\hat{x}) \) such that

\[
\nabla_x c(\hat{x}) + \int_a^b \nabla_x \Gamma(\hat{x})(\eta) \, d\mu(\eta) = -q.
\]

Applying both sides of this equation to the direction \( d \) and using the fact that \( q \in \mathcal{N}_X(\hat{x}) \) and \( d \in T_X(\hat{x}) \), we obtain:

\[
\nabla_x c(\hat{x})(d) + \int_a^b \left( \nabla_x \Gamma(\hat{x})(\eta) \right)(d) \, d\mu(\eta) \geq 0.
\] (4.107)
Condition (4.102), relation (4.106), and the nonnegativity of $\mu$ imply that
\[ \int_a^b \left( \nabla_x \Gamma(\hat{x})(\eta) \right) (d) \, d\mu(\eta) = \int_a^b \left[ (\nabla_x \Gamma(\hat{x})(\eta)) (d) + t(\Gamma(\hat{x})(\eta)) \right] \, d\mu(\eta) \leq 0. \]
Substituting into (4.107) we conclude that
\[ d^T \nabla_x c(\hat{x}) \geq 0, \]
for every feasible direction $d$ at $\hat{x}$. By the convexity of $c$, for every feasible point $x$ we obtain the inequality
\[ c(x) \geq c(\hat{x}) + d^T \nabla_x c(\hat{x}) \geq c(\hat{x}), \]
as stated. \(\Box\)

**Exercises**

4.1. Are the following density functions $\alpha$-concave and do they define a $\gamma$-concave probability measure? What are $\alpha$ and $\gamma$?

(a) If the $m$-dimensional random vector $Z$ has the normal distribution with expected value $\mu = 0$ and covariance matrix $\Sigma$, the random variable $Y$ is independent of $Z$ and has the $\chi^2_k$ distribution, then the distribution of the vector $X$ with components
\[ X_i = \frac{Z_i}{\sqrt{Y/k}}, \quad i = 1, \ldots, m, \]
is called a *multivariate Student distribution*. Its density function is defined as follows:
\[ \theta_m(x) = \frac{\Gamma(\frac{m+k}{2})}{\Gamma(\frac{k}{2}) \sqrt{(2\pi)^m \det(\Sigma)}} \left( 1 + \frac{1}{k} x^T \Sigma^{-1} x \right)^{-\frac{(m+k)/2}{2}}, \]
If $m = k = 1$, then this function reduces to the well-known univariate Cauchy density
\[ \theta_1(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad -\infty < x < \infty. \]

(b) The density function of the $m$-dimensional $F$-distribution with parameters $n_0, \ldots, n_m$, and $n = \sum_{i=1}^m n_i$, defined as follows:
\[ \theta(x) = c \prod_{i=1}^m x_i^{n_i/2 - 1} \left( n_0 + \sum_{i=1}^m n_i x_i \right)^{-n/2}, \quad x_i \geq 0, \quad i = 1, \ldots, m, \]
where $c$ is an appropriate normalizing constant.
Exercises 153

(c) Consider another multivariate generalization of the beta distribution, which is obtained in the following way. Let $S_1$ and $S_2$ be two independent sampling covariance matrices corresponding to two independent samples of sizes $s_1 + 1$ and $s_2 + 1$, respectively, taken from the same $q$-variate normal distribution with covariance matrix $\Sigma$. The joint distribution of the elements on and above the main diagonal of the random matrix

$$(S_1 + S_2)^{\frac{1}{2}} S_2 (S_1 + S_2)^{-\frac{1}{2}}$$

is continuous if $s_1 \geq q$ and $s_2 \geq q$. The probability density function of this distribution is defined by

$$\theta(X) = \begin{cases} 
\frac{c(s_1, q)c(s_2, q)}{c(s_1 + s_2, q)} \det(X)^{\frac{1}{2}(s_2 - q - 1)} \det(I - X)^{\frac{1}{2}(s_1 - q - 1)}, & \text{for } X, I - X \text{ positive definite}, \\
0, & \text{otherwise}.
\end{cases}$$

Here $I$ stands for the identity matrix and the function $c(\cdot, \cdot)$ is defined as follows:

$$\frac{1}{c(k, q)} = \frac{2^{qk/2} \pi^{q(q-1)/2}}{\prod_{i=1}^{q} \Gamma \left( \frac{k - i + 1}{2} \right)}.$$

The number of independent variables in $X$ is $s = \frac{1}{2} q(q + 1)$.

(d) The probability density function of the Pareto distribution is

$$\theta(x) = a(a + 1) \cdots (a + s - 1) \left( \prod_{j=1}^{s} \Theta_j \right)^{-(a+s)} \left( \sum_{j=1}^{s} \Theta_j^{-1} x_j - s + 1 \right)^{-(a+s)}$$

for $x_i > \Theta_i, \ i = 1, \ldots, s$ and $\theta(x) = 0$ otherwise. Here $\Theta_i, i = 1, \ldots, s$ are positive constants.

4.2. Assume that $P$ is an $\alpha$-concave probability distribution and $A \subset \mathbb{R}^n$ is a convex set. Prove that the function $f(x) = P(A + x)$ is $\alpha$-concave.

4.3. Prove that if $\theta : \mathbb{R} \rightarrow \mathbb{R}$ is a log-concave probability density function, then the functions

$$F(x) = \int_{t \leq x} \theta(t) dt \quad \text{and} \quad \bar{F}(x) = 1 - F(x)$$

are log-concave as well.

4.4. Check that the binomial, the Poisson, the geometric, and the hypergeometric one-dimensional probability distributions satisfy the conditions of Theorem 4.38 and are, therefore, log-concave.

4.5. Let $Z_1$, $Z_2$, and $Z_3$ be independent exponentially distributed random variables with parameters $\lambda_1$, $\lambda_2$, and $\lambda_3$, respectively. We define $Y_1 = \min\{Z_1, Z_3\}$ and $Y_2 = \min\{Z_2, Z_3\}$. Describe $G(\eta_1, \eta_2) = P(Y_1 \geq \eta_1, Y_2 \geq \eta_2)$ for nonnegative scalars $\eta_1$ and $\eta_2$ and prove that $G(\eta_1, \eta_2)$ is log-concave on $\mathbb{R}^2$. 
4.6. Let $Z$ be a standard normal random variable, $W$ be a $\chi^2$-random variable with one degree of freedom, and $A$ be an $n \times n$ positive definite matrix. Is the set
\[
\left\{ x \in \mathbb{R}^n : \Pr( Z - \sqrt{(x^T A x)} W \geq 0 ) \geq 0.9 \right\}
\]
convex?

4.7. If $Y$ is an $m$-dimensional random vector with a log-normal distribution, and $g : \mathbb{R}^n \to \mathbb{R}^m$ is such that each component $g_i$ is a concave function, the show that the set
\[
C = \left\{ x \in \mathbb{R}^n : \Pr( g(x) \geq Y ) \geq 0.9 \right\}
\]
is convex.

(a) Find the set of $p$-efficient points for $m = 1$, $p = 0.9$ and write an equivalent algebraic description of $C$.

(b) Assume that $m = 2$ and the components of $Y$ are independent. Find a disjunctive algebraic formulation for the set $C$.

4.8. Consider the following optimization problem:
\[
\begin{align*}
\text{Min} & \quad c^T x \\
\text{s.t.} & \quad \Pr\{ g_i(x) \geq Y_i, i = 1, 2 \} \geq 0.9, \\
& \quad x \geq 0.
\end{align*}
\]
Here $c \in \mathbb{R}^n$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2$, is a concave function, $Y_1$ and $Y_2$ are independent random variables that have the log-normal distribution with parameters $\mu = 0$, $\sigma = 2$.

Formulate necessary and sufficient optimality conditions for this problem.

4.9. Assuming that $Y$ and $Z$ are independent exponentially distributed random variables, show that the following set is convex:
\[
\left\{ x \in \mathbb{R}^3 : \Pr( x_1^2 + x_2^2 + Y x_2 + x_2 x_3 + Y x_3 \leq Z ) \geq e^{-y}, \forall y \in [1, 4] \right\}.
\]

4.10. Assume that the random variable $Z$ is uniformly distributed in the interval $[-1, 1]$ and $c = (1, \ldots, 1)^T$. Prove that the following set is convex:
\[
\left\{ x \in \mathbb{R}^n : \Pr( \exp(x^T y) \geq (c^T y) Z, \forall y \in \mathbb{R}^n : \|y\| \leq 1 ) \geq 0.95 \right\}.
\]

4.11. Let $Z$ be a two-dimensional random vector with Dirichlet distribution. Show that the following set is convex:
\[
\left\{ x \in \mathbb{R}^2 : \Pr( \min(x_1 + 2x_2 + Z_1, x_1 Z_2 - x_1^2 - Z_2^2) \geq y) \geq e^{-y}, \forall y \in [1, 4] \right\}.
\]

4.12. Let $Z$ be an $n$-dimensional random vector uniformly distributed on a set $A$. Check whether the set
\[
\left\{ x \in \mathbb{R}^n : \Pr( x^T Z \leq 1 ) \geq 0.95 \right\}
\]
is convex for the following cases:
Exercises 155

(a) \( A = \{ z \in \mathbb{R}^n : \| z \| \leq 1 \} \).

(b) \( A = \{ z \in \mathbb{R}^n : 0 \leq z_i \leq i, i = 1, \ldots, m \} \).

(c) \( A = \{ z \in \mathbb{R}^n : Tz \leq 0, -1 \leq z_i \leq 1, i = 1, \ldots, m \} \), where \( T \) is an \( (n-1) \times n \) matrix of form

\[
T = \begin{pmatrix}
1 & -1 & 0 & \cdots & 0 \\
0 & 1 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -1
\end{pmatrix}.
\]

4.13. Assume that the two-dimensional random vector \( Z \) has independent components, which have the Poisson distribution with parameters \( \lambda_1 = \lambda_2 = 2 \). Find all \( p \)-efficient points of \( F_Z \) for \( p = 0.8 \).
Chapter 5

Statistical Inference

Alexander Shapiro

5.1 Statistical Properties of SAA Estimators

Consider the following stochastic programming problem

\[
\min_{x \in X} \{ f(x) := \mathbb{E}[F(x, \xi)] \}.
\] (5.1)

Here \( X \) is a nonempty closed subset of \( \mathbb{R}^n \), \( \xi \) is a random vector whose probability distribution \( P \) is supported on a set \( \Xi \subset \mathbb{R}^d \), and \( F : X \times \Xi \rightarrow \mathbb{R} \). In the framework of two-stage stochastic programming the objective function \( F(x, \xi) \) is given by the optimal value of the corresponding second stage problem. Unless stated otherwise we assume in this chapter that the expectation function \( f(x) \) is well defined and finite valued for all \( x \in X \). This implies, of course, that for every \( x \in X \) the value \( F(x, \xi) \) is finite for a.e. \( \xi \in \Xi \). In particular, for two-stage programming this implies that the recourse is relatively complete.

Suppose that we have a sample \( \xi_1, \ldots, \xi_N \) of \( N \) realizations of the random vector \( \xi \). This random sample can be viewed as historical data of \( N \) observations of \( \xi \), or it can be generated in the computer by Monte Carlo sampling techniques. For any \( x \in X \) we can estimate the expected value \( f(x) \) by averaging values \( F(x, \xi_j), j = 1, \ldots, N \). This leads to the following, so-called sample average approximation (SAA):

\[
\min_{x \in X} \left\{ \hat{f}_N(x) := \frac{1}{N} \sum_{j=1}^{N} F(x, \xi_j) \right\}
\] (5.2)

of the “true” problem (5.1). Let us observe that we can write the sample average function as the expectation

\[
\hat{f}_N(x) = \mathbb{E}_{P_N}[F(x, \xi)]
\] (5.3)
taken with respect to the empirical distribution\(^1\) (measure) \(P_N := N^{-1} \sum_{i=1}^{N} \Delta(\xi_i)\).

Therefore, for a given sample, the SAA problem (5.2) can be considered as a stochastic programming problem with respective scenarios \(\xi^1, \ldots, \xi^N\), each taken with probability \(1/N\).

As with data vector \(\xi\), the sample \(\xi^1, \ldots, \xi^N\) can be considered from two points of view, as a sequence of random vectors or as a particular realization of that sequence. Which one of these two meanings will be used in a particular situation will be clear from the context. The SAA problem is a function of the considered sample, and in that sense is random. For a particular realization of the random sample the corresponding SAA problem is a stochastic programming problem with respective scenarios \(\xi^1, \ldots, \xi^N\) each taken with probability \(1/N\). We always assume that each random vector \(\xi^j\), in the sample, has the same (marginal) distribution \(P\) as the data vector \(\xi\). If, moreover, each \(\xi^j, j = 1, \ldots, N\), is distributed independently of other sample vectors, we say that the sample is independently identically distributed (iid).

By the Law of Large Numbers we have that, under some regularity conditions, \(\hat{f}_N(x)\) converges pointwise w.p.1 to \(f(x)\) as \(N \to \infty\). In particular, by the classical LLN this holds if the sample is iid. Moreover, under mild additional conditions the convergence is uniform (see section 7.2.5). We also have that \(E[\hat{f}_N(x)] = f(x)\), i.e., \(\hat{f}_N(x)\) is an unbiased estimator of \(f(x)\). Therefore, it is natural to expect that the optimal value and optimal solutions of the SAA problem (5.2) converge to their counterparts of the true problem (5.1) as \(N \to \infty\). We denote by \(\vartheta^*\) and \(S\) the optimal value and the set of optimal solutions, respectively, of the true problem (5.1), and by \(\hat{\vartheta}_N\) and \(\hat{S}_N\) the optimal value and the set of optimal solutions, respectively, of the SAA problem (5.2).

We can view the sample average functions \(f_N(x)\) as defined on a common probability space \((\Omega, \mathcal{F}, P)\). For example, in the case of iid sample, a standard construction is to consider the set \(\Omega := \Xi^\infty\) of sequences \(\{(\xi_1, \ldots)\}_{\xi \in \Xi, i \in \mathbb{N}}\), equipped with the product of the corresponding probability measures. Assume that \(F(x, \xi)\) is a Carathéodory function, i.e., continuous in \(x\) and measurable in \(\xi\). Then \(\hat{f}_N(x) = f_N(x, \omega)\) is also a Carathéodory function and hence is a random lsc function. It follows (see section 7.2.3 and Theorem 7.37 in particular) that \(\vartheta_N = \vartheta_N(\omega)\) and \(\hat{S}_N = \hat{S}_N(\omega)\) are measurable. We also consider a particular optimal solution \(\bar{x}_N\) of the SAA problem, and view it as a measurable selection \(\bar{x}_N(\omega) \in \hat{S}_N(\omega)\) of \(\hat{S}_N(\omega)\). Existence of such measurable selection is ensured by the Measurable Selection Theorem (Theorem 7.34). This takes care of the measurability questions.

We are going to discuss next statistical properties of the SAA estimators \(\hat{\vartheta}_N\) and \(\hat{S}_N\). Let us make the following useful observation.

**Proposition 5.1.** Let \(f : X \to \mathbb{R}\) and \(f_N : X \to \mathbb{R}\) be a sequence of (deterministic) real valued functions. Then the following two properties are equivalent: (i) for any \(\bar{x} \in X\) and any sequence \(\{x_N\} \subset X\) converging to \(\bar{x}\) it follows that \(f_N(x_N)\) converges to \(f(\bar{x})\), (ii) the function \(f(\cdot)\) is continuous on \(X\) and \(f_N(\cdot)\) converges to \(f(\cdot)\) uniformly on any compact subset of \(X\).

**Proof.** Suppose that property (i) holds. Consider a point \(\bar{x} \in X\), a sequence \(\{x_N\} \subset X\) converging to \(\bar{x}\) and a number \(\varepsilon > 0\). By taking a sequence with each element equal \(x_1\), we

\(^1\)Recall that \(\Delta(\xi)\) denotes measure of mass one at the point \(\xi\).
have by (i) that $f_N(x_1) \to f(x_1)$. Therefore there exists $N_1$ such that $|f_N(x_1) - f(x_1)| < \epsilon/2$. Similarly there exists $N_2 > N_1$ such that $|f_N(x_2) - f(x_2)| < \epsilon/2$, and so on. Consider now a sequence, denoted $x_N$, constructed as follows, $x'_1 = x_1$, $i = 1, \ldots, N_1$, $x'_i = x_2$, $i = N_1 + 1, \ldots, N_2$, and so on. We have that this sequence $x'_N$ converges to $x$ and hence $|f_N(x'_N) - f(x)| < \epsilon/2$ for all $N$ large enough. We also have that $|f_N(x'_N) - f(x_k)| < \epsilon/2$, and hence $|f(x_k) - f(x)| < \epsilon$ for all $k$ large enough. This shows that $f(x_k) \to f(x)$ and hence $f(\cdot)$ is continuous at $x$.

Now let $C$ be a compact subset of $X$. Arguing by a contradiction suppose that $f_N(\cdot)$ does not converge to $f(\cdot)$ uniformly on $C$. Then there exists a sequence $\{x_N\} \subset C$ and $\epsilon > 0$ such that $|f_N(x_N) - f(x_N)| \geq \epsilon$ for all $N$. Since $C$ is compact, we can assume that $\{x_N\}$ converges to a point $\bar{x} \in C$. We have

$$|f_N(x_N) - f(x_N)| \leq |f_N(x_N) - f(\bar{x})| + |f(x_N) - f(\bar{x})|. \quad (5.4)$$

The first term in the right hand side of (5.4) tends to zero by (i) and the second term tends to zero since $f(\cdot)$ is continuous, and hence these terms are less that $\epsilon/2$ for $N$ large enough. This gives a desired contradiction.

Conversely, suppose that property (ii) holds. Consider a sequence $\{x_N\} \subset X$ converging to a point $\bar{x} \in X$. We can assume that this sequence is contained in a compact subset of $X$. By employing the inequality

$$|f_N(x_N) - f(x_N)| \leq |f_N(x_N) - f(\bar{x})| + |f(x_N) - f(\bar{x})| \quad (5.5)$$

and noting that the first term in the right hand side of this inequality tends to zero because of the uniform convergence of $f_N$ to $f$ and the second term tends to zero by continuity of $f$, we obtain that property (i) holds.

### 5.1.1 Consistency of SAA Estimators

In this section we discuss convergence properties of the SAA estimators $\hat{\vartheta}_N$ and $\hat{S}_N$. It is said that an estimator $\hat{\vartheta}_N$ of a parameter $\vartheta$ is consistent if $\hat{\vartheta}_N$ converges w.p.1 to $\vartheta$ as $N \to \infty$. Let us consider first consistency of the SAA estimator of the optimal value. We have that for any fixed $x \in X$, $\hat{\vartheta}_N \leq \hat{f}_N(x)$ and hence if the pointwise LLN holds, then $\limsup_{N \to \infty} \hat{\vartheta}_N \leq f(x)$ w.p.1. It follows that if the pointwise LLN holds, then

$$\limsup_{N \to \infty} \hat{\vartheta}_N \leq \vartheta^* \text{ w.p.1.} \quad (5.6)$$

Without some additional conditions, the inequality in (5.6) can be strict.

**Proposition 5.2.** Suppose that $\hat{f}_N(x) \to f(x)$ w.p.1, as $N \to \infty$, uniformly on $X$. Then $\hat{\vartheta}_N$ converges to $\vartheta^*$ w.p.1 as $N \to \infty$.

**Proof.** The uniform convergence w.p.1 of $\hat{f}_N(x) = \hat{f}_N(x, \omega)$ to $f(x)$ means that for any $\epsilon > 0$ and a.e. $\omega \in \Omega$ there is $N^* = N^*(\epsilon, \omega)$ such that the following inequality holds for all $N \geq N^*$:

$$\sup_{x \in X} |\hat{f}_N(x, \omega) - f(x)| \leq \epsilon. \quad (5.7)$$
It follows then that $|\hat{\theta}_N(\omega) - \theta^*| \leq \varepsilon$ for all $N \geq N^*$, which completes the proof. \hfill \Box

In order to establish consistency of the SAA estimators of optimal solutions we need slightly stronger conditions. Recall that $\mathbb{D}(A, B)$ denotes the deviation of set $A$ from set $B$ (see equation (7.4) for the corresponding definition).

**Theorem 5.3.** Suppose that there exists a compact set $C \subset \mathbb{R}^n$ such that: (i) the set $S$ of optimal solutions of the true problem is nonempty and is contained in $C$, (ii) the function $f(x)$ is finite valued and continuous on $C$, (iii) $\hat{f}_N(x)$ converges to $f(x)$ w.p.1, as $N \to \infty$, uniformly in $x \in C$, (iv) w.p.1 for $N$ large enough the set $S_N$ is nonempty and $\hat{S}_N \subset C$. Then $\hat{\theta}_N \to \theta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.

**Proof.** Assumptions (i) and (iv) imply that both the true and SAA problems can be restricted to the set $X \cap C$. Therefore we can assume without loss of generality that the set $X$ is compact. The assertion that $\hat{\theta}_N \to \theta^*$ w.p.1 follows by Proposition 5.2. It suffices show now that $\mathbb{D}(\hat{S}_N(\omega), S) \to 0$ for every $\omega \in \Omega$ such that $\hat{\theta}_N(\omega) \to \theta^*$ and assumptions (iii) and (iv) hold. This basically a deterministic result, therefore we omit $\omega$ for the sake of notational convenience.

We argue now by a contradiction. Suppose that $\mathbb{D}(\hat{S}_N, S) \nrightarrow 0$. Since $X$ is compact, by passing to a subsequence if necessary, we can assume that there exists $\hat{x}_N \in \hat{S}_N$ such that $\text{dist}(\hat{x}_N, S) \geq \varepsilon$ for some $\varepsilon > 0$, and that $\hat{x}_N$ tends to a point $x^* \in X$. It follows that $x^* \notin S$ and hence $f(x^*) > \theta^*$. Moreover, $\hat{\theta}_N = \hat{f}_N(\hat{x}_N)$ and

$$
\hat{f}_N(\hat{x}_N) - f(x^*) = [\hat{f}_N(\hat{x}_N) - f(\hat{x}_N)] + [f(\hat{x}_N) - f(x^*)].
$$

(5.8)

The first term in the right hand side of (5.8) tends to zero by the assumption (iii) and the second term by continuity of $f(x)$. That is, we obtain that $\hat{\theta}_N$ tends to $f(x^*) > \theta^*$, a contradiction. \hfill \Box

Recall that, by Proposition 5.1, assumptions (ii) and (iii) in the above theorem are equivalent to the condition that for any sequence $\{x_N\} \subset C$ converging to a point $\hat{x}$ it follows that $\hat{f}_N(x_N) \to f(\hat{x})$ w.p.1. The last assumption (iv) in the above theorem holds, in particular, if the feasible set $X$ is closed, the functions $\hat{f}_N(x)$ are lower semicontinuous and for some $\alpha > \theta^*$ the level sets $\{x \in X : \hat{f}_N(x) \leq \alpha\}$ are uniformly bounded w.p.1. This condition is often referred to as the inf-compactness condition. Conditions ensuring the uniform convergence of $\hat{f}_N(x)$ to $f(x)$ (assumption (iii)) are given in Theorems 7.48 and 7.50, for example.

The assertion that $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 means that for any (measurable) selection $\hat{x}_N \in \hat{S}_N$, of an optimal solution of the SAA problem, it holds that $\text{dist}(\hat{x}_N, S) \to 0$ w.p.1. If, moreover, $S = \{\hat{x}\}$ is a singleton, i.e., the true problem has unique optimal solution $\hat{x}$, then this means that $\hat{x}_N \to \hat{x}$ w.p.1. The inf-compactness condition ensures that $\hat{x}_N$ cannot escape to infinity as $N$ increases.

If the problem is convex, it is possible to relax the required regularity conditions. In the following theorem we assume that the integrand function $F(x, \xi)$ is an extended real valued function, i.e., can also take values $\pm \infty$. Denote

$$
\bar{F}(x, \xi) := F(x, \xi) + \mathbb{I}_X(x), \quad \bar{f}(x) := f(x) + \mathbb{I}_X(x), \quad \bar{f}_N(x) := \bar{f}_N(x) + \mathbb{I}_X(x),
$$

(5.9)
i.e., \( f(x) = f(x) \) if \( x \in X \), and \( \bar{f}(x) = +\infty \) if \( x \not\in X \), and similarly for functions \( F(\cdot, \xi) \) and \( \bar{f}_N(\cdot) \). Clearly \( \bar{f}(x) = \mathbb{E}[F(x, \xi)] \) and \( \bar{f}_N(x) = N^{-1} \sum_{j=1}^N F(x, \xi^j) \). Note that if the set \( X \) is convex, then the above ‘penalization’ operation preserves convexity of respective functions.

**Theorem 5.4.** Suppose that: (i) the integrand function \( F \) is random lower semicontinuous, (ii) for almost every \( \xi \in \Xi \) the function \( F(\cdot, \xi) \) is convex, (iii) the set \( X \) is closed and convex, (iv) the expected value function \( f \) is lower semicontinuous and there exists a point \( \bar{x} \in X \) such that \( f(\bar{x}) < +\infty \) for all \( x \) in a neighborhood of \( \bar{x} \), (v) the set \( S \) of optimal solutions of the true problem is nonempty and bounded, (vi) the LLN holds pointwise. Then \( \Theta_N \rightarrow \theta^* \) and \( \mathbb{D}(\bar{S}_N, S) \rightarrow 0 \) w.p.1 as \( N \rightarrow \infty \).

**Proof.** Clearly we can restrict both the true and the SAA problems to the affine space generated by the convex set \( X \). Relative to that affine space the set \( \bar{x} \) has a nonempty interior. Therefore, without loss of generality we can assume that the set \( X \) has a nonempty interior. Since it is assumed that \( f(x) \) possesses an optimal solution, we have that \( \theta^* \) is finite and hence \( f(x) \geq \theta^* > -\infty \) for all \( x \in X \). Since \( f(x) \) is convex and greater than \( -\infty \) on an open set (e.g., interior of \( X \)), it follows that \( f(\cdot) \) is subdifferentiable at any point \( x \in \text{int}(X) \) such that \( f(x) \) is finite. Consequently \( f(x) > -\infty \) for all \( x \in \mathbb{R}^n \), and hence \( f \) is proper.

Observe that the pointwise LLN for \( F(x, \xi) \) (assumption (vi)) implies the corresponding pointwise LLN for \( \bar{F}(x, \xi) \). Since \( X \) is convex and closed, it follows that \( \bar{f} \) is convex and lower semicontinuous. Moreover, because of the assumption (iv) and since the interior of \( X \) is nonempty, we have that \( \text{dom} \bar{f} \) has a nonempty interior. By Theorem 7.49 it follows then that \( \bar{f}_N \rightarrow \bar{f} \) w.p.1. Consider a compact set \( K \) with a nonempty interior and such that it does not contain a boundary point of \( \text{dom} \bar{f} \), and \( \bar{f}(x) \) is finite valued on \( K \). Since \( \text{dom} \bar{f} \) has a nonempty interior such set exists. Then it follows from \( \bar{f}_N \rightarrow \bar{f} \), that \( \bar{f}_N(\cdot) \) converge to \( \bar{f}(\cdot) \) uniformly on \( K \), all w.p.1 (see Theorem 7.27). It follows that w.p.1 for \( N \) large enough the functions \( \bar{f}_N(x) \) are finite valued on \( K \), and hence \( f \) is proper.

Now let \( C \) be a compact subset of \( \mathbb{R}^m \) such that the set \( S \) is contained in the interior of \( C \). Such set exists since it is assumed that the set \( S \) is bounded. Consider the set \( \bar{S}_N \) of minimizers of \( \bar{f}_N(x) \) over \( C \). Since \( C \) is nonempty and compact and \( \bar{f}_N(x) \) is lower semicontinuous and proper for \( N \) large enough, and because by the pointwise LLN we have that for any \( x \in S \), \( f_N(x) \) is finite w.p.1 for \( N \) large enough, the set \( \bar{S}_N \) is nonempty w.p.1 for \( N \) large enough. Let us show that \( \mathbb{D}(\bar{S}_N, S) \rightarrow 0 \) w.p.1. Let \( \omega \in \Omega \) be such that \( \bar{f}_N(\cdot, \omega) \rightarrow \bar{f}(\cdot) \). We have that this happens for a.e. \( \omega \in \Omega \). We argue now by a contradiction. Suppose that there exists a minimizer \( \bar{x}_N = \bar{x}_N(\omega) \) of \( \bar{f}_N(x, \omega) \) over \( C \) such that \( \text{dist}(\bar{x}_N, S) \geq \varepsilon \) for some \( \varepsilon > 0 \). Since \( C \) is compact, by passing to a subsequence if necessary, we can assume that \( \bar{x}_N \) tends to a point \( x^* \in C \). It follows that \( x^* \not\in S \). On the other hand we have by Proposition 7.26 that \( x^* \in \arg\min_{x \in C} f(x) \). Since \( \arg\min_{x \in C} f(x) = S \), we obtain a contradiction.

Now because of the convexity assumptions, any minimizer of \( \bar{f}_N(x) \) over \( C \) which lies inside the interior of \( C \), is also an optimal solution of the SAA problem (5.2). Therefore, w.p.1 for \( N \) large enough we have that \( \bar{S}_N = \bar{S}_N \). Consequently we can restrict both the true and the SAA optimization problems to the compact set \( C \), and hence the assertions
of the above theorem follow. \( \square \)

Let us make the following observations. Lower semicontinuity of \( f(\cdot) \) follows from lower semicontinuity \( F(\cdot, \xi) \), provided that \( F(x, \cdot) \) is bounded from below by an integrable function (see Theorem 7.42 for a precise formulation of this result). It was assumed in the above theorem that the LLN holds pointwise for all \( x \in \mathbb{R}^n \). Actually it suffices to assume that this holds for all \( x \) in some neighborhood of the set \( S \). Under the assumptions of the above theorem we have that \( f(x) > -\infty \) for every \( x \in \mathbb{R}^n \). The above assumptions do not prevent, however, for \( f(x) \) to take value \( +\infty \) at some points \( x \in X \). Nevertheless, it was possible to push the proof through because in the considered convex case local optimality implies global optimality. There are two possible reasons why \( f(x) \) can be \(+\infty\). Namely, it can be that \( F(x, \cdot) \) is finite valued but grows sufficiently fast so that its integral is \(+\infty\), or it can be that \( F(x, \cdot) \) is equal \(+\infty\) on a set of positive measure, and of course it can be both. For example, in the case of two-stage programming it may happen that for some \( x \in X \) the corresponding second stage problem is infeasible with a positive probability \( p \). Then w.p.1 for \( N \) large enough, for at least one of the sample points \( \xi \) the corresponding second stage problem will be infeasible, and hence \( f_N(x) = +\infty \). Of course, if the probability \( p \) is very small, then the required sample size for such event to happen could be very large.

We assumed so far that the feasible set \( X \) of the SAA problem is fixed, i.e., independent of the sample. However, in some situations it also should be estimated. Then the corresponding SAA problem takes the form

\[
\min_{x \in X_N} f_N(x), \tag{5.10}
\]

where \( X_N \) is a subset of \( \mathbb{R}^n \) depending on the sample, and therefore is random. As before we denote by \( \hat{\theta}_N \) and \( \hat{S}_N \) the optimal value and the set of optimal solutions, respectively, of the SAA problem (5.10).

**Theorem 5.5.** Suppose that in addition to the assumptions of Theorem 5.3 the following conditions hold:

(a) If \( x_N \in X_N \) and \( x_N \) converges w.p.1 to a point \( x \), then \( x \in X \).

(b) For some point \( x \in S \) there exists a sequence \( x_N \in X_N \) such that \( x_N \to x \) w.p.1. Then \( \hat{\theta}_N \to \theta^* \)

and \( D(\hat{S}_N, S) \to 0 \) w.p.1 as \( N \to \infty \).

**Proof.** Consider an \( \hat{x}_N \in \hat{S}_N \). By compactness arguments we can assume that \( \hat{x}_N \) converges w.p.1 to a point \( x^* \in \mathbb{R}^n \). Since \( \hat{S}_N \subset X_N \), we have that \( \hat{x}_N \in X_N \), and hence it follows by condition (a) that \( x^* \in X \). We also have (see Proposition 5.1) that \( \hat{\theta}_N = f_N(\hat{x}_N) \) tends w.p.1 to \( f(x^*) \), and hence \( \lim\inf_{N \to \infty} \hat{\theta}_N \geq \theta^* \) w.p.1. On the other hand, by condition (b), there exists a sequence \( x_N \in X_N \) converging to a point \( x \in S \) w.p.1. Consequently, \( \hat{\theta}_N \leq f_N(\hat{x}_N) \to f(x) = \theta^* \) w.p.1, and hence \( \lim\sup_{N \to \infty} \hat{\theta}_N \leq \theta^* \). It follows that \( \hat{\theta}_N \to \theta^* \) w.p.1. The remainder of the proof can be completed by the same arguments as in the proof of Theorem 5.3. \( \square \)

The SAA problem (5.10) is convex if the functions \( f_N(\cdot) \) and the sets \( X_N \) are convex w.p.1. It is also possible to show consistency of the SAA estimators of problem (5.10) under
the assumptions of Theorem 5.4 together with conditions (a) and (b), of the above theorem, and convexity of the set $X_N$.

Suppose, for example, that the set $X$ is defined by the constraints

$$X := \{ x \in X_0 : g_i(x) \leq 0, \; i = 1, \ldots, p \},$$

(5.11)

where $X_0$ is a nonempty closed subset of $\mathbb{R}^n$ and the constraint functions are given as the expected value functions

$$g_i(x) := \mathbb{E}[G_i(x, \xi)], \quad i = 1, \ldots, p,$$

(5.12)

with $G_i(x, \xi), \; i = 1, \ldots, p,$ being random lsc functions. Then the set $X$ can be estimated by

$$X_N := \{ x \in X_0 : \hat{g}_i(x) \leq 0, \; i = 1, \ldots, p \},$$

(5.13)

where $\hat{g}_i(x) := N^{-1} \sum_{j=1}^N g_i(x, \xi^j)$. If for a given point $x \in X_0$ we have that $\hat{g}_i(x)$ converge uniformly to $g_i$ w.p.1 on a neighborhood of $x$ and the functions $g_i$ are continuous, then the above condition (a) holds.

In order to ensure condition (b) (of Theorem 5.5) one needs to impose a constraint qualification (on the true problem). Consider, for example, $X := \{ x \in \mathbb{R} : g(x) \leq 0 \}$ with $g(x) := x^2$. Clearly $X = \{0\}$, while an arbitrary small perturbation of the function $g(\cdot)$ can result in the corresponding set $X_N$ being empty. It is possible to show that if a constraint qualification for the true problem is satisfied at $x$, then condition (b) follows. For instance, if the set $X_0$ is convex and for every $\xi \in \Xi$ the functions $G_i(\cdot, \xi)$ are convex, and hence the corresponding expected value functions $g_i(\cdot), \; i = 1, \ldots, p,$ are also convex, then such a simple constraint qualification is the Slater condition. Recall that it is said that the Slater condition holds if there exists a point $x^* \in X_0$ such that $g_i(x^*) < 0, \; i = 1, \ldots, p$.

As another example suppose that the feasible set is given by probabilistic (chance) constraints in the form

$$X = \{ x \in \mathbb{R}^n : \Pr(C_i(x, \xi) \leq 0) \geq 1 - \alpha_i, \; i = 1, \ldots, p \},$$

(5.14)

where $\alpha_i \in (0, 1)$ and $C_i : \mathbb{R}^n \times \Xi \to \mathbb{R}, \; i = 1, \ldots, p,$ are Carathéodory functions. Of course, we have that

$$\Pr(C_i(x, \xi) \leq 0) = \mathbb{E} \left[ 1_{(-\infty, 0]}(C_i(x, \xi)) \right].$$

(5.15)

Consequently, we can write the above set $X$ in the form (5.11)–(5.12) with $X_0 := \mathbb{R}^n$ and

$$G_i(x, \xi) := 1 - \alpha_i - 1_{(-\infty, 0]}(C_i(x, \xi)).$$

(5.16)

The corresponding set $X_N$ can be written as

$$X_N = \left\{ x \in \mathbb{R}^n : \frac{1}{N} \sum_{j=1}^N 1_{(-\infty, 0]}(C_i(x, \xi^j)) \geq 1 - \alpha_i, \; i = 1, \ldots, p \right\}.$$

(5.17)

\[3\] Recall that $1_{(-\infty, 0]}(t) = 1$ if $t \leq 0$, and $1_{(-\infty, 0]}(t) = 0$ if $t > 0$. 


Note that \( \sum_{j=1}^{N} 1_{(-\infty, 0)}(C_i(x, \xi^j)) \), in the above formula, counts the number of times that the event “\( C_i(x, \xi^j) \leq 0 \)”, \( i = 1, ..., N \), happens. The additional difficulty here is that the (step) function \( 1_{(-\infty, 0)}(t) \) is discontinuous at \( t = 0 \). Nevertheless, suppose that the sample is iid and for every \( x \) in a neighborhood of the set \( X \) and \( i = 1, ..., p \), the event “\( C_i(x, \xi) = 0 \)” happens with probability zero, and hence \( G_i(\cdot, \xi) \) is continuous at \( x \) for a.e. \( \xi \). By Theorem 7.48 this implies that the expectation function \( g_i(x) \) is continuous and \( g_i(x) \) converge uniformly w.p.1 on compact neighborhoods to \( g_i(x) \), and hence condition (a) holds. Condition (b) could be verified by ad hoc methods.

### 5.1.2 Asymptotics of the SAA Optimal Value

Consistency of the SAA estimators gives a certain assurance that the error of the estimation approaches zero in the limit as the sample size grows to infinity. Although important conceptually, this does not give any indication of the magnitude of the error for a given sample. Suppose for the moment that the sample is iid and let us fix a point \( x \in X \). Then we have that the sample average estimator \( \hat{f}_N(x) \), of \( f(x) \), is unbiased and has variance \( \sigma^2(x)/N \), where \( \sigma^2(x) := \text{Var}[F(x, \xi)] \) is supposed to be finite. Moreover, by the Central Limit Theorem (CLT) we have that

\[
N^{1/2} \left[ \hat{f}_N(x) - f(x) \right] \xrightarrow{D} \mathcal{N}_x,
\]

where “\( \xrightarrow{D} \)” denotes convergence in distribution and \( \mathcal{N}_x \) has a normal distribution with mean 0 and variance \( \sigma^2(x) \), written \( \mathcal{N}_x \sim \mathcal{N}(0, \sigma^2(x)) \). That is, \( \hat{f}_N(x) \) has asymptotically normal distribution, i.e., for large \( N \), \( \hat{f}_N(x) \) has approximately normal distribution with mean \( f(x) \) and variance \( \sigma^2(x)/N \).

This leads to the following (approximate) 100(1 - \( \alpha \))% confidence interval for \( f(x) \):

\[
\left[ \hat{f}_N(x) - \frac{z_{\alpha/2} \hat{\sigma}(x)}{\sqrt{N}}, \hat{f}_N(x) + \frac{z_{\alpha/2} \hat{\sigma}(x)}{\sqrt{N}} \right],
\]

where \( z_{\alpha/2} := \Phi^{-1}(1 - \alpha/2) \) and \(^{3}\)

\[
\hat{\sigma}^2(x) := \frac{1}{N-1} \sum_{j=1}^{N} \left[ F(x, \xi^j) - \hat{f}_N(x) \right]^2
\]

is the sample variance estimate of \( \sigma^2(x) \). That is, the error of estimation of \( f(x) \) is (stochastically) of order \( O_p(N^{-1/2}) \).

Consider now the optimal value \( \hat{\vartheta}_N \) of the SAA problem (5.2). Clearly we have that for any \( x' \in X \) the inequality \( \hat{f}_N(x') \geq \inf_{x \in X} \hat{f}_N(x) \) holds. By taking the expected value of both sides of this inequality and minimizing the left hand side over all \( x' \in X \) we obtain

\[
\inf_{x \in X} \mathbb{E} \left[ \hat{f}_N(x') \right] \geq \mathbb{E} \left[ \inf_{x \in X} \hat{f}_N(x) \right].
\]

\(^{3}\)Here \( \Phi(\cdot) \) denotes the cdf of the standard normal distribution. For example, to 95% confidence intervals corresponds \( z_{0.025} = 1.96 \).
5.1. Statistical Properties of SAA Estimators

Note that the inequality (5.21) holds even if \( f(x) = +\infty \) or \( f(x) = -\infty \) for some \( x \in X \). Since \( E[f_N(x)] = f(x) \), it follows that \( \vartheta^* \geq E[\hat{\vartheta}_N] \). In fact, typically, \( E[\hat{\vartheta}_N] \) is strictly less than \( \vartheta^* \), i.e., \( \hat{\vartheta}_N \) is a downwards biased estimator of \( \vartheta^* \). As the following result shows this bias decreases monotonically with increase of the sample size \( N \).

**Proposition 5.6.** Let \( \hat{\vartheta}_N \) be the optimal value of the SAA problem (5.2), and suppose that the sample is iid. Then \( E[\hat{\vartheta}_N] \leq E[\hat{\vartheta}_{N+1}] \leq \vartheta^* \) for any \( N \in \mathbb{N} \).

**Proof.** It was already shown above that \( E[\hat{\vartheta}_N] \leq \vartheta^* \) for any \( N \in \mathbb{N} \). We can write

\[
\hat{f}_{N+1}(x) = \frac{1}{N+1} \sum_{i=1}^{N+1} \left[ \frac{1}{N} \sum_{j \neq i} F(x, \xi_j) \right].
\]

Moreover, since the sample is iid we have

\[
E[\hat{\vartheta}_{N+1}] = E\left[ \inf_{x \in X} \hat{f}_{N+1}(x) \right]
\]

\[
= E\left[ \inf_{x \in X} \frac{1}{N+1} \sum_{i=1}^{N+1} \left( \frac{1}{N} \sum_{j \neq i} F(x, \xi_j) \right) \right]
\]

\[
\geq E\left[ \frac{1}{N+1} \sum_{i=1}^{N+1} \left( \inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi_j) \right) \right]
\]

\[
= \frac{1}{N+1} \sum_{i=1}^{N+1} E\left[ \inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi_j) \right]
\]

\[
= \frac{1}{N+1} \sum_{i=1}^{N+1} E[\hat{\vartheta}_N] = E[\hat{\vartheta}_N],
\]

which completes the proof. \( \square \)

**First Order Asymptotics of the SAA Optimal Value**

We use the following assumptions about the integrand \( F \):

(A1) For some point \( \bar{x} \in X \) the expectation \( E[F(\bar{x}, \xi)^2] \) is finite.

(A2) There exists a measurable function \( C : \Xi \to \mathbb{R}_+ \) such that \( E[C(\xi)^2] \) is finite and

\[
|F(x, \xi) - F(x', \xi)| \leq C(\xi)\|x - x'\|, \quad (5.22)
\]

for all \( x, x' \in X \) and a.e. \( \xi \in \Xi \).

The above assumptions imply that the expected value \( f(x) \) and variance \( \sigma^2(x) \) are finite valued for all \( x \in X \). Moreover, it follows from (5.22) that

\[
|f(x) - f(x')| \leq \kappa \|x - x'\|, \quad \forall x, x' \in X,
\]

where \( \kappa := E[C(\xi)] \), and hence \( f(x) \) is Lipschitz continuous on \( X \). If \( X \) is compact, we have then that the set \( S \), of minimizers of \( f(x) \) over \( X \), is nonempty.

Let \( Y_x \) be random variables defined in (5.18). These variables depend on \( x \in X \) and we also use notation \( Y(x) = Y_x \). By the (multivariate) CLT we have that for any finite
set \( \{x_1, ..., x_m\} \subset X \), the random vector \( (Y(x_1), ..., Y(x_m)) \) has a multivariate normal distribution with zero mean and the same covariance matrix as the covariance matrix of \( (F(x_1, \xi), ..., F(x_m, \xi)) \). Moreover, by assumptions (A1) and (A2), compactness of \( X \) and since the sample is iid, we have that \( N^{1/2}(\hat{f}_N - f) \) converges in distribution to \( Y \) viewed as a random element\(^4\) of \( C(X) \). This is a so-called functional CLT (see, e.g., Araujo and Giné [4, Corollary 7.17]).

**Theorem 5.7.** Let \( \hat{\vartheta}_N \) be the optimal value of the SAA problem (5.2). Suppose that the sample is iid, the set \( X \) is compact and assumptions (A1) and (A2) are satisfied. Then the following holds:

\[
\hat{\vartheta}_N = \inf_{x \in S} \hat{f}_N(x) + o_p(N^{-1/2}), \tag{5.23}
\]

\[
N^{1/2} \left( \hat{\vartheta}_N - \vartheta^* \right) \xrightarrow{D} \inf_{x \in S} Y(x). \tag{5.24}
\]

If, moreover, \( S = \{\bar{x}\} \) is a singleton, then

\[
N^{1/2} \left( \hat{\vartheta}_N - \vartheta^* \right) \xrightarrow{D} \mathcal{N}(0, \sigma^2(\bar{x})). \tag{5.25}
\]

**Proof.** Proof is based on the functional CLT and Delta Theorem (Theorem 7.59). Consider Banach space \( C(X) \) of continuous functions \( \psi : X \to \mathbb{R} \) equipped with the sup-norm \( \|\psi\| := \sup_{x \in X} |\psi(x)| \). Define the min-value function \( V(\psi) := \inf_{x \in X} \psi(x) \).

Since \( X \) is compact, the function \( V : C(X) \to \mathbb{R} \) is real valued and measurable (with respect to the Borel sigma algebra of \( C(X) \)). Moreover, it is not difficult to see that \( |V(\psi_1) - V(\psi_2)| \leq \|\psi_1 - \psi_2\| \) for any \( \psi_1, \psi_2 \in C(X) \), i.e., \( V(\cdot) \) is Lipschitz continuous with Lipschitz constant one. By Danskin Theorem (Theorem 7.21), \( V(\cdot) \) is directionally differentiable at any \( \mu \in C(X) \) and

\[
V'_\mu(\delta) = \inf_{x \in X(\mu)} \delta(x), \quad \forall \delta \in C(X), \tag{5.26}
\]

where \( X(\mu) := \arg \min_{x \in X} \mu(x) \). Since \( V(\cdot) \) is Lipschitz continuous, directional differentiability in the Hadamard sense follows (see Proposition 7.57). As it was discussed above, we also have here under assumptions (A1) and (A2) and since the sample is iid that \( N^{1/2}(\hat{f}_N - f) \) converges in distribution to the random element \( Y \) of \( C(X) \). Noting that \( \hat{\vartheta}_N = V(\hat{f}_N) \), \( \vartheta^* = V(f) \) and \( X(f) = S \), and by applying Delta Theorem to the min-function \( V(\cdot) \) at \( \mu := f \) and using (5.26) we obtain (5.24) and that

\[
\hat{\vartheta}_N - \vartheta^* = \inf_{x \in S} \left[ \hat{f}_N(x) - f(x) \right] + o_p(N^{-1/2}). \tag{5.27}
\]

Since \( f(x) = \vartheta^* \) for any \( x \in S \), we have that assertions (5.23) and (5.27) are equivalent. Finally, (5.25) follows from (5.24). \( \square \)

\(^4\)Recall that \( C(X) \) denotes the space of continuous functions equipped with the sup-norm. A random element of \( C(X) \) is a mapping \( Y : \Omega \to C(X) \) from a probability space \( (\Omega, \mathcal{F}, P) \) into \( C(X) \) which is measurable with respect to the Borel sigma algebra of \( C(X) \), i.e., \( Y(x) = Y(x, \omega) \) can be viewed as a random function.
5.1. Statistical Properties of SAA Estimators

Under mild additional conditions (see Remark 29 in section 7.2.7, page 385) it follows from (5.24) that $N^{-1/2}E[\hat{\vartheta} - \vartheta^*]$ tends to $E[\inf_{x \in S} Y(x)]$ as $N \to \infty$, that is

$$
E[\hat{\vartheta}_N] - \vartheta^* = N^{-1/2}E[\inf_{x \in S} Y(x)] + o(N^{-1/2}).
$$

(5.28)

In particular, if $S = \{\bar{x}\}$ is a singleton, then by (5.25) the SAA optimal value $\hat{\vartheta}_N$ has asymptotically normal distribution and, since $E[Y(\bar{x})] = 0$, we obtain that in this case the bias $E[\hat{\vartheta}_N] - \vartheta^*$ is of order $o(N^{-1/2})$. On the other hand, if the true problem has more than one optimal solution, then the right hand side of (5.24) is given by the minimum of a number of random variables. Although each $Y(x)$ has mean zero, their minimum $\inf_{x \in S} Y(x)$ typically has a negative mean if the set $S$ has more than one element. Therefore, if $S$ is not a singleton, then the bias $E[\hat{\vartheta}_N] - \vartheta^*$ typically is strictly less than zero and is of order $O(N^{-1/2})$. Moreover, the bias tends to be bigger the larger the set $S$ is. For a further discussion of the bias issue see Remark 5, page 170, of the next section.

5.1.3 Second Order Asymptotics

Formula (5.23) gives a first order expansion of the SAA optimal value $\hat{\vartheta}_N$. In this section we discuss a second order term in an expansion of $\hat{\vartheta}_N$. It turns out that the second order analysis of $\hat{\vartheta}_N$ is closely related to deriving (first order) asymptotics of optimal solutions of the SAA problem. We assume in this section that the true (expected value) problem (5.1) has unique optimal solution $\bar{x}$ and denote by $\hat{x}_N$ an optimal solution of the corresponding SAA problem. In order to proceed with the second order analysis we need to impose considerably stronger assumptions.

Our analysis is based on the second order Delta Theorem 7.62 and second order perturbation analysis of section 7.1.5. As in section 7.1.5 we consider a convex compact set $U \subset \mathbb{R}^n$ such that $X \subset \text{int}(U)$, and work with the space $W^{1,\infty}(U)$ of Lipschitz continuous functions $\psi: U \to \mathbb{R}$ equipped with the norm

$$
\|\psi\|_{1,U} := \sup_{x \in U} |\psi(x)| + \sup_{x \in U'}\|\nabla \psi(x)\|,
$$

(5.29)

where $U' \subset \text{int}(U)$ is the set of points where $\psi(\cdot)$ is differentiable.

We make the following assumptions about the true problem.

(S1) The function $f(x)$ is Lipschitz continuous on $U$, has unique minimizer $\bar{x}$ over $x \in X$, and is twice continuously differentiable at $\bar{x}$.

(S2) The set $X$ is second order regular at $\bar{x}$.

(S3) The quadratic growth condition (7.70) holds at $\bar{x}$.

Let $K$ be the subset of $W^{1,\infty}(U)$ formed by differentiable at $\bar{x}$ functions. Note that the set $K$ forms a closed (in the norm topology) linear subspace of $W^{1,\infty}(U)$. Assumption (S1) ensures that $f \in K$. In order to ensure that $f_N \in K$ w.p.1, we make the following assumption.
(S4) Function $F(\cdot, \xi)$ is Lipschitz continuous on $U$ and differentiable at $\bar{x}$ for a.e. $\xi \in \Xi$.

We view $\hat{f}_N$ as a random element of $W^{1,\infty}(U)$, and assume, further, that $N^{1/2}(\hat{f}_N - f)$ converges in distribution to a random element $Y$ of $W^{1,\infty}(U)$.

Consider the min-function $V : W^{1,\infty}(U) \to \mathbb{R}$ defined as

$$V(\psi) := \inf_{x \in X} \psi(x), \quad \psi \in W^{1,\infty}(U).$$

By Theorem 7.23, under the above assumptions (S1)–(S3), the min-function $V(\cdot)$ is second order Hadamard directionally differentiable at $f$ tangentially to the set $K$ and we have the following formula for the second order directional derivative in a direction $\delta \in K$:

$$V''(\delta) = \inf_{h \in C(\bar{x})} \{2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 f(\bar{x}) h - s(\nabla f(\bar{x}), T_K^2(\bar{x}, h))\}. \tag{5.30}$$

Here $C(\bar{x})$ is the critical cone of the true problem, $T_K^2(\bar{x}, h)$ is the second order tangent set to $X$ at $\bar{x}$ and $s(\cdot, A)$ denotes the support function of set $A$ (see page 390 for the definition of second order directional derivatives).

Moreover, suppose that the set $X$ is given in the form

$$X := \{x \in \mathbb{R}^n : G(x) \in K\}, \tag{5.31}$$

where $G : \mathbb{R}^n \to \mathbb{R}^m$ is a twice continuously differentiable mapping and $K \subset \mathbb{R}^m$ is a closed convex cone. Then, under Robinson constraint qualification, the optimal value of the right hand side of (5.30) can be written in a dual form (compare with (7.84)), which results in the following formula for the second order directional derivative in a direction $\delta \in K$:

$$V''(\delta) = \inf_{h \in C(\bar{x})} \sup_{\lambda \in \Lambda(\bar{x})} \{2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 L(\bar{x}, \lambda) h - s(\lambda, T(h))\}. \tag{5.32}$$

Here

$$T(h) := T_K^2(G(\bar{x}), [\nabla G(\bar{x})] h), \tag{5.33}$$

and $L(x, \lambda)$ is the Lagrangian and $\Lambda(\bar{x})$ is the set of Lagrange multipliers of the true problem.

**Theorem 5.8.** Suppose that the assumptions (S1)–(S4) hold and $N^{1/2}(\hat{f}_N - f)$ converges in distribution to a random element $Y$ of $W^{1,\infty}(U)$. Then

$$\hat{\theta}_N = \hat{f}_N(\bar{x}) + \frac{1}{2} V''(\hat{f}_N - f) + o_p(N^{-1}), \tag{5.34}$$

and

$$N \left[ \hat{\theta}_N - \hat{f}_N(\bar{x}) \right] \overset{D}{\to} \frac{1}{2} V''(Y). \tag{5.35}$$

Moreover, suppose that for every $\delta \in K$ the problem in the right hand side of (5.30) has unique optimal solution $\hat{h} = h(\delta)$. Then

$$N^{1/2}(\hat{f}_N - \bar{x}) \overset{D}{\to} h(Y). \tag{5.36}$$
5.1. Statistical Properties of SAA Estimators

Proof. By the second order Delta Theorem 7.62 we have that
\[ \hat{\theta}_N = \theta^* + V'_f(\hat{x}_N - f) + \frac{1}{2}V''_f(\hat{x}_N - f) + o_p(N^{-1}), \]
and
\[ N[\hat{\theta}_N - \theta^* - V'_f(\hat{x}_N - f)] \xrightarrow{D} \frac{1}{2}V''_f(Y). \]
We also have (compare with formula (5.26)) that
\[ V'_f(\hat{x}_N - f) = \hat{x}_N - f(\bar{x}) = \hat{x}_N - \theta^*, \]
and hence (5.34) and (5.35) follow.

Now consider a (measurable) mapping \( \bar{x} : W^{1,\infty}(U) \to \mathbb{R}^n \), such that
\[ \bar{x}(\psi) \in \arg \min_{\psi \in X} \psi(x), \quad \psi \in W^{1,\infty}(U). \]
We have that \( \bar{x}(f) = \bar{x} \) and by equation (7.82) of Theorem 7.23, we have that \( \bar{x}(\cdot) \) is Hadamard directionally differentiable at \( f \) tangentially to \( K \), and for \( \delta \in K \) the directional derivative \( \bar{x}'(f, \delta) \) is equal to the optimal solution in the right hand side of (5.30), provided that it is unique. By applying Delta Theorem 7.61 this completes the proof of (5.36).

One of the difficulties in applying the above theorem is verification of convergence in distribution of \( N^{1/2}(\hat{x}_N - f) \) in the space \( W^{1,\infty}(X) \). Actually it could be easier to prove asymptotic results (5.34) – (5.36) by direct methods. Note that formulas (5.30) and (5.32), for the second order directional derivatives \( V''_f(\hat{x}_N - f) \), involve statistical properties of \( \hat{x}_N(x) \) only at the (fixed) point \( \bar{x} \).

Let \( Z \) be a random vector having normal distribution, \( Z \sim N(0, \Sigma) \), with covariance matrix \( \Sigma \) defined in (5.37), and let the set \( X \) be given in the form (5.31). Then by the above discussion and formula (5.32), we have that under appropriate regularity conditions,
\[ N[\hat{\theta}_N - \hat{x}_N(\bar{x})] \xrightarrow{D} \frac{1}{2}v(Z), \quad (5.38) \]
where \( v(Z) \) is the optimal value of the problem:
\[ \min_{h \in C(x)} \sup_{\lambda \in \Lambda(x)} \{ 2h^TZ + h^T\nabla^2_{xx} L(\bar{x}, \lambda) h - s(\lambda, T(h)) \}, \quad (5.39) \]
with \( T(h) \) being the second order tangent set defined in (5.33). Moreover, if for all \( Z \) problem (5.39) possesses unique optimal solution \( h = h(Z) \), then
\[ N^{1/2}(\hat{x}_N - \bar{x}) \xrightarrow{D} h(Z). \quad (5.40) \]
Remark 5. Note that \( \mathbb{E}[\hat{f}_N(\bar{x})] = f(\bar{x}) = \vartheta^* \). Therefore, under the respective regularity conditions, in particular under the assumption that the true problem has unique optimal solution \( \bar{x} \), we have by (5.38) that the expected value of the term \( \frac{1}{2}N^{-1}w(Z) \) can be viewed as the asymptotic bias of \( \hat{\vartheta}_N \). This asymptotic bias is of order \( O(N^{-1}) \). This can be compared with formula (5.28) for the asymptotic bias of order \( O(N^{-1/2}) \) when the set of optimal solutions of the true problem is not a singleton. Note also that \( \sigma(\cdot) \) is nonpositive; in order to see that just take \( h = 0 \) in (5.39).

As an example, consider the case where the set \( X \) is defined by a finite number of constraints:

\[
X := \{ x \in \mathbb{R}^n : g_i(x) = 0, \ i = 1, ..., q, \ g_i(x) \leq 0, \ i = q + 1, ..., p \}, \tag{5.41}
\]

with the functions \( g_i(x) \), \( i = 1, ..., p \), being twice continuously differentiable. This is a particular form of (5.31) with \( G(x) := (g_1(x), ..., g_p(x)) \) and \( K := \{0_q\} \times \mathbb{R}_-^{p-q} \). Denote

\[
I(x) := \{ i : g_i(\bar{x}) = 0, \ i = q + 1, ..., p \}
\]

the index set of active at \( \bar{x} \) inequality constraints. Suppose that the linear independence constraint qualification (LICQ) holds at \( \bar{x} \), i.e., the gradient vectors \( \nabla g_i(\bar{x}) \), \( i \in \{1, ..., q\} \cup I(\bar{x}) \), are linearly independent. Then the corresponding set of Lagrange multipliers is a singleton, \( \Lambda(\bar{x}) = \{ \lambda \} \). In that case

\[
C(\bar{x}) = \{ h : h^T \nabla g_i(\bar{x}) = 0, \ i \in \{1, ..., q\} \cup I_+(\bar{\lambda}), \ h^T \nabla g_i(\bar{x}) \leq 0, i \in I_0(\bar{\lambda}) \},
\]

where

\[
I_0(\bar{\lambda}) := \{ i \in I(\bar{x}) : \bar{\lambda}_i = 0 \} \quad \text{and} \quad I_+(\bar{\lambda}) := \{ i \in I(\bar{x}) : \bar{\lambda}_i > 0 \}.
\]

Consequently problem (5.39) takes the form

\[
\begin{align*}
\min_{h \in \mathbb{R}^n} & \quad 2h^T Z + h^T \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) h \\
\text{s.t.} & \quad h^T \nabla g_i(\bar{x}) = 0, \ i \in \{1, ..., q\} \cup I_+(\bar{\lambda}), \ h^T \nabla g_i(\bar{x}) \leq 0, i \in I_0(\bar{\lambda}).
\end{align*} \tag{5.42}
\]

This is a quadratic programming problem. The linear independence constraint qualification implies that problem (5.42) has a unique vector \( \alpha(Z) \) of Lagrange multipliers, and that it has a unique optimal solution \( h(Z) \) if the Hessian matrix \( H := \nabla^2_{xx} L(\bar{x}, \bar{\lambda}) \) is positive definite over the linear space defined by the first \( q + |I_+(\bar{\lambda})| \) (equality) linear constraints in (5.42).

If, furthermore, the strict complementarity condition holds, i.e., \( \bar{\lambda}_i > 0 \) for all \( i \in I_+(\bar{\lambda}) \), or in other words \( I_0(\bar{\lambda}) = \emptyset \), then \( h = h(Z) \) and \( \alpha = \alpha(Z) \) can be obtained as solutions of the following system of linear equations

\[
\begin{bmatrix}
H \\ A^T \\
0
\end{bmatrix}
\begin{bmatrix}
h \\ \alpha
\end{bmatrix} =
\begin{bmatrix}
Z \\ 0
\end{bmatrix}. \tag{5.43}
\]
5.1. Statistical Properties of SAA Estimators

Here \( H = \nabla^2_{x,x}L(\bar{x}, \bar{\lambda}) \) and \( A \) is the \( n \times (q + |I(\bar{x})|) \) matrix whose columns are formed by vectors \( \nabla g_i(\bar{x}), i \in \{1, \ldots, q\} \cup I(\bar{x}) \). Then

\[
N^{1/2} \begin{bmatrix} \hat{x}_N - \bar{x} \\ \hat{\lambda}_N - \bar{\lambda} \end{bmatrix} \overset{D}{\to} \mathcal{N}(0, J^{-1} \Upsilon J^{-1}),
\]

(5.44)

where

\[
J := \begin{bmatrix} H & A \\ A^T & 0 \end{bmatrix} \quad \text{and} \quad \Upsilon := \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix},
\]

provided that the matrix \( J \) is nonsingular.

Under the linear independence constraint qualification and strict complementarity condition, we have by the second order necessary conditions that the Hessian matrix \( H = \nabla^2_{x,x}L(\bar{x}, \bar{\lambda}) \) is positive semidefinite over the linear space \( \{ h : A^T h = 0 \} \). Note that this linear space coincides here with the critical cone \( C(\bar{x}) \). It follows that the matrix \( J \) is nonsingular iff \( H \) is positive definite over this linear space. That is, here the nonsingularity of the matrix \( J \) is equivalent to the second order sufficient conditions at \( \bar{x} \).

**Remark 6.** As it was mentioned earlier, the curvature term \( s(\lambda, T(h)) \) in the auxiliary problem (5.39) vanishes if the cone \( K \) is polyhedral. In particular, this happens if \( K = \{0_q\} \times \mathbb{R}^{n-q} \), and hence the feasible set \( X \) is given in the form (5.41). This curvature term can be also written in an explicit form for some nonpolyhedral cones; in particular for the cone of positive semidefinite matrices (see [22, section 5.3.6]).

### 5.1.4 Minimax Stochastic Programs

Sometimes it is worthwhile to consider minimax stochastic programs of the form

\[
\min_{x \in X} \sup_{y \in Y} \left\{ f(x, y) := \mathbb{E}[F(x, y, \xi)] \right\},
\]

(5.45)

where \( X \subseteq \mathbb{R}^n \) and \( Y \subseteq \mathbb{R}^m \) are closed sets, \( F : X \times Y \times \Xi \to \mathbb{R} \) and \( \xi = \xi(\omega) \) is a random vector whose probability distribution is supported on set \( \Xi \subseteq \mathbb{R}^d \). The corresponding SAA problem is obtained by using the sample average as an approximation of the expectation \( f(x, y) \), that is

\[
\min_{x \in X} \sup_{y \in Y} \left\{ \hat{f}_N(x, y) := \frac{1}{N} \sum_{j=1}^{N} F(x, y, \xi_j) \right\}.
\]

(5.46)

As before denote by \( \vartheta^* \) and \( \hat{\vartheta}_N \) the optimal values of (5.45) and (5.46), respectively, and by \( S_\vartheta \subset X \) and \( S_{\vartheta,N} \subset X \) the respective sets of optimal solutions. Recall that \( F(x, y, \xi) \) is said to be a Carathéodory function, if \( F(x, y, \xi(\cdot)) \) is measurable for every \((x, y)\) and \( F(\cdot, \cdot, \xi) \) is continuous for a.e. \( \xi \in \Xi \). We make the following assumptions.

(A/1) \( F(x, y, \xi) \) is a Carathéodory function.

(A/2) The sets \( X \) and \( Y \) are nonempty and compact.
Chapter 5. Statistical Inference

(A'3) \( F(x, y, \xi) \) is dominated by an integrable function, i.e., there is an open set \( N \subset \mathbb{R}^{n+m} \), containing the set \( X \times Y \), and an integrable, with respect to the probability distribution of the random vector \( \xi \), function \( h(\xi) \) such that \( |F(x, y, \xi)| \leq h(\xi) \) for all \( (x, y) \in N \) and a.e. \( \xi \in \Xi \).

By Theorem 7.43 it follows that the expected value function \( f(x, y) \) is continuous on \( X \times Y \). Since \( Y \) is compact, this implies that the max-function

\[
\phi(x) := \sup_{y \in Y} f(x, y)
\]

is continuous on \( X \). It also follows that the function \( \hat{f}_N(x, y) = \hat{f}_N(x, y, \omega) \) is a Carathéodory function. Consequently, the sample average max-function

\[
\hat{\phi}_N(x, \omega) := \sup_{y \in Y} \hat{f}_N(x, y, \omega)
\]

is a Carathéodory function. Since \( \hat{\theta}_N = \hat{\theta}_N(\omega) \) is given by the minimum of the Carathéodory function \( \hat{\phi}_N(x, \omega) \), it follows that it is measurable.

Theorem 5.9. Suppose that assumptions (A'1) - (A'3) hold and the sample is iid. Then \( \hat{\theta}_N \rightarrow \theta^* \) and \( D(\hat{S}_{x,N}, S_x) \rightarrow 0 \) w.p.1 as \( N \rightarrow \infty \).

Proof. By Theorem 7.48 we have that, under the specified assumptions, \( \hat{f}_N(x, y) \) converges to \( f(x, y) \) w.p.1 uniformly on \( X \times Y \). That is, \( \Delta_N \rightarrow 0 \) w.p.1 as \( N \rightarrow \infty \), where

\[
\Delta_N := \sup_{(x,y) \in X \times Y} \left| \hat{f}_N(x, y) - f(x, y) \right|.
\]

Consider \( \hat{\phi}_N(x) := \sup_{y \in Y} \hat{f}_N(x, y) \) and \( \phi(x) := \sup_{y \in Y} f(x, y) \). We have that

\[
\sup_{x \in X} \left| \hat{\phi}_N(x) - \phi(x) \right| \leq \Delta_N,
\]

and hence \( |\hat{\theta}_N - \theta^*| \leq \Delta_N \). It follows that \( \hat{\theta}_N \rightarrow \theta^* \) w.p.1.

The function \( \phi(x) \) is continuous and \( \hat{\phi}_N(x) \) is continuous w.p.1. Consequently, the set \( S_x \) is nonempty and \( \hat{S}_{x,N} \) is nonempty w.p.1. Now to prove that \( D(\hat{S}_{x,N}, S_x) \rightarrow 0 \) w.p.1, one can proceed exactly in the same way as in the proof of Theorem 5.3.

We discuss now asymptotics of \( \hat{\theta}_N \) in the convex-concave case. We make the following additional assumptions.

(A'4) The sets \( X \) and \( Y \) are convex, and for a.e. \( \xi \in \Xi \) the function \( F(\cdot, \cdot, \xi) \) is convex-concave on \( X \times Y \), i.e., the function \( F(\cdot, y, \xi) \) is convex on \( X \) for every \( y \in Y \), and the function \( F(x, \cdot, \xi) \) is concave on \( Y \) for every \( x \in X \).

It follows that the expected value function \( f(x, y) \) is convex concave and continuous on \( X \times Y \). Consequently, problem (5.45) and its dual

\[
\max_{y \in Y} \inf_{x \in X} f(x, y)
\]

(5.47)
5.1. Statistical Properties of SAA Estimators

have nonempty and bounded sets of optimal solutions $S_x \subset X$ and $S_y \subset Y$, respectively. Moreover, the optimal values of problems (5.45) and (5.47) are equal to each other and $S_x \times S_y$ forms the set of saddle points of these problems.

(A'5) For some point $(x, y) \in X \times Y$, the expectation $\mathbb{E}[F(x, y, \xi)^2]$ is finite, and there exists a measurable function $C : \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[C(\xi)^2]$ is finite and the inequality

$$|F(x, y, \xi) - F(x', y', \xi)| \leq C(\xi) \left(\|x - x'\| + \|y - y'\|\right) \quad (5.48)$$

holds for all $(x, y), (x', y') \in X \times Y$ and a.e. $\xi \in \Xi$.

The above assumption implies that $f(x, y)$ is Lipschitz continuous on $X \times Y$ with Lipschitz constant $\kappa = \mathbb{E}[C(\xi)]$.

**Theorem 5.10.** Consider the minimax stochastic problem (5.45) and the SAA problem (5.46) based on an iid sample. Suppose that assumptions (A'1) - (A'2) and (A'4) - (A'5) hold. Then

$$\hat{N} = \inf_{x \in S_x} \sup_{y \in S_y} f_N(x, y) + o_p(N^{-1/2}). \quad (5.49)$$

Moreover, if the sets $S_x = \{\bar{x}\}$ and $S_y = \{\bar{y}\}$ are singletons, then $N^{1/2}(\hat{N} - \hat{\vartheta})$ converges in distribution to normal with zero mean and variance $\sigma^2 = \text{Var}[F(\bar{x}, \bar{y}, \xi)]$.

**Proof.** Consider the space $C(X, Y)$, of continuous functions $\psi : X \times Y \to \mathbb{R}$ equipped with the sup-norm $\|\psi\| = \sup_{x \in X, y \in Y} |\psi(x, y)|$, and set $K \subset C(X, Y)$ formed by convex-concave on $X \times Y$ functions. It is not difficult to see that the set $K$ is a closed (in the norm topology of $C(X, Y)$) and convex cone. Consider the optimal value function $V : C(X, Y) \to \mathbb{R}$ defined as

$$V(\psi) := \inf_{x \in X} \sup_{y \in Y} \psi(x, y), \quad \text{for } \psi \in C(X, Y). \quad (5.50)$$

Recall that it is said that $V(\cdot)$ is Hadamard directionally differentiable at $f \in K$, tangentially to the set $K$, if the following limit exists for any $\gamma \in T_K(f)$:

$$V'_f(\gamma) := \lim_{\epsilon \to 0} \frac{V(f + \epsilon \gamma) - V(f)}{\epsilon}. \quad (5.51)$$

By Theorem 7.24 we have that the optimal value function $V(\cdot)$ is Hadamard directionally differentiable at $f$ tangentially to the set $K$ and

$$V'_f(\gamma) = \inf_{x \in S_x} \sup_{y \in S_y} \gamma(x, y), \quad (5.52)$$

for any $\gamma \in T_K(f)$.

By the assumption (A'5) we have that $N^{1/2}(\hat{N} - f)$, considered as a sequence of random elements of $C(X, Y)$, converges in distribution to a random element of $C(X, Y)$. Then by noting that $\vartheta^* = f(x^*, y^*)$ for any $(x^*, y^*) \in S_x \times S_y$ and using Hadamard

$$\hat{N} = \inf_{x \in S_x} \sup_{y \in S_y} f_N(x, y) + o_p(N^{-1/2}),$$

Moreover, if the sets $S_x = \{\bar{x}\}$ and $S_y = \{\bar{y}\}$ are singletons, then $N^{1/2}(\hat{N} - \hat{\vartheta})$ converges in distribution to normal with zero mean and variance $\sigma^2 = \text{Var}[F(\bar{x}, \bar{y}, \xi)]$.
directional differentiability of the optimal value function, tangentially to the set $K$, together with formula (5.52) and a version of the Delta method given in Theorem 7.61, we can complete the proof. □

Suppose now that the feasible set $X$ is defined by constraints in the form (5.11). The Lagrangian function of the true problem is

$$L(x, \lambda) := f(x) + \sum_{i=1}^{p} \lambda_i g_i(x).$$

Suppose also that the problem is convex, that is the set $X_0$ is convex and for all $\xi \in \Xi$ the functions $F(\cdot, \xi)$ and $G_i(\cdot, \xi), i = 1, \ldots, p$, are convex. Suppose, further, that the functions $f(x)$ and $g_i(x)$ are finite valued on a neighborhood of the set $S$ (of optimal solutions of the true problem) and the Slater condition holds. Then with every optimal solution $\bar{x} \in S$ is associated a nonempty and bounded set $\Lambda$ of Lagrange multipliers vectors $\lambda = (\lambda_1, \ldots, \lambda_p)$ satisfying the optimality conditions:

$$\bar{x} \in \arg \min_{x \in X_0} L(x, \lambda), \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i g_i(\bar{x}) = 0, \quad i = 1, \ldots, p.$$  (5.53)

The set $\Lambda$ coincides with the set of optimal solutions of the dual of the true problem, and therefore is the same for any optimal solution $\bar{x} \in S$.

Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem (5.10) with $X_N$ given in the form (5.13). We assume that conditions (A1) and (A2), formulated on page 165, are satisfied for the integrands $F$ and $G_i, i = 1, \ldots, p$, i.e., finiteness of the corresponding second order moments and the Lipschitz continuity condition of assumption (A2) hold for each function. It follows that the functions $f(x)$ and $g_i(x)$ are finite valued and continuous on $X$. As in Theorem 5.7, we denote by $Y(x)$ random variables which are normally distributed and have the same covariance structure as $F(x, \xi)$. We also denote by $Y_i(x)$ random variables which are normally distributed and have the same covariance structure as $G_i(x, \xi), i = 1, \ldots, p$.

**Theorem 5.11.** Let $\hat{\vartheta}_N$ be the optimal value of the SAA problem (5.10) with $X_N$ given in the form (5.13). Suppose that the sample is iid, the problem is convex and the following conditions are satisfied: (i) the set $S$, of optimal solutions of the true problem, is nonempty and bounded, (ii) the functions $f(x)$ and $g_i(x)$ are finite valued on a neighborhood of $S$, (iii) the Slater condition for the true problem holds, (iv) the assumptions (A1) and (A2) hold for the integrands $F$ and $G_i, i = 1, \ldots, p$. Then

$$N^{1/2} \left( \hat{\vartheta}_N - \vartheta^* \right) \overset{D}{\to} \inf_{x \in S} \sup_{\lambda \in \Lambda} \left[ Y(x) + \sum_{i=1}^{p} \lambda_i Y_i(x) \right].$$  (5.54)

If, moreover, $S = \{ \bar{x} \}$ and $\Lambda = \{ \bar{\lambda} \}$ are singletons, then

$$N^{1/2} \left( \hat{\vartheta}_N - \vartheta^* \right) \overset{D}{\to} \mathcal{N}(0, \sigma^2),$$  (5.55)

with $\sigma^2 := \text{Var} \left[ F(\bar{x}, \xi) + \sum_{i=1}^{p} \bar{\lambda}_i G_i(\bar{x}, \xi) \right]$. 

5.2. Stochastic Generalized Equations

Proof. Since the problem is convex and the Slater condition (for the true problem) holds, we have that \( \partial^* \) is equal to the optimal value of the (Lagrangian) dual

\[
\text{Max} \inf_{\lambda \geq 0} L(x, \lambda),
\]

and the set of optimal solutions of (5.56) is nonempty and compact and coincides with the set of Lagrange multipliers \( \Lambda \). Since the problem is convex and \( S \) is nonempty and bounded, the problem can be considered on a bounded neighborhood of \( S \), i.e., without loss of generality it can be assumed that the set \( X \) is compact. The proof can be now completed by applying Theorem 5.10.

5.2 Stochastic Generalized Equations

In this section we discuss the following so-called stochastic generalized equations. Consider a random vector \( \xi \), whose distribution is supported on a set \( \Xi \subset \mathbb{R}^n \), a mapping \( \Phi : \mathbb{R}^n \times \Xi \to \mathbb{R}^n \) and a multifunction \( \Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n \). Suppose that the expectation \( \phi(x) := \mathbb{E}[\Phi(x, \xi)] \) is well defined and finite valued. We refer to

\[
\phi(x) \in \Gamma(x)
\]

as true, or expected value, generalized equation and say that a point \( \bar{x} \in \mathbb{R}^n \) is a solution of (5.57) if \( \phi(\bar{x}) \in \Gamma(\bar{x}) \).

The above abstract setting includes the following cases. If \( \Gamma(x) = \{0\} \) for every \( x \in \mathbb{R}^n \), then (5.57) becomes the ordinary equation \( \phi(x) = 0 \). As another example, let \( \Gamma(\cdot) := \mathcal{N}_X(\cdot) \), where \( X \) is a nonempty closed convex subset of \( \mathbb{R}^n \) and \( \mathcal{N}_X(x) \) denotes the (outwards) normal cone to \( X \) at \( x \). Recall that, by the definition, \( \mathcal{N}_X(x) = \emptyset \) if \( x \notin X \). In that case \( \bar{x} \) is a solution of (5.57) iff \( \bar{x} \in X \) and the following, so-called variational inequality, holds

\[
(x - \bar{x})^T \phi(x) \leq 0, \quad \forall x \in X.
\]

Since the mapping \( \phi(x) \) is given in the form of the expectation, we refer to such variational inequalities as stochastic variational inequalities. Note that if \( X = \mathbb{R}^n \), then \( \mathcal{N}_X(x) = \{0\} \) for any \( x \in \mathbb{R}^n \), and hence in that case the above variational inequality is reduced to the equation \( \phi(x) = 0 \). Let us also remark that if \( \Phi(x, \xi) := -\nabla_x F(x, \xi) \), for some real valued function \( F(x, \xi) \), and the interchangeability formula \( \mathbb{E}[\nabla_x F(x, \xi)] = \nabla f(x) \) holds, i.e., \( \phi(x) = -\nabla f(x) \) where \( f(x) := \mathbb{E}[F(x, \xi)] \), then (5.58) represents first order necessary, and if \( f(x) \) is convex, sufficient conditions for \( \bar{x} \) to be an optimal solution for the optimization problem (5.1).

If the feasible set \( X \) of the optimization problem (5.1) is defined by constraints in the form

\[
X := \{x \in \mathbb{R}^n : g_i(x) = 0, \; i = 1, ..., q, \; g_i(x) \leq 0, \; i = q + 1, ..., p\},
\]

with \( g_i(x) := \mathbb{E}[G_i(x, \xi)], \; i = 1, ..., p \), then the corresponding first-order (KKT) optimality conditions can be written in a form of variational inequality. That is, let \( z := (x, \lambda) \in \mathbb{R}^{n+p} \) and

\[
L(z, \xi) := F(x, \xi) + \sum_{i=1}^{p} \lambda_i G_i(x, \xi), \quad \ell(z) := \mathbb{E}[L(z, \xi)] = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x)
\]
be the corresponding Lagrangians. Define

$$\Phi(z, \xi) := \begin{bmatrix} \nabla_x L(z, \xi) \\ G_1(x; \xi) \\ \vdots \\ G_p(x; \xi) \end{bmatrix}$$

and $$\Gamma(z) := \mathcal{N}_K(z),$$

(5.60)

where $$K := \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{p-q}_+ \subset \mathbb{R}^{n+p}.$$ Note that if $$z \in K,$$ then

$$\mathcal{N}_K(z) = \{(v, \gamma) \in \mathbb{R}^{n+p} : v = 0 \text{ and } \gamma_i = 0, \ i = 1, \ldots, q, \ \gamma_i = 0, \ i \in \mathcal{I}_0(\lambda), \ \gamma_i \leq 0, \ i \in \mathcal{I}_+(\lambda)\}, \quad (5.61)$$

and $$\mathcal{N}_K(z) = \emptyset$$ if $$z \notin K.$$ Consequently, assuming that the interchangeability formula holds, and hence

$$\mathbb{E}[\nabla_x L(z, \xi)] = \nabla \ell_x(z) = \nabla f(x) + \sum_{i=1}^p \lambda_i \nabla g_i(x),$$

we have that

$$\phi(z) := \mathbb{E}[\Phi(z, \xi)] = \begin{bmatrix} \nabla_x \ell(z) \\ g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}$$

(5.63)

and variational inequality $$\phi(z) \in \mathcal{N}_K(z)$$ represents the KKT optimality conditions for the true optimization problem.

We make the following assumption about the multifunction $$\Gamma(x).$$

(E1) The multifunction $$\Gamma(x)$$ is closed, that is, the following holds: if $$x_k \to x,$$ $$y_k \in \Gamma(x_k)$$ and $$y_k \to y,$$ then $$y \in \Gamma(x).$$

The above assumption implies that the multifunction $$\Gamma(x)$$ is closed valued, i.e., for any $$x \in \mathbb{R}^n$$ the set $$\Gamma(x)$$ is closed. For variational inequalities assumption (E1) always holds, i.e., the multifunction $$x \mapsto \mathcal{N}_X(x)$$ is closed.

Now let $$\xi_1, \ldots, \xi^N$$ be a random sample of $$N$$ realizations of the random vector $$\xi,$$ and $$\hat{\phi}_N(x) := N^{-1} \sum_{j=1}^N \Phi(x, \xi^j)$$ be the corresponding sample average estimate of $$\phi(x).$$ We refer to

$$\hat{\phi}_N(x) \in \Gamma(x)$$

(5.64)

as the SAA generalized equation. There are standard numerical algorithms for solving nonlinear equations which can be applied to (5.64) in the case $$\Gamma(x) \equiv \{0\},$$ i.e., when (5.64) is reduced to the ordinary equation $$\phi_N(x) = 0.$$ There are also numerical procedures for solving variational inequalities. We are not going to discuss such numerical algorithms but rather concentrate on statistical properties of solutions of SAA equations. We denote by $$\mathcal{S}$$ and $$\hat{\mathcal{S}}_N$$ the sets of (all) solutions of the true (5.57) and SAA (5.64) generalized equations, respectively.
5.2.1 Consistency of Solutions of the SAA Generalized Equations

In this section we discuss convergence properties of the SAA solutions.

**Theorem 5.12.** Let $C$ be a compact subset of $\mathbb{R}^n$ such that $S \subset C$. Suppose that: (i) the multifunction $\Gamma(x)$ is closed (assumption (E1)), (ii) the mapping $\phi(x)$ is continuous on $C$, (iii) w.p.1 for $N$ large enough the set $\hat{S}_N$ is nonempty and $\hat{S}_N \subset C$, (iv) $\hat{\phi}_N(x)$ converges to $\phi(x)$ w.p.1 uniformly on $C$ as $N \to \infty$. Then $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.

**Proof.** The above result basically is deterministic in the sense that if we view $\hat{\phi}_N(x) = \phi_N(x, \omega)$ as defined on a common probability space, then it should be verified for a.e. $\omega$. Therefore we omit saying “w.p.1”. Consider a sequence $\hat{x}_N \in \hat{S}_N$. Because of the assumption (iii), by passing to a subsequence if necessary, we only need to show that if $\hat{x}_N \to \hat{x}^*$, then $\hat{x}_N \subset S$ (compare with the proof of Theorem 5.3). Now since it is assumed that $\phi(\cdot)$ is continuous and $\phi_N(x)$ converges to $\phi(x)$ uniformly, it follows that $\hat{\phi}_N(\hat{x}_N) \to \phi(\hat{x}^*)$ (see Proposition 5.1). Since $\hat{\phi}_N(\hat{x}_N) \in \Gamma(\hat{x}_N)$, it follows by assumption (E1) that $\phi(\hat{x}^*) \in \Gamma(\hat{x}^*)$, which completes the proof. \qed

A few remarks about the assumptions involved in the above consistency result are now in order. By Theorem 7.48 we have that, in the case of iid sampling, the assumptions (ii) and (iv) of the above proposition are satisfied for any compact set $C$ if the following assumption holds.

**(E2)** For every $\xi \in \Xi$ the function $\Phi(\cdot, \xi)$ is continuous on $C$ and $\|\Phi(x, \xi)\|_{x \in C}$ is dominated by an integrable function.

There are two parts to the assumption (iii) of Theorem 5.12, namely, that the SAA generalized equations do not have a solution which escapes to infinity, and that they possess at least one solution w.p.1 for $N$ large enough. The first of these assumptions can be often verified by ad hoc methods. The second assumption is more subtle. We are going to discuss it next. The following concept of strong regularity is due to Robinson [170].

**Definition 5.13.** Suppose that the mapping $\phi(x)$ is continuously differentiable. We say that a solution $\hat{x} \in S$ is strongly regular if there exist neighborhoods $\mathcal{N}_1$ and $\mathcal{N}_2$ of $0 \in \mathbb{R}^n$ and $\hat{x}$, respectively, such that for every $\delta \in \mathcal{N}_1$ the following (linearized) generalized equation

$$\delta + \phi'(\hat{x})\delta \in \Gamma'(\hat{x})$$

(5.65)

has a unique solution in $\mathcal{N}_2$, denoted $\hat{x} = \hat{x}(\delta)$, and $\hat{x}(\cdot)$ is Lipschitz continuous on $\mathcal{N}_1$.

Note that it follows from the above conditions that $\hat{x}(0) = \hat{x}$. In case $\Gamma'(x) \equiv \{0\}$, strong regularity simply means that the $n \times n$ Jacobian matrix $J := \nabla \phi(x)$ is invertible, or in other words nonsingular. Also in the case of variational inequalities, the strong regularity condition was investigated extensively, we discuss this later.

Let $\mathcal{V}$ be a convex compact neighborhood of $\hat{x}$, i.e., $\hat{x} \in \text{int}(\mathcal{V})$. Consider the space...
Let \( \hat{x}(u) = \tilde{x}(u(x) - \phi(\tilde{x})) + o(\|u - \phi\|_{1, V}) \). \hspace{1cm} (5.66)

Clearly, we have that \( \hat{x}(\phi) = \tilde{x} \) and \( \hat{x}(\hat{x}_N) \) is a solution, in a neighborhood of \( \tilde{x} \), of the SAA generalized equation provided that \( \|\hat{x}_N - \phi\|_{1, V} \leq \epsilon \). Therefore, by employing the above results for the mapping \( u(\cdot) := \hat{x}_N(\cdot) \) we immediately obtain the following.

**Theorem 5.14.** Let \( \tilde{x} \) be a strongly regular solution of the true generalized equation (5.57), and suppose that \( \phi(x) \) and \( \hat{x}_N(x) \) are continuously differentiable in a neighborhood \( V \) of \( \tilde{x} \) and \( \|\hat{x}_N - \phi\|_{1, V} \rightarrow 0 \) w.p.1 as \( N \rightarrow \infty \). Then w.p.1 for \( N \) large enough the SAA generalized equation (5.64) possesses a unique solution \( \hat{x}_N \) in a neighborhood of \( \tilde{x} \), and \( \hat{x}_N \rightarrow \tilde{x} \) w.p.1 as \( N \rightarrow \infty \).

The assumption that \( \|\hat{x}_N - \phi\|_{1, V} \rightarrow 0 \) w.p.1, in the above theorem, means that \( \hat{x}_N(x) \) and \( \nabla \hat{x}_N(x) \) converge w.p.1 to \( \phi(x) \) and \( \nabla \phi(x) \), respectively, uniformly on \( V \). By Theorem 7.48, in the case of iid sampling this is ensured by the following assumption.

(E3) For a.e. \( \xi \) the mapping \( \Phi(\cdot, \xi) \) is continuously differentiable on \( V \), and \( \|\Phi(x, \xi)\|_{x \in V} \) and \( \|\nabla_x \Phi(x, \xi)\|_{x \in V} \) are dominated by an integrable function.

Note that the assumption that \( \Phi(\cdot, \xi) \) is continuously differentiable on a neighborhood of \( \tilde{x} \) is essential in the above analysis. By combining Theorems 5.12 and 5.14 we obtain the following result.

**Theorem 5.15.** Let \( C \) be a compact subset of \( \mathbb{R}^n \) and \( \tilde{x} \) be a unique in \( C \) solution of the true generalized equation (5.57). Suppose that: (i) the multifunction \( \Gamma(x) \) is closed (assumption (E1)), (ii) for a.e. \( \xi \) the mapping \( \Phi(\cdot, \xi) \) is continuously differentiable on \( C \), and \( \|\Phi(x, \xi)\|_{x \in C} \) and \( \|\nabla_x \Phi(x, \xi)\|_{x \in C} \) are dominated by an integrable function, (iii) the solution \( \tilde{x} \) is strongly regular, (iv) \( \hat{x}_N(x) \) and \( \nabla \hat{x}_N(x) \) converge w.p.1 to \( \phi(x) \) and \( \nabla \phi(x) \), respectively, uniformly on \( C \). Then w.p.1 for \( N \) large enough the SAA generalized equation possesses unique in \( C \) solution \( \hat{x}_N \) converging to \( \tilde{x} \) w.p.1 as \( N \rightarrow \infty \).

Note again that if the sample is iid, then the assumption (iv) in the above theorem is implied by the assumption (ii) and hence is redundant.
5.2.2 Asymptotics of SAA Generalized Equations Estimators

By using the first order approximation (5.66) it is also possible to derive asymptotics of \( \hat{x}_N \). Suppose for the moment that \( \Gamma(\hat{x}) \equiv \{0\} \). Then strong regularity means that the Jacobian matrix
\[
J := \nabla \phi(\hat{x})
\]
is nonsingular, and \( \hat{x}(\delta) \) is the solution of the corresponding linear equations and hence can be written in the form
\[
\hat{x}(\delta) = \hat{x} - J^{-1}\delta.
\]  
(5.67)

By using (5.67) and (5.66) with \( u(\cdot) := \hat{\phi}_N(\cdot) \), we obtain under certain regularity conditions which ensure that the remainder in (5.66) is of order \( o_p(N^{-1/2}) \), that
\[
N^{1/2}(\hat{x}_N - \hat{x}) = -J^{-1}Y_N + o_p(1),
\]  
(5.68)

where \( Y_N := N^{1/2}[\hat{\phi}_N(\hat{x}) - \phi(\hat{x})] \). Moreover, in the case of iid sample, we have by the CLT that \( Y_N \overset{D}{\to} N(0, \Sigma) \), where \( \Sigma \) is the covariance matrix of the random vector \( \Phi(\hat{x}, \xi) \). Consequently, \( \hat{x}_N \) has asymptotically normal distribution with mean vector \( \hat{x} \) and the covariance matrix \( N^{-1}J^{-1}\Sigma J^{-1} \).

Suppose now that \( \Gamma(\cdot) := N_X(\cdot) \), with the set \( X \) being nonempty closed convex and polyhedral, and let \( \hat{x} \) be a strongly regular solution of (5.57). Let \( \hat{x}(\delta) \) be the (unique) solution, of the corresponding linearized variational inequality (5.65), in a neighborhood of \( \hat{x} \). Consider the cone
\[
C_X(\hat{x}) := \{y \in T_X(\hat{x}) : y^T \phi(\hat{x}) = 0\},
\]  
(5.69)
called the critical cone, and the Jacobian matrix \( J := \nabla \phi(\hat{x}) \). Then for all \( \delta \) sufficiently close to \( 0 \in \mathbb{R}^n \), we have that \( \hat{x}(\delta) - \hat{x} \) coincides with the solution \( \hat{d}(\delta) \) of the variational inequality
\[
\delta + Jd \in N_{C_X(\hat{x})}(d).
\]  
(5.70)

Note that the mapping \( \hat{d}(\cdot) \) is positively homogeneous, i.e., for any \( \delta \in \mathbb{R}^n \) and \( t \geq 0 \), it follows that \( \hat{d}(t\delta) = t\hat{d}(\delta) \). Consequently, under the assumption that the solution \( \hat{x} \) is strongly regular, we obtain by (5.66) that \( \hat{d}(\cdot) \) is the directional derivative of \( \hat{x}(u) \), at \( u = \phi \), in the Hadamard sense. Therefore, under appropriate regularity conditions ensuring functional CLT for \( N^{1/2}(\hat{\phi}_N - \phi) \) in the space \( C(\mathcal{Y}, \mathbb{R}^n) \), it follows by the Delta Theorem that
\[
N^{1/2}(\hat{x}_N - \hat{x}) \overset{D}{\to} \hat{d}(Y),
\]  
(5.71)

where \( Y \sim N(0, \Sigma) \) and \( \Sigma \) is the covariance matrix of \( \Phi(\hat{x}, \xi) \). Consequently, \( \hat{x}_N \) is asymptotically normal iff the mapping \( \hat{d}(\cdot) \) is linear. This, in turn, holds if the cone \( C_X(\hat{x}) \) is a linear space.

In the case \( \Gamma(\cdot) := N_X(\cdot) \), with the set \( X \) being nonempty closed convex and polyhedral, there is a complete characterization of the strong regularity in terms of the so-called coherent orientation associated with the matrix (mapping) \( J := \nabla \phi(\hat{x}) \) and the critical cone \( C_X(\hat{x}) \). The interested reader is referred to [172],[79] for a discussion of this topic. Let us just remark that if \( C_X(\hat{x}) \) is a linear subspace of \( \mathbb{R}^n \), then the variational inequality (5.70) can be written in the form
\[
P\delta + PJd = 0,
\]  
(5.72)
where $P$ denotes the orthogonal projection matrix onto the linear space $C_X(\bar{x})$. Then $\bar{x}$ is strongly regular iff the matrix (mapping) $PJ$ restricted to the linear space $C_X(\bar{x})$ is invertible, or in other words nonsingular.

Suppose now that $S = \{\bar{x}\}$ is such that $\phi(\bar{x})$ belongs to the interior of the set $\Gamma(\bar{x})$. Then, since $\phi_N(\bar{x})$ converges w.p.1 to $\phi(\bar{x})$, it follows that the event “$\phi_N(\bar{x}) \in \Gamma(\bar{x})$” happens w.p.1 for $N$ large enough. Moreover, by the LD principle (see (7.197)) we have that this event happens with probability approaching one exponentially fast. Of course, $\phi_N(\bar{x}) \in \Gamma(\bar{x})$ means that $\bar{x}_N = \bar{x}$ is a solution of the SAA generalized equation (5.64). Therefore, in such case one may compute an exact solution of the true problem (5.57), by solving the SAA problem, with probability approaching one exponentially fast with increase of the sample size. Note that if $\Gamma(\cdot) := \mathcal{N}_X(\cdot)$ and $\bar{x} \in S$, then $\phi(\bar{x}) \in \text{int} \Gamma(\bar{x})$ iff the critical cone $C_X(\bar{x})$ is equal to $\{0\}$. In that case the variational inequality (5.70) has solution $d = 0$ for any $\delta$, i.e., $d(\delta) \equiv 0$.

The above asymptotics can be applied, in particular, to the generalized equation (variational inequality) $\phi(z) \in \mathcal{N}_K(z)$, where $K := \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{p-q}$ and $\mathcal{N}_K(z)$ and $\phi(z)$ are given in (5.61) and (5.63), respectively. Recall that this variational inequality represents the KKT optimality conditions of the expected value optimization problem (5.1) with the feasible set $X$ given in the form (5.59). (We assume that the expectation functions $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$, are continuously differentiable.) Let $\bar{x}$ be an optimal solution of the (expected value) problem (5.1). It is said that the linear independence constraint qualification (LICQ) holds at the point $\bar{x}$ if the gradient vectors $\nabla g_i(\bar{x})$, $i \in \{i : g_i(\bar{x}) = 0, i = 1, \ldots, p\}$, (of active at $\bar{x}$ constraints) are linearly independent.

Under the LICQ, to $\bar{x}$ corresponds a unique vector $\lambda$ of Lagrange multipliers, satisfying the KKT optimality conditions. Let $\bar{z} = (\bar{x}, \bar{\lambda})$ and $I_0(\lambda)$ and $I_+ (\lambda)$ be the index sets defined in (5.62). Then

$$T_K(\bar{z}) = \mathbb{R}^n \times \mathbb{R}^q \times \{\gamma \in \mathbb{R}^{p-q} : \gamma_i \geq 0, \ i \in I_0(\bar{\lambda})\}. \quad (5.73)$$

In order to simplify notation let us assume that all constraints are active at $\bar{x}$, i.e., $g_i(\bar{x}) = 0$, $i = 1, \ldots, p$. Since for sufficiently small perturbations of $x$, inactive constraints remain inactive, we do not lose generality in the asymptotic analysis by considering only active at $\bar{x}$ constraints. Then $\phi(\bar{z}) = 0$, and hence $C_K(\bar{z}) = T_K(\bar{z})$.

Assuming, further, that $f(x)$ and $g_i(x)$, $i = 1, \ldots, p$ are twice continuously differentiable, we have that the following second order necessary conditions hold at $\bar{x}$:

$$h^T \nabla^2_x \ell(\bar{z}) h \geq 0, \ \forall h \in C_X(\bar{x}), \quad (5.74)$$

where

$$C_X(\bar{x}) := \{h : h^T \nabla g_i(\bar{x}) = 0, \ i \in \{1, \ldots, q\} \cup I_+(\bar{\lambda}), \ h^T \nabla g_i(\bar{x}) \leq 0, \ i \in I_0(\bar{\lambda})\}. \quad (5.75)$$

The corresponding second order sufficient conditions are

$$h^T \nabla^2_x \ell(\bar{z}) h > 0, \ \forall h \in C_X(\bar{x}) \setminus \{0\}. \quad (5.76)$$

Moreover, $\bar{z}$ is a strongly regular solution of the corresponding generalized equation iff the LICQ holds at $\bar{x}$ and the following (strong) form of second order sufficient conditions is satisfied:

$$h^T \nabla^2_x \ell(\bar{z}) h > 0, \ \forall h \in \text{lin}(C_X(\bar{x})) \setminus \{0\}, \quad (5.77)$$
where
\[
\text{lin}(C_X(\bar{x})) := \{ h : h^T \nabla g_i(\bar{x}) = 0, \ i \in \{1, \ldots, q\} \cup I_+(\bar{\lambda}) \}.
\] (5.77)

Under the LICQ, the set defined in the right hand side of (5.77) is, indeed, the linear space generated by the cone \( C_X(\bar{x}) \). We also have here
\[
J := \nabla \phi(\bar{z}) = \begin{bmatrix} H & A \\ A^T & 0 \end{bmatrix},
\] (5.78)

where \( H := \nabla^2_{xx} \ell(\bar{z}) \) and \( A := [\nabla g_1(\bar{x}), \ldots, \nabla g_p(\bar{x})] \).

It is said that the strict complementarity condition holds at \( \bar{x} \) if the index set \( I_0(\bar{\lambda}) \) is empty, i.e., all Lagrange multipliers corresponding to active at \( \bar{x} \) inequality constraints are strictly positive. We have here that \( C_K(\bar{z}) \) is a linear space, and hence the SAA estimator \( \hat{z}_N = \left[ \frac{\hat{z}_N}{\hat{\lambda}_N} \right] \) is asymptotically normal, if the strict complementarity condition holds. If the strict complementarity condition holds, then \( C_K(\bar{z}) = \mathbb{R}^{n+p} \) (recall that it is assumed that all constraints are active at \( \bar{x} \)), and hence the normal cone to \( C_K(\bar{z}) \), at every point, is \{0\}. Consequently, the corresponding variational inequality (5.70) takes the form \( \delta^T + Jd = 0 \).

Under the strict complementarity condition, \( \bar{z} \) is strongly regular iff the matrix \( J \) is nonsingular. It follows that under the above assumptions together with the strict complementarity condition, the following asymptotics hold (compare with (5.44))
\[
N^{1/2}(\hat{z}_N - \bar{z}) \overset{D}{\to} \mathcal{N}(0, J^{-1} \Sigma J^{-1}),
\] (5.79)

where \( \Sigma \) is the covariance matrix of the random vector \( \Phi(\bar{z}, \xi) \) defined in (5.60).

### 5.3 Monte Carlo Sampling Methods

In this section we assume that a random sample \( \xi_1, \ldots, \xi_N \), of \( N \) realizations of the random vector \( \xi \), can be generated in the computer. In the Monte Carlo sampling method this is accomplished by generating a sequence \( U^1, U^2, \ldots \) of independent random (or rather pseudorandom) numbers uniformly distributed on the interval \([0,1]\), and then constructing the sample by an appropriate transformation. In that way we can consider the sequence \( \omega := \{U^1, U^2, \ldots\} \) as an element of the probability space equipped with the corresponding product probability measure, and the sample \( \xi_i = \xi_i(\omega), \ i = 1, 2, \ldots \), as a function of \( \omega \). Since computer is a finite deterministic machine, sooner or later the generated sample will start to repeat itself. However, modern random numbers generators have a very large cycle period and this method was tested in numerous applications. We view now the corresponding SAA problem (5.2) as a way of approximating the true problem (5.1) while drastically reducing the number of generated scenarios. For a statistical analysis of the constructed SAA problems a particular numerical algorithm applied to solve these problems is irrelevant.

Let us also remark that values of the sample average function \( \hat{f}_N(x) \) can be computed in two somewhat different ways. The generated sample \( \xi_1, \ldots, \xi^N \) can be stored in the computer memory, and called every time a new value (at a different point \( x \)) of the sample average function should be computed. Alternatively, the same sample can be generated by using a common seed number in an employed pseudorandom numbers generator (this is why this approach is called the common random number generation method).
The idea of common random number generation is well known in simulation. That is, suppose that we want to compare values of the objective function at two points \(x_1, x_2 \in X\). In that case we are interested in the difference \(f(x_1) - f(x_2)\) rather than in the individual values \(f(x_1)\) and \(f(x_2)\). If we use sample average estimates \(\hat{f}_N(x_1)\) and \(\hat{f}_N(x_2)\) based on independent samples, both of size \(N\), then \(\hat{f}_N(x_1)\) and \(\hat{f}_N(x_2)\) are uncorrelated and

\[
\mathbb{Var}[\hat{f}_N(x_1) - \hat{f}_N(x_2)] = \mathbb{Var}[\hat{f}_N(x_1)] + \mathbb{Var}[\hat{f}_N(x_2)].
\]

(5.80)

On the other hand, if we use the same sample for the estimators \(\hat{f}_N(x_1)\) and \(\hat{f}_N(x_2)\), then

\[
\mathbb{Var}[\hat{f}_N(x_1) - \hat{f}_N(x_2)] = \mathbb{Var}[\hat{f}_N(x_1)] + \mathbb{Var}[\hat{f}_N(x_2)] - 2\mathbb{Cov}(\hat{f}_N(x_1), \hat{f}_N(x_2)).
\]

(5.81)

In both cases, \(\hat{f}_N(x_1) - \hat{f}_N(x_2)\) is an unbiased estimator of \(f(x_1) - f(x_2)\). However, in the case of the same sample the estimators \(\hat{f}_N(x_1)\) and \(\hat{f}_N(x_2)\) tend to be positively correlated with each other, in which case the variance in (5.81) is smaller than the one in (5.80). The difference between the independent and the common random number generated estimators of \(f(x_1) - f(x_2)\) can be especially dramatic when the points \(x_1\) and \(x_2\) are close to each other and hence the common random number generated estimators are highly positively correlated.

By the results of section 5.1.1 we have that, under mild regularity conditions, the optimal value and optimal solutions of the SAA problem (5.2) converge w.p.1, as the sample size increases, to their true counterparts. These results, however, do not give any indication of quality of solutions for a given sample of size \(N\). In the next section we discuss exponential rates of convergence of optimal and nearly optimal solutions of the SAA problem (5.2). This allows to give an estimate of the sample size which is required to solve the true problem with a given accuracy by solving the SAA problem. Although such estimates of the sample size typically are too conservative for a practical use, they give an insight into the complexity of solving the true (expected value) problem.

Unless stated otherwise, we assume in this section that the random sample \(\xi^1, ..., \xi^N\) is iid, and make the following assumption:

(M1) The expectation function \(f(x)\) is well defined and finite valued for all \(x \in X\).

For \(\varepsilon \geq 0\) we denote by

\[
S^\varepsilon := \{ x \in X : f(x) \leq \hat{\theta} + \varepsilon \} \quad \text{and} \quad \hat{S}^\varepsilon_N := \{ x \in X : \hat{f}_N(x) \leq \hat{\theta}_N + \varepsilon \}
\]

the sets of \(\varepsilon\)-optimal solutions of the true and the SAA problems, respectively.

### 5.3.1 Exponential Rates of Convergence and Sample Size Estimates in Case of Finite Feasible Set

In this section we assume that the feasible set \(X\) is finite, although its cardinality \(|X|\) can be very large. Since \(X\) is finite, the sets \(S^\varepsilon\) and \(\hat{S}^\varepsilon_N\) are nonempty and finite. For parameters \(\varepsilon \geq 0\) and \(\delta \in [0, \varepsilon]\), consider the event \(\{ \hat{S}^\delta_N \subset S^\varepsilon \}\). This event means that any \(\delta\)-optimal solution of the SAA problem is an \(\varepsilon\)-optimal solution of the true problem. We estimate now the probability of that event.
5.3. Monte Carlo Sampling Methods

We can write
\[
\left\{ \hat{S}_N^\delta \not\subset S^\varepsilon \right\} = \bigcup_{x \in X \setminus S^\varepsilon} \bigcap_{y \in X} \left\{ \hat{f}_N(x) \leq \hat{f}_N(y) + \delta \right\}, \tag{5.82}
\]
and hence
\[
\Pr\left( \hat{S}_N^\delta \not\subset S^\varepsilon \right) \leq \sum_{x \in X \setminus S^\varepsilon} \Pr\left( \bigcap_{y \in X} \left\{ \hat{f}_N(x) \leq \hat{f}_N(y) + \delta \right\} \right). \tag{5.83}
\]

Consider a mapping \( u : X \setminus S^\varepsilon \to X \). If the set \( X \setminus S^\varepsilon \) is empty, then any feasible point \( x \in X \) is an \( \varepsilon \)-optimal solution of the true problem. Therefore we assume that this set is nonempty. It follows from (5.83) that
\[
\Pr\left( \hat{S}_N^\delta \not\subset S^\varepsilon \right) \leq \sum_{x \in X \setminus S^\varepsilon} \Pr\left\{ \hat{Y}_N(x) \leq -\delta \right\}. \tag{5.84}
\]

We assume that the mapping \( u(\cdot) \) is chosen in such a way that
\[
f(u(x)) \leq f(x) - \varepsilon^* \quad \text{for all} \quad x \in X \setminus S^\varepsilon, \tag{5.85}
\]
and for some \( \varepsilon^* \geq \varepsilon \). Note that such a mapping always exists. For example, if we use a mapping \( u : X \setminus S^\varepsilon \to S^\varepsilon \), then (5.85) holds with
\[
\varepsilon^* := \min_{x \in X \setminus S^\varepsilon} f(x) - \theta^*, \tag{5.86}
\]
and that \( \varepsilon^* > \varepsilon \) since the set \( X \) is finite. Different choices of \( u(\cdot) \) give a certain flexibility to the following derivations.

For each \( x \in X \setminus S^\varepsilon \), define
\[
Y(x, \xi) := F(u(x), \xi) - F(x, \xi). \tag{5.87}
\]
Note that \( \mathbb{E}[Y(x, \xi)] = f(u(x)) - f(x) \), and hence \( \mathbb{E}[Y(x, \xi)] \leq -\varepsilon^* \) for all \( x \in X \setminus S^\varepsilon \). The corresponding sample average is
\[
\hat{Y}_N(x) := \frac{1}{N} \sum_{j=1}^{N} Y(x, \xi_j) = \hat{f}_N(u(x)) - \hat{f}_N(x). \tag{5.88}
\]

By (5.84) we have
\[
\Pr\left( \hat{S}_N^\delta \not\subset S^\varepsilon \right) \leq \sum_{x \in X \setminus S^\varepsilon} \Pr\left\{ \hat{Y}_N(x) \geq -\delta \right\}. \tag{5.88}
\]
Let \( I_x(\cdot) \) denote the (large deviations) rate function of the random variable \( Y(x, \xi) \). The inequality (5.88) together with the LD upper bound (7.179) implies
\[
1 - \Pr\left( \hat{S}_N^\delta \subset S^\varepsilon \right) \leq \sum_{x \in X \setminus S^\varepsilon} e^{-NI_x(-\delta)}. \tag{5.89}
\]
Note that the above inequality (5.89) is valid for any random sample of size \( N \). Let us make the following assumption.
(M2) For every \( x \in X \setminus S^c \) the moment generating function \( E \left[ e^{tY(x,\xi)} \right] \), of the random variable \( Y(x,\xi) = F(u(x),\xi) - F(x,\xi) \), is finite valued in a neighborhood of \( t = 0 \).

The above assumption (M2) holds, for example, if the support \( \Xi \) of \( \xi \) is a bounded subset of \( \mathbb{R}^d \), or if \( Y(x,\cdot) \) grows at most linearly and \( \xi \) has a distribution from an exponential family.

**Theorem 5.16.** Let \( \varepsilon \) and \( \delta \) be nonnegative numbers. Then

\[
1 - \Pr \left( \hat{S}_N^\delta \subseteq S^\varepsilon \right) \leq |X| e^{-N\eta(\delta,\varepsilon)},
\]

where

\[
\eta(\delta,\varepsilon) := \min_{x \in X \setminus S^c} I_x(-\delta).
\]

Moreover, if \( \delta < \varepsilon^* \) and assumption (M2) holds, then \( \eta(\delta,\varepsilon) > 0 \).

**Proof.** The inequality (5.90) is an immediate consequence of the inequality (5.89). If \( \delta < \varepsilon^* \), then \(-\delta > -\varepsilon^* \geq \mathbb{E}[Y(x,\xi)]\), and hence it follows by assumption (M2) that \( I_x(-\delta) > 0 \) for every \( x \in X \setminus S^c \) (see discussion above equation (7.184)). This implies that \( \eta(\delta,\varepsilon) > 0 \).

The following asymptotic result is an immediate consequence of inequality (5.90),

\[
\limsup_{N \to \infty} \frac{1}{N} \ln \left[ 1 - \Pr \left( \hat{S}_N^\delta \subseteq S^\varepsilon \right) \right] \leq -\eta(\delta,\varepsilon).
\]

It means that the probability of the event that any \( \delta \)-optimal solution of the SAA problem provides an \( \varepsilon \)-optimal solution of the true problem approaches one exponentially fast as \( N \to \infty \). Note that since it is possible to employ a mapping \( u : X \setminus S^c \to S \), with \( \varepsilon^* > \varepsilon \) (see (5.86)), this exponential rate of convergence holds even if \( \delta = \varepsilon \), and in particular if \( \delta = \varepsilon = 0 \). However, if \( \delta = \varepsilon \) and the difference \( \varepsilon^* - \varepsilon \) is small, then the constant \( \eta(\delta,\varepsilon) \) could be close to zero. Indeed, for \( \delta \) close to \(-\mathbb{E}[Y(x,\xi)]\), we can write by (7.184) that

\[
I_x(-\delta) \approx \frac{(-\delta - \mathbb{E}[Y(x,\xi)])^2}{2\sigma_x^2} \geq \frac{(\varepsilon^* - \delta)^2}{2\sigma_x^2},
\]

where

\[
\sigma_x^2 := \text{Var}[Y(x,\xi)] = \text{Var}[F(u(x),\xi) - F(x,\xi)].
\]

Let us make now the following assumption.

(M3) There is a constant \( \sigma > 0 \) such that for any \( x \in X \setminus S^c \) the moment generating function \( M_x(t) \) of the random variable \( Y(x,\xi) - \mathbb{E}[Y(x,\xi)] \) satisfies

\[
M_x(t) \leq \exp \left( \sigma^2 t^2 / 2 \right), \quad \forall t \in \mathbb{R}.
\]

It follows from the above assumption (M3) that

\[
\ln E \left[ e^{tY(x,\xi)} \right] - t\mathbb{E}[Y(x,\xi)] = \ln M_x(t) \leq \sigma^2 t^2 / 2,
\]
and hence the rate function $I_x(\cdot)$, of $Y(x, \xi)$, satisfies
\[
I_x(z) = \sup_{t \in \mathbb{R}} \left\{ t\left( z - \mathbb{E}[Y(x, \xi)] \right) - \frac{\sigma^2 t^2}{2} \right\} = \frac{\left( z - \mathbb{E}[Y(x, \xi)] \right)^2}{2\sigma^2}, \quad \forall z \in \mathbb{R}. \tag{5.97}
\]
In particular, it follows that
\[
I_x(-\delta) \geq \frac{(\varepsilon^* - \delta)^2}{2\sigma^2} \geq \frac{(\varepsilon - \delta)^2}{2\sigma^2}. \tag{5.98}
\]
Consequently the constant $\eta(\delta, \varepsilon)$ satisfies
\[
\eta(\delta, \varepsilon) \geq \frac{(\varepsilon - \delta)^2}{2\sigma^2}, \tag{5.99}
\]
and hence the bound (5.90) of Theorem 5.16 takes the form
\[
1 - \Pr(\hat{S}_N^\delta \subset S^\varepsilon) \leq |X| e^{-N(\varepsilon^* - \delta)^2/(2\sigma^2)}. \tag{5.100}
\]
This leads to the following result giving an estimate of the sample size which guarantees that any $\delta$-optimal solution of the SAA problem is an $\varepsilon$-optimal solution of the true problem with probability at least $1 - \alpha$.

**Theorem 5.17.** Suppose that assumptions (M1) and (M3) hold. Then for $\varepsilon > 0$, $0 \leq \delta < \varepsilon$ and $\alpha \in (0, 1)$, and for the sample size $N$ satisfying
\[
N \geq \frac{2\sigma^2}{(\varepsilon - \delta)^2} \ln \left( \frac{|X|}{\alpha} \right), \tag{5.101}
\]
it follows that
\[
\Pr(\hat{S}_N^\delta \subset S^\varepsilon) \geq 1 - \alpha. \tag{5.102}
\]

**Proof.** By setting the right hand side of the estimate (5.100) to $\leq \alpha$ and solving the obtained inequality, we obtain (5.101). \qed

**Remark 7.** A key characteristic of the estimate (5.101) is that the required sample size $N$ depends logarithmically both on the size (cardinality) of the feasible set $X$ and on the tolerance probability (significance level) $\alpha$. The constant $\sigma$, postulated in assumption (M3), measures, in a sense, variability of a considered problem. If, for some $x \in X$, the random variable $Y(x, \xi)$ has a normal distribution with mean $\mu_x$ and variance $\sigma^2_x$, then its moment generating function is equal to $\exp \left( \mu_x t + \frac{\sigma^2_x t^2}{2} \right)$, and hence the moment generating function $M_x(t)$, specified in assumption (M3), is equal to $\exp \left( \frac{\sigma^2_x t^2}{2} \right)$. In that case $\sigma^2 := \max_{x \in X \setminus S^\varepsilon} \sigma^2_x$ gives the smallest possible value for the corresponding constant in assumption (M3). If $Y(x, \xi)$ is bounded w.p.1, i.e., there is constant $b > 0$ such that
\[
|Y(x, \xi) - \mathbb{E}[Y(x, \xi)]| \leq b \text{ for all } x \in X \text{ and a.e. } \xi \in \Xi,
\]
then by Hoeffding inequality (see Proposition 7.63 and estimate (7.192)) we have that
\[ M_x(t) \leq \exp \left( b^2 t^2 / 2 \right). \]
In that case we can take \( \sigma^2 := b^2 \).

In any case for small \( \varepsilon > 0 \) we have by (5.93) that \( I_x(\delta) \) can be approximated from
down by \( (\varepsilon - \delta)^2 / (2\sigma^2) \).

**Remark 8.** For, say, \( \delta := \varepsilon / 2 \), the right hand side of the estimate (5.101) is proportional
to \((\sigma / \varepsilon)^2\). For Monte Carlo sampling based methods such dependence on \( \sigma \) and \( \varepsilon \) seems
to be unavoidable. In order to see that consider a simple case when the feasible set \( X \) consists of just two elements, i.e., \( X = \{ x_1, x_2 \} \), with \( f(x_2) - f(x_1) > \varepsilon > 0 \). By
solving the corresponding SAA problem we make the (correct) decision that \( x_1 \) is the \( \varepsilon \)-optimal solution if \( f_N(x_2) - f_N(x_1) > 0 \). If the random variable \( F(x_2, \xi) - F(x_1, \xi) \)
has a normal distribution with mean \( \mu = f(x_2) - f(x_1) \) and variance \( \sigma^2 \), then \( f_N(x_2) - f_N(x_1) \sim N(\mu, \sigma^2 / N) \) and the probability of the event \( \{| f_N(x_2) - f_N(x_1) > 0 \} \) (i.e., of the
correct decision) is \( \Phi(\sqrt{N}/\sigma) \), where \( \Phi(z) \) is the cumulative distribution function of \( N(0, 1) \). We have that \( \Phi(\varepsilon \sqrt{N}/\sigma) < \Phi(\mu \sqrt{N}/\sigma) \), and in order to make the probability of the incorrect decision less than \( \alpha \) we have to take the sample size \( N > z_\alpha^2 \sigma^2 / \varepsilon^2 \), where \( z_\alpha := \Phi^{-1}(1 - \alpha) \). Even if \( F(x_2, \xi) - F(x_1, \xi) \) is not normally distributed, the sample size of \( \sigma^2 / \varepsilon^2 \) could be justified asymptotically, say by applying the CLT. It also
could be mentioned that if \( F(x_2, \xi) - F(x_1, \xi) \) has a normal distribution (with known variance),
then the uniformly most powerful test for testing \( H_0 : \mu = 0 \) versus \( H_A : \mu > 0 \) is of the form: “reject \( H_0 \) if \( f_N(x_2) - f_N(x_1) \) is bigger than a specified critical value” (this is a consequence of the Neyman-Pearson Lemma). In other words, in such situations if we
only have an access to a random sample, then solving the corresponding SAA problem is in
a sense a best way to proceed.

**Remark 9.** Condition (5.95) of assumption (M3) can be replaced by a more general condition
\[ M_x(t) \leq \exp \left( \psi(t) \right), \quad \forall t \in \mathbb{R}, \tag{5.103} \]
where \( \psi(t) \) is a convex even function with \( \psi(0) = 0 \). Then, similar to (5.97), we have
\[ I_x(z) \geq \sup_{t \in \mathbb{R}} \{ t(z - \mathbb{E}[Y(x, \xi)]) - \psi(t) \} = \psi^* \left( z - \mathbb{E}[Y(x, \xi)] \right), \quad \forall z \in \mathbb{R}, \tag{5.104} \]
where \( \psi^* \) is the conjugate of function \( \psi \). Consequently the estimate (5.90) takes the form
\[ 1 - \Pr(\Delta \subset S^c) \leq \sqrt{N} e^{-N \psi^*(\varepsilon - \delta)}, \tag{5.105} \]
and hence the estimate (5.101) takes the form
\[ N \geq \frac{1}{\psi^*(\varepsilon - \delta)} \ln \left( \frac{|X|}{\alpha} \right). \tag{5.106} \]
For example, instead of assuming that condition (5.95) of assumption (M3) holds for all \( t \in \mathbb{R} \), we may assume that this holds for all \( t \) in a finite interval \([-a, a] \), where \( a > 0 \) is a given constant. That is, we can take \( \psi(t) := \sigma^2 t^2 / 2 \) if \( |t| \leq a \), and \( \psi(t) := +\infty \) otherwise. In that case \( \psi^*(z) = z^2 / (2\sigma^2) \) for \( |z| \leq a\sigma^2 \), and \( \psi^*(z) = a|z| - a^2 \sigma^2 \) for
\( |z| > a\sigma^2 \). Consequently the estimate (5.101) of Theorem 5.17 still holds provided that
\( 0 < \varepsilon - \delta \leq a\sigma^2 \).
5.3. Monte Carlo Sampling Methods

5.3.2 Sample Size Estimates in General Case

Suppose now that $X$ is a bounded, not necessarily finite, subset of $\mathbb{R}^n$. and that $f(x)$ is finite valued for all $x \in X$. Then we can proceed in a way similar to the derivations of section 7.2.9. Let us make the following assumptions.

(M4) For any $x', x \in X$ there exists constant $\sigma_{x',x} > 0$ such that the moment generating function $M_{x',x}(t) = \mathbb{E}[e^{tY_{x',x}]}$, of random variable $Y_{x',x} := [F(x',\xi) - f(x')] - [F(x,\xi) - f(x)]$, satisfies

$$M_{x',x}(t) \leq \exp\left(\frac{\sigma^2_{x',x}t^2}{2}\right), \quad \forall t \in \mathbb{R}. \quad (5.107)$$

(M5) There exists a (measurable) function $\kappa : \Xi \to \mathbb{R}_+$ such that its moment generating function $M_{x}(t)$ is finite valued for all $t$ in a neighborhood of zero and

$$|F(x',\xi) - F(x,\xi)| \leq \kappa(\xi)\|x' - x\| \quad (5.108)$$

for a.e. $\xi \in \Xi$ and all $x', x \in X$.

Of course, it follows from (5.107) that

$$M_{x',x}(t) \leq \exp\left(\frac{\sigma^2t^2}{2}\right), \quad \forall x', x \in X, \quad \forall t \in \mathbb{R}, \quad (5.109)$$

where

$$\sigma^2 := \sup_{x', x \in X} \sigma^2_{x',x}. \quad (5.110)$$

The above assumption (M4) is slightly stronger than assumption (M3), i.e., assumption (M3) follows from (M4) by taking $x' = u(x)$. Note that $\mathbb{E}[Y_{x',x}] = 0$ and recall that if $Y_{x',x}$ has a normal distribution, then equality in (5.107) holds with $\sigma^2_{x',x} := \text{Var}(Y_{x',x})$.

The assumption (M5) implies that the expectation $\mathbb{E}[\kappa(\xi)]$ is finite and the function $f(x)$ is Lipschitz continuous on $X$ with Lipschitz constant $L = \mathbb{E}[\kappa(\xi)]$. It follows that the optimal value $\theta^*$ of the true problem is finite, provided the set $X$ is bounded (recall that it was assumed that $X$ is nonempty and closed). Moreover, by Cramér’s Large Deviation Theorem we have that for any $L' > \mathbb{E}[\kappa(\xi)]$ there exists a positive constant $\beta = \beta(L')$ such that

$$\Pr(\hat{\kappa}_N > L') \leq \exp(-N\beta), \quad (5.111)$$

where $\hat{\kappa}_N := N^{-1}\sum_{j=1}^N \kappa(\xi_j)$. Note that it follows from (5.108) that w.p.1

$$|\hat{f}_N(x') - \hat{f}_N(x)| \leq \hat{\kappa}_N\|x' - x\|, \quad \forall x', x \in X, \quad (5.112)$$

i.e., $\hat{f}_N(\cdot)$ is Lipschitz continuous on $X$ with Lipschitz constant $\hat{\kappa}_N$.

By $D := \sup_{x,x' \in X}\|x' - x\|$ we denote the diameter of the set $X$. Of course, the set $X$ is bounded if its diameter is finite. We also use notation $a \vee b := \max\{a, b\}$ for numbers $a, b \in \mathbb{R}$.

Theorem 5.18. Suppose that assumptions (M1) and (M4)–(M5) hold, with the corresponding constant $\sigma^2$ defined in (5.110) being finite, the set $X$ has a finite diameter $D$, and let $\varepsilon > 0$, $\delta \in [0, \varepsilon)$, $\alpha \in (0, 1)$, $L' > L := \mathbb{E}[\kappa(\xi)]$ and $\beta = \beta(L')$ be the corresponding
constants and $\varrho > 0$ be a constant specified below in (5.115). Then for the sample size $N$ satisfying

$$N \geq \frac{8\sigma^2}{(\varepsilon - \delta)^2} \left[ n \ln \left( \frac{8\sigma L' D}{\varepsilon - \delta} \right) + \ln \left( \frac{2}{\alpha} \right) \right] \vee \left[ \beta^{-1} \ln \left( \frac{2}{\alpha} \right) \right],$$

(5.113)

it follows that

$$\Pr(\tilde{S}_N^\delta \subset S^\delta) \geq 1 - \alpha.$$  

(5.114)

**Proof.** Let us set $\nu := (\varepsilon - \delta)/(4L')$, $\varepsilon' := \varepsilon - L' \nu$ and $\delta' := \delta + L' \nu$. Note that $\nu > 0$, $\varepsilon' = 3\varepsilon/4 + \delta/4 > 0$, $\delta' = \varepsilon/4 + 3\delta/4 > 0$ and $\varepsilon' - \delta' = (\varepsilon - \delta)/2 > 0$. Let $\bar{x}_1, ..., \bar{x}_M \in X$ be such that for every $x \in X$ there exists $\bar{x}_i$, $i \in \{1, ..., M\}$, such that $\|x - \bar{x}_i\| \leq \nu$, i.e., the set $X' := \{\bar{x}_1, ..., \bar{x}_M\}$ forms a $\nu$-net in $X$. We can choose this net in such a way that

$$M \leq (\varrho D/\nu)^n$$

(5.115)

for a constant $\varrho > 0$. If the $X' \setminus S^\varepsilon'$ is empty, then any point of $X'$ is an $\varepsilon'$-optimal solution of the true problem. Otherwise, choose a mapping $u : X' \setminus S^\varepsilon' \rightarrow S$ and consider the sets $\tilde{S} := \cup_{x \in X'} \{u(x)\}$ and $\tilde{X} := X' \cup \tilde{S}$. Note that $\tilde{X} \subset X$ and $|\tilde{X}| \leq (2\varrho D/\nu)^n$. Now let us replace the set $X$ by its subset $\tilde{X}$. We refer to the obtained true and SAA problems as respective reduced problems. We have that $\tilde{S} \subset S$, any point of the set $\tilde{S}$ is an optimal solutions of the true reduced problem and the optimal value of the true reduced problem is equal to the optimal value of the true (unreduced) problem. By Theorem 5.17 we have that with probability at least $1 - \alpha/2$ any $\delta'$-optimal solution of the reduced SAA problem is an $\varepsilon'$-optimal solutions of the reduced (and hence unreduced) true problem provided that

$$N \geq \frac{8\sigma^2}{(\varepsilon' - \delta')^2} \left[ n \ln \left( \frac{8\sigma L' D}{\varepsilon' - \delta'} \right) + \ln \left( \frac{2}{\alpha} \right) \right],$$

(note that the right hand side of (5.116) is greater than or equal to the estimate

$$\frac{2\sigma^2}{(\varepsilon' - \delta')^2} \ln \left( \frac{2|\tilde{X}|}{\alpha} \right)$$

required by Theorem 5.17). We also have by (5.111) that for

$$N \geq \beta^{-1} \ln \left( \frac{2}{\alpha} \right),$$

(5.117)

the Lipschitz constant $\tilde{k}_N$ of the function $\tilde{f}_N(x)$ is less than or equal to $L'$ with probability at least $1 - \alpha/2$.

Now let $\tilde{x}$ be a $\delta$-optimal solution of the (unreduced) SAA problem. Then there is a point $x' \in \tilde{X}$ such that $\|\tilde{x} - x'\| \leq \nu$, and hence $\tilde{f}_N(x') \leq \tilde{f}_N(\tilde{x}) + L' \nu$, provided that $\tilde{k}_N \leq L'$. We also have that the optimal value of the (unreduced) SAA problem is smaller than or equal to the optimal value of the reduced SAA problem. It follows that $x'$ is a $\delta'$-optimal solution of the reduced SAA problem, provided that $\tilde{k}_N \leq L'$. Consequently,
5.3. Monte Carlo Sampling Methods

we have that \( x' \) is an \( \epsilon' \)-optimal solution of the true problem with probability at least \( 1 - \alpha \) provided that \( N \) satisfies both inequalities (5.116) and (5.117). It follows that

\[
 f(x) \leq f(x') + L\nu \leq f(x') + L'\nu \leq \theta^* + \epsilon' + L'\nu = \theta^* + \epsilon.
\]

We obtain that if \( N \) satisfies both inequalities (5.116) and (5.117), then with probability at least \( 1 - \alpha \) any \( \delta \)-optimal solution of the SAA problem is an \( \epsilon \)-optimal solution of the true problem. The required estimate (5.113) follows.

It is also possible to derive sample size estimates of the form (5.113) directly from the uniform exponential bounds derived in section 7.2.9, see Theorem 7.67 in particular.

**Remark 10.** If instead of assuming that condition (5.107), of assumption (M4), holds for all \( t \in \mathbb{R} \), we assume that it holds for all \( t \in [-a, a] \), where \( a > 0 \) is a given constant, then the estimate (5.113) of the above theorem still holds provided that \( 0 < \epsilon - \delta \leq a\sigma^2 \) (see Remark 9, page 186).

In a sense the above estimate (5.113) of the sample size gives an estimate of complexity of solving the corresponding true problem by the SAA method. Suppose, for instance, that the true problem represents the first stage of a two-stage stochastic programming problem. For decomposition type algorithms the total number of iterations, required to solve the SAA problem, typically is independent of the sample size \( N \) (this is an empirical observation) and the computational effort at every iteration is proportional to \( N \). Anyway size of the SAA problem grows linearly with increase of \( N \). For \( \delta \in [0, \epsilon/2] \), say, the right hand side of (5.113) is proportional to \( \sigma^2/\epsilon^2 \), which suggests complexity of order \( \sigma^2/\epsilon^2 \) with respect to the desirable accuracy. This is in a sharp contrast with deterministic (convex) optimization where complexity usually is bounded in terms of \( \ln(\epsilon^{-1}) \). It seems that such dependence on \( \sigma \) and \( \epsilon \) is unavoidable for Monte Carlo sampling based methods. On the other hand, the estimate (5.113) is linear in the dimension \( n \) of the first-stage problem. It also depends linearly on \( \ln(a^{-1}) \). This means that by increasing confidence, say, from 99% to 99.99% we need to increase the sample size by the factor of \( \ln 100 \approx 4.6 \) at most. Assumption (M4) requires for the probability distribution of the random variable \( F(x, \xi) - F(x', \xi) \) to have sufficiently light tails. In a sense, the constant \( \sigma^2 \) can be viewed as a bound reflecting variability of the random variables \( F(x, \xi) - F(x', \xi) \), for \( x, x' \in X \). Naturally, larger variability of the data should result in more difficulty in solving the problem (see Remark 8, page 186).

This suggests that by using Monte Carlo sampling techniques one can solve two-stage stochastic programs with a reasonable accuracy, say with relative accuracy of 1% or 2%, in a reasonable time, provided that: (a) its variability is not too large, (b) it has relatively complete recourse, and (c) the corresponding SAA problem can be solved efficiently. And, indeed, this was verified in numerical experiments with two-stage problems having a linear second stage recourse. Of course, the estimate (5.113) of the sample size is far too conservative for actual calculations. For practical applications there are techniques which allow to estimate (statistically) the error of a considered feasible solution \( \bar{x} \) for a chosen sample size \( N \), we will discuss this in section 5.6.

We are going to discuss next some modifications of the sample size estimate. It will be convenient in the following estimates to use notation \( O(1) \) for a generic constant.
There exists constant $\lambda > 0$ such that for any $x', x \in X$ the moment generating function $M_{x',x}(t)$, of random variable $Y_{x',x} := [F(x', \xi) - f(x')] - [F(x, \xi) - f(x)]$, satisfies

$$M_{x',x}(t) \leq \exp \left( \lambda^2 \|x' - x\|^2 t^2 / 2 \right), \quad \forall t \in \mathbb{R}. \quad (5.118)$$

The above assumption (M6) is a particular case of assumption (M4) with

$$\sigma^2_{x',x} = \lambda^2 \|x' - x\|^2,$$

and we can set the corresponding constant $\sigma^2 = \lambda^2 D^2$. The following corollary follows from Theorem 5.18.

**Corollary 5.19.** Suppose that assumptions (M1) and (M5)–(M6) hold, the set $X$ has a finite diameter $D$, and let $\varepsilon > 0$, $\delta \in [0, \varepsilon)$, $\alpha \in (0, 1)$ and $L = \mathbb{E}[\kappa(\xi)]$ be the corresponding constants. Then for the sample size $N$ satisfying

$$N \geq \frac{O(1)\lambda^2 D^2}{(\varepsilon - \delta)^2} \left[ n \ln \left( \frac{O(1)L D}{\varepsilon - \delta} \right) + \ln \left( \frac{1}{\alpha} \right) \right], \quad (5.119)$$

it follows that

$$\Pr(\hat{S}_N \subset S^*) \geq 1 - \alpha. \quad (5.120)$$

For example, suppose that the Lipschitz constant $\kappa(\xi)$, in the assumption (M5), can be taken independent of $\xi$. That is, there exists a constant $L > 0$ such that

$$|F(x', \xi) - F(x, \xi)| \leq L \|x' - x\| \quad (5.121)$$

for a.e. $\xi \in \Xi$ and all $x', x \in X$. It follows that the expectation function $f(x)$ is also Lipschitz continuous on $X$ with Lipschitz constant $L$, and hence the random variable $Y_{x',x}$, of assumption (M6), can be bounded as $|Y_{x',x}| \leq 2L \|x' - x\|$ w.p.1. Moreover, we have that $\mathbb{E}[Y_{x',x}] = 0$, and hence it follows by Hoeffding’s inequality (see the estimate (7.192)) that

$$M_{x',x}(t) \leq \exp \left( 2L^2 \|x' - x\|^2 t^2 \right), \quad \forall t \in \mathbb{R.} \quad (5.122)$$

Consequently we can take $\lambda = 2L$ in (5.118) and the estimate (5.119) takes the form

$$N \geq \left( \frac{O(1)L D}{\varepsilon - \delta} \right)^2 n \ln \left( \frac{O(1)L D}{\varepsilon - \delta} \right) + \ln \left( \frac{1}{\alpha} \right). \quad (5.123)$$

**Remark 11.** It was assumed in Theorem 5.18 that the set $X$ has a finite diameter, i.e., that $X$ is bounded. For convex problems this assumption can be relaxed. Assume that the problem is convex, the optimal value $\vartheta^*$ of the true problem is finite and for some $a > \varepsilon$ the set $S^a$ has a finite diameter $D^*_a$ (recall that $S^a := \{x \in X : f(x) \leq \vartheta^* + a\}$). We refer here to the respective true and SAA problems, obtained by replacing the feasible set $X$ by its subset $S^a$, as reduced problems. Note that the set $S^\varepsilon$, of $\varepsilon$-optimal solutions,
of the reduced and original true problems are the same. Let \( N^* \) be an integer satisfying the inequality (5.113) with \( D \) replaced by \( D^*_\star \). Then, under the assumptions of Theorem 5.18, we have that with probability at least \( 1 - \alpha \) all \( \delta \)-optimal solutions of the reduced SAA problem are \( \epsilon \)-optimal solutions of the true problem. Let us observe now that in this case the set of \( \delta \)-optimal solutions of the reduced SAA problem coincides with the set of \( \delta \)-optimal solutions of the original SAA problem. Indeed, suppose that the original SAA problem has a \( \delta \)-optimal solution \( x^* \in X \setminus S^a \). Let \( \tilde{x} \in \arg\min_{x \in S^a} f_N(x) \), such a minimizer does exist since \( S^a \) is compact and \( f_N(x) \) is real valued convex and hence continuous. Then \( \tilde{x} \in S^a \) and \( f_N(x^*) \leq f_N(\tilde{x}) + \delta \). By convexity of \( f_N(x) \) it follows that \( f_N(x) \leq \max \{ f_N(\tilde{x}), f_N(x^*) \} \) for all \( x \) on the segment joining \( \tilde{x} \) and \( x^* \). This segment has a common point \( \tilde{x} \) with the set \( S^a \setminus S^\epsilon \). We obtain that \( \tilde{x} \in S^a \setminus S^\epsilon \) is a \( \delta \)-optimal solutions of the reduced SAA problem, a contradiction.

That is, with such sample size \( N^* \) we are guaranteed with probability at least \( 1 - \alpha \) that any \( \delta \)-optimal solution of the SAA problem is an \( \epsilon \)-optimal solution of the true problem. Also assumptions (M4) and (M5) should be verified for \( x, x' \) in the set \( S^a \) only.

**Remark 12.** Suppose that the set \( S \), of optimal solutions of the true problem, is nonempty. Then it follows from the proof of Theorem 5.18 that it suffices in the assumption (M4) to verify condition (5.107) only for every \( x \in X \setminus S^\epsilon \) and \( x' := u(x) \), where \( u : X \setminus S^\epsilon \to S \) and \( \epsilon' := 3/4\epsilon + \delta/4 \). If the set \( S \) is closed, we can use, for instance, a mapping \( u(x) \) assigning to each \( x \in X \setminus S^\epsilon \) a point of \( S \) closest to \( x \). If, moreover, the set \( S \) is convex and the employed norm is strictly convex (e.g., the Euclidean norm), then such mapping (called metric projection onto \( S \)) is defined uniquely.

If, moreover, the assumption (M6) holds, then for such \( x \) and \( x' \) we have \( \sigma_{x',x}^2 \leq \lambda^2 D^2 \), where \( D := \sup_{x \in X \setminus S^\epsilon} \text{dist}(x,S) \). Suppose, further, that the problem is convex. Then (see Remark 11, page 190), for any \( a > \epsilon \), we can use \( S^a \) instead of \( X \). Therefore, if the problem is convex and the assumption (M6) holds, we can write the following estimate of the required sample size:

\[
N \geq \frac{O(1)\lambda^2 D_{a,\epsilon}^2}{\frac{c}{\delta} - \frac{\delta}{\epsilon}} \left[ n \ln \left( \frac{O(1)LD_{a,\epsilon}^*}{\frac{c}{\delta} - \frac{\delta}{\epsilon}} \right) + \ln \left( \frac{1}{\alpha} \right) \right],
\]

(5.124)

where \( D_{a,\epsilon}^* \) is the diameter of \( S^a \) and \( D_{a,\epsilon} := \sup_{x \in S^a \setminus S^\epsilon} \text{dist}(x,S) \).

**Corollary 5.20.** Suppose that assumptions (M1) and (M5)–(M6) hold, the problem is convex, the “true” optimal set \( S \) is nonempty and for some \( \gamma \geq 1, c > 0 \) and \( r > 0 \), the following growth condition holds

\[
f(x) \geq \vartheta^* + c \text{[dist}(x,S)]^\gamma, \quad \forall x \in S^r.
\]

(5.125)

Let \( \alpha \in (0,1) \), \( \epsilon \in (0,r) \) and \( \delta \in [0,\epsilon/2] \) and suppose, further, that for \( \alpha := \min\{2\epsilon, r\} \) the diameter \( D_{a}^* \) of \( S^a \) is finite.

Then for the sample size \( N \) satisfying

\[
N \geq \frac{O(1)\lambda^2}{c^\gamma/\gamma \epsilon^{2(\gamma - 1)/\gamma}} \left[ n \ln \left( \frac{O(1)LD_{a,\epsilon}^*}{\epsilon} \right) + \ln \left( \frac{1}{\alpha} \right) \right],
\]

(5.126)

it follows that

\[
\Pr(\hat{S}_N^a \subset S^\epsilon) \geq 1 - \alpha.
\]

(5.127)
Consider the true problem with singleton, then it follows from (5.125) that 
\[ D(a/c)^{1/\gamma} \leq (a/c) \] holds. Consequently, for any \( \varepsilon \in (0, r) \), by taking \( a := \min\{2c, r\} \) and \( \delta \in [0, \varepsilon/2] \) we obtain from (5.124) the required sample size estimate (5.126).

Note that since \( a = \min\{2c, r\} \leq r \), we have that \( S^a \subset S^r \), and if \( S = \{x^*\} \) is a singleton, then it follows from (5.125) that \( D(a/c)^{1/\gamma} \leq (a/c) \). In particular, if \( \gamma = 1 \) and \( S = \{x^*\} \) is a singleton (in that case it is said that the optimal solution \( x^* \) is sharp), then \( D(a/c)^{1/\gamma} \leq (a/c) \) can be bounded by \( 4c^{-1} \varepsilon \) and hence we obtain the following estimate

\[ N \geq O(1)e^{-2\lambda^2 \left[ n \ln \left( O(1)c^{-1}L \right) + \ln \left( \alpha^{-1} \right) \right],} \tag{5.128} \]

which does not depend on \( \varepsilon \). For \( \gamma = 2 \) condition (5.125) is called the “second-order” or “quadratic” growth condition. Under the quadratic growth condition the first term in the right hand side of (5.126) becomes of order \( c^{-1} \varepsilon^{-1} \lambda^2 \).

The following example shows that the estimate (5.113) of the sample size cannot be significantly improved for the class of convex stochastic programs.

**Example 5.21** Consider the true problem with \( F(x, \xi) := \|x\|^{2m} - 2m \xi^T x \), where \( m \) is a positive constant, \( \| \cdot \| \) is the Euclidean norm and \( X := \{ x \in \mathbb{R}^n : \|x\| \leq 1 \} \). Suppose, further, that random vector \( \xi \) has normal distribution \( \mathcal{N}(0, \sigma^2 I_n) \), where \( \sigma^2 \) is a positive constant and \( I_n \) is the \( n \times n \) identity matrix, i.e., components \( \xi_i \) of \( \xi \) are independent and \( \xi_i \sim \mathcal{N}(0, \sigma^2), i = 1, ..., n \). It follows that \( f(x) = \|x\|^{2m} \), and hence for \( \varepsilon \in [0, 1] \) the set of \( \varepsilon \)-optimal solutions of the true problem is given by \( \{ x : \|x\|^{2m} \leq \varepsilon \} \). Now let \( \xi_1, ..., \xi_N \) be an iid random sample of \( \xi \) and \( \bar{\xi}_N := (\xi_1 + ... + \xi_N)/N \). The corresponding sample average function is

\[ \bar{f}_N(x) = \|x\|^{2m} - 2m \bar{\xi}_N^T x, \tag{5.129} \]

and the optimal solution \( \hat{x}_N \) of the SAA problem is \( \hat{x}_N = \|\bar{\xi}_N\|^{-b} \bar{\xi}_N \), where

\[ b := \begin{cases} \frac{2m-2}{2m-1}, & \text{if } \|\bar{\xi}_N\| \leq 1, \\ 1, & \text{if } \|\bar{\xi}_N\| > 1. \end{cases} \]

It follows that, for \( \varepsilon \in (0, 1) \), the optimal solution of the corresponding SAA problem is an \( \varepsilon \)-optimal solution of the true problem if \( \|\xi_N\|^{2m-2}\bar{\xi}_N \leq \varepsilon \), where \( \nu := \frac{2m-2}{2m-1} \). We have that \( \xi_N \sim \mathcal{N}(0, \sigma^2 N^{-1} I_n) \), and hence \( N\|\xi_N\|^2/\sigma^2 \) has a chi-square distribution with \( n \) degrees of freedom. Consequently, the probability that \( \|\bar{\xi}_N\|^{2m-2}\bar{\xi}_N \leq \varepsilon \) is equal to the probability \( \Pr(\chi^2_n > N\varepsilon^{2m-2}/\sigma^2) \). Moreover, \( \mathbb{E}[\chi^2_n] = n \) and the probability \( \Pr(\chi^2_n > n) \) increases and tends to 1/2 as \( n \) increases. Consequently, for \( \alpha \in (0, 0.3) \) and \( \varepsilon \in (0, 1) \), for example, the sample size \( N \) should satisfy

\[ N > \frac{n \sigma^2}{\varepsilon^{2m-2}} \tag{5.130} \]

in order to have the property: “with probability \( 1 - \alpha \) an (exact) optimal solution of the SAA problem is an \( \varepsilon \)-optimal solution of the true problem”. Compared with (5.113), the lower bound (5.130) also grows linearly in \( n \) and is proportional to \( \sigma^2/\varepsilon^{2m-2} \). It remains to note that the constant \( \nu \) decreases to one as \( m \) increases.
5.3. Monte Carlo Sampling Methods

Note that in this example the growth condition (5.125) holds with $\gamma = 2m$, and that the power constant of $\varepsilon$ in the estimate (5.130) is in accordance with the estimate (5.126). Note also that here

$$[F(x', \xi) - f(x')] - [F(x, \xi) - f(x)] = 2m \xi^T (x - x')$$

has normal distribution with zero mean and variance $4m^2 \sigma^2 \|x' - x\|^2$. Consequently, assumption (M6) holds with $\lambda^2 = 4m^2 \sigma^2$.

Of course, in this example the “true” optimal solution is $\bar{x} = 0$, and one does not need sampling in order to solve this problem. Note, however, that the sample average function $\hat{f}_N(x)$ here depends on the random sample only through the data average vector $\bar{\xi}_N$. Therefore, any numerical procedure based on averaging will need a sample of size $N$ satisfying the estimate (5.130) in order to produce an $\varepsilon$-optimal solution.

5.3.3 Finite Exponential Convergence

We assume in this section that the problem is convex and the expectation function $f(x)$ is finite valued.

**Definition 5.22.** It is said that $x^* \in X$ is a sharp (optimal) solution, of the true problem (5.1), if there exists constant $c > 0$ such that

$$f(x) \geq f(x^*) + c\|x - x^*\|, \ \forall x \in X.$$ (5.131)

The above condition (5.131) corresponds to the growth condition (5.125) with the power constant $\gamma = 1$ and $S = \{x^*\}$. Since $f(\cdot)$ is convex finite valued we have that the directional derivatives $f'(x^*, h)$ exist for all $h \in \mathbb{R}^n$, $f'(x^*, \cdot)$ is (locally Lipschitz) continuous and formula (7.17) holds. Also by convexity of the set $X$ we have that the tangent cone $T_X(x^*)$, to $X$ at $x^*$, is given by the topological closure of the corresponding radial cone. By using these facts it is not difficult to show that condition (5.131) is equivalent to:

$$f'(x^*, h) \geq c\|h\|, \ \forall h \in T_X(x^*).$$ (5.132)

Since the above condition (5.132) is local in nature, we have that it actually suffices to verify (5.131) for all $x \in X$ in a neighborhood of $x^*$.

**Theorem 5.23.** Suppose that the problem is convex and assumption (M1) holds, and let $x^* \in X$ be a sharp optimal solution of the true problem. Then $S_N = \{x^*\}$ w.p.1 for $N$ large enough. Suppose, further, that assumption (M4) holds. Then there exist constants $C > 0$ and $\beta > 0$ such that

$$1 - \Pr(S_N = \{x^*\}) \leq Ce^{-N\beta},$$ (5.133)

i.e., probability of the event that “$x^*$ is the unique optimal solution of the SAA problem” converges to one exponentially fast with increase of the sample size $N$. 

Proof. By convexity of $F(\cdot, \xi)$ we have that $\hat{f}_N(x^*, \cdot)$ converges to $f'(x^*, \cdot)$ w.p.1 uniformly on the unit sphere (see the proof of Theorem 7.54). It follows w.p.1 for $N$ large enough that

$$\hat{f}_N(x^*, h) \geq (c/2)\|h\|, \quad \forall h \in T_X(x^*),$$

(5.134)

which implies that $x^*$ is the sharp optimal solution of the corresponding SAA problem.

Now under the assumptions of convexity and (M1) and (M4), we have that $\hat{f}_N(x^*, \cdot)$ converges to $f'(x^*, \cdot)$ exponentially fast on the unit sphere (see inequality (7.225) of Theorem 7.69). By taking $\varepsilon := c/2$ in (7.225), we can conclude that (5.133) follows. □

It is also possible to consider the growth condition (5.125) with $\gamma = 1$ and the set $S$ not necessarily being a singleton. That is, it is said that the set $S$ of optimal solutions of the true problem is sharp if for some $c > 0$ the following condition holds

$$f(x) \geq \vartheta^* + c[\text{dist}(x, S)], \quad \forall x \in X.$$  

(5.135)

Of course, if $S = \{x^*\}$ is a singleton, then conditions (5.131) and (5.135) do coincide. The set of optimal solutions of the true problem is always nonempty and sharp if its optimal value is finite and the problem is piecewise linear in the sense that the following conditions hold:

(P1) The set $X$ is a convex closed polyhedron.

(P2) The support set $\Xi = \{\xi_1, \ldots, \xi_K\}$ is finite.

(P3) For every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is polyhedral.

The above conditions (P1)–(P3) hold in case of two-stage linear stochastic programming problems with a finite number of scenarios.

Under conditions (P1)–(P3) the true and SAA problems are polyhedral, and hence their sets of optimal solutions are polyhedral. By using polyhedral structure and finiteness of the set $\Xi$ it is possible to show the following result (cf., [208]).

Theorem 5.24. Suppose that conditions (P1)–(P3) hold and the set $S$ is nonempty and bounded. Then $S$ is polyhedral and there exist constants $C > 0$ and $\beta > 0$ such that

$$1 - \Pr(\hat{S}_N \neq \emptyset \text{ and } \hat{S}_N \text{ is a face of } S) \leq C e^{-N^{1/3}},$$

(5.136)

i.e., probability of the event that “$\hat{S}_N$ is nonempty and forms a face of the set $S$” converges to one exponentially fast with increase of the sample size $N$.

5.4 Quasi-Monte Carlo Methods

In the previous section we discussed an approach to evaluating (approximating) expectations by employing random samples generated by Monte Carlo techniques. It should be understood, however, that when dimension $d$ (of the random data vector $\xi$) is small, the Monte Carlo approach may not be a best way to proceed. In this section we give a brief discussion of the so-called quasi-Monte Carlo methods. It is beyond the scope of this book.
5.4. Quasi-Monte Carlo Methods

to give a detailed discussion of that subject. This section is based on Niederreiter [138], to
which the interested reader is referred for a further reading on that topic. Let us start our
discussion by considering a one dimensional case (of $d = 1$).

Let $\xi$ be a real valued random variable having cdf $H(z) = \Pr(\xi \leq z)$. Suppose that
we want to evaluate the expectation

$$E[F(\xi)] = \int_{-\infty}^{+\infty} F(z) dH(z), \quad (5.137)$$

where $F : \mathbb{R} \to \mathbb{R}$ is a measurable function. Let $U \sim U[0,1]$, i.e., $U$ is a random
variable uniformly distributed on $[0,1]$. Then random variable $5 \cdot H^{-1}(U)$ has cdf $H(\cdot)$. Therefore,
by making change of variables we can write the expectation (5.137) as

$$E[\psi(U)] = \int_{0}^{1} \psi(u) du, \quad (5.138)$$

where $\psi(u) := F(H^{-1}(u))$.

Evaluation of the above expectation by Monte Carlo method is based on generating an
iid sample $U^1, ..., U^N$, of $N$ replications of $U \sim U[0,1]$, and consequently approximating
$E[\psi(U)]$ by the average $\hat{\psi}_N := N^{-1} \sum_{j=1}^{N} \psi(U^j)$. Alternatively, one can employ the
Riemann sum approximation

$$\int_{0}^{1} \psi(u) du \approx \frac{1}{N} \sum_{j=1}^{N} \psi(u_j) \quad (5.139)$$

by using some points $u_j \in [(j-1)/N, j/N]$, $j = 1, ..., N$, e.g., taking midpoints $u_j :=
(2j - 1)/(2N)$ of equally spaced partition intervals $[(j-1)/N, j/N]$, $j = 1, ..., N$. If the
function $\psi(u)$ is Lipschitz continuous on $[0,1]$, then the error of the Riemann sum approxima-
tion$^6$ is of order $O(N^{-1})$, while the Monte Carlo sample average error is of (stochas-
tic) order $O_p(N^{-1/2})$. An explanation of this phenomenon is rather clear, an iid sample
$U^1, ..., U^N$ will tend to cluster in some areas while leaving other areas of the interval $[0,1]$
uncovered.

One can argue that the Monte Carlo sampling approach has an advantage of the possi-
bility of estimating the approximation error by calculating the sample variance

$$s^2 := (N - 1)^{-1} \sum_{j=1}^{N} [\psi(U^j) - \hat{\psi}_N]^2,$$

and consequently constructing a corresponding confidence interval. It is possible, however,
to employ a similar procedure for the Riemann sums by making them random. That is,
each point $u_j$, in the right hand side of (5.139), is generated randomly, say uniformly

---

$^3$Recall that $H^{-1}(u) := \inf \{ z : H(z) \geq u \}$.

$^6$If $\psi(u)$ is continuously differentiable, then, e.g., the trapezoidal rule gives even a slightly better approximation error of order $O(N^{-2})$. Also one should be careful in making the assumption of Lipschitz continuity of $\psi(u)$. If the distribution of $\xi$ is supported on the whole real line, e.g., is normal, then $H^{-1}(u)$ tends to $\infty$ as $u$ tends to 0 or 1. In that case $\psi(u)$ typically will be discontinuous at $u = 0$ and $u = 1$. 

distributed, on the corresponding interval \([j - 1]/N, j/N\], independently of other points \(u_k, k \neq j\). This will make the right hand side of (5.139) a random variable. Its variance can be estimated by using several independently generated batches of such approximations.

It does not make sense to use Monte Carlo sampling methods in case of one dimensional random data. The situation starts to change quickly with increase of the dimension \(d\). By making an appropriate transformation we may assume that the random data vector is distributed uniformly on the \(d\)-dimensional cube \(I^d = [0, 1]^d\). For \(d > 1\) we denote by \((\text{bold-faced})\) \(U\) a random vector uniformly distributed on \(I^d\). Suppose that we want to evaluate the expectation \(E[\psi(U)] = \int_{I^d} \psi(u) du\), where \(\psi : I^d \to \mathbb{R}\) is a measurable function. We can partition each coordinate of \(I^d\) into \(M\) equally spaced intervals, and hence partition \(I^d\) into the corresponding \(N = M^d\) subintervals\(^7\) and use a corresponding Riemann sum approximation \(N^{-1} \sum_{j=1}^{N} \psi(u_j)\). The resulting error is of order \(O(M^{-1})\), provided that the function \(\psi(u)\) is Lipschitz continuous. In terms of the total number \(N\) of function evaluations, this error is of order \(O(N^{-1/d})\). For \(d = 2\) it is still compatible with the Monte Carlo sample average approximation approach. However, for larger values of \(d\) the Riemann sums approach quickly becomes unacceptable. On the other hand, the rate of convergence (error bounds) of the Monte Carlo sample average approximation of \(E[\psi(U)]\) does not depend directly on dimensionality \(d\), but only on the corresponding variance \(\text{Var}[\psi(U)]\). Yet the problem of uneven covering of \(I^d\) by an iid sample \(U^N\), \(j = 1, \ldots, N\), remains persistent.

Quasi-Monte Carlo methods employ the approximation

\[
E[\psi(U)] \approx \frac{1}{N} \sum_{j=1}^{N} \psi(u_j)
\]

for a carefully chosen (deterministic) sequence of points \(u_1, \ldots, u_N \in I^d\). From the numerical point of view it is important to be able to generate such a sequence iteratively as an infinite sequence of points \(u_j, j = 1, \ldots, \) in \(I^d\). In that way one does not need to recalculate already calculated function values \(\psi(u_j)\) with increase of \(N\). A basic requirement for this sequence is that the right hand side of (5.140) converges to \(E[\psi(U)]\) as \(N \to \infty\). It is not difficult to show that this holds (for any Riemann-integrable function \(\psi(u)\)) if

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} 1_A(u_j) = V_d(A)
\]

for any interval \(A \subset I^d\). Here \(V_d(A)\) denotes the \(d\)-dimensional Lebesgue measure (volume) of set \(A \subset \mathbb{R}^d\).

**Definition 5.25.** The star discrepancy of a point set \(\{u_1, \ldots, u_N\} \subset I^d\) is defined by

\[
D^*(u_1, \ldots, u_N) := \sup_{A \in \mathcal{I}} \left| \frac{1}{N} \sum_{j=1}^{N} 1_A(u_j) - V_d(A) \right|,
\]

where \(\mathcal{I}\) is the family of all subintervals of \(I^d\) of the form \(\prod_{i=1}^{d} [0, b_i]\).

\(^{7}\)A set \(A \subset \mathbb{R}^d\) is said to be a \((d\text{-dimensional})\) interval if \(A = [a_1, b_1] \times \cdots \times [a_d, b_d]\).
5.4. Quasi-Monte Carlo Methods

It is possible to show that for a sequence $u_j \in I_d$, $j = 1, \ldots$, condition (5.141) holds iff $\lim_{N \to \infty} D^*(u_1, \ldots, u_N) = 0$. A more important property of the star discrepancy is that it is possible to give error bounds in terms of $D^*(u_1, \ldots, u_N)$ for quasi-Monte Carlo approximations. Let us start with one dimensional case. Recall that variation of a function $\psi : [0, 1] \to \mathbb{R}$ is the $\sup \sum_{i=1}^{m} |\psi(t_i) - \psi(t_{i-1})|$, where the supremum is taken over all partitions $0 = t_0 < t_1 < \ldots < t_m = 1$ of the interval $[0,1]$. It is said that $\psi$ has bounded variation if its variation is finite.

**Theorem 5.26 (Koksma).** If $\psi : [0, 1] \to \mathbb{R}$ has bounded variation $V(\psi)$, then for any $u_1, \ldots, u_N \in [0, 1]$ we have

$$\left| \frac{1}{N} \sum_{j=1}^{N} \psi(u_j) - \int_{0}^{1} \psi(u)du \right| \leq V(\psi)D^*(u_1, \ldots, u_N).$$ (5.143)

**Proof.** We can assume that the sequence $u_1, \ldots, u_N$ is arranged in the increasing order, and set $u_0 = 0$ and $u_{N+1} = 1$. That is, $0 = u_0 \leq u_1 \leq \ldots \leq u_{N+1} = 1$. Using integration by parts we have

$$\int_{0}^{1} \psi(u)du = u\psi(u)|_{0}^{1} - \int_{0}^{1} u\psi'(u) = \psi(1) - \int_{0}^{1} u\psi'(u),$$

and summation by parts

$$\frac{1}{N} \sum_{j=1}^{N} \psi(u_j) = \psi(u_{N+1}) - \sum_{j=0}^{N} \frac{j}{N} [\psi(u_{j+1}) - \psi(u_j)],$$

we can write

$$\frac{1}{N} \sum_{j=1}^{N} \psi(u_j) - \int_{0}^{1} \psi(u)du = -\sum_{j=0}^{N} \frac{j}{N} [\psi(u_{j+1}) - \psi(u_j)] + \int_{0}^{1} u\psi'(u)$$

$$= \sum_{j=0}^{N} \int_{u_j}^{u_{j+1}} (u - \frac{j}{N}) d\psi(u).$$

Also for any $u \in [u_j, u_{j+1}], j = 0, \ldots, N$, we have

$$\left| u - \frac{j}{N} \right| \leq D^*(u_1, \ldots, u_N).$$

It follows that

$$\left| \frac{1}{N} \sum_{j=1}^{N} \psi(u_j) - \int_{0}^{1} \psi(u)du \right| \leq \sum_{j=0}^{N} \int_{u_j}^{u_{j+1}} |u - \frac{j}{N}| d\psi(u)$$

$$\leq D^*(u_1, \ldots, u_N) \sum_{j=0}^{N} |\psi(u_{j+1}) - \psi(u_j)|,$$

and, of course, $\sum_{j=0}^{N} |\psi(u_{j+1}) - \psi(u_j)| \leq V(\psi)$. This completes the proof. ☐

This can be extended to a multidimensional setting as follows. Consider a function $\psi : I^d \to \mathbb{R}$. The variation of $\psi$, in the sense of Vitali, is defined as

$$V^{(d)}(\psi) := \sup_{P \in J} \sum_{A \in P} |\Delta \psi(A)|,$$ (5.144)
where \( \mathcal{F} \) denotes the family of all partitions \( \mathcal{P} \) of \( I^d \) into subintervals, and for \( A \in \mathcal{P} \) the notation \( \Delta_\psi(A) \) stands for an alternating sum of the values of \( \psi \) at the vertices of \( A \) (i.e., function values at adjacent vertices have opposite signs). The variation of \( \psi \), in the sense of Hardy and Krause, is defined as

\[
V(\psi) := \sum_{k=1}^{d} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq d} V^{(k)}(\psi; i_1, \ldots, i_k),
\]  

(5.145)

where \( V^{(k)}(\psi; i_1, \ldots, i_k) \) is the variation in the sense of Vitali of restriction of \( \psi \) to the \( k \)-dimensional face of \( I^d \) defined by \( u_j = 1 \) for \( j \notin \{i_1, \ldots, i_k\} \).

**Theorem 5.27 (Hlawka).** If \( \psi : I^d \to \mathbb{R} \) has bounded variation \( V(\psi) \) on \( I^d \) in the sense of Hardy and Krause, then for any \( u_1, \ldots, u_N \in I^d \) we have

\[
\left| \frac{1}{N} \sum_{j=1}^{N} \psi(u_j) - \int_{I^d} \psi(u) du \right| \leq V(\psi)D^*(u_1, \ldots, u_N).
\]

(5.146)

In order to see how good the above error estimates could be let us consider the one dimensional case with \( u_j := (2j - 1)/(2N) \), \( j = 1, \ldots, N \). Then \( D^*(u_1, \ldots, u_N) = 1/(2N) \), and hence the estimate (5.143) leads to the error bound \( V(\psi)/2 \). This error bound gives the correct order \( O(N^{-1}) \) for the error estimates (provided that \( \psi \) has bounded variation), but the involved constant \( V(\psi)/2 \) typically is far too large for practical calculations. Even worse, the inverse function \( H^{-1}(u) \) is monotonically nondecreasing and hence its variation is given by the difference of the limits \( \lim_{u \to +\infty} H^{-1}(u) \) and \( \lim_{u \to -\infty} H^{-1}(u) \). Therefore, if one of these limits is infinite, i.e., the support of the corresponding random variable is unbounded, then the associated variation is infinite. Typically this variation unboundedness will carry over to the function \( \psi(u) = F(H^{-1}(u)) \). For example, if the function \( F(\cdot) \) is monotonically nondecreasing, then

\[
V(\psi) = F\left( \lim_{u \to +\infty} H^{-1}(u) \right) - F\left( \lim_{u \to -\infty} H^{-1}(u) \right).
\]

This overestimation of the corresponding constant becomes even worse with increase of the dimension \( d \).

A sequence \( \{u_j\}_{j \in \mathbb{N}} \subseteq I^d \) is called a low-discrepancy sequence if \( D^*(u_1, \ldots, u_N) \) is “small” for all \( N \geq 1 \). We proceed now to a description of classical constructions of low-discrepancy sequences. Let us start with the one dimensional case. It is not difficult to show that \( D^*(u_1, \ldots, u_N) \) always greater than or equal to \( 1/(2N) \) and this lower bound is attained for \( u_j := (2j - 1)/2N \), \( j = 1, \ldots, N \). While the lower bound of order \( O(N^{-1}) \) is attained for some \( N \)-element point sets from \( [0,1] \), there does not exist a sequence \( u_1, \ldots, \) in \( [0,1] \) such that \( D^*(u_1, \ldots, u_N) \leq c/N \) for some \( c > 0 \) and all \( N \in \mathbb{N} \). It is possible to show that a best possible rate for \( D^*(u_1, \ldots, u_N) \), for a sequence of points \( u_j \in [0,1] \), \( j = 1, \ldots, \) is of order \( O \left( N^{-1} \ln N \right) \). We are going to construct now a sequence for which this rate is attained.

For any integer \( n \geq 0 \) there is a unique digit expansion

\[
n = \sum_{i \geq 0} a_i(n)b^i
\]

(5.147)
5.4. Quasi-Monte Carlo Methods

in integer base $b \geq 2$, where $a_i(n) \in \{0, 1, ..., b-1\}$, $i = 0, 1, ..., \text{and } a_i(n) = 0$ for all $i$ large enough, i.e., the sum (5.147) is finite. The associated radical-inverse function $\phi_b(n)$, in base $b$, is defined by

$$\phi_b(n) := \sum_{i \geq 0} a_i(n)b^{-i-1}.$$  \hfill (5.148)

Note that

$$\phi_b(n) \leq (b-1) \sum_{i=0}^{\infty} b^{-i-1} = 1,$$

and hence $\phi_b(n) \in [0, 1]$ for any integer $n \geq 0$.

**Definition 5.28.** For an integer $b \geq 2$, the van der Corput sequence in base $b$ is the sequence $u_j := \phi_b(j), j = 0, 1, ...$

It is possible to show that to every van der Corput sequence $u_1, ..., u_n$, in base $b$, corresponds constant $C_b$ such that

$$D^*(u_1, ..., u_n) \leq C_b N^{-1-\ln N}, \text{ for all } N \in \mathbb{N}.$$  \hfill (5.149)

A classical extension of van der Corput sequences to multidimensional settings is the following. Let $p_1 = 2, p_2 = 3, ..., p_d$ be the first $d$ prime numbers. Then the Halton sequence, in the bases $p_1, ..., p_d$, is defined as

$$u_j := (\phi_{p_1}(j), ..., \phi_{p_d}(j)) \in I^d, j = 0, 1, ...$$  \hfill (5.150)

It is possible to show that for that sequence

$$D^*(u_1, ..., u_N) \leq A_d N^{-1} (\ln N)^d + O(N^{-1} (\ln N)^{d-1}), \text{ for all } N \geq 2,$$  \hfill (5.151)

where $A_d = \prod_{i=1}^{d} \frac{p_{i+1}-1}{2p_i p_{i+1}}$. By the bound (5.146) of Theorem 5.27, this implies that the error of the corresponding quasi-Monte Carlo approximation is of order $O(N^{-1}(\ln N)^d)$, provided that variation $V(\psi)$ is finite. This compares favorably with the bound $O_{\psi}(N^{-1/2})$ of the Monte Carlo sampling. Note, however, that by the prime number theorem we have that $\frac{\ln N}{\ln d}$ tends to 1 as $d \to \infty$. That is, the coefficient $A_d$, of the leading term in the right hand side of (5.150), grows superexponentially with increase of the dimension $d$. This makes the corresponding error bounds useless for larger values of $d$. It should be noticed that the above are upper bounds for the rates of convergence and in practice convergence rates could be much better. It seems that for low dimensional problems, say $d \leq 20$, quasi-Monte Carlo methods are advantageous over Monte Carlo methods. With increase of the dimension $d$ this advantage becomes less apparent. Of course, all this depends on a particular class of problems and applied quasi-Monte Carlo method. This issue requires a further investigation.

A drawback of (deterministic) quasi-Monte Carlo sequences $\{u_j\}_{j \in \mathbb{N}}$ is that there is no easy way to estimate the error of the corresponding approximations $N^{-1} \sum_{j=1}^{N} \psi(u_j)$. In that respect bounds like (5.146) typically are too loose and impossible to calculate anyway. A way of dealing with this problem is to use a randomization of the set $\{u_1, ..., u_N\}$,
of generating points in $I^d$, without destroying its regular structure. Such simple randomization procedure was suggested by Cranley and Patterson [39]. That is, generate a random point $\mathbf{u}$ uniformly distributed over $I^d$, and use the randomization $\tilde{\mathbf{u}}_j := (u_j + \mathbf{u}) \mod 1$, $j = 1, ..., N$. It is not difficult to show that (marginal) distribution of each random vector $\tilde{\mathbf{u}}_j$ is uniform on $I^d$. Therefore each $\psi(\tilde{\mathbf{u}}_j)$, and hence $N^{-1} \sum_{j=1}^{N} \psi(\tilde{\mathbf{u}}_j)$, is an unbiased estimator of the corresponding expectation $E[\psi(U)]$. Variance of the estimator $N^{-1} \sum_{j=1}^{N} \psi(\tilde{\mathbf{u}}_j)$ can be significantly smaller than variance of the corresponding Monte Carlo estimator based on samples of the same size. This randomization procedure can be applied in batches. That is, it can be repeated $M$ times for independently generated uniformly distributed vectors $\mathbf{u} = \mathbf{u}_i$, $i = 1, ..., M$, and consequently averaging the obtained replications of $N^{-1} \sum_{j=1}^{N} \psi(\tilde{\mathbf{u}}_j)$. Simultaneously, variance of this estimator can be evaluated by calculating the sample variance of the obtained $M$ independent replications of $N^{-1} \sum_{j=1}^{N} \psi(\tilde{\mathbf{u}}_j)$.

5.5 Variance Reduction Techniques

Consider the sample average estimators $\hat{f}_N(x)$. We have that if the sample is iid, then the variance of $\hat{f}_N(x)$ is equal to $\sigma^2(x)/N$, where $\sigma^2(x) := \text{Var}[F(x, \xi)]$. In some cases it is possible to reduce the variance of generated sample averages, which in turn enhances convergence of the corresponding SAA estimators. In section 5.4 we discussed Quasi-Monte Carlo techniques for enhancing rates of convergence of sample average approximations. In this section we briefly discuss some other variance reduction techniques which seem to be useful in the SAA method.

5.5.1 Latin Hypercube Sampling

Suppose that the random data vector $\xi = \xi(\omega)$ is one dimensional with the corresponding cumulative distribution function (cdf) $H(\cdot)$. We can write then

$$E[F(x, \xi)] = \int_{-\infty}^{+\infty} F(x, \xi) dH(\xi). \quad (5.151)$$

In order to evaluate the above integral numerically it will be much better to generate sample points evenly distributed than to use an iid sample (we already discussed this in section 5.4). That is, we can generate independent random points $^9$

$$U^j \sim U[(j - 1)/N, j/N], \quad j = 1, ..., N, \quad (5.152)$$

and then to construct the random sample of $\xi$ by the inverse transformation $\xi^j := H^{-1}(U^j)$, $j = 1, ..., N$ (compare with (5.138)).

Now suppose that $j$ is chosen at random from the set $\{1, ..., N\}$ (with equal probability for each element of that set). Then conditional on $j$, the corresponding random

---

^9For a number $a \in \mathbb{R}$ the notation “$a \mod 1$” denotes the fractional part of $a$, i.e., $a \mod 1 = a - \lfloor a \rfloor$, where $\lfloor a \rfloor$ denotes the largest integer less than or equal to $a$. In vector case the “modulo 1” reduction is understood coordinatewise.

^For an interval $[a, b] \subset \mathbb{R}$, we denote by $U[a, b]$ the uniform probability distribution on that interval.
5.5. Variance Reduction Techniques

variable \( U^j \) is uniformly distributed on the interval \([ (j-1)/N, j/N] \), and the unconditional distribution of \( U^j \) is uniform on the interval \([ 0, 1] \). Consequently, let \( \{ j_1, \ldots, j_N \} \) be a random permutation of the set \( \{ 1, \ldots, N \} \). Then the random variables \( \xi^{j_1}, \ldots, \xi^{j_N} \) have the same marginal distribution, with the same cdf \( H(\cdot) \), and are negatively correlated with each other. Therefore the expected value of

\[
\hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^{N} F(x, \xi^i) = \frac{1}{N} \sum_{s=1}^{N} F(x, \xi^{j_s}) \tag{5.153}
\]

is \( f(x) \), while

\[
\text{Var} \left[ \hat{f}_N(x) \right] = N^{-1} \sigma^2(x) + 2N^{-2} \sum_{s<t} \text{Cov} \left( F(x, \xi^{j_s}), F(x, \xi^{j_t}) \right). \tag{5.154}
\]

If the function \( F(x, \cdot) \) is monotonically increasing or decreasing, than the random variables \( F(x, \xi^{j_s}) \) and \( F(x, \xi^{j_t}) \), \( s \neq t \), are also negatively correlated. Therefore, the variance of \( \hat{f}_N(x) \) tends to be smaller, and in some cases much smaller, than \( \sigma^2(x)/N \).

Suppose now that the random vector \( \xi = (\xi_1, \ldots, \xi_d) \) is \( d \)-dimensional, and that its components \( \xi_i, i = 1, \ldots, d \), are distributed independently of each other. Then we can use the above procedure for each component \( \xi_i \). That is, a random sample \( U^j \) of the form (5.152) is generated, and consequently \( N \) replications of the first component of \( \xi \) are computed by the corresponding inverse transformation applied to randomly permuted \( U^j \). The same procedure is applied to every component of \( \xi \) with the corresponding random samples of the form (5.152) and random permutations generated independently of each other. This sampling scheme is called the Latin Hypercube (LH) sampling.

If the function \( F(x, \cdot) \) is decomposable, i.e., \( F(x, \xi) := F_1(x, \xi_1) + \ldots + F_d(x, \xi_d) \), then \( \mathbb{E}[F(x, \xi)] = \mathbb{E}[F_1(x, \xi_1)] + \ldots + \mathbb{E}[F_d(x, \xi_d)] \), where each expectation is calculated with respect to a one dimensional distribution. In that case the LH sampling ensures that each expectation \( \mathbb{E}[F_i(x, \xi_i)] \) is estimated in nearly optimal way. Therefore, the LH sampling works especially well in cases where the function \( F(x, \cdot) \) has a somewhat decomposable structure. In any case the LH sampling procedure is easy to implement and can be applied to SAA optimization procedures in a straightforward way. Since in LH sampling the random replications of \( F(x, \xi) \) are correlated with each other, one cannot use variance estimates like (5.20). Therefore the LH method usually is applied in several independent batches in order to estimate variance of the corresponding estimators.

5.5.2 Linear Control Random Variables Method

Suppose that we have a measurable function \( A(x, \xi) \) such that \( \mathbb{E}[A(x, \xi)] = 0 \) for all \( x \in X \). Then, for any \( t \in \mathbb{R} \), the expected value of \( F(x, \xi) + tA(x, \xi) \) is \( f(x) \), while

\[
\text{Var} \left[ F(x, \xi) + tA(x, \xi) \right] = \text{Var} \left[ F(x, \xi) \right] + t^2 \text{Var} \left[ A(x, \xi) \right] + 2t \text{Cov} \left( F(x, \xi), A(x, \xi) \right). \]

It follows that the above variance attains its minimum, with respect to \( t \), for

\[
t^*: = -\rho_{F,A}(x) \left[ \frac{\text{Var}(F(x, \xi))}{\text{Var}(A(x, \xi))} \right]^{1/2}, \tag{5.155}
\]
where \( \rho_{F,A}(x) := \text{Corr}(F(x, \xi), A(x, \xi)) \), and with
\[
\text{Var}[F(x, \xi) + t^* A(x, \xi)] = \text{Var}[F(x, \xi)] \left[ 1 - \rho_{F,A}(x)^2 \right].
\] (5.156)

For a given \( x \in X \) and generated sample \( \xi^1, \ldots, \xi^N \), one can estimate, in the standard way, the covariance and variances appearing in the right hand side of (5.155), and hence to construct an estimate \( \hat{t} \) of \( t^* \). Then \( f(x) \) can be estimated by
\[
\hat{f}_N^A(x) := \frac{1}{N} \sum_{j=1}^{N} \left[ F(x, \xi^j) + \hat{t} A(x, \xi^j) \right].
\] (5.157)

By (5.156), the linear control estimator \( \hat{f}_N^A(x) \) has a smaller variance than \( \hat{f}_N(x) \) if \( F(x, \xi) \) and \( A(x, \xi) \) are highly correlated with each other.

Let us make the following observations. The estimator \( \hat{t} \), of the optimal value \( t^* \), depends on \( x \) and the generated sample. Therefore, it is difficult to apply linear control estimators in an SAA optimization procedure. That is, linear control estimators are mainly suitable for estimating expectations at a fixed point. Also if the same sample is used in estimating \( \hat{t} \) and \( \hat{f}_N^A(x) \), then \( \hat{f}_N^A(x) \) can be a slightly biased estimator of \( f(x) \).

Of course, the above Linear Control procedure can be successful only if a function \( A(x, \xi) \), with mean zero and highly correlated with \( F(x, \xi) \), is available. Choice of such a function is problem dependent. For instance, one can use a linear function \( A(x, \xi) := \lambda(\xi)^T x \). Consider, for example, two-stage stochastic programming problems with recourse of the form (2.1)–(2.2). Suppose that the random vector \( h = h(\omega) \) and matrix \( T = T(\omega) \), in the second stage problem (2.2), are independently distributed, and let \( \mu := E[h] \). Then
\[
E \left[ (h - \mu)^T T \right] = E \left[ (h - \mu)^T \right] E \left[ T \right] = 0,
\]
and hence one can use \( A(x, \xi) := (h - \mu)^T T x \) as the control variable.

Let us finally remark that the above procedure can be extended in a straightforward way to a case where several functions \( A_1(x, \xi), \ldots, A_m(x, \xi) \), each with zero mean and highly correlated with \( F(x, \xi) \), are available.

### 5.5.3 Importance Sampling and Likelihood Ratio Methods

Suppose that \( \xi \) has a continuous distribution with probability density function (pdf) \( h(\cdot) \). Let \( \psi(\cdot) \) be another pdf such that the so-called likelihood ratio function \( L(\cdot) := \frac{h(\cdot)}{\psi(\cdot)} \) is well defined. That is, if \( \psi(z) = 0 \) for some \( z \in \mathbb{R}^d \), then \( h(z) = 0 \), and by the definition, \( 0/0 = 0 \), i.e., we do not divide a positive number by zero. Then we can write
\[
f(x) = \int F(x, \xi) h(\xi) \, d\xi = \int F(x, \xi) L(\xi) \psi(\xi) \, d\xi = E_{\psi}[F(x, Z) L(Z)],
\] (5.158)
where the integration is performed over the space \( \mathbb{R}^d \) and the notation \( E_{\psi} \) emphasizes that the expectation is taken with respect to the random vector \( Z \) having pdf \( \psi(\cdot) \).

Let us show that, for a fixed \( x \), the variance of \( F(x, Z) L(Z) \) attains its minimal value for \( \psi(\cdot) \) proportional to \( |F(x, \cdot) h(\cdot)| \), i.e., for
\[
\psi^*(\cdot) := \frac{|F(x, \cdot) h(\cdot)|}{\int |F(x, \zeta) h(\zeta)| d\zeta^*}.
\] (5.159)
Since $\mathbb{E}_\psi[F(x, Z)L(Z)] = f(x)$ and does not depend on $\psi(\cdot)$, we have that the variance of $F(x, Z)L(Z)$ is minimized if

$$
\mathbb{E}_\psi[F(x, Z)^2L(Z)^2] = \int \frac{F(x, \zeta)^2 h(\zeta)^2}{\psi(\zeta)} d\zeta,
$$

is minimized. Furthermore, by Cauchy inequality we have

$$
\left( \int |F(x, \zeta) h(\zeta)| d\zeta \right)^2 \leq \left( \int \frac{F(x, \zeta)^2 h(\zeta)^2}{\psi(\zeta)} d\zeta \right) \left( \int \psi(\zeta) d\zeta \right).
$$

It remains to note that $\int \psi(\zeta) d\zeta = 1$ and the left hand side of (5.161) is equal to the expected value of squared $F(x, Z)L(Z)$ for $\psi(\cdot) = \psi^*(\cdot)$.

Note that if $F(x, \cdot)$ is nonnegative valued, then $\psi^*(\cdot) = F(x, \cdot) h(\cdot)/f(x)$ and for that choice of the pdf $\psi(\cdot)$, the function $F(x, \cdot)L(\cdot)$ is identically equal to $f(x)$. Of course, in order to achieve such absolute variance reduction to zero we need to know the expectation $f(x)$ which was our goal in the first place. Nevertheless, it gives the idea that if we can construct a pdf $\psi(\cdot)$ roughly proportional to $|F(x, \cdot) h(\cdot)|$, then we may achieve a considerable variance reduction by generating a random sample $\zeta^1, \ldots, \zeta^N$ from the pdf $\psi(\cdot)$, and then estimating $f(x)$ by

$$
\tilde{f}_{N}(x) := \frac{1}{N} \sum_{j=1}^{N} F(x, \zeta^j) L(\zeta^j).
$$

The estimator $\tilde{f}_{N}(x)$ is an unbiased estimator of $f(x)$ and may have significantly smaller variance than $\tilde{f}(x)$ depending on a successful choice of the pdf $\psi(\cdot)$.

Similar analysis can be performed in cases where $\xi$ has a discrete distribution by replacing the integrals with the corresponding summations.

Let us remark that the above approach, called importance sampling, is extremely sensitive to a choice of the pdf $\psi(\cdot)$ and is notorious for its instability. This is understandable since the likelihood ratio function in the tail is the ratio of two very small numbers. For a successful choice of $\psi(\cdot)$, the method may work very well while even a small perturbation of $\psi(\cdot)$ may be disastrous. This is why a single choice of $\psi(\cdot)$ usually does not work for different points $x$, and consequently cannot be used for a whole optimization procedure.

Note also that $\mathbb{E}_\psi[L(Z)] = 1$. Therefore, $L(\zeta) - 1$ can be used as a linear control variable for the likelihood ratio estimator $\tilde{f}_{N}(x)$.

In some cases it is also possible to use the likelihood ratio method for estimating first and higher order derivatives of $f(x)$. Consider, for example, the optimal value $Q(x, \xi)$ of the second stage linear program (2.2). Suppose that the vector $q$ and matrix $W$ are fixed, i.e., not stochastic, and for the sake of simplicity that $h = h(\omega)$ and $T = T(\omega)$ are distributed independently of each other. We have then that $Q(x, \xi) = Q(h - T x)$, where

$$
Q(z) := \inf \left\{ q^T y : W y = z, \ y \geq 0 \right\}.
$$

Suppose, further, that $h$ has a continuous distribution with pdf $\eta(\cdot)$. We have that

$$
\mathbb{E}[Q(x, \xi)] = \mathbb{E}_T \left\{ \mathbb{E}_{h \mid T}[Q(x, \xi)] \right\},
$$
and by using the transformation \( z = h - T x \), since \( h \) and \( T \) are independent we obtain

\[
\begin{align*}
\mathbb{E}_{h|T}[Q(x, \xi)] &= \mathbb{E}_h[Q(x, \xi)] \\
&= \int Q(h - T x) \eta(h) dh = \int Q(z) \eta(z + T x) dz \\
&= \int Q(\zeta) L(x, \zeta) \psi(\zeta) d\zeta = \mathbb{E}_\psi [L(x, Z) Q(Z)],
\end{align*}
\]

where \( \psi(\cdot) \) is a chosen pdf and \( L(x, \zeta) := \eta(\zeta + T x)/\psi(\zeta) \). If the function \( \eta(\cdot) \) is smooth, then the likelihood ratio function \( L(\cdot, \zeta) \) is also smooth. In that case, under mild additional conditions, first and higher order derivatives can be taken inside the expected value in the right hand side of (5.163), and consequently can be estimated by sampling. Note that the first order derivatives of \( Q(\cdot, \xi) \) are piecewise constant, and hence its second order derivatives are zeros whenever defined. Therefore, second order derivatives cannot be taken inside the expectation \( \mathbb{E}[Q(x, \xi)] \) even if \( \xi \) has a continuous distribution.

### 5.6 Validation Analysis

Suppose that we are given a feasible point \( \bar{x} \in X \) as a candidate for an optimal solution of the true problem. For example, \( \bar{x} \) can be an output of a run of the corresponding SAA problem. In this section we discuss ways to evaluate quality of this candidate solution. This is important, in particular, for a choice of the sample size and stopping criteria in simulation based optimization. There are basically two approaches to such validation analysis. We can either try to estimate the optimality gap \( f(\bar{x}) - \vartheta^* \) between the objective value at the considered point \( \bar{x} \) and the optimal value of the true problem, or to evaluate first order (KKT) optimality conditions at \( \bar{x} \).

Let us emphasize that the following analysis is designed for the situations where the value \( f(\bar{x}) \), of the true objective function at the considered point, is finite. In the case of two stage programming this requires, in particular, that the second stage problem, associated with first stage decision vector \( \bar{x} \), is feasible for almost every realization of the random data.

#### 5.6.1 Estimation of the Optimality Gap

In this section we consider the problem of estimating the optimality gap

\[
gap(\bar{x}) := f(\bar{x}) - \vartheta^*
\]

associated with the candidate solution \( \bar{x} \). Clearly, for any feasible \( \bar{x} \in X \), \( \gap(\bar{x}) \) is nonnegative and \( \gap(\bar{x}) = 0 \) iff \( \bar{x} \) is an optimal solution of the true problem.

Consider the optimal value \( \vartheta_N \) of the SAA problem (5.2). We have that \( \vartheta^* \geq \mathbb{E}[\vartheta_N] \) (see the discussion following equation (5.21)). This means that \( \vartheta_N \) provides a valid statistical lower bound for the optimal value \( \vartheta^* \) of the true problem. The expectation \( \mathbb{E}[\vartheta_N] \) can be estimated by averaging. That is, one can solve \( M \) times sample average approximation problems based on independently generated samples each of size \( N \). Let \( \vartheta_{N,1}, \ldots, \vartheta_{N,M} \) be the computed optimal values of these SAA problems. Then

\[
\bar{\vartheta}_{N,M} := \frac{1}{M} \sum_{m=1}^{M} \vartheta_{N,m}
\]
5.6. Validation Analysis

is an unbiased estimator of \( \mathbb{E}[\hat{\vartheta}_N] \). Since the samples, and hence \( \hat{\vartheta}_N, ..., \hat{\vartheta}_M \), are independent and have the same distribution, we have that \( \mathbb{V} \mathbb{A} [\bar{v}_{N,M}] = M^{-1} \mathbb{V} \mathbb{A} [\hat{\vartheta}_N] \), and hence we can estimate variance of \( \bar{v}_{N,M} \) by

\[
\hat{\sigma}^2_{N,M} := \frac{1}{M} \left[ \frac{1}{M-1} \sum_{m=1}^{M} (\hat{\vartheta}_N^m - \bar{v}_{N,M})^2 \right].
\]  

(5.166)

Note that the above make sense only if the optimal value \( \vartheta^* \) of the true problem is finite. Note also that the inequality \( \vartheta^* \geq \mathbb{E}[\hat{\vartheta}_N] \) holds and \( \hat{\vartheta}_N \) gives a valid statistical lower bound even if \( f(x) = +\infty \) for some \( x \in X \). Note finally that the samples do not need to be iid (for example, one can use Latin Hypercube sampling), they only should be independent of each other in order to use estimate (5.166) of the corresponding variance.

In general the random variable \( \hat{\vartheta}_N \), and hence its replications \( \hat{\vartheta}_N^m \), does not have a normal distribution, even approximately (see Theorem 5.7 and the following afterwards discussion). However, by the CLT, the probability distribution of the average \( \bar{v}_{N,M} \) becomes approximately normal as \( M \) increases. Therefore, we can use

\[
L_{N,M} := \bar{v}_{N,M} - t_{\alpha,M-1} \hat{\sigma}_{N,M}
\]  

(5.167)

as an approximate \( 100(1-\alpha)\% \) confidence\(^{10} \) lower bound for the expectation \( \mathbb{E}[\hat{\vartheta}_N] \).

We can also estimate \( f(x) \) by sampling. That is, let \( \bar{f}_{N'}(\bar{x}) \) be the sample average estimate of \( f(x) \), based on a sample of size \( N' \) generated independently of samples involved in computing \( \bar{x} \). Let \( \hat{\sigma}^2_{N'}(\bar{x}) \) be an estimate of the variance of \( \bar{f}_{N'}(\bar{x}) \). In the case of iid sample, one can use the sample variance estimate

\[
\hat{\sigma}^2_{N'}(\bar{x}) := \frac{1}{N'(N'-1)} \sum_{j=1}^{N'} \left[ F(\bar{x}, \xi^j) - \bar{f}_{N'}(\bar{x}) \right]^2.
\]  

(5.168)

Then

\[
U_{N'}(\bar{x}) := \bar{f}_{N'}(\bar{x}) + z_{\alpha} \hat{\sigma}_{N'}(\bar{x})
\]  

(5.169)

gives an approximate \( 100(1-\alpha)\% \) confidence upper bound for \( f(\bar{x}) \). Note that since \( N' \) typically is large, we use here critical value \( z_{\alpha} \) from the standard normal distribution rather than a \( t \)-distribution.

We have that

\[
\mathbb{E} \left[ \bar{f}_{N'}(\bar{x}) - \bar{v}_{N,M} \right] = f(\bar{x}) - \mathbb{E}[\hat{\vartheta}_N] = \text{gap}(\bar{x}) - \mathbb{E}[\hat{\vartheta}_N] \geq \text{gap}(\bar{x}),
\]

i.e., \( \bar{f}_{N'}(\bar{x}) - \bar{v}_{N,M} \) is a biased estimator of the gap(\( \bar{x} \)). Also the variance of this estimator is equal to the sum of the variances of \( \bar{f}_{N'}(\bar{x}) \) and \( \bar{v}_{N,M} \), and hence

\[
\bar{f}_{N'}(\bar{x}) - \bar{v}_{N,M} + z_{\alpha} \sqrt{\hat{\sigma}^2_{N'}(\bar{x}) + \hat{\sigma}^2_{N,M}}
\]  

(5.170)

\(^{10}\)Here \( t_{\alpha,\nu} \) is the \( \alpha \)-critical value of \( t \)-distribution with \( \nu \) degrees of freedom. This critical value is slightly bigger than the corresponding standard normal critical value \( z_{\alpha} \), and \( t_{\alpha,\nu} \) quickly approaches \( z_{\alpha} \) as \( \nu \) increases.
provides a conservative 100(1 − α)% confidence upper bound for the gap(\bar{x}). We say that this upper bound is “conservative” since in fact it gives a 100(1 − α)% confidence upper bound for the gap(\bar{x}) + \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] and we have that \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] ≥ 0.

In order to calculate the estimate \hat{f}_N(\bar{x}) one needs to compute the value \hat{F}(\bar{x}, \xi^j) of the objective function for every generated sample realization \xi^j, j = 1, ..., N'. Typically it is much easier to compute \hat{F}(\bar{x}, \xi), for a given \xi ∈ Ξ, than to solve the corresponding SAA problem. Therefore, often one can use a relatively large sample size \hat{N}', and hence to estimate \hat{f}(\bar{x}) quite accurately. Evaluation of the optimal value \vartheta^* by employing the estimator \hat{v}_{N,M} is a more delicate problem.

There are two types of errors in using \hat{v}_{N,M} as an estimator of \vartheta^*, namely the bias \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] and variability of \hat{v}_{N,M} measured by its variance. Both errors can be reduced by increasing \hat{N}, and the variance can be reduced by increasing \hat{N} and \hat{M}. Note, however, that the computational effort in computing \hat{v}_{N,M} is proportional to \hat{M}, since the corresponding SAA problems should be solved \hat{M} times, and to the computational time for solving a single SAA problem based on a sample of size \hat{N}. Naturally one may ask what is the best way of distributing computational resources between increasing the sample size \hat{N} and the number of repetitions \hat{M}. This question is, of course, problem dependent. In cases where computational complexity of SAA problems grows fast with increase of the sample size \hat{N}, it may be more advantageous to use a larger number of repetitions \hat{M}. On the other hand it was observed empirically that the computational effort in solving SAA problems by ‘good’ subgradient algorithms grows only linearly with the sample size \hat{N}. In such cases one can use a larger \hat{N} and make only a few repetitions \hat{M} in order to estimate the variance of \hat{v}_{N,M}.

The bias \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] does not depend on \hat{M}, of course. It was shown in Proposition 5.6 that if the sample is iid, then \mathbb{E}[\hat{\vartheta}_N] ≤ \mathbb{E}[\hat{\vartheta}_{N+1}] for any \hat{N} ∈ \mathbb{N}. It follows that the bias \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] decreases monotonically with increase of the sample size \hat{N}. By Theorem 5.7 we have that, under mild regularity conditions,

\[ \hat{\vartheta}_N = \inf_{x ∈ \mathcal{B}} \hat{f}_N(x) + o_p(\hat{N}^{-1/2}). \] (5.171)

Consequently, if the set \mathcal{S} of optimal solutions of the true problem is not a singleton, then the bias \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] typically converges to zero, as \hat{N} increases, at a rate of \mathcal{O}(\hat{N}^{-1/2}), and tends to be bigger for a larger set \mathcal{S} (see equation (5.28) and the following afterwards discussion). On the other hand, in well conditioned problems, where the optimal set \mathcal{S} is a singleton, the bias typically is of order \mathcal{O}(\hat{N}^{-1}) (see Theorem 5.8), and the bias tends to be of a lesser problem. Moreover, if the true problem has a sharp optimal solution \check{x}^*, then the event “\check{x}^* = \check{x}^*”, and hence the event “\hat{\vartheta}_N = \hat{f}_N(\check{x}^*)”, happens with probability approaching one exponentially fast (see Theorem 5.23). Since \mathbb{E}[\hat{f}_N(\check{x}^*)] = f(\check{x}^*), in such cases the bias \vartheta^* − \mathbb{E}[\hat{\vartheta}_N] = f(\check{x}^*) − \mathbb{E}[\hat{\vartheta}_N] tends to be much smaller.

In the above approach the upper and lower statistical bounds were computed independently of each other. Alternatively, it is possible to use the same sample for estimating \hat{f}(\bar{x}) and \mathbb{E}[\hat{\vartheta}_N]. That is, for \hat{M} generated samples each of size \hat{N}, the gap is estimated by

\[ \text{gap}_{N,M}(\bar{x}) := \frac{1}{\hat{M}} \sum_{m=1}^{\hat{M}} \left[ f_N^m(\bar{x}) - \hat{\vartheta}_N \right], \] (5.172)
where \( \hat{f}_N^m(\bar{x}) \) and \( \hat{\vartheta}_N^m \) are computed from the same sample \( m = 1, \ldots, M \). We have that the expected value of \( \hat{g}_N^{P_{N,M}}(\bar{x}) \) is \( f(\bar{x}) - E[\hat{\vartheta}_N] \), i.e., the estimator \( \hat{g}_N^{P_{N,M}}(\bar{x}) \) has the same bias as \( \hat{f}_N(\bar{x}) - \bar{v}_{N,M} \). On the other hand, for a problem with sharp optimal solution \( x^* \) it happens with high probability that \( \hat{\vartheta}_N^m = \hat{f}_N^m(x^*) \), and as a consequence \( \hat{f}_N^m(\bar{x}) \) tends to be highly positively correlated with \( \hat{\vartheta}_N^m \), provided that \( \bar{x} \) is close to \( x^* \). In such cases variability of \( \hat{g}_N^{P_{N,M}}(\bar{x}) \) can be considerably smaller than variability of \( \hat{f}_{N'}(\bar{x}) - \bar{v}_{N,M} \). This is the idea of common random number generated estimators.

**Remark 13.** Of course, in order to obtain a valid statistical lower bound for the optimal value \( \vartheta^* \) we can use any (deterministic) lower bound for the optimal value \( \vartheta_N \) of the corresponding SAA problem instead of \( \vartheta_N \) itself. For example, suppose that the problem is convex. By convexity of \( f_N(\cdot) \) we have that for any \( x' \in X \) and \( \gamma \in \partial \hat{f}_N(x') \) it holds that

\[
\hat{f}_N(x) \geq \hat{f}_N(x') + \gamma^T(x - x'), \quad \forall x \in \mathbb{R}^n. \tag{5.173}
\]

Therefore, we can proceed as follows. Choose points \( x_1, \ldots, x_r \in X \), calculate subgradients \( \hat{\gamma}_{iN} \in \partial f_N(x_i), i = 1, \ldots, r \), and solve the problem

\[
\min_{x \in X} \max_{1 \leq i \leq r} \left\{ \hat{f}_N(x_i) + \hat{\gamma}_{iN}^T(x - x_i) \right\}. \tag{5.174}
\]

Denote by \( \hat{\lambda}_N \) the optimal value of (5.174). By (5.173) we have that \( \hat{\lambda}_N \) is less than or equal to the optimal value \( \vartheta_N \) of the corresponding SAA problem, and hence gives a valid statistical lower bound for \( \vartheta^* \). A possible advantage of \( \hat{\lambda}_N \) over \( \vartheta_N \) is that it could be easier to solve (5.174) than the corresponding SAA problem. For instance, if the set \( X \) is polyhedral, then (5.174) can be formulated as a linear programming problem.

Of course, this approach raises the question of how to choose the points \( x_1, \ldots, x_r \in X \). Suppose that the expectation function \( f(x) \) is differentiable at the points \( x_1, \ldots, x_r \). Then for any choice of \( \hat{\gamma}_{iN} \in \partial \hat{f}_N(x_i) \) we have that subgradients \( \hat{\gamma}_{iN} \) converge to \( \nabla f(x_i) \) w.p.1. Therefore \( \hat{\lambda}_N \) converges w.p.1 to the optimal value of the problem

\[
\min_{x \in X} \max_{1 \leq i \leq r} \left\{ f(x_i) + \nabla f(x_i)^T(x - x_i) \right\}, \tag{5.175}
\]

provided that the set \( X \) is bounded. Again by convexity arguments the optimal value of (5.175) is less than or equal to the optimal value \( \vartheta^* \) of the true problem. If we can find such points \( x_1, \ldots, x_r \in X \) that the optimal value of (5.175) is less than \( \vartheta^* \) by a ‘small’ amount, then it could be advantageous to use \( \hat{\lambda}_N \) instead of \( \vartheta_N \). We also should keep in mind that the number \( r \) should be relatively small, otherwise we may loose the advantage of solving the ‘easier’ problem (5.174).

A natural approach to choosing the required points and hence to apply the above procedure is the following. By solving (once) an SAA problem, find points \( x_1, \ldots, x_r \in X \) such that the optimal value of the corresponding problem (5.174) provides with a high accuracy an estimate of the optimal value of this SAA problem. Use some (all) these points to calculate lower bound estimates \( \hat{\lambda}_N^m, m = 1, \ldots, M \), probably with a larger sample size \( N \). Calculate the average \( \hat{\lambda}_{N,M} \) together with the corresponding sample variance and construct the associated \( 100(1 - \alpha)\% \) confidence lower bound similar to (5.167).
Chapter 5. Statistical Inference

Estimation of Optimality Gap of Minimax and Expectation-Constrained Problems

Consider a minimax problem of the form (5.45). Let $\vartheta^*$ be the optimal value of this (true) minimax problem. Clearly for any $\bar{y} \in Y$ we have that

$$\vartheta^* \geq \inf_{x \in X} f(x, \bar{y}).$$  \hfill (5.176)

Now for the optimal value of the right hand side of (5.176) we can construct a valid statistical lower bound, and hence a valid statistical lower bound for $\vartheta^*$, as before by solving the corresponding SAA problems several times and averaging calculated optimal values. Suppose, further, that the minimax problem (5.45) has a nonempty set $S_x \times S_y \subset X \times Y$ of saddle points, and hence its optimal value is equal to the optimal value of its dual problem (5.47). Then for any $\bar{x} \in X$ we have that

$$\vartheta^* \leq \sup_{y \in Y} f(\bar{x}, y),$$  \hfill (5.177)

and the equalities in (5.176) and/or (5.177) hold iff $\bar{y} \in S_y$ and/or $\bar{x} \in S_x$. By (5.177) we can construct a valid statistical upper bound for $\vartheta^*$ by averaging optimal values of sample average approximations of the right hand side of (5.177). Of course, quality of these bounds will depend on a good choice of the points $\bar{y}$ and $\bar{x}$. A natural construction for the candidate solutions $\bar{y}$ and $\bar{x}$ will be to use optimal solutions of a run of the corresponding minimax SAA problem (5.46).

Similar ideas can be applied to validation of stochastic problems involving constraints given as expected value functions (see equations (5.11)–(5.13)). That is, consider the following problem

$$\text{Min } f(x) \text{ subject to } g_i(x) \leq 0, \ i = 1, ..., p, \quad (5.178)$$

where $X_0$ is a nonempty subset of $\mathbb{R}^n$, $f(x) := E[F(x, \xi)]$ and $g_i(x) := E[G_i(x, \xi)]$, $i = 1, ..., p$. We have that

$$\vartheta^* = \inf_{x \in X_0} \sup_{\lambda \geq 0} L(x, \lambda),$$  \hfill (5.179)

where $\vartheta^*$ is the optimal value and $L(x, \lambda) := f(x) + \sum_{i=1}^{p} \lambda_i g_i(x)$ is the Lagrangian of problem (5.178). Therefore, for any $\lambda \geq 0$, we have that

$$\vartheta^* \geq \inf_{x \in X_0} L(x, \lambda),$$  \hfill (5.180)

and the equality in (5.180) is attained if the problem (5.176) is convex and $\bar{\lambda}$ is a Lagrange multipliers vector satisfying the corresponding first-order optimality conditions. Of course, a statistical lower bound for the optimal value of the problem in the right hand side of (5.180), is also a statistical lower bound for $\vartheta^*$.

Unfortunately an upper bound which can be obtained by interchanging the ‘inf’ and ‘sup’ operators in (5.179) cannot be used in a straightforward way. This is because if, for a chosen $\bar{x} \in X_0$, it happens that $\hat{g}_{iN}(\bar{x}) > 0$ for some $i \in \{1, ..., p\}$, then

$$\sup_{\lambda \geq 0} \left\{ \hat{f}_{N}(\bar{x}) + \sum_{i=1}^{p} \lambda_i \hat{g}_{iN}(\bar{x}) \right\} = +\infty.$$  \hfill (5.181)
Of course, in such a case the obtained upper bound $+\infty$ is useless. This typically will be the case if $\bar{x}$ is constructed as a solution of an SAA problem and some of the SAA constraints are active at $\bar{x}$. Note also that if $\hat{g}_i(\bar{x}) \leq 0$ for all $i \in \{1, \ldots, p\}$, then the supremum in the left hand side of (5.181) is equal to $\hat{f}_N(\bar{x})$.

If we can ensure, with a high probability $1 - \alpha$, that a chosen point $\bar{x}$ is a feasible point of the true problem 5.176, then we can construct an upper bound by estimating $f(\bar{x})$ using a relatively large sample. This, in turn, can be approached by verifying, for an independent sample of size $N'$, that $\hat{g}_i(\bar{x}) + \kappa \hat{\sigma}_i N(\bar{x}) \leq 0$, $i = 1, \ldots, p$, where $\hat{\sigma}_i N(\bar{x})$ is a sample variance of $g_i(\bar{x})$ and $\kappa$ is a positive constant chosen in such a way that the probability of $g_i(\bar{x})$ being bigger that $\hat{g}_i(\bar{x}) + \kappa \hat{\sigma}_i N(\bar{x})$ is less than $\alpha/p$ for all $i \in \{1, \ldots, p\}$.

5.6.2 Statistical Testing of Optimality Conditions

Suppose that the feasible set $X$ is defined by (equality and inequality) constraints in the form

$$X := \{ x \in \mathbb{R}^n : g_i(x) = 0, \ i = 1, \ldots, q; \ g_i(x) \leq 0, \ i = q + 1, \ldots, p \} \tag{5.182}$$

where $g_i(x)$ are smooth (at least continuously differentiable) deterministic functions. Let $x^* \in X$ be an optimal solution of the true problem and suppose that the expected value function $f(\cdot)$ is differentiable at $x^*$. Then, under a constraint qualification, first order (KKT) optimality conditions hold at $x^*$. That is, there exist Lagrange multipliers $\lambda_i$ such that $\lambda_i \geq 0$, $i \in I(x^*)$ and

$$\nabla f(x^*) + \sum_{i \in J(x^*)} \lambda_i \nabla g_i(x^*) = 0, \tag{5.183}$$

where $I(x) := \{ i : g_i(x) = 0, \ i = q + 1, \ldots, p \}$ denotes the index set of inequality constraints active at a point $x \in \mathbb{R}^n$, and $J(x) := \{ 1, \ldots, q \} \cup I(x)$. Note that if the constraint functions are linear, say $g_i(x) := a_i^T x + b_i$, then $\nabla g_i(x) = a_i$ and the above KKT conditions hold without a constraint qualification. Consider the (polyhedral) cone

$$K(x) := \left\{ z \in \mathbb{R}^n : z = \sum_{i \in J(x)} \alpha_i \nabla g_i(x), \ \alpha_i \leq 0, \ i \in I(x) \right\}. \tag{5.184}$$

Then the KKT optimality conditions can be written in the form $\nabla f(x^*) \in K(x^*)$.

Suppose now that $f(\cdot)$ is differentiable at the candidate point $\bar{x} \in X$, and that the gradient $\nabla f(\bar{x})$ can be estimated by a (random) vector $\gamma_N(\bar{x})$. In particular, if $F(\cdot, \xi)$ is differentiable at $\bar{x}$ w.p.1, then we can use the estimator

$$\gamma_N(\bar{x}) := \frac{1}{N} \sum_{j=1}^N \nabla_x F(\bar{x}, \xi^j) = \hat{f}_N(\bar{x}) \tag{5.185}$$

associated with the generated random sample. Note that if, moreover, the derivatives can be taken inside the expectation, that is,

$$\nabla f(\bar{x}) = \mathbb{E}[\nabla_x F(\bar{x}, \xi)], \tag{5.186}$$

\footnote{We emphasize that the random sample in (5.185) is generated independently of the sample used to compute the candidate point $\bar{x}$.}
then the above estimator is unbiased, i.e., \( \mathbb{E}[\gamma_N(\bar{x})] = \nabla f(\bar{x}) \). In the case of two-stage linear stochastic programming with recourse, formula (5.186) typically holds if the corresponding random data have a continuous distribution. On the other hand if the random data have a discrete distribution with a finite support, then the expected value function \( f(x) \) is piecewise linear and typically is nondifferentiable at an optimal solution.

Suppose, further, that \( V_N := N^{1/2} [\gamma_N(\bar{x}) - \nabla f(\bar{x})] \) converges in distribution, as \( N \) tends to infinity, to multivariate normal with zero mean vector and covariance matrix \( \Sigma \), written \( V_N \overset{D}{\rightarrow} \mathcal{N}(0, \Sigma) \). For the estimator \( \gamma_N(\bar{x}) \) defined in (5.185), this holds by the CLT if the interchangeability formula (5.186) holds, the sample is iid, and \( \nabla x F(\bar{x}, \xi) \) has finite second order moments. Moreover, in that case the covariance matrix \( \Sigma \) can be estimated by the corresponding sample covariance matrix

\[
\hat{\Sigma}_N := \frac{1}{N-1} \sum_{j=1}^{N} \left[ \nabla_x F(\bar{x}, \xi^j) - \nabla \hat{f}(\bar{x}) \right] \left[ \nabla_x F(\bar{x}, \xi^j) - \nabla \hat{f}(\bar{x}) \right]^T. \tag{5.187}
\]

Under the above assumptions, the sample covariance matrix \( \hat{\Sigma}_N \) is an unbiased and consistent estimator of \( \Sigma \).

We have that if \( V_N \overset{D}{\rightarrow} \mathcal{N}(0, \Sigma) \) and the covariance matrix \( \Sigma \) is nonsingular, then (given a consistent estimator \( \hat{\Sigma}_N \) of \( \Sigma \)) the following holds

\[
N (\gamma_N(\bar{x}) - \nabla f(\bar{x}))^T \hat{\Sigma}_N^{-1} (\gamma_N(\bar{x}) - \nabla f(\bar{x})) \overset{D}{\rightarrow} \chi^2_n, \tag{5.188}
\]

where \( \chi^2_n \) denotes chi-square distribution with \( n \) degrees of freedom. This allows to construct the following (approximate) \( 100(1 - \alpha)\% \) confidence region\(^{12} \) for \( \nabla f(\bar{x}) \):

\[
\left\{ z \in \mathbb{R}^n : (\gamma_N(\bar{x}) - z)^T \hat{\Sigma}_N^{-1} (\gamma_N(\bar{x}) - z) \leq \frac{\chi^2_{\alpha,n}}{N} \right\}. \tag{5.189}
\]

Consider the statistic

\[
T_N := N \inf_{z \in K(\bar{x})} (\gamma_N(\bar{x}) - z)^T \hat{\Sigma}_N^{-1} (\gamma_N(\bar{x}) - z). \tag{5.190}
\]

Note that since the cone \( K(\bar{x}) \) is polyhedral and \( \hat{\Sigma}_N^{-1} \) is positive definite, the minimization in the right hand side of (5.190) can be formulated as a quadratic programming problem, and hence can be solved by standard quadratic programming algorithms. We have that the confidence region, defined in (5.189), does not have common points with the cone \( K(\bar{x}) \) iff \( T_N > \chi^2_{\alpha,n} \). We can also use the statistic \( T_N \) for testing the hypothesis:

\[
H_0 : \ nabla f(\bar{x}) \in K(\bar{x}) \text{ against the alternative } H_1 : \ nabla f(\bar{x}) \notin K(\bar{x}). \tag{5.191}
\]

The \( T_N \) statistic represents the squared distance, with respect to the norm\(^{13} \| \cdot \|^2_{\hat{\Sigma}_N^{-1}} \), from \( N^{1/2} \gamma_N(\bar{x}) \) to the cone \( K(\bar{x}) \). Suppose for the moment that only equality constraints are

\(^{12}\)Here \( \chi^2_{\alpha,n} \) denotes the \( \alpha \)-critical value of chi-square distribution with \( n \) degrees of freedom. That is, if \( Y \sim \chi^2_n \), then \( \Pr \{ Y \geq \chi^2_{\alpha,n} \} = \alpha \).

\(^{13}\)For a positive definite matrix \( A \), the norm \( \| \cdot \|_A \) is defined as \( \| z \|_A := (z^T A z)^{1/2} \).
present in the definition (5.182) of the feasible set, and that the gradient vectors \( \nabla g_i(\bar{x}) \), \( i = 1, \ldots, q \), are linearly independent. Then the set \( K(\bar{x}) \) forms a linear subspace of \( \mathbb{R}^n \) of dimension \( q \), and the optimal value of the right hand side of (5.190) can be written in a closed form. Consequently, it is possible to show that \( T_N \) has asymptotically noncentral chi-square distribution with \( n - q \) degrees of freedom and the noncentrality parameter\(^{14}\)

\[
\delta := N \inf_{z \in K(\bar{x})} (\nabla f(\bar{x}) - z)^T \Sigma^{-1} (\nabla f(\bar{x}) - z). \tag{5.192}
\]

In particular, under \( H_0 \) we have that \( \delta = 0 \), and hence the null distribution of \( T_N \) is asymptotically central chi-square with \( n - q \) degrees of freedom.

Consider now the general case where the feasible set is defined by equality and inequality constraints as in (5.182). Suppose that the gradient vectors \( \nabla g_i(\bar{x}) \), \( i \in \mathcal{J}(\bar{x}) \), are linearly independent and that the strict complementarity condition holds at \( \bar{x} \), that is, the Lagrange multipliers \( \lambda_i \), \( i \in \mathcal{J}(\bar{x}) \), corresponding to the active at \( \bar{x} \) inequality constraints, are positive. Then for \( \gamma_N(\bar{x}) \) sufficiently close to \( \nabla f(\bar{x}) \) the minimizer in the right hand side of (5.190) will be lying in the linear space generated by vectors \( \nabla g_i(\bar{x}) \), \( i \in \mathcal{J}(\bar{x}) \). Therefore, in such case the null distribution of \( T_N \) is asymptotically central chi-square with \( n - |\mathcal{J}(\bar{x})| \) degrees of freedom. Consequently, for a computed value \( T_N^* \) of the statistic \( T_N \) we can calculate (approximately) the corresponding \( p \)-value, which is equal to \( \Pr \{ Y \geq T_N^* \} \), where \( Y \sim \chi^2_n \). This \( p \)-value gives an indication of the quality of the candidate solution \( \bar{x} \) with respect to the stochastic precision.

It should be understood that by accepting (i.e., failing to reject) \( H_0 \), we do not claim that the KKT conditions hold exactly at \( \bar{x} \). By accepting \( H_0 \) we rather assert that we cannot separate \( \nabla f(\bar{x}) \) from \( K(\bar{x}) \), given precision of the generated sample. That is, statistical error of the estimator \( \gamma_N(\bar{x}) \) is bigger than the squared \( \| \cdot \|_{\Sigma^{-1}} \)-norm distance between \( \nabla f(\bar{x}) \) and \( K(\bar{x}) \). Also rejecting \( H_0 \) does not necessarily mean that \( \bar{x} \) is a poor candidate for an optimal solution of the true problem. The calculated value of \( T_N \) statistic can be large, i.e., the \( p \)-value can be small, simply because the estimated covariance matrix \( N^{-1} \Sigma_N \) of \( \gamma_N(\bar{x}) \) is “small”. In such cases, \( \gamma_N(\bar{x}) \) provides an accurate estimator of \( \nabla f(\bar{x}) \) with the corresponding confidence region (5.189) being “small”. Therefore, the above \( p \)-value should be compared with the size of the confidence region (5.189), which in turn is defined by the size of the matrix \( N^{-1} \Sigma_N \) measured, for example, by its eigenvalues. Note also that it may happen that \( |\mathcal{J}(\bar{x})| = n \), and hence \( \nu = 0 \). Under the strict complementarity condition, this means that \( \nabla f(\bar{x}) \) lies in the interior of the cone \( K(\bar{x}) \), which in turn is equivalent to the condition that \( f'(\bar{x}, d) \geq c\|d\| \) for some \( c > 0 \) and all \( d \in \mathbb{R}^n \). Then, by the LD principle (see (7.198) in particular), the event \( \gamma_N(\bar{x}) \in K(\bar{x}) \) happens with probability approaching one exponentially fast.

Let us remark again that the above testing procedure is applicable if \( F(\cdot, \xi) \) is differentiable at \( \bar{x} \) w.p.1 and the interchangeability formula (5.186) holds. This typically happens in cases where the corresponding random data have a continuous distribution.

\(^{14}\)Note that under the alternative (i.e., if \( \nabla f(\bar{x}) \notin K(\bar{x}) \)), the noncentrality parameter \( \delta \) tends to infinity as \( N \to \infty \). Therefore, in order to justify the above asymptotics one needs a technical assumption known as Pitman’s parameter drift.
5.7 Chance Constrained Problems

Consider a chance constrained problem of the form

$$\min_{x \in X} f(x) \quad \text{subject to} \quad p(x) \leq \alpha,$$  \hspace{1cm} (5.193)

where $X \subset \mathbb{R}^n$ is a closed set, $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous function, $\alpha \in (0, 1)$ is a given significance level and

$$p(x) := \Pr\{C(x, \xi) > 0\}$$  \hspace{1cm} (5.194)

is the probability that constraint is violated at point $x \in X$. We assume that $\xi$ is a random vector, whose probability distribution $P$ is supported on set $\Xi \subset \mathbb{R}^d$, and the function $C : \mathbb{R}^n \times \Xi \to \mathbb{R}$ is a Carathéodory function. The chance constraint "$p(x) \leq \alpha$" can be written equivalently in the form

$$\Pr\{C(x, \xi) \leq 0\} \geq 1 - \alpha.$$  \hspace{1cm} (5.195)

Let us also remark that several chance constraints

$$\Pr\{C_i(x, \xi) \leq 0, \ i = 1, ..., q\} \geq 1 - \alpha,$$  \hspace{1cm} (5.196)

can be reduced to one chance constraint (5.195) by employing the max-function $C(x, \xi) := \max_{1 \leq i \leq q} C_i(x, \xi)$. Of course, in some cases this may destroy a ‘nice’ structure of considered functions. At this point, however, this is not important.

5.7.1 Monte Carlo Sampling Approach

We discuss now a way of solving problem (5.193) by Monte Carlo sampling. For the sake of simplicity we assume that the objective function $f(x)$ is given explicitly and only the chance constraints should be approximated.

We can write the probability $p(x)$ in the form of the expectation,

$$p(x) = \mathbb{E} \left[ 1_{(0, \infty)}(C(x, \xi)) \right],$$

and to estimate this probability by the corresponding SAA function (compare with (5.14)–(5.16))

$$\hat{p}_N(x) := N^{-1} \sum_{j=1}^{N} 1_{(0, \infty)}(C(x, \xi_j)).$$  \hspace{1cm} (5.197)

Recall that $1_{(0, \infty)}(C(x, \xi))$ is equal to 1 if $C(x, \xi) > 0$, and it is equal 0 otherwise. Therefore, $\hat{p}_N(x)$ is equal to the proportion of times that $C(x, \xi_j) > 0$, $j = 1, ..., N$. Consequently we can write the corresponding SAA problem as

$$\min_{x \in X} f(x) \quad \text{subject to} \quad \hat{p}_N(x) \leq \alpha.$$  \hspace{1cm} (5.198)

**Proposition 5.29.** Let $C(x, \xi)$ be a Carathéodory function. Then the functions $p(x)$ and $\hat{p}_N(x)$ are lower semicontinuous. Suppose, further, that the sample is iid. Then $\hat{p}_N \xrightarrow{w.p.1} p$.

Moreover, suppose that for every $x \in X$ it holds that
\[
\Pr \{ \xi \in \Xi : C(x, \xi) = 0 \} = 0, \tag{5.199}
\]
i.e., $C(x, \xi) \neq 0$ w.p.1. Then the function $p(x)$ is continuous on $X$ and $\hat{p}_N(x)$ converges to $p(x)$ w.p.1 uniformly on any compact subset of $X$.

**Proof.** Consider function $\psi(x, \xi) := 1_{(0, \infty)}(C(x, \xi))$. Recall that $p(x) = \mathbb{E}_P[\psi(x, \xi)]$ and $\hat{p}_N(x) = \mathbb{E}_{P_N}[\psi(x, \xi)]$, where $P_N := N^{-1} \sum_{j=1}^{N} \Delta(\xi^j)$ is the respective empirical measure. Since the function $1_{(0, \infty)}(\cdot)$ is lower semicontinuous and $C(x, \xi)$ is a Carathéodory function, it follows that the function $\psi(x, \xi)$ is random lower semicontinuous. Lower semicontinuity of $p(x)$ and $\hat{p}_N(x)$ follows by Fatou’s lemma (see Theorem 7.42). If the sample is iid, the epiconvergence $\hat{p}_N \rightharpoonup p$ w.p.1 follows by Theorem 7.51. Note that the dominance condition, from below and from above, holds here automatically since $|\psi(x, \xi)| \leq 1$.

Suppose, further, that condition (5.199) holds. Then for every $x \in X$, $\psi(\cdot, \xi)$ is continuous at $x$ w.p.1. It follows by the Lebesgue Dominated Convergence Theorem that $p(\cdot)$ is continuous at $x$ (see Theorem 7.43). Finally, the uniform convergence w.p.1 follows by Theorem 7.48. \qed

Since the function $p(x)$ is lower semicontinuous and the set $X$ is closed, it follows that the feasible set of problem (5.193) is closed. If, moreover, it is nonempty and bounded, then problem (5.193) has a nonempty set $S$ of optimal solutions (recall that the objective function $f(x)$ is assumed to be continuous here). The same applies to the corresponding SAA problem (5.198). We have here the following consistency properties of the optimal value $\hat{\vartheta}_N$ and the set $\hat{S}_N$ of optimal solutions of the SAA problem (5.198) (compare with Theorem 5.5).

**Proposition 5.30.** Suppose that the set $X$ is compact, the function $f(x)$ is continuous, $C(x, \xi)$ is a Carathéodory function, the sample is iid and that the following condition holds: (a) there is an optimal solution $\bar{x}$ of the true problem such that for any $\epsilon > 0$ there is $x \in X$ with $\|x - \bar{x}\| \leq \epsilon$ and $p(x) < \alpha$. Then $\hat{\vartheta}_N \to \vartheta^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.

**Proof.** By condition (a), the set $S$ is nonempty and there is $x' \in X$ such that $p(x') < \alpha$. By the LLN we have that $\hat{p}_N(x')$ converges to $p(x')$ w.p.1. Consequently $\hat{\vartheta}_N(x') < \alpha$, and hence the SAA problem has a feasible solution, w.p.1 for $N$ large enough. Since $\hat{p}_N(\cdot)$ is lower semicontinuous, the feasible set of SAA problem is closed and hence compact, and thus $\hat{S}_N$ is nonempty w.p.1 for $N$ large enough. Of course, if $x'$ is a feasible solution of an SAA problem, then $f(x') \geq \hat{\vartheta}_N$, where $\hat{\vartheta}_N$ is the optimal value of that SAA problem.

For a given $\epsilon > 0$ let $x' \in X$ be a point sufficiently close to $\bar{x} \in S$ such that $\hat{p}_N(x') < \alpha$ and $f(x') \leq f(\bar{x}) + \epsilon$. Since $f(\cdot)$ is continuous, existence of such point is ensured by condition (a). Consequently
\[
\limsup_{N \to \infty} \hat{\vartheta}_N \leq f(x') \leq f(\bar{x}) + \epsilon = \vartheta^* + \epsilon \quad \text{w.p.1.} \tag{5.200}
\]
Chapter 5. Statistical Inference

Since \( \varepsilon > 0 \) is arbitrary, it follows that

\[
\limsup_{N \to \infty} \hat{\vartheta}_N \leq \vartheta^* \quad \text{w.p.1.} \tag{5.201}
\]

Now let \( \hat{x}_N \in \hat{S}_N \), i.e., \( \hat{x}_N \in X, \hat{p}_N(\hat{x}_N) \leq \alpha \) and \( \hat{\vartheta}_N = f(\hat{x}_N) \). Since the set \( X \) is compact, we can assume by passing to a subsequence if necessary that \( \hat{x}_N \) converges to a point \( \bar{x} \in X \) w.p.1. Also by Proposition 5.29 we have that \( \hat{p}_N \to p \) w.p.1, and hence

\[
\liminf_{N \to \infty} \hat{p}_N(x_N) \geq p(\bar{x}) \quad \text{w.p.1.} \tag{5.202}
\]

It follows from (5.201) and (5.202) that \( \hat{\vartheta}_N \to \vartheta^* \) w.p.1. It also follows that the point \( \bar{x} \) is an optimal solution of the true problem and consequently we obtain that \( D(\hat{S}_N, S) \to 0 \) w.p.1.

The above condition (a) is essential for the consistency of \( \hat{\vartheta}_N \) and \( \hat{S}_N \). Think, for example, about a situation where the constraint \( p(\bar{x}) \leq \alpha \) defines just one feasible point \( \bar{x} \) such that \( p(\bar{x}) = \alpha \). Then arbitrary small changes in the constraint \( \hat{p}_N(x) \leq \alpha \) may result in that the feasible set of the corresponding SAA problem becomes empty. Note also that condition (a) was not used in the proof of inequality (5.202). Verification of this condition (a) can be done by ad hoc methods.

We have that, under mild regularity conditions, optimal value and optimal solutions of the SAA problem (5.198) converge w.p.1, as \( N \to \infty \), to their counterparts of the true problem (5.193). There are, however, several potential problems with the SAA approach here. In order for \( \hat{p}_N(x) \) to be a reasonably accurate estimate of \( p(x) \), the sample size \( N \) should be significantly bigger than \( \alpha^{-1} \). For small \( \alpha \) this may result in a large sample size. Another problem is that typically the function \( \hat{p}_N(x) \) is discontinuous and the SAA problem (5.198) is a combinatorial problem which could be difficult to solve. Therefore we consider the following approach.

For a generated sample \( \xi^1, \ldots, \xi^N \) consider the problem:

\[
\min_{x \in X} f(x) \quad \text{subject to} \quad C(x, \xi^j) \leq 0, \quad j = 1, \ldots, N. \tag{5.203}
\]

Note that for \( \alpha = 0 \) the SAA problem (5.198) coincides with the above problem (5.203). If the set \( X \) and functions \( f(\cdot) \) and \( C(\cdot, \xi), \xi \in \Xi \), are convex, then (5.203) is a convex problem and could be efficiently solved provided that the involved functions are given in a closed form and the sample size \( N \) is not too large. Clearly, as \( N \to \infty \) the feasible set of the above problem (5.203) will shrink to the set of \( x \in X \) determined by the constraints \( C(x, \xi) \leq 0 \) for a.e. \( \xi \in \Xi \), and hence for large \( N \) will be overly conservative for the true chance constrained problem (5.193). Nevertheless, it makes sense to ask the question for what sample size \( N \) an optimal solution of problem (5.203) is guaranteed to be a feasible
5.7. Chance Constrained Problems 215

point of problem (5.193).

We need the following auxiliary result.

**Lemma 5.31.** Suppose that the set $X$ and functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex and let $\bar{x}_N$ be an optimal solution of problem (5.203). Then there exists an index set $J \subset \{1, ..., N\}$ such that $|J| \leq n$ and $\bar{x}_N$ is an optimal solution of the problem

$$\min_{x \in X} f(x) \text{ subject to } C(x, \xi^j) \leq 0, \ j \in J. \quad (5.204)$$

**Proof.** Consider sets $A_0 := \{x \in X : f(x) < f(\bar{x}_N)\}$ and $A_j := \{x \in X : C(x, \xi^j) \leq 0\}$, $j = 1, ..., N$. Since $X$, $f(\cdot)$ and $C(\cdot, \xi)$ are convex, these sets are convex. Now we argue by a contradiction. Suppose that the assertion of this lemma is not correct. Then the intersection of all sets $A_j$, $j \in \{1, ..., N\}$, is nonempty (these sets have at least one common element $\bar{x}_N$), it follows that the intersection of any $n + 1$ sets of the family $A_j$, $j \in \{0, 1, ..., N\}$, is nonempty. By Helly’s theorem (Theorem 7.3) this implies that the intersection of all sets $A_j$, $j \in \{0, 1, ..., N\}$, is nonempty. This, in turn, implies existence of a feasible point $\bar{x}$ of problem (5.203) such that $f(\bar{x}) < f(\bar{x}_N)$, which contradicts optimality of $\bar{x}_N$. \qed

We will use the following assumptions.

**(F1)** For any $N \in \mathbb{N}$ and any $(\xi_1, ..., \xi_N) \in \Xi^N$, problem (5.203) attains unique optimal solution $\bar{x}_N = \bar{x}(\xi_1, ..., \xi_N)$.

Recall that sometimes we use the same notation for a random vector and its particular value (realization). In the above assumption we view $\xi_1, ..., \xi_N$ as an element of the set $\Xi^N$, and $\bar{x}_N$ as a function of $\xi_1, ..., \xi_N$. Of course, if $\xi_1, ..., \xi_N$ is a random sample, then $\bar{x}_N$ becomes a random vector.

Let $J = J(\xi^1, ..., \xi^N) \subset \{1, ..., N\}$ be an index set such that $\bar{x}_N$ is an optimal solution of the problem (5.204) for $J = J$. Moreover, let the index set $J$ be minimal in the sense that if any of the constraints $C(x, \xi^j) \leq 0$, $j \in J$, is removed, then $\bar{x}_N$ is not an optimal solution of the obtained problem. We assume that w.p.1 such minimal index set is unique. By Lemma 5.31, we have that $|J| \leq n$. By $P^N$ we denote here the product probability measure on the set $\Xi^N$, i.e., $P^N$ is the probability distribution of the iid sample $\xi^1, ..., \xi^N$.

**(F2)** There is an integer $n \in \mathbb{N}$ such that, for any $N \geq n$, w.p.1 the minimal set $J = J(\xi^1, ..., \xi^N)$ is uniquely defined and has constant cardinality $n$, i.e., $P^N \{ |J| = n \} = 1$.

By Lemma 5.31, we have that $n \leq n$.

Assumption (F1) holds, for example, if the set $X$ is compact and convex, functions $f(\cdot)$ and $C(\cdot, \xi)$, $\xi \in \Xi$, are convex, and either $f(\cdot)$ or the feasible set of problem (5.203) is strictly convex. Assumption (F2) is more involved, it is needed in order to show an equality in the estimate (5.206) of the following theorem.
The following result is due to Campi and Garatti [30], building on work of Calafiore and Campi [29]. Denote
\[
b(k; \alpha, N) := \sum_{i=0}^{k} \binom{N}{i} \alpha^i (1 - \alpha)^{N-i}, \quad k = 0, \ldots, N. \tag{5.205}
\]
That is, \( b(k; \alpha, N) = \Pr(W \leq k) \), where \( W \sim B(N, \alpha) \) is a random variable having binomial distribution.

**Theorem 5.32.** Suppose that the set \( X \) and functions \( f(\cdot) \) and \( C(\cdot, \xi) \), \( \xi \in \Xi \), are convex and conditions (F1) and (F2) hold. Then for \( \alpha \in (0, 1) \) and for an iid sample \( \xi_1, \ldots, \xi_N \) of size \( N \geq n \) we have that
\[
\Pr\{ p(\bar{x}_N) > \alpha \} = b(n - 1; \alpha, N). \tag{5.206}
\]

**Proof.** Let \( \mathcal{J}_n \) be the family of all sets \( J \subset \{1, \ldots, N\} \) of cardinality \( n \). We have that \( |\mathcal{J}_n| = \binom{N}{n} \). For \( J \in \mathcal{J}_n \) define the set
\[
\Sigma_J := \{ (\xi^1, \ldots, \xi^N) \in \Xi^N : \mathcal{J}(\xi^1, \ldots, \xi^N) = J \}, \tag{5.207}
\]
and denote by \( \hat{x}_J = \hat{x}_J(\xi^1, \ldots, \xi^N) \) an optimal solution of problem (5.204) for \( J = J \). By condition (F1), such optimal solution \( \hat{x}_J \) exists and is unique, and hence
\[
\Sigma_J = \{ (\xi^1, \ldots, \xi^N) \in \Xi^N : \hat{x}_J = \bar{x}_N \}. \tag{5.208}
\]
Note that for any permutation of vectors \( \xi^1, \ldots, \xi^N \), problem (5.203) remains the same. Therefore, any set from the family \( \{ \Sigma_J \}_{J \in \mathcal{J}_n} \) can be obtained from another set of that family by an appropriate permutation of its components. Since \( P^N \) is the direct product probability measure, it follows that the probability measure of each set \( \Sigma_J, J \in \mathcal{J}_n, \) is the same. The sets \( \Sigma_J \) are disjoint and, because of condition (F2), union of all these sets is equal to \( \Xi^N \) up to a set of \( P^N \)-measure zero. Since there are \( \binom{N}{n} \) such sets, we obtain that
\[
P^N(\Sigma_J) = \frac{1}{\binom{N}{n}}. \tag{5.209}
\]
Consider the optimal solution \( \bar{x}_n = \bar{x}(\xi^1, \ldots, \xi^n) \) for \( N = n \), and let \( H(z) \) be the cdf of the random variable \( p(\bar{x}_n) \), i.e.,
\[
H(z) := P^n \{ p(\bar{x}_n) \leq z \}. \tag{5.210}
\]
Let us show that for \( N \geq n \),
\[
P^N(\Sigma_J) = \int_{0}^{1} (1 - z)^{N-n} dH(z). \tag{5.211}
\]
5.7. Chance Constrained Problems

Indeed, for $z \in [0, 1]$ and $J \in \mathcal{J}_n$ consider the sets

\[
\begin{align*}
\Delta_z &:= \{(\xi_1, \ldots, \xi_N) : p(\hat{x}_N) \in [z, z + dz]\}, \\
\Delta_{J,z} &:= \{(\xi_1, \ldots, \xi_N) : p(\hat{x}_J) \in [z, z + dz]\}.
\end{align*}
\]  
(5.212)

By (5.208) we have that $\Delta_z \cap \Sigma_J = \Delta_{J,z} \cap \Sigma_J$. For $J \in \mathcal{J}_n$ let us evaluate probability of the event $\Delta_{J,z} \cap \Sigma_J$. For the sake of notational simplicity let us take $J = \{1, \ldots, n\}$. Note that $\hat{x}_J$ depends on $(\xi^1, \ldots, \xi^n)$ only. Therefore $\Delta_{J,z} = \Delta_z \times \Xi^{n-1}$, where $\Delta_z$ is a subset of $\Xi^n$ corresponding to the event $\{p(\hat{x}_J) \in [z, z + dz]\}$. Conditional on $(\xi^1, \ldots, \xi^n)$, the event $\Delta_{J,z} \cap \Sigma_J$ happens iff the point $\hat{x}_J = \hat{x}_J(\xi^1, \ldots, \xi^n)$ remains feasible for the remaining $N - n$ constraints, i.e., iff $C(\hat{x}_J, \xi^1) \leq 0$ for all $j = n + 1, \ldots, N$. If $p(\hat{x}_J) = z$, then probability of each event $\{C(\hat{x}_J, \xi^1) \leq 0\}$ is equal to $1 - z$. Since the sample is iid, we obtain that conditional on $(\xi^1, \ldots, \xi^n) \in \Delta_z$, probability of the event $\Delta_{J,z} \cap \Sigma_J$ is equal to $(1 - z)^{N-n}$. Consequently, the unconditional probability

\[
P^N(\Delta_z \cap \Sigma_J) = P^N(\Delta_{J,z} \cap \Sigma_J) = (1 - z)^{N-n}dH(z),
\]  
(5.213)

and hence equation (5.211) follows.

It follows from (5.209) and (5.211) that

\[
\binom{N}{n} \int_0^1 (1 - z)^{N-n}dH(z) = 1, \quad N \geq n.
\]  
(5.214)

Let us observe that $H(z) := z^n$ satisfies equations (5.214) for all $N \geq n$. Indeed, using integration by parts, we have

\[
\left(\binom{N}{n}\right) \int_0^1 (1 - z)^{N-n}dz^n = -\binom{N}{n} \frac{n}{N-n+1} \int_0^1 (1 - z)^{N-n+1}dz^{n+1} = \binom{N}{n-1} \int_0^1 (1 - z)^{N-n+1}dz^{n-1} = \cdots = 1.
\]  
(5.215)

We also have that equations (5.214) determine respective moments of random variable $1 - Z$, where $Z \sim H(z)$, and hence (since random variable $p(\hat{x}_N)$ has a bounded support) by the general theory of moments these equations have unique solution. Therefore we conclude that $H(z) = z^n, 0 \leq z \leq 1$, is the cdf of $p(\hat{x}_N)$.

We also have by (5.213) that

\[
P^N\{p(\hat{x}_N) \in [z, z + dz]\} = \sum_{J \in \mathcal{J}_n} P^N(\Delta_z \cap \Sigma_J) = \binom{N}{n} (1 - z)^{N-n}dH(z).
\]  
(5.216)

Therefore, since $H(z) = z^n$ and using integration by parts similar to (5.215), we can write

\[
P^N\{p(\hat{x}_N) > a\} = \binom{N}{n} \int_0^1 (1 - z)^{N-n}dH(z) = \binom{N}{n} n \int_{a}^{1} (1 - z)^{N-n-1}dz^n = \binom{N}{n-1} (1 - a)^{N-n+1} + \frac{N}{n-1} \int_{0}^{1} (1 - z)^{N-n+1}dz^n = \cdots = \sum_{i=0}^{n-1} \binom{N}{i} a^i(1 - a)^{N-i}.
\]  
(5.217)
Since $\Pr\{p(\bar{x}_N) > \alpha\} = P^N\{p(\bar{x}_N) > \alpha\}$, this completes the proof. □

Of course, the event “$p(\bar{x}_N) > \alpha$” means that $\bar{x}_N$ is not a feasible point of the true problem (5.193). Recall that $n \leq n$. Therefore, given $\beta \in (0, 1)$, the inequality (5.206) implies that for sample size $N \geq n$ such that

$$b(n-1; \alpha, N) \leq \beta,$$

we have with probability at least $1 - \beta$ that $\bar{x}_N$ is a feasible solution of the true problem (5.193).

Recall that

$$b(n-1; \alpha, N) = \Pr(W \leq n-1),$$

where $W \sim B(N, \alpha)$ is a random variable having binomial distribution. For “not too small” $\alpha$ and large $N$, good approximation of that probability is suggested by the Central Limit Theorem. That is, $W$ has approximately normal distribution with mean $N\alpha$ and variance $N\alpha(1 - \alpha)$, and hence

$$b(n-1; \alpha, N) \approx \Phi\left(\frac{n-1-N\alpha}{\sqrt{N\alpha(1-\alpha)}}\right).$$

(5.219)

For $N\alpha \geq n-1$, Hoeffding inequality (7.194) gives the estimate

$$b(n-1; \alpha, N) \leq \exp\left\{-\frac{2(N\alpha - n+1)^2}{N}\right\},$$

(5.220)

and Chernoff inequality (7.196) gives

$$b(n-1; \alpha, N) \leq \exp\left\{-\frac{(N\alpha - n+1)^2}{2\alpha N}\right\}.$$ 

(5.221)

The estimates (5.218) and (5.221) show that the required sample size $N$ should be of order $O(\alpha^{-1})$. This, of course, is not surprising since just to estimate the probability $p(x)$, for a given $x$, by Monte Carlo sampling we will need a sample size of order $O(1/p(x))$. For example, for $n = 100$ and $\alpha = \beta = 0.01$, bound (5.218) suggests estimate $N = 12460$ for the required sample size. Normal approximation (5.219) gives practically the same estimate of $N$. The estimate derived from the bound (5.220) gives significantly bigger estimate of $N = 40372$. The estimate derived from Chernoff inequality (5.221) gives much better estimate of $N = 13410$.

Anyway this indicates that the “guaranteed” estimates like (5.218) could be too conservative for practical calculations. Note also that Theorem 5.32 does not make any claims about quality of $\bar{x}_N$ as a candidate for an optimal solution of the true problem (5.193), it only guarantees its feasibility.

15Recall that $\Phi(\cdot)$ is the cdf of standard normal distribution.
5.7.2 Validation of an Optimal Solution

We discuss now an approach to a practical validation of a candidate \( \bar{x} \) for an optimal solution of the true problem. This task is twofold, namely we need to verify feasibility and optimality of \( \bar{x} \). Let us start with verification of feasibility of \( \bar{x} \). For that we need to estimate the probability \( p(\bar{x}) = \Pr\{C(\bar{x}, \xi) > 0\} \). We proceed by employing Monte Carlo sampling techniques. For a generated iid random sample \( \xi^1, \ldots, \xi^N \), let \( m \) be the number of times that the constraints \( C(\bar{x}, \xi^j) \leq 0, j = 1, \ldots, N \), are violated, i.e.,

\[
m := \sum_{j=1}^N 1_{(0, \infty)}(C(\bar{x}, \xi^j)).
\]

Then \( \hat{p}(\bar{x}) = m/N \) is an unbiased estimator of \( p(\bar{x}) \), and \( m \) has Binomial distribution \( B(N, p(\bar{x})) \). For a given \( \beta \in (0, 1) \) consider

\[
U_N^{feas}(\bar{x}) := \sup_{\rho \in [0, 1]} \{ \rho : b(m; \rho, N) \geq \beta \}.
\]

We have that \( U_N^{feas}(\bar{x}) \) is a function of \( m \) and hence is a random variable. Note that \( b(m; \rho, N) \) is continuous and monotonically decreasing in \( \rho \in (0, 1) \). Therefore, in fact, the supremum in the right hand side of (5.222) is attained, and \( U_N^{feas}(\bar{x}) \) is equal to such \( \hat{\rho} \) that \( b(m; \hat{\rho}, N) = \beta \). Denoting \( V := b(m; \hat{\rho}(\bar{x}), N) \), we have that

\[
\Pr\left\{ p(\bar{x}) < U_N^{feas}(\bar{x}) \right\} = \Pr\left\{ V > \frac{\beta}{\hat{\rho}(\bar{x})} \right\}
= 1 - \Pr\left\{ V \leq \beta \right\} = 1 - \sum_{k=0}^N \Pr\left\{ V \leq \beta \mid m = k \right\} \Pr(m = k).
\]

Since

\[
\Pr\left\{ V \leq \beta \mid m = k \right\} = \begin{cases} 1, & \text{if } b(k; \hat{\rho}(\bar{x}), N) \leq \beta, \\ 0, & \text{otherwise}, \end{cases}
\]

and \( \Pr(m = k) = \binom{N}{k} \hat{\rho}(\bar{x})^k (1 - \hat{\rho}(\bar{x}))^{N-k} \), it follows that

\[
\sum_{k=0}^N \Pr\left\{ V \leq \beta \mid m = k \right\} \Pr(m = k) \leq \beta,
\]

and hence

\[
\Pr\left\{ p(\bar{x}) < U_N^{feas}(\bar{x}) \right\} \geq 1 - \beta.
\]

That is, \( p(\bar{x}) < U_N^{feas}(\bar{x}) \) with probability at least \( 1 - \beta \). Therefore we can take \( U_N^{feas}(\bar{x}) \) as an upper \((1 - \beta)\)-confidence bound for \( p(\bar{x}) \). In particular, if \( m = 0 \), then

\[
U_N^{feas}(\bar{x}) = 1 - \beta^1/N < N^{-1} \ln(\beta^{-1}).
\]

We obtain that if \( U_N^{feas}(\bar{x}) \leq \alpha \), then \( \bar{x} \) is a feasible solution of the true problem with probability at least \( 1 - \beta \). Since this procedure involves only calculations of \( C(\bar{x}, \xi^j) \), it
can be performed with a large sample size \( N \), and hence feasibility of \( \bar{x} \) can be verified with a high accuracy provided that \( \alpha \) is not too small.

Of course, if we have verified feasibility of \( \bar{x} \), with confidence \( 1 - \beta \), then \( f(\bar{x}) \) gives an upper bound, with confidence \( 1 - \beta \), for the optimal value \( \vartheta^* \) of the true problem (5.193). It is more tricky to construct a valid lower statistical bound for \( \vartheta^* \). One possible approach is to apply a general methodology of the SAA method (see discussion at the end of section 5.6.1). We have that for any \( \lambda \geq 0 \) the following inequality holds (compare with (5.180)):

\[
\vartheta^* \geq \inf_{x \in X} \left[ f(x) + \lambda(p(x) - \alpha) \right].
\]

(5.224)

We also have that expectation of \( \hat{\vartheta}(N) := \inf_{x \in X} \left[ f(x) + \lambda(p(x) - \alpha) \right] \)

(5.225)
gives a valid lower bound for the right hand side of (5.224), and hence for \( \vartheta^* \). An unbiased estimate of \( E[\hat{\vartheta}(N)] \) can be obtained by solving the right hand side problem of (5.225) several times and averaging calculated optimal values. Note, however, that there are two difficulties with applying this approach. First, recall that typically the function \( \hat{\vartheta}(x) \) is discontinuous and hence it could be difficult to solve these optimization problems. Second, it may happen that for any choice of \( \lambda \geq 0 \) the optimal value of the right hand side of (5.224) is smaller than \( \vartheta^* \), i.e., there is a gap between problem (5.193) and its (Lagrangian) dual.

We discuss now an alternative approach to construction statistical lower bounds. Let us choose three positive integers \( M, N, L \), with \( L \leq M \), and constant \( \gamma \in [0, 1) \), and let us generate \( M \) independent samples \( \xi_1, \ldots, \xi_N, m = 1, \ldots, M \), each of size \( N \), of random vector \( \xi \). For each sample we solve the associated optimization problem

\[
\text{Min}_{x \in X} f(x) \quad \text{subject to} \quad \sum_{j=1}^{N} 1_{(0, \infty)} \left( C(x, \xi_{j,m}) \right) \leq \gamma N,
\]

(5.226)

and hence calculate its optimal value \( \hat{\vartheta}_{\gamma,N}^{m}, m = 1, \ldots, M \). That is, we solve \( M \) times the corresponding SAA problem at the significance level \( \gamma \). In particular, for \( \gamma = 0 \) problem (5.226) takes the form

\[
\text{Min}_{x \in X} f(x) \quad \text{subject to} \quad C(x, \xi_{j,m}) \leq 0, j = 1, \ldots, N.
\]

(5.227)

It may happen that problem (5.226) is either infeasible or unbounded from below, in which case we assign its optimal value as +\( \infty \) or -\( \infty \), respectively. We can view \( \hat{\vartheta}_{\gamma,N}^{m}, m = 1, \ldots, M \), as an iid sample of the random variable \( \hat{\vartheta}_{\gamma,N} \), where \( \hat{\vartheta}_{\gamma,N} \) is the optimal value of the respective SAA problem at significance level \( \gamma \). Next we rearrange the calculated optimal values in the nondecreasing order as follows \( \hat{\vartheta}_{\gamma,N}^{(1)} \leq \ldots \leq \hat{\vartheta}_{\gamma,N}^{(M)} \), i.e., \( \hat{\vartheta}_{\gamma,N}^{(1)} \) is the smallest, \( \hat{\vartheta}_{\gamma,N}^{(2)} \) is the second smallest etc, among the values \( \hat{\vartheta}_{\gamma,N}^{m}, m = 1, \ldots, M \). By definition, we use the random quantity \( \hat{\vartheta}_{\gamma,N}^{(L)} \) as a lower bound of the true optimal value \( \vartheta^* \).

Let us analyze the resulting bounding procedure. Let \( x \) be a feasible point of the true problem, i.e., \( \Pr \{ C(x, \xi) > 0 \} \leq \alpha \). Since \( \sum_{j=1}^{N} 1_{(0, \infty)} \left( C(x, \xi_{j,m}) \right) \) has binomial
distribution with probability of success equal to the probability of the event \( C(x, \xi) > 0 \), it follows that \( x \) is feasible for problem (5.226) with probability at least\(^{16}\)

\[
\sum_{i=0}^{\lfloor \gamma N \rfloor} \binom{N}{i} \alpha^i(1-\alpha)^{N-i} = b \left( \lfloor \gamma N \rfloor; \alpha, N \right) =: \theta_N.
\]

When \( x \) is feasible for (5.227), we of course have \( \hat{\vartheta}^m_{\gamma,N} \leq f(x) \). Thus, for every \( m \in \{1, \ldots, M\} \) and for every \( \epsilon > 0 \) we have

\[
\theta := \Pr \{ \hat{\vartheta}^m_{\gamma,N} \leq \vartheta^* + \epsilon \} \geq \theta_N.
\]

Now, in the case of \( \hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^* + \epsilon \), the corresponding realization of the random sequence \( \hat{\vartheta}^1_{\gamma,N}, \ldots, \hat{\vartheta}^M_{\gamma,N} \) contains less than \( L \) elements which are less than or equal to \( \vartheta^* + \epsilon \). Since the elements of the sequence are independent, the probability of the latter event is \( b(L - 1; \theta; M) \). Since \( \theta \geq \theta_N \), we have that \( b(L - 1; \theta; M) \leq b(L - 1; \theta_N; M) \). Thus, \( \Pr \{ \hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^* + \epsilon \} \leq b(L - 1; \theta_N; M) \). Since the resulting inequality is valid for all \( \epsilon > 0 \), we arrive at the bound

\[
\Pr \{ \hat{\vartheta}^{(L)}_{\gamma,N} > \vartheta^* \} \leq b(L - 1; \theta_N, M).
\] (5.228)

We have arrived at the following result.

**Proposition 5.33.** Given \( \beta \in (0, 1) \) and \( \gamma \in [0, 1) \), let us choose positive integers \( M, N, L \) in such a way that

\[
b(L - 1; \theta_N, M) \leq \beta,
\] (5.229)

where \( \theta_N := b \left( \lfloor \gamma N \rfloor; \alpha, N \right) \). Then with probability at least \( 1 - \beta \), the random quantity \( \hat{\vartheta}^{(L)}_{\gamma,N} \) gives a lower bound for the true optimal value \( \vartheta^* \).

Note that for \( \gamma = 0 \), we have that \( \theta_N = (1 - \alpha)^N \) and the bound (5.229) takes the form

\[
\sum_{k=0}^{L-1} \binom{M}{k} (1-\alpha)^{Nk}[1-(1-\alpha)^N]^{M-k} \leq \beta.
\] (5.230)

The question arising in connection with the outlined bounding scheme is how to choose \( M, N, L \) and \( \gamma \). In the convex case it is advantageous to take \( \gamma = 0 \), since then we need to solve convex problems (5.227), rather than combinatorial problems (5.226). Given a desired reliability \( 1 - \beta \) of the bound, and \( M \) and \( N \), it is easy to specify \( L \) : this should be just the largest \( L > 0 \) satisfying condition (5.229). (If no \( L > 0 \) satisfying (5.229) exists, the lower bound, by definition, is \( -\infty \).) We end up with a question of how to choose \( M \) and \( N \). For \( N \) given, the larger is \( M \), the better. For given \( N \), the “ideal” bound yielded by our scheme as \( M \) tends to infinity, is the (lower) \( \theta_N \)-quantile of the true distribution of the random variable \( \hat{\vartheta}_N \). The larger is \( M \), the better can we estimate this quantile from a sample of \( M \) independent realizations of this random variable. In reality, however, \( M \)

\(^{16}\)Recall that the notation \( \lfloor a \rfloor \) stands for the largest integer less than or equal to \( a \in \mathbb{R} \).
is bounded by the computational effort required to solve $M$ problems (5.227). Note that
the effort per problem is larger the larger is the sample size $N$. For $L = 1$ (which is the
smallest value of $L$) and $\gamma = 0$ the left hand side of (5.230) is equal to $[1 - (1 - \alpha)^N]^M$.
Note that $(1 - \alpha)^N \approx e^{-\alpha N}$ for small $\alpha > 0$. Therefore if $\alpha N$ is large, then one will need
a very large $M$ to make $[1 - (1 - \alpha)^N]^M$ smaller than, say, $\beta = 0.01$, and hence to get a
meaningful lower bound. For example, for $\alpha N = 7$ we have that $e^{-\alpha N} = 0.0009$, and we
will need $M > 5000$ to make $[1 - (1 - \alpha)^N]^M$ smaller than 0.01. Therefore, for $\gamma = 0$ it
is recommended to take $N$ not larger than, say, $2/\alpha$.

5.8 SAA Method Applied to Multistage Stochastic Programming

Consider a multistage stochastic programming problem, in the general form (3.1), driven
by the random data process $\xi_1, \xi_2, ..., \xi_T$. The exact meaning of this formulation was dis-
cussed in section 3.1.1. In this section we discuss application of the SAA method to such
multistage problems.

Consider the following sampling scheme. Generate a sample $\xi_{ij}^1, ..., \xi_{ij}^{N_i}$ of $N_i$
realizations of random vector $\xi_2$. Conditional on each $\xi_{2j}^i, j = 1, ..., N_1$, generate a random
sample $\xi_{3j}^i$, $j = 1, ..., N_2$, of $N_2$ realizations of $\xi_3$ according to conditional distribution
of $\xi_3$ given $\xi_2 = \xi_{2j}^i$. Conditional on each $\xi_{3j}^i$ generate a random sample of size $N_3$ of $\xi_4$
conditional on $\xi_3 = \xi_{3j}^i$, and so on for later stages. (Although we do not consider such
possibility here, it is also possible to generate at each stage conditional samples of different
sizes.) In that way we generate a scenario tree with $N = \prod_{t=1}^{T-1} N_t$ number of scenarios
each taken with equal probability $1/N$. We refer to this scheme as conditional sampling. Unless stated otherwise\(^\text{17}\), we assume that, at the first stage, the sample $\xi_{21}^1, ..., \xi_{21}^{N_1}$ is iid and the following samples, at each stage $t = 2, ..., T - 1$, are conditionally iid. If, moreover, all conditional samples at each stage are independent of each other, we refer to such conditional sampling as the independent conditional sampling. The multistage stochastic
programming problem induced by the original problem (3.1) on the scenario tree generated by conditional sampling is viewed as the sample average approximation (SAA) of the
“true” problem (3.1).

It could be noted that in case of stagewise independent process $\xi_1, ..., \xi_T$, the inde-
pendent conditional sampling destroys the stagewise independence structure of the original
process. This is because at each stage conditional samples are independent of each other and hence are different. In the stagewise independence case an alternative approach is to use the same sample at each stage. That is, independent of each other random samples $\xi_{1t}^1, ..., \xi_{1t}^{N_{t-1}}$ of respective $\xi_t, t = 2, ..., T$, are generated and the corresponding scenario
tree is constructed by connecting every ancestor node at stage $t - 1$ with the same set of
children nodes $\xi_{1t}, ..., \xi_{1t}^{N_{t-1}}$. In that way stagewise independence is preserved in the sce-
nario tree generated by conditional sampling. We refer to this sampling scheme as the
identical conditional sampling.

\(^\text{17}\)It is also possible to employ Quasi-Monte Carlo sampling in constructing conditional sampling. In some situations this may reduce variability of the corresponding SAA estimators. In the following analysis we assume independence in order to simplify statistical analysis.
5.8. SAA Method Applied to Multistage Stochastic Programming

5.8.1 Statistical Properties of Multistage SAA Estimators

Similar to two stage programming it makes sense to discuss convergence of the optimal value and first stage solutions of multistage SAA problems to their true counterparts as sample sizes \( N_1, \ldots, N_{T-1} \) tend to infinity. We denote \( \mathcal{N} := \{ N_1, \ldots, N_{T-1} \} \) and by \( \theta^* \) and \( \theta_N^* \) the optimal values of the true and the corresponding SAA multistage programs, respectively.

In order to simplify the presentation let us consider now three stage stochastic programs, i.e., \( T = 3 \). In that case conditional sampling consists of sample \( \xi_i, i = 1, \ldots, N_1 \), of \( \xi_2 \) and for each \( i = 1, \ldots, N_1 \) of conditional samples \( \xi_{ij}, j = 1, \ldots, N_2 \), of \( \xi_3 \) given \( \xi_{2}, \xi_{2} \). Let us write dynamic programming equations for the true problem. We have

\[
Q_3(x_2, \xi_3) = \inf_{x_3 \in X_3(x_2, \xi_3)} f_3(x_3, \xi_3),
\]

\[
Q_2(x_1, \xi_2) = \inf_{x_2 \in X_2(x_1, \xi_2)} \{ f_2(x_2, \xi_2) + E \left[ Q_3(x_2, \xi_3) | \xi_2 \right] \},
\]

and at the first stage we solve the problem

\[
\min_{x_1 \in X_1} \left\{ f_1(x_1) + E \left[ Q_2(x_1, \xi_2) \right] \right\}.
\]

If we could calculate values \( Q_2(x_1, \xi_2) \), we could approximate problem (5.233) by the sample average problem

\[
\min_{x_1 \in X_1} \left\{ \hat{f}_{N_1}(x_1) := f_1(x_1) + \frac{1}{N_1} \sum_{i=1}^{N_1} Q_2(x_1, \xi_{2i}) \right\}.
\]

However, values \( Q_2(x_1, \xi_{2i}) \) are not given explicitly and are approximated by

\[
\hat{Q}_{2,N_2}(x_1, \xi_{2i}) := \inf_{x_2 \in X_2(x_1, \xi_{2i})} \left\{ f_2(x_2, \xi_{2i}) + \frac{1}{N_2} \sum_{j=1}^{N_2} Q_3(x_2, \xi_{3ij}) \right\},
\]

\( i = 1, \ldots, N_1 \). That is, the SAA method approximates the first stage problem (5.233) by the problem

\[
\min_{x_1 \in X_1} \left\{ \hat{f}_{N_1,N_2}(x_1) := f_1(x_1) + \frac{1}{N_1} \sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_{2i}) \right\}.
\]

In order to verify consistency of the SAA estimators, obtained by solving problem (5.236), we need to show that \( \hat{f}_{N_1,N_2}(x_1) \) converges to \( f_1(x_1) + E \left[ Q_2(x_1, \xi_2) \right] \) w.p.1 uniformly on any compact subset \( X \) of \( X_1 \) (compare with analysis of section 5.1.1). That is, we need to show that

\[
\lim_{N_1,N_2 \to \infty} \sup_{x_1 \in X} \left| \frac{1}{N_1} \sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_{2i}) - E \left[ Q_2(x_1, \xi_2) \right] \right| = 0 \text{ w.p.1}. \tag{5.237}
\]

For that it suffices to show that

\[
\lim_{N_1 \to \infty} \sup_{x_1 \in X} \left| \frac{1}{N_1} \sum_{i=1}^{N_1} Q_2(x_1, \xi_{2i}) - E \left[ Q_2(x_1, \xi_2) \right] \right| = 0 \text{ w.p.1}, \tag{5.238}
\]
and

\[ \lim_{N_1,N_2 \to \infty} \sup_{x_1 \in X} \left| \frac{1}{N_1} \sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi^*_i) - \frac{1}{N_2} \sum_{i=1}^{N_2} Q_2(x_1, \xi^*_i) \right| = 0 \text{ w.p.1.} \quad (5.239) \]

Condition (5.238) can be verified by applying a version of uniform Law of Large Numbers (see section 7.2.5). Condition (5.239) is more involved. Of course, we have that 

\[ \left| \frac{1}{N} \sum_{i=1}^{N} \hat{Q}_{2,N_2}(x_1, \xi^*_i) - \frac{1}{N} \sum_{i=1}^{N} Q_2(x_1, \xi^*_i) \right| \]

and hence condition (5.239) holds if \( \hat{Q}_{2,N_2}(x_1, \xi^*_i) \) converges to \( Q_2(x_1, \xi^*_i) \) w.p.1 as \( N_2 \to \infty \) in a certain uniform way. Unfortunately an exact mathematical analysis of such condition could be quite involved. The analysis simplifies considerably if the underline random \( X \) is of the true problem.

Let us make the following observations. By increasing sample sizes \( N_1, \ldots, N_{T-1} \), of conditional sampling, we eventually reconstruct the scenario tree structure of the original multistage problem. Therefore it should be expected that in the limit, as these sample sizes tend (simultaneously) to infinity, the corresponding SAA estimators of the optimal value and first stage solutions are consistent, i.e., converge w.p.1 to their true counterparts. And, indeed, this can be shown under certain regularity conditions. However, consistency alone does not justify the SAA method since in reality sample sizes are always finite and are constrained by available computational resources. Similar to the two stage case we have here that (for minimization problems)

\[ \vartheta^* \geq \mathbb{E} [\hat{\vartheta}_N^*]. \quad (5.240) \]

That is, the SAA optimal value \( \hat{\vartheta}_N^* \) is a downwards biased estimator of the true optimal value \( \vartheta^* \).

Suppose now that the data process \( \xi_1, \ldots, \xi_T \) is stagewise independent. As it was discussed above in that case it is possible to use two different approaches to conditional sampling, namely to use at every stage independent or the same samples for every ancestor node at the previous stage. These approaches were referred to as the independent and identical conditional samplings, respectively. Consider, for instance, the three stage stochastic programming problem (5.231)–(5.233). In the second approach of identical conditional sampling we have sample \( \xi^*_i, i = 1, \ldots, N_1 \), of \( \xi_2 \) and sample \( \xi^*_j, j = 1, \ldots, N_2 \), of \( \xi_3 \).
5.8. SAA Method Applied to Multistage Stochastic Programming

Consider the example of multistage portfolio selection discussed in section 1.4.2. Suppose for the moment that the problem has three stages $t = 0, 1, 2$. In the SAA approach we generate sample $\xi^i_1, i = 1, ..., N_0$, of returns at stage $t = 0$, and conditional samples $\xi^j_2, j = 1, ..., N_1$, of returns at stage $t = 2$. The dynamic programming equations for the SAA problem can be written as follows (see equations (1.50)–(1.52)). At stage $t = 1$ for $i = 1, ..., N_0$, we have

$$\hat{Q}_{1,N_1}(W_1, \xi^i_1) = \sup_{x_1 \geq 0} \left\{ \frac{1}{N_1} \sum_{j=1}^{N_1} U((\xi^j_2)^T x_1) : e^T x_1 = W_1 \right\},$$

(5.243)

where $e \in \mathbb{R}^n$ is vector of ones, and at stage $t = 0$ we solve the problem

$$\max_{x_0 \geq 0} \frac{1}{N_0} \sum_{i=1}^{N_0} \hat{Q}_{1,N_1}(e^T x_0, \xi^i_1) \text{ s.t. } e^T x_0 = W_0.$$  

(5.244)

Now let $U(W) := \ln W$ be the logarithmic utility function. Suppose that the data process is stagewise independent. Then the optimal value $\hat{\vartheta}^*$ of the true problem is (see (1.58))

$$\hat{\vartheta}^* = \ln W_0 + \sum_{t=0}^{T-1} \nu_t,$$

(5.245)

where $\nu_t$ is the optimal value of the problem

$$\max_{x_t \geq 0} \mathbb{E} \left[ \ln (\xi^T_{t+1} x_t) \right] \text{ s.t. } e^T x_t = 1.$$  

(5.246)

Because of independence of $\xi_2$ and $\xi_3$ we have that conditional distribution of $\xi_3$ given $\xi_2$ is the same as its unconditional distribution, and hence in both sampling approaches $\hat{Q}_{2,N_2}(x_1, \xi^i_2)$ has the same distribution independent of $i$. Therefore in both sampling schemes $\frac{1}{N_2} \sum_{i=1}^{N_2} \hat{Q}_{2,N_2}(x_1, \xi^i_2)$ has the same expectation, and hence we may expect that in both cases the estimator $\hat{\vartheta}_N^*$ has a similar bias. Variance of $\hat{\vartheta}_N^*$, however, could be quite different. In the case of independent conditional sampling we have that $\hat{Q}_{2,N_2}(x_1, \xi^i_2), i = 1, ..., N_1$, are independent of each other, and hence

$$\text{Var} \left[ \frac{1}{N_1} \sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi^i_2) \right] = \frac{1}{N_1} \text{Var} \left[ \hat{Q}_{2,N_2}(x_1, \xi^i_2) \right].$$

(5.242)

On the other hand in the case of identical conditional sampling the right hand side of (5.241) has the same component $\frac{1}{N_2} \sum_{i=1}^{N_2} Q_3(x_2, \xi^i_2)$ for all $i = 1, ..., N_1$. Consequently $\hat{Q}_{2,N_2}(x_1, \xi^i_2)$ would tend to be positively correlated for different values of $i$, and as a result $\hat{\vartheta}_N^*$ will have a higher variance than in the case of independent conditional sampling. Therefore from a statistical point of view it is advantageous to use the independent conditional sampling.

**Example 5.34 (Portfolio Selection)** Consider the example of multistage portfolio selection discussed in section 1.4.2. Suppose for the moment that the problem has three stages $t = 0, 1, 2$. In the SAA approach we generate sample $\xi^i_1, i = 1, ..., N_0$, of returns at stage $t = 0$, and conditional samples $\xi^j_2, j = 1, ..., N_1$, of returns at stage $t = 2$. The dynamic programming equations for the SAA problem can be written as follows (see equations (1.50)–(1.52)). At stage $t = 1$ for $i = 1, ..., N_0$, we have

$$\hat{Q}_{1,N_1}(W_1, \xi^i_1) = \sup_{x_1 \geq 0} \left\{ \frac{1}{N_1} \sum_{j=1}^{N_1} U((\xi^j_2)^T x_1) : e^T x_1 = W_1 \right\},$$

(5.243)
Let the SAA method be applied with the identical conditional sampling, with respective sample \( \xi^t_j, j = 1, \ldots, N_{t-1} \) of \( \xi_t, t = 1, \ldots, T \). In that case the corresponding SAA problem is also stagewise independent and the optimal value of the SAA problem

\[
\hat{\theta}_N = \ln W_0 + \sum_{t=0}^{T-1} \hat{\nu}_{t,N},
\]

(5.247)

where \( \hat{\nu}_{t,N} \) is the optimal value of the problem

\[
\max_{x_t \geq 0} \frac{1}{N_t} \sum_{j=1}^{N_t} \ln \left( (\xi^t_{j+1})^T x_t \right) \quad \text{s.t.} \quad e^T x_t = 1.
\]

(5.248)

We can view \( \hat{\nu}_{t,N} \) as an SAA estimator of \( \nu_t \). Since here we solve a maximization rather than a minimization problem, \( \hat{\nu}_{t,N} \) is an upwards biased estimator of \( \nu_t \), i.e., \( \mathbb{E}[\hat{\nu}_{t,N}] \geq \nu_t \).

We also have that \( \mathbb{E}[\hat{\theta}_N] = \ln W_0 + \sum_{t=0}^{T-1} \mathbb{E}[\hat{\nu}_{t,N}] \), and hence

\[
\mathbb{E}[\hat{\theta}_N] - \theta^* = \sum_{t=0}^{T-1} \left( \mathbb{E}[\hat{\nu}_{t,N}] - \nu_t \right).
\]

(5.249)

That is, for the logarithmic utility function and identical conditional sampling, bias of the SAA estimator of the optimal value grows additively with increase of the number of stages. Also because the samples at different stages are independent of each other, we have that

\[
\mathbb{V}_{a}[\hat{\theta}_N] = \sum_{t=0}^{T-1} \mathbb{V}_{a}[\hat{\nu}_{t,N}].
\]

(5.250)

Let now \( U(W) := W^\gamma \), with \( \gamma \in (0, 1] \), be the power utility function and suppose that the data process is stagewise independent. Then (see (1.61))

\[
\theta^* = W_0^{\gamma} \prod_{t=0}^{T-1} \eta_t,
\]

(5.251)

where \( \eta_t \) is the optimal value of problem

\[
\max_{x_t \geq 0} \mathbb{E} \left[ (\xi^T_{t+1} x_t)^\gamma \right] \quad \text{s.t.} \quad e^T x_t = 1.
\]

(5.252)

For the corresponding SAA method with the identical conditional sampling, we have that

\[
\hat{\theta}_N = W_0^{\gamma} \prod_{t=0}^{T-1} \hat{\eta}_{t,N},
\]

(5.253)

where \( \hat{\eta}_{t,N} \) is the optimal value of problem

\[
\max_{x_t \geq 0} \frac{1}{N_t} \sum_{j=1}^{N_t} \left( (\xi^T_{j+1} x_t)^\gamma \right) \quad \text{s.t.} \quad e^T x_t = 1.
\]

(5.254)
5.8. SAA Method Applied to Multistage Stochastic Programming

Because of the independence of the samples, and hence independence of \( \hat{\rho}_{t,Nt} \), we can write
\[
E[\hat{\rho}_N] = W_0^\gamma \prod_{t=0}^{T-1} E[\hat{\rho}_{t,Nt}],
\]
and hence
\[
E[\hat{\rho}_N] = \vartheta^* \prod_{t=0}^{T-1} (1 + \beta_{t,Nt}),
\tag{5.255}
\]
where \( \beta_{t,Nt} := \frac{E[\rho_{t,Nt}] - \rho_t}{\rho_t} \) is the relative bias of \( \hat{\rho}_{t,Nt} \). That is, bias of \( \hat{\rho}_N \) grows with increase of the number of stages in a multiplicative way. In particular, if the relative biases \( \beta_{t,Nt} \) are constant, then bias of \( \hat{\rho}_N \) grows exponentially fast with increase of the number of stages.

**Statistical Validation Analysis**

By (5.240) we have that the optimal value \( \hat{\vartheta}_N \) of SAA problem gives a valid statistical lower bound for the optimal value \( \vartheta^* \). Therefore in order to construct a lower bound for \( \vartheta^* \) one can proceed exactly in the same way as it was discussed in section 5.6.1. Unfortunately, typically the bias and variance of \( \hat{\vartheta}_N \) grow fast with increase of the number of stages, which makes the corresponding statistical lower bounds quite inaccurate already for a mild number of stages.

In order to construct an upper bound we proceed as follows. Let \( x_t(\xi_{[t]}) \) be a feasible policy. Recall that a policy is feasible if it satisfies the feasibility constraints (3.2). Since the multistage problem can be formulated as the minimization problem (3.3) we have that
\[
E\left[ f_1(x_1) + f_2(x_2(\xi_{[2]}), \xi_2) + \ldots + f_T(x_T(\xi_{[T]}), \xi_T) \right] \geq \vartheta^*,
\tag{5.256}
\]
and equality in (5.256) holds iff the policy \( x_t(\xi_{[t]}) \) is optimal. The expectation in the left hand side of (5.256) can be estimated in a straightforward way. That is, generate random sample \( \xi_1^j, \ldots, \xi_T^j, j = 1, \ldots, N \), of \( N \) realizations (scenarios) of the random data process \( \xi_1, \ldots, \xi_T \) and estimate this expectation by the average
\[
\frac{1}{N} \sum_{j=1}^{N} \left[ f_1(x_1) + f_2(x_2(\xi_2^j), \xi_2) + \ldots + f_T(x_T(\xi_T^j), \xi_T) \right].
\tag{5.257}
\]

Note that in order to construct the above estimator we do not need to generate a scenario tree, say by conditional sampling, we only need to generate a sample of single scenarios of the data process. The above estimator (5.257) is an unbiased estimator of the expectation in the left hand side of (5.256), and hence is a valid statistical upper bound for \( \vartheta^* \). Of course, quality of this upper bound depends on a successful choice of the feasible policy, i.e., on how small is the optimality gap between the left and right hand sides of (5.256). It also depends on variability of the estimator (5.257), which unfortunately often grows fast with increase of the number of stages.

We also may address the problem of validating a given feasible first stage solution \( x_1 \in \mathcal{X}_1 \). The value of the multistage problem at \( x_1 \) is given by the optimal value of the problem
\[
\min_{x_2, \ldots, x_T} \quad f_1(x_1) + E\left[ f_2(x_2(\xi_{[2]}), \xi_2) + \ldots + f_T(x_T(\xi_{[T]}), \xi_T) \right] \\
\text{s.t.} \quad x_t(\xi_{[t]}) \in \mathcal{X}_t(x_{t-1}(\xi_{[t-1]}), \xi_t), \quad t = 2, \ldots, T.
\tag{5.258}
\]
Recall that the optimization in (5.258) is performed over feasible policies. That is, in order to validate \( \bar{x}_1 \) we basically need to solve the corresponding \( T-1 \) stage problems. Therefore, for \( T > 2 \) validation of \( \bar{x}_1 \) can be almost as difficult as solving the original problem.

### 5.8.2 Complexity Estimates of Multistage Programs

In order compute value of two stage stochastic program \( \min_{x \in X} \mathbb{E}[F(x, \xi)] \), where \( F(x, \xi) \) is the optimal value of the corresponding second stage problem, at a feasible point \( \bar{x} \in X \) we need to calculate the expectation \( \mathbb{E}[F(\bar{x}, \xi)] \). This, in turn, involves two difficulties. First, the objective value \( F(\bar{x}, \xi) \) is not given explicitly, its calculation requires solution of the associated second stage optimization problem. Second, the multivariate integral \( \mathbb{E}[F(\bar{x}, \xi)] \) cannot be evaluated with a high accuracy even for moderate values of dimension \( d \) of the random data vector \( \xi \). Monte Carlo techniques allow to evaluate \( \mathbb{E}[F(x, \xi)] \) with accuracy \( \varepsilon > 0 \) by employing samples of size \( N = O(\varepsilon^{-2}) \). The required sample size \( N \) gives, in a sense, an estimate of complexity of evaluation of \( \mathbb{E}[F(\bar{x}, \xi)] \) since this is how many times we will need to solve the corresponding second stage problem. It is remarkable that in order to solve the two stage stochastic program with accuracy \( \varepsilon > 0 \), say by the SAA method, we need a sample size basically of the same order \( N = O(\varepsilon^{-2}) \). These complexity estimates were analyzed in details in section 5.3. Two basic conditions required for such analysis are that the problem has relatively complete recourse and that for given \( x \) and \( \xi \) the optimal value \( F(x, \xi) \) of the second stage problem can be calculated with a high accuracy.

In this section we discuss analogous estimates of complexity of the SAA method applied to multistage stochastic programming problems. From the point of view of the SAA method it is natural to evaluate complexity of a multistage stochastic program in terms of the total number of scenarios required to find a first stage solution with a given accuracy \( \varepsilon > 0 \).

In order to simplify the presentation we consider three stage stochastic programs, say of the form (5.231)–(5.233). Assume that for every \( x_1 \in X_1 \) the expectation \( \mathbb{E}[Q_2(x_1, \xi_2)] \) is well defined and finite valued. In particular, this assumption implies that the problem has relatively complete recourse. Let us look at the problem of computing value of the first stage problem (5.233) at a feasible point \( \bar{x}_1 \in X_1 \). Apart from the problem of evaluating the expectation \( \mathbb{E}[Q_2(x_1, \xi_2)] \), we also face here the problem of computing \( Q_2(\bar{x}_1, \xi_2) \) for different realizations of random vector \( \xi_2 \). For that we need to solve the two stage stochastic programming problem given in the right hand side of (5.232). As it was already discussed, in order to evaluate \( Q_2(\bar{x}_1, \xi_2) \) with accuracy \( \varepsilon > 0 \) by solving the corresponding SAA problem, given in the right hand side of (5.235), we also need a sample of size \( N_2 = O(\varepsilon^{-2}) \). Recall that the total number of scenarios involved in evaluation of the sample average \( \bar{f}_{N_1, N_2}(\bar{x}_1) \), defined in (5.236), is \( N = N_1 N_2 \). Therefore we will need \( N = O(\varepsilon^{-4}) \) scenarios just to compute value of the first stage problem at a given feasible point with accuracy \( \varepsilon \) by the SAA method. This indicates that complexity of the SAA method, applied to multistage stochastic programs, grows exponentially with increase of the number of stages.

We discuss now in details sample size estimates of the three stage SAA program (5.234)–(5.236). For the sake of simplicity we assume that the data process is stagewise independent, i.e., random vectors \( \xi_2 \) and \( \xi_3 \) are independent. Also similar to assumptions...
5.8. SAA Method Applied to Multistage Stochastic Programming \(229\)

(M1)–(M5) of section 5.3 let us make the following assumptions.

(M'1) For every \(x_1 \in X_1\) the expectation \(\mathbb{E}[Q_2(x_1, \xi_2)]\) is well defined and finite valued.

(M'2) The random vectors \(\xi_2\) and \(\xi_3\) are independent.

(M'3) The set \(X_1\) has finite diameter \(D_1\).

(M'4) There is a constant \(L_1 > 0\) such that
\[
|Q_2(x'_1, \xi_2) - Q_2(x_1, \xi_2)| \leq L_1\|x'_1 - x_1\|, 
\tag{5.259}
\]

for all \(x'_1, x_1 \in X_1\) and a.e. \(\xi_2\).

(M'5) There exists a constant \(\sigma_1 > 0\) such that for any \(x_1 \in X_1\) it holds that
\[
M_{1,x_1}(t) \leq \exp \left\{ \frac{\sigma_1^2 t^2}{2} \right\}, \quad \forall t \in \mathbb{R}, \tag{5.260}
\]

where \(M_{1,x_1}(t)\) is the moment generating function of \(Q_2(x_1, \xi_2) - \mathbb{E}[Q_2(x_1, \xi_2)]\).

(M'6) There is a set \(C\) of finite diameter \(D_2\) such that for every \(x_1 \in X_1\) and a.e. \(\xi_2\), the set \(X_2(x_1, \xi_2)\) is contained in \(C\).

(M'7) There is a constant \(L_2 > 0\) such that
\[
|Q_3(x'_2, \xi_3) - Q_3(x_2, \xi_3)| \leq L_2\|x'_2 - x_2\|, 
\tag{5.261}
\]

for all \(x'_2, x_2 \in C\) and a.e. \(\xi_3\).

(M'8) There exists a constant \(\sigma_2 > 0\) such that for any \(x_2 \in X_2(x_1, \xi_2)\) and all \(x_1 \in X_1\) and a.e. \(\xi_2\) it holds that
\[
M_{2,x_2}(t) \leq \exp \left\{ \frac{\sigma_2^2 t^2}{2} \right\}, \quad \forall t \in \mathbb{R}, \tag{5.262}
\]

where \(M_{2,x_2}(t)\) is the moment generating function of \(Q_3(x_2, \xi_3) - \mathbb{E}[Q_3(x_2, \xi_3)]\).

**Theorem 5.35.** Under assumptions (M'1)–(M'8) and for \(\varepsilon > 0\) and \(\alpha \in (0, 1)\), and the sample sizes \(N_1\) and \(N_2\) (using either independent or identical conditional sampling schemes) satisfying
\[
\left[ \frac{O(1)D_1L_1}{\varepsilon} \right]^{N_1} \exp \left\{ - \frac{O(1)N_1\varepsilon^2}{\sigma_1^2} \right\} + \left[ \frac{O(1)D_2L_2}{\varepsilon} \right]^{N_2} \exp \left\{ - \frac{O(1)N_2\varepsilon^2}{\sigma_2^2} \right\} \leq \alpha, \tag{5.263}
\]

we have that any \(\varepsilon/2\)-optimal solution of the SAA problem (5.236) is an \(\varepsilon\)-optimal solution of the first stage (5.233) of the true problem with probability at least \(1 - \alpha\).

**Proof.** Proof of this theorem is based on the uniform exponential bound of Theorem 7.67. Let us sketch the arguments. Assume that the conditional sampling is identical. We have that for every \(x_1 \in X_1\) and \(i = 1, ..., N_1\),
\[
|Q_{2,N_i}(x_1, \xi'_2) - Q_2(x_1, \xi_2)| \leq \sup_{x_2 \in C} \left| \frac{1}{N_2} \sum_{j=1}^{N_2} Q_3(x_2, \xi'_3) - \mathbb{E}[Q_3(x_2, \xi_3)] \right|, 
\]

\[
\text{Proof.}
\]
Therefore bound (5.268) also implies that
\[
\Pr\{\left| \hat{\vartheta}_N - \vartheta^* \right| > \varepsilon \} \leq \left[ \frac{O(1)}{\varepsilon} \right]^{n_1} \exp \left\{ \frac{-O(1)N_1 \varepsilon^2}{\sigma_f^2} \right\} + \left[ \frac{O(1)}{\varepsilon} \right]^{n_2} \exp \left\{ \frac{-O(1)N_2 \varepsilon^2}{\sigma_f^2} \right\},
\]

which implies the assertion of the theorem.

In the case of the independent conditional sampling the proof can be completed in a similar way. \qed

**Remark 14.** We have, of course, that
\[
\left| \hat{\vartheta}_N - \vartheta^* \right| \leq \sup_{x_1 \in X_1} \left| \hat{f}_{N_1, N_2}(x_1) - f_1(x_1) - \mathbb{E}[Q_2(x_1, \xi_2)] \right|.
\]

Therefore bound (5.269) also implies that
\[
\Pr\{\left| \hat{\vartheta}_N - \vartheta^* \right| > \varepsilon \} \leq \left[ \frac{O(1)}{\varepsilon} \right]^{n_1} \exp \left\{ \frac{-O(1)N_1 \varepsilon^2}{\sigma_f^2} \right\} + \left[ \frac{O(1)}{\varepsilon} \right]^{n_2} \exp \left\{ \frac{-O(1)N_2 \varepsilon^2}{\sigma_f^2} \right\}.
\]
5.9. Stochastic Approximation Method

In particular, suppose that $N_1 = N_2$. Then for

$$n := \max\{n_1, n_2\}, \quad L := \max\{L_1, L_2\}, \quad D := \max\{D_1, D_2\}, \quad \sigma := \max\{\sigma_1, \sigma_2\},$$

the estimate (5.263) implies the following estimate of the required sample size $N_1 = N_2$:

$$\left(\frac{O(1)DL}{\varepsilon}\right)^n \exp\left\{-\frac{O(1)N_1^2}{\sigma^2}\right\} \leq \alpha,$$

which is equivalent to

$$N_1 \geq \frac{O(1)\sigma^2}{\varepsilon^2} \left[ n \ln \left(\frac{O(1)DL}{\varepsilon}\right) + \ln \left(\frac{1}{\alpha}\right)\right].$$

The estimate (5.272), for 3-stage programs, looks similar to the estimate (5.113), of Theorem 5.18, for two-stage programs. Recall, however, that if we use the SAA method with conditional sampling and respective sample sizes $N_1$ and $N_2$, then the total number of scenarios is $N = N_1N_2$. Therefore, our analysis indicates that for 3-stage problems we need random samples with the total number of scenarios of order of the square of the corresponding sample size for two-stage problems. This analysis can be extended to $T$-stage problems with the conclusion that the total number of scenarios needed to solve the true problem with a reasonable accuracy grows exponentially with increase of the number of stages $T$. Some numerical experiments seem to confirm this conclusion. Of course, it should be mentioned that the above analysis does not prove in a rigorous mathematical sense that complexity of multistage programming grows exponentially with increase of the number of stages. It only indicates that the SAA method, which showed a considerable promise for solving two-stage problems, could be practically inapplicable for solving multistage problems with a large (say greater than 4) number of stages.

5.9 Stochastic Approximation Method

To an extent this section is based on Nemirovski et al [133]. Consider stochastic optimization problem (5.1). We assume that the expected value function $f(x) = \mathbb{E}[F(x, \xi)]$ is well defined, finite valued and continuous at every $x \in X$, and that the set $X \subset \mathbb{R}^n$ is nonempty, closed and bounded. We denote by $\bar{x}$ an optimal solution of problem (5.1). Such an optimal solution does exist since the set $X$ is compact and $f(x)$ is continuous. Clearly, $\vartheta^* = f(\bar{x})$ (recall that $\vartheta^*$ denotes the optimal value of problem (5.1)). We also assume throughout this section that the set $X$ is convex and the function $f(\cdot)$ is convex. Of course, if $F(\cdot, \xi)$ is convex for every $\xi \in \Xi$, then convexity of $f(\cdot)$ follows. We assume availability of the following stochastic oracle:

- There is a mechanism which for every given $x \in X$ and $\xi \in \Xi$ returns value $F(x, \xi)$ and a stochastic subgradient, a vector $G(x, \xi)$ such that $g(x) := \mathbb{E}[G(x, \xi)]$ is well defined and is a subgradient of $f(\cdot)$ at $x$, i.e., $g(x) \in \partial f(x)$. 
Remark 15. Recall that if $F(\cdot, \xi)$ is convex for every $\xi \in \Xi$, and $x$ is an interior point of $X$, i.e., $f(\cdot)$ is finite valued in a neighborhood of $x$, then
\begin{equation}
\partial f(x) = \mathbb{E}[\partial_x F(x, \xi)] \tag{5.273}
\end{equation}
(see Theorem 7.47). Therefore, in that case we can employ a measurable selection $G(x, \xi) \in \partial_x F(x, \xi)$ as a stochastic subgradient. Note also that for an implementation of a stochastic approximation algorithm we only need to employ stochastic subgradients, while objective values $F(x, \xi)$ are used for accuracy estimates in section 5.9.4.

We also assume that we can generate, say by Monte Carlo sampling techniques, an iid sequence $\xi^j, j = 1, \ldots$, of realizations of the random vector $\xi$, and hence to compute a stochastic subgradient $G(x_j, \xi^j)$ at an iterate point $x_j \in X$.

### 5.9.1 Classical Approach

We denote by $\|x\|_2 = (x^T x)^{1/2}$ the Euclidean norm of vector $x \in \mathbb{R}^n$, and by
\begin{equation}
\Pi_X(x) := \arg\min_{z \in X} \|x - z\|_2 \tag{5.274}
\end{equation}
the metric projection of $x$ onto the set $X$. Since $X$ is convex and closed, the minimizer in the right hand side of (5.274) exists and is unique. Note that $\Pi_X$ is a nonexpanding operator, i.e.,
\begin{equation}
\|\Pi_X(x') - \Pi_X(x)\|_2 \leq \|x' - x\|_2, \quad \forall x', x \in \mathbb{R}^n. \tag{5.275}
\end{equation}

The classical Stochastic Approximation (SA) algorithm solves problem (5.1) by mimicking a simple subgradient descent method. That is, for chosen initial point $x_1 \in X$ and a sequence $\gamma^j > 0, j = 1, \ldots$, of stepsizes, it generates the iterates by the formula
\begin{equation}
x_{j+1} = \Pi_X (x_j - \gamma_j G(x_j, \xi^j)). \tag{5.276}
\end{equation}

The crucial question of that approach is how to choose the stepsizes $\gamma_j$. Also the set $X$ should be simple enough so that the corresponding projection can be easily calculated. We are going now to analyze convergence of the iterates, generated by this procedure, to an optimal solution $\bar{x}$ of problem (5.1). Note that the iterate $x_{j+1} = x_{j+1}(\xi^j), j = 1, \ldots$, is a function of the history $\xi^j = (\xi^1, \ldots, \xi^j)$ of the generated random process and hence is random, while the initial point $x_1$ is given (deterministic). We assume that there is number $M > 0$ such that
\begin{equation}
\mathbb{E} \left[\|G(x, \xi)\|_2^2\right] \leq M^2, \quad \forall x \in X. \tag{5.277}
\end{equation}

Note that since for a random variable $Z$ it holds that $\mathbb{E}(Z^2) \geq (\mathbb{E}|Z|)^2$, it follows from (5.277) that $\mathbb{E}\|G(x, \xi)\| \leq M$.

Denote
\begin{equation}
A_j := \frac{1}{2}\|x_j - \bar{x}\|_2^2 \quad \text{and} \quad a_j := \mathbb{E}[A_j] = \frac{1}{2}\mathbb{E}[\|x_j - \bar{x}\|_2^2]. \tag{5.278}
\end{equation}

By (5.275) and since $\bar{x} \in X$ and hence $\Pi_X(\bar{x}) = \bar{x}$, we have
\begin{equation}
A_{j+1} = \frac{1}{2}\|\Pi_X (x_j - \gamma_j G(x_j, \xi^j)) - \bar{x}\|_2^2 \\
= \frac{1}{2}\|\Pi_X (x_j - \gamma_j G(x_j, \xi^j)) - \Pi_X(\bar{x})\|_2^2 \\
\leq \frac{1}{2}\|x_j - \gamma_j G(x_j, \xi^j) - \bar{x}\|_2^2 \\
= A_j + \frac{1}{2}\gamma_j^2\|G(x_j, \xi^j)\|_2^2 - \gamma_j(x_j - \bar{x})^T G(x_j, \xi^j). \tag{5.279}
\end{equation}
Since $x_j = x_j(\xi_{j-1})$ is independent of $\xi_j$, we have

$$
E[(x_j - \bar{x})^T G(x_j, \xi)] = E\{E[(x_j - \bar{x})^T G(x_j, \xi) | \xi_{j-1}]\}
= E\{E[(x_j - \bar{x})^T E[G(x_j, \xi) | \xi_{j-1}]]\}
= E\{(x_j - \bar{x})^T g(x_j)\}.
$$

Therefore, by taking expectation of both sides of (5.279) and since (5.277) we obtain

$$
a_{j+1} \leq a_j - \gamma_j E[(x_j - \bar{x})^T g(x_j)] + \frac{1}{2} \gamma_j^2 M^2. \quad (5.280)
$$

Suppose, further, that the expectation function $f(x)$ is differentiable and strongly convex on $X$ with parameter $c > 0$, i.e.,

$$
(x' - x)^T (\nabla f(x') - \nabla f(x)) \geq c\|x' - x\|^2_2, \quad \forall x, x' \in X. \quad (5.281)
$$

Note that strong convexity of $f(x)$ implies that the minimizer $\bar{x}$ is unique, and that because of differentiability of $f(x)$ it follows that $\partial f(x) = \{\nabla f(x)\}$ and hence $g(x) = \nabla f(x)$.

By optimality of $\bar{x}$ we have that

$$
(x - \bar{x})^T \nabla f(\bar{x}) \geq 0, \quad \forall x \in X, \quad (5.282)
$$

which together with (5.281) implies that

$$
E[(x_j - \bar{x})^T \nabla f(x_j)] \geq cE[\|x_j - \bar{x}\|^2_2] = 2ca_j. \quad (5.283)
$$

Therefore it follows from (5.280) that

$$
a_{j+1} \leq (1 - 2c\gamma_j)a_j + \frac{1}{2} \gamma_j^2 M^2. \quad (5.284)
$$

In the classical approach to stochastic approximation the employed stepsizes are $\gamma_j := \theta/j$ for some constant $\theta > 0$. Then by (5.284) we have

$$
a_{j+1} \leq (1 - 2\theta/j)a_j + \frac{1}{2} \theta^2 M^2/j^2. \quad (5.285)
$$

Suppose now that $\theta > 1/(2c)$. Then it follows from (5.285) by induction that for $j = 1, \ldots,$

$$
2a_j \leq \max \left\{ \frac{\theta^2 M^2(2\theta - 1)^{-1}}{j}, 2a_1 \right\}. \quad (5.286)
$$

Recall that $2a_j = E[\|x_j - \bar{x}\|^2_2]$ and, since $x_1$ is deterministic, $2a_1 = \|x_1 - \bar{x}\|^2_2$. Therefore, by (5.286) we have that

$$
E[\|x_j - \bar{x}\|^2_2] \leq \frac{Q(\theta)}{j}, \quad (5.287)
$$

where

$$
Q(\theta) := \max \left\{ \theta^2 M^2(2\theta - 1)^{-1}, \|x_1 - \bar{x}\|^2_2 \right\}. \quad (5.288)
$$

The constant $Q(\theta)$ attains its optimal (minimal) value at $\theta = 1/c$. 

Suppose, further, that \( \bar{x} \) is an interior point of \( X \) and \( \nabla f(x) \) is Lipschitz continuous, i.e., there is constant \( L > 0 \) such that
\[
\| \nabla f(x') - \nabla f(x) \|_2 \leq L \| x' - x \|_2, \quad \forall x', x \in X.
\] (5.289)

Then
\[
f(x) \leq f(\bar{x}) + \frac{1}{2} L \| x - \bar{x} \|_2^2, \quad \forall x \in X,
\] (5.290)

and hence by (5.287)
\[
E \left[ f(x_j) - f(\bar{x}) \right] \leq \frac{1}{2} L E \left[ \| x_j - \bar{x} \|_2^2 \right] \leq \frac{Q(\theta)L}{2j}.
\] (5.291)

We obtain that under the specified assumptions, after \( j \) iterations the expected error of the current solution in terms of the distance to the true optimal solution \( \bar{x} \) is of order \( O(j^{-1/2}) \), and the expected error in terms of the objective value is of order \( O(j^{-1}) \), provided that \( \theta > 1/(2c) \). Note, however, that the classical stepsize rule \( \gamma_j = \theta/j \) could be very dangerous if the parameter \( c \) of strong convexity is overestimated, i.e., if \( \theta < 1/(2c) \).

**Example 5.36** As a simple example consider \( f(x) := \frac{1}{2} \kappa x^2 \), with \( \kappa > 0 \) and \( X := [-1,1] \subset \mathbb{R} \), and assume that there is no noise, i.e., \( G(x, \xi) \equiv \nabla f(x) \). Clearly \( \bar{x} = 0 \) is the optimal solution and zero is the optimal value of the corresponding optimization (minimization) problem. Let us take \( \theta = 1 \), i.e., use stepsizes \( \gamma_j = 1/j \), in which case the iteration process becomes
\[
x_{j+1} = x_j - f'(x_j)/j = \left( 1 - \frac{\kappa}{j} \right) x_j.
\] (5.292)

For \( \kappa = 1 \) the above choice of the stepsizes is optimal and the optimal solution is obtained in one iteration.

Suppose now that \( \kappa < 1 \). Then starting with \( x_1 > 0 \), we have
\[
x_{j+1} = x_1 \prod_{s=1}^{j} \left( 1 - \frac{\kappa}{s} \right) = x_1 \exp \left\{ - \sum_{s=1}^{j} \ln \left( 1 + \frac{\kappa}{s - \kappa} \right) \right\} > x_1 \exp \left\{ - \sum_{s=1}^{j} \frac{\kappa}{s - \kappa} \right\}.
\]

Moreover,
\[
\sum_{s=1}^{j} \frac{\kappa}{s - \kappa} \leq \frac{\kappa}{1 - \kappa} + \int_{1}^{j} \frac{\kappa}{t - \kappa} dt < \frac{\kappa}{1 - \kappa} + \kappa \ln j - \kappa \ln(1 - \kappa).
\]

It follows that
\[
x_{j+1} > O(1) j^{-\kappa} \quad \text{and} \quad f(x_{j+1}) > O(1) j^{-2\kappa}, \quad j = 1, \ldots.
\] (5.293)

(In the first of the above inequalities the constant \( O(1) = x_1 \exp\{-\kappa/(1-\kappa)+\kappa \ln(1-\kappa)\} \), and in the second inequality the generic constant \( O(1) \) is obtained from the first one by taking square and multiplying it by \( \kappa/2 \).) That is, the convergence becomes extremely slow for small \( \kappa \) close to zero. In order to reduce the value \( x_j \) (the objective value \( f(x_j) \))
by factor 10, i.e., to improve the error of current solution by one digit, we will need to increase the number of iterations \( j \) by factor \( 10^{1/\kappa} \) (by factor \( 10^{1/(2\kappa)} \)). For example, for \( \kappa = 0.1 \), \( x_1 = 1 \) and \( j = 10^5 \) we have that \( x_j > 0.28 \). In order to reduce the error of the iterate to 0.028 we will need to increase the number of iterations by factor \( 10^{10} \), i.e., to \( j = 10^{15} \).

It could be added that if \( f(x) \) loses strong convexity, i.e., the parameter \( \epsilon \) degenerates to zero, and hence no choice of \( \theta > 1/(2\epsilon) \) is possible, then the stepsizes \( \gamma_j = \theta/j \) may become completely unacceptable for any choice of \( \theta \).

### 5.9.2 Robust SA Approach

It was argued in the previous section 5.9.1 that the classical stepsizes \( \gamma_j = O(j^{-1}) \) can be too small to ensure a reasonable rate of convergence even in the “no noise” case. An important improvement of the SA method was developed by Polyak [152] and Polyak and Juditsky [153], where longer stepsizes were suggested with consequent averaging of the obtained iterates. Under the outlined “classical” assumptions, the resulting algorithm exhibits the same optimal \( O(j^{-1}) \) asymptotical convergence rate, while using an easy to implement and “robust” stepsize policy. The main ingredients of Polyak’s scheme (long steps and averaging) were, in a different form, proposed already in Nemirovski and Yudin [135] for problems with general type Lipschitz continuous convex objectives and for convex-concave saddle point problems. Results of this section go back to Nemirovski and Yudin [135], [136].

Recall that \( g(x) \in \partial f(x) \) and \( \alpha_j = \frac{1}{\gamma_j} \mathbb{E} [\|x_j - \bar{x}\|^2] \), and we assume the boundedness condition (5.277). By convexity of \( f(x) \) we have that \( f(x) \geq f(x_j) + (x - x_j)^T g(x_j) \) for any \( x \in X \), and hence

\[
\mathbb{E}[(x_j - \bar{x})^T g(x_j)] \geq \mathbb{E}[f(x_j) - f(\bar{x})].
\]  

(5.294)

Together with (5.280) this implies

\[
\gamma_j \mathbb{E}[f(x_j) - f(\bar{x})] \leq \alpha_j - \alpha_{j+1} + \frac{1}{2} \gamma_j^2 M^2.
\]

It follows that whenever \( 1 \leq i \leq j \), we have

\[
\sum_{t=1}^{j} \gamma_t \mathbb{E}[f(x_t) - f(\bar{x})] \leq \sum_{t=1}^{j} \alpha_t - \alpha_{t+1} + \frac{1}{2} M^2 \sum_{t=1}^{j} \gamma_t^2 \leq \alpha_i + \frac{1}{2} M^2 \sum_{t=1}^{j} \gamma_t^2.
\]

(5.295)

Denote

\[
\nu_t := \frac{\gamma_t}{\sum_{\tau=i}^{j} \gamma_\tau} \text{ and } D_X := \max_{x \in X} \|x - x_1\|_2.
\]

(5.296)

Clearly \( \nu_t \geq 0 \) and \( \sum_{t=1}^{j} \nu_t = 1 \). By (5.295) we have

\[
\mathbb{E} \left[ \sum_{t=1}^{j} \nu_t f(x_t) - f(\bar{x}) \right] \leq \frac{\alpha_i + \frac{1}{2} M^2 \sum_{t=1}^{j} \gamma_t^2}{\sum_{t=1}^{j} \gamma_t}.
\]  

(5.297)
Consider points
\[ \tilde{x}_{i,j} := \sum_{t=i}^{j} \nu_t x_t. \] (5.298)

Since \( X \) is convex it follows that \( \tilde{x}_{i,j} \in X \) and by convexity of \( f(\cdot) \) we have
\[ f(\tilde{x}_{i,j}) \leq \sum_{t=i}^{j} \nu_t f(x_t). \]

Thus, by (5.297) and in view of \( a_i \leq D_X^2 \) and \( a_i \leq 4D_X^2, i > 1 \), we get
\[ E[f(\tilde{x}_{1,j}) - f(\bar{x})] \leq \frac{D_X^2 + M^2 \sum_{t=1}^{j} \gamma_t^2}{2 \sum_{t=1}^{j} \gamma_t}, \quad \text{for } 1 \leq j, \] (5.299)
\[ E[f(\tilde{x}_{i,j}) - f(\bar{x})] \leq \frac{4D_X^2 + M^2 \sum_{t=i}^{j} \gamma_t^2}{2 \sum_{t=i}^{j} \gamma_t}, \quad \text{for } 1 < i \leq j. \] (5.300)

Based of the above bounds on the expected accuracy of approximate solutions \( \tilde{x}_{i,j} \), we can now develop "reasonable" stepsize policies along with the associated efficiency estimates.

**Constant stepsizes and error estimates**

Assume now that the number of iterations of the method is fixed in advance, say equal to \( N \), and that we use the constant stepsize policy, i.e., \( \gamma_t = \gamma, t = 1, \ldots, N \). It follows then from (5.299) that
\[ E[f(\tilde{x}_{1,N}) - f(\bar{x})] \leq \frac{D_X^2 + M^2 N \gamma^2}{2N \gamma}. \] (5.301)

Minimizing the right hand side of (5.301) over \( \gamma > 0 \), we arrive at the constant stepsize policy
\[ \gamma_t = \frac{D_X}{M \sqrt{N}}, \quad t = 1, \ldots, N, \] (5.302)
along with the associated efficiency estimate
\[ E[f(\tilde{x}_{1,N}) - f(\bar{x})] \leq \frac{D_X M}{\sqrt{N}}. \] (5.303)

By (5.300), with the constant stepsize policy (5.302), we also have for \( 1 \leq K \leq N \),
\[ E[f(\tilde{x}_{K,N}) - f(\bar{x})] \leq \frac{C_{N,K} D_X M}{\sqrt{N}}. \] (5.304)

where
\[ C_{N,K} := \frac{2N}{N-K+1} + \frac{1}{2}. \]
5.9. Stochastic Approximation Method

When $K/N \leq 1/2$, the right hand side of (5.304) coincides, within an absolute constant factor, with the right hand side of (5.303). If we change the stepsizes (5.302) by a factor of $\theta > 0$, i.e., use the stepsizes

$$
\gamma_t = \frac{\theta D_X M}{\sqrt{N}}, \quad t = 1, \ldots, N,
$$

then the efficiency estimate (5.304) becomes

$$
\mathbb{E} [f(\bar{x}_{K,N}) - f(\bar{x})] \leq \max \{\theta, \theta^{-1}\} \frac{C_{N,K} D_X M}{\sqrt{N}}.
$$

The expected error of the iterates (5.298), with constant stepsize policy (5.305), after $N$ iterations is $O(N^{-1/2})$. Of course, this is worse than the rate $O(N^{-1})$ for the classical SA algorithm as applied to a smooth strongly convex function attaining minimum at an interior point of the set $X$. However, the error bound (5.306) is guaranteed independently of any smoothness and/or strong convexity assumptions on $f(\cdot)$. Moreover, changing the stepsizes by factor $\theta$ results just in rescaling of the corresponding error estimate (5.306). This is in a sharp contrast with the classical approach discussed in the previous section, when such change of stepsizes can be a disaster. These observations, in particular the fact that there is no necessity in “fine tuning” the stepsizes to the objective function $f(\cdot)$, explains the adjective “robust” in the name of the method.

It can be interesting to compare sample size estimates derived from the error bounds of the (robust) SA approach with respective sample size estimates of the SAA method discussed in section 5.3.2. By Chebyshev (Markov) inequality we have that for $\varepsilon > 0$,

$$
\Pr \{f(\bar{x}_{1,N}) - f(\bar{x}) \geq \varepsilon\} \leq \varepsilon^{-1} \mathbb{E} [f(\bar{x}_{1,N}) - f(\bar{x})].
$$

Together with (5.303) this implies that, for the constant stepsize policy (5.302),

$$
\Pr \{f(\bar{x}_{1,N}) - f(\bar{x}) \geq \varepsilon\} \leq \frac{D_X M}{\varepsilon \sqrt{N}}.
$$

It follows that for $\alpha \in (0,1)$ and sample size

$$
N \geq \frac{D_X^2 M^2}{\varepsilon^2 \alpha^2}
$$

we are guaranteed that $\bar{x}_{1,N}$ is an $\varepsilon$-optimal solution of the “true” problem (5.1) with probability at least $1 - \alpha$.

Compared with the corresponding estimate (5.123) for the sample size by the SAA method, the above estimate (5.309) is of the same order with respect to parameters $D_X, M$ and $\varepsilon$. On the other hand, the dependence on the significance level $\alpha$ is different, in (5.123) it is of order $O(\ln(\alpha^{-1}))$, while in (5.309) it is of order $O(\alpha^{-2})$. It is possible to derive better estimates, similar to the respective estimates of the SAA method, of the required sample size by using Large Deviations theory, we will discuss this further in the next section (see Theorem 5.41 in particular).
5.9.3 Mirror Descent SA Method

The robust SA approach discussed in the previous section is tailored to Euclidean structure of the space $\mathbb{R}^n$. In this section we discuss a generalization of the Euclidean SA approach allowing to adjust, to some extent, the method to the geometry, not necessary Euclidean, of the problem in question. A rudimentary form of the following generalization can be found in Nemirovski and Yudin [136], from where the name “Mirror Descent” originates.

In this section we denote by $\| \cdot \|$ a general norm on $\mathbb{R}^n$. Its dual norm is defined as

$$
\|x\|_* := \sup_{\|y\| \leq 1} y^T x.
$$

By $\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ we denote the $\ell_p$, $p \in [1, \infty)$, norm on $\mathbb{R}^n$. In particular, $\| \cdot \|_2$ is the Euclidean norm. Recall that the dual of $\| \cdot \|_p$ is the norm $\| \cdot \|_q$, where $q > 1$ is such that $1/p + 1/q = 1$. The dual norm of $\ell_1$ norm $\|x\|_1 = |x_1| + \cdots + |x_n|$ is the $\ell_\infty$ norm $\|x\|_\infty = \max \{|x_1|, \ldots, |x_n|\}$.

**Definition 5.37.** We say that a function $\vartheta : X \to \mathbb{R}$ is a distance generating function with modulus $\kappa > 0$ with respect to norm $\| \cdot \|$ if the following holds: $\vartheta(\cdot)$ is convex continuous on $X$, the set

$$
X^* := \{x \in X : \vartheta(x) \neq 0\}
$$

is convex, $\vartheta(\cdot)$ is continuously differentiable on $X^*$, and

$$
(x' - x)^T (\nabla \vartheta(x') - \nabla \vartheta(x)) \geq \kappa \|x' - x\|^2, \quad \forall x, x' \in X^*. \tag{5.311}
$$

Note that the set $X^*$ includes the relative interior of the set $X$, and hence condition (5.311) implies that $\vartheta(\cdot)$ is strongly convex on $X$ with the parameter $\kappa$ taken with respect to the considered norm $\| \cdot \|$.

A simple example of a distance generating function (with modulus 1 with respect to the Euclidean norm) is $\vartheta(x) := \frac{1}{2} x^T x$. Of course, this function is continuously differentiable at every $x \in \mathbb{R}^n$. Another interesting example is the entropy function

$$
\vartheta(x) := \sum_{i=1}^n x_i \ln x_i, \tag{5.312}
$$

defined on the standard simplex $X := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, \ x \geq 0\}$ (note that by continuity, $x \ln x = 0$ for $x = 0$). Here the set $X^*$ is formed by points $x \in X$ having all coordinates different from zero. The set $X^*$ is the subset of $X$ of those points at which the entropy function is differentiable with $\nabla \vartheta(x) = (1 + \ln x_1, \ldots, 1 + \ln x_n)$. The entropy function is strongly convex with modulus 1 on standard simplex with respect to $\| \cdot \|_1$ norm.

Indeed, it suffices to verify that $h^T \nabla^2 \vartheta(x) h \geq \|h\|_1^2$ for every $h \in \mathbb{R}^n$ and $x \in X^*$. This, in turn, is verified by

$$
\left[\sum_i h_i^2\right]^2 = \left[\sum_i (x_i^2 h_i^2) x_i^2\right] \leq \left[\sum_i h_i^2 x_i^{-1}\right] \left[\sum_i x_i\right] = \sum_i h_i^2 x_i^{-1} = h^T \nabla^2 \vartheta(x) h, \tag{5.313}
$$

where the inequality follows by Cauchy inequality.
5.9. Stochastic Approximation Method

Let us define function \( V : X^* \times X \to \mathbb{R}_+ \) as follows
\[
V(x, z) := \vartheta(z) - [\vartheta(x) + \nabla \vartheta(x)^T(z - x)]. \tag{5.314}
\]

In what follows we refer to \( V(\cdot, \cdot) \) as prox-function\(^{18}\) associated with distance generating function \( \vartheta(x) \). Note that \( V(x, \cdot) \) is nonnegative and is strongly convex with modulus \( \kappa \) with respect to the norm \( \| \cdot \| \). Let us define prox-mapping \( P_x : \mathbb{R}^n \to X^* \), associated with the distance generating function and a point \( x \in X^* \), viewed as a parameter, as follows
\[
P_x(y) := \arg \min_{z \in X} \{ y^T(z - x) + V(x, z) \}. \tag{5.315}
\]

Observe that the minimum in the right hand side of (5.315) is attained since \( \vartheta(\cdot) \) is continuous on \( X \) and \( X \) is compact, and a corresponding minimizer is unique since \( V(x, \cdot) \) is strongly convex on \( X \). Moreover, by the definition of the set \( X^* \), all these minimizers belong to \( X^* \). Thus, the prox-mapping is well defined.

For the (Euclidean) distance generating function \( \vartheta(x) := \frac{1}{2} x^T x \), we have that \( P_x(y) = \Pi_X (x - y) \). In that case the iteration formula (5.276) of the SA algorithm can be written as
\[
x_{j+1} = P_x(\gamma_j G(x_j, \xi_j)), \quad x_1 \in X^*. \tag{5.316}
\]

Our goal is to demonstrate that the main properties of the recurrence (5.276) are inherited by (5.316) for any distance generating function \( \vartheta(x) \).

**Lemma 5.38.** For every \( u \in X, x \in X^* \) and \( y \in \mathbb{R}^n \) one has
\[
V(P_x(y), u) \leq V(x, u) + y^T(u - x) + (2\kappa)^{-1}\| y \|^2_x. \tag{5.317}
\]

**Proof.** Let \( x \in X^* \) and \( v := P_x(y) \). Note that \( v \) is of the form \( \arg \min_{z \in X} \{ h^T z + \vartheta(z) \} \) and thus \( v \in X^* \), so that \( \vartheta(\cdot) \) is differentiable at \( v \). Since \( \nabla_v V(x, v) = \nabla \vartheta(v) - \nabla \vartheta(x) \), the optimality conditions for (5.315) imply that
\[
(\nabla \vartheta(v) - \nabla \vartheta(x) + y)^T(v - u) \leq 0, \quad \forall u \in X. \tag{5.318}
\]

Therefore, for \( u \in X \) we have
\[
V(v, u) - V(x, u) = [\vartheta(u) - \nabla \vartheta(x)^T(u - v) - \vartheta(v)] - [\vartheta(u) - \nabla \vartheta(x)^T(u - x) - \vartheta(x)]
\leq y^T(u - v) - V(x, v),
\]

where the last inequality follows by (5.318).

For any \( a, b \in \mathbb{R}^n \) we have by the definition of the dual norm that \( \| a \|_\ast \| b \| \geq a^T b \) and hence
\[
(\| a \|^2_\ast / \kappa + \kappa \| b \|^2)/2 \geq \| a \|_\ast \| b \| \geq a^T b. \tag{5.319}
\]

Applying this inequality with \( a = y \) and \( b = x - v \) we obtain
\[
y^T(x - v) \leq \frac{\| y \|^2_\ast}{2\kappa} + \kappa \frac{\| x - v \|^2}{2}.
\]

\(^{18}\)The function \( V(\cdot, \cdot) \) is also called Bregman divergence.
Also due to the strong convexity of \(V(x, \cdot)\) and since \(V(x, x) = 0\) we have
\[
V(x, v) \geq V(x, x) + (x - v)\nabla_x V(x, v) + \frac{1}{2\kappa}||x - v||^2
\]
\[
= (x - v)^T (\nabla \theta(v) - \nabla \theta(x)) + \frac{1}{2\kappa}||x - v||^2
\]
(5.320)
where the last inequality holds by convexity of \(\theta(\cdot)\). We get
\[
V(v, u) - V(x, u) \leq y^T (u - v) - V(x, v) = y^T (u - x) + y^T (x - v) - V(x, v)
\]
\[
\leq y^T (u - x) + (2\kappa)^{-1} ||y||^2,
\]
as required in (5.317).

Using (5.317) with \(x = x_j, y = \gamma_j G(x_j, \xi^j)\) and \(u = \bar{x}\), and noting that by (5.316) \(x_{j+1} = P_{\bar{x}}(y)\), we get
\[
\gamma_j (x_j - \bar{x})^T G(x_j, \xi^j) \leq V(x_j, \bar{x}) - V(x_{j+1}, \bar{x}) + \frac{\gamma_j^2}{2\kappa} ||G(x_j, \xi^j)||^2.
\]
(5.321)

Let us observe that for the Euclidean distance generating function \(\theta(x) = \frac{1}{2} x^T x\), one has \(V(x, z) = \frac{1}{2} ||x - z||^2\) and \(\kappa = 1\). That is, in the Euclidean case (5.321) becomes
\[
\frac{1}{2} ||x_{j+1} - \bar{x}||^2 \leq \frac{1}{2} ||x_j - \bar{x}||^2 + \frac{1}{2} \gamma_j^2 ||G(x_j, \xi^j)||^2 - \gamma_j (x_j - \bar{x})^T G(x_j, \xi^j).
\]
(5.322)
The above inequality is exactly the relation (5.279), which played a crucial role in the developments related to the Euclidean SA. We are about to process, in a similar way, the relation (5.321) in the case of a general distance generating function, thus arriving at the Mirror Descent SA.

Specifically, setting
\[
\Delta_j := G(x_j, \xi^j) - g(x_j),
\]
we can rewrite (5.321), with \(j\) replaced by \(t\), as
\[
\gamma_t (x_t - \bar{x})^T g(x_t) \leq V(x_t, \bar{x}) - V(x_{t+1}, \bar{x}) - \gamma_t \Delta_t^T (x_t - \bar{x}) + \frac{\gamma_t^2}{2\kappa} ||G(x_t, \xi^t)||^2.
\]
(5.324)

Summing up over \(t = 1, \ldots, j\), and taking into account that \(V(x_{j+1}, u) \geq 0, u \in X\), we get
\[
\sum_{t=1}^j \gamma_t (x_t - \bar{x})^T g(x_t) \leq V(x_1, \bar{x}) + \sum_{t=1}^j \frac{\gamma_t^2}{2\kappa} ||G(x_t, \xi^t)||^2 - \sum_{t=1}^j \gamma_t \Delta_t^T (x_t - \bar{x}).
\]
(5.325)

Set \(\nu_t := \frac{\gamma_t}{\sum_{r=1}^j \gamma_r}, t = 1, \ldots, j\), and
\[
\tilde{x}_{1,j} := \sum_{t=1}^j \nu_t x_t.
\]
(5.326)

By convexity of \(f(\cdot)\) we have \(f(x_t) - f(\bar{x}) \leq (x_t - \bar{x})^T g(x_t)\), and hence
\[
\sum_{t=1}^j \gamma_t (x_t - \bar{x})^T g(x_t) \geq \sum_{t=1}^j \gamma_t [f(x_t) - f(\bar{x})]
\]
\[
= \left(\sum_{t=1}^j \gamma_t\right) \left[\sum_{t=1}^j \nu_t f(x_t) - f(\bar{x})\right]
\]
\[
\geq \left(\sum_{t=1}^j \gamma_t\right) [f(\tilde{x}_{1,j}) - f(\bar{x})].
\]
5.9. Stochastic Approximation Method

Combining this with (5.325) we obtain
\[
f(\tilde{x}_{1,j}) - f(\bar{x}) \leq \frac{V(x_1, \bar{x}) + \sum_{t=1}^{j} (2\kappa)^{-1} \gamma_t^2 \|G(x_t, \xi_t')\|_*^2 - \sum_{t=1}^{j} \gamma_t \Delta_t^T (x_t - \bar{x})}{\sum_{t=1}^{j} \gamma_t}. \quad (5.327)
\]

- Assume from now on that the procedure starts with the minimizer of \(d(\cdot, \cdot)\), that is
  \[
x_1 := \arg\min_{x \in X} d(x).
  \]

Since by the optimality of \(x_1\) we have that \((u - x_1)^T \nabla d(x_1) \geq 0\) for any \(u \in X\), it follows from the definition (5.314) of the function \(V(\cdot, \cdot)\) that
\[
\max_{u \in X} V(x_1, u) \leq D_{d,X}^2, \quad (5.329)
\]
where
\[
D_{d,X} := \left[ \max_{u \in X} d(u) - \min_{x \in X} d(x) \right]^{1/2}.
\]
(5.330)
Together with (5.327) this implies
\[
f(\tilde{x}_{1,j}) - f(\bar{x}) \leq \frac{D_{d,X}^2 + \sum_{t=1}^{j} (2\kappa)^{-1} \gamma_t^2 \|G(x_t, \xi_t')\|_*^2 - \sum_{t=1}^{j} \gamma_t \Delta_t^T (x_t - \bar{x})}{\sum_{t=1}^{j} \gamma_t}. \quad (5.331)
\]
We also have (see (5.320)) that \(V(x_1, u) \geq \frac{1}{2} \kappa \|x_1 - u\|^2\), and hence it follows from (5.329) that for all \(u \in X\):
\[
\|x_1 - u\| \leq \sqrt{\frac{2}{\kappa} D_{d,X}}. \quad (5.332)
\]
Let us assume, as in the previous section (see (5.277)), that there is a positive number \(M\) such that
\[
E[\|G(x, \xi)\|_*^2] \leq M^2, \quad \forall x \in X. \quad (5.333)
\]

**Proposition 5.39.** Let \(x_1 := \arg\min_{x \in X} d(x)\) and suppose that condition (5.333) holds. Then
\[
E[f(\tilde{x}_{1,j}) - f(\bar{x})] \leq \frac{D_{d,X}^2 + (2\kappa)^{-1} M^2 \sum_{t=1}^{j} \gamma_t^2}{\sum_{t=1}^{j} \gamma_t}. \quad (5.334)
\]

**Proof.** Taking expectations of both sides of (5.331) and noting that: (i) \(x_t\) is a deterministic function of \(\xi_{[t-1]} = (\xi_1, ..., \xi^{t-1})\), (ii) conditional on \(\xi_{[t-1]}\), the expectation of \(\Delta_t\) is 0, and (iii) the expectation of \(\|G(x_t, \xi_t')\|_*^2\) does not exceed \(M^2\), we obtain (5.334). \(\square\)
Chapter 5. Statistical Inference

Constant stepsize policy

Assume that the total number of steps \( N \) is given in advance and the constant stepsize policy \( \gamma_t = \gamma, t = 1, ..., N \), is employed. Then (5.334) becomes

\[
E \left[ f(\tilde{x}_{1,j}) - f(\bar{x}) \right] \leq \frac{D_{b,X}^2 + (2\kappa)^{-1}M_*^2N\gamma^2}{N\gamma}. \tag{5.335}
\]

Minimizing the right hand side of (5.335) over \( \gamma > 0 \) we arrive at the constant stepsize policy

\[
\gamma_t = \frac{\sqrt{2\kappa}D_{b,X}}{M_* \sqrt{N}}, \quad t = 1, ..., N, \tag{5.336}
\]

and the associated efficiency estimate

\[
E \left[ f(\tilde{x}_{1,N}) - f(\bar{x}) \right] \leq D_{b,X}M_* \sqrt{\frac{2}{\kappa N}}. \tag{5.337}
\]

This can be compared with the respective stepsize (5.302) and efficiency estimate (5.303) for the robust Euclidean SA method. Passing from the stepsizes (5.336) to the stepsizes

\[
\gamma_t = \frac{\theta \sqrt{2\kappa}D_{b,X}}{M_* \sqrt{N}}, \quad t = 1, ..., N, \tag{5.338}
\]

with rescaling parameter \( \theta > 0 \), the efficiency estimate becomes

\[
E \left[ f(\tilde{x}_{1,N}) - f(\bar{x}) \right] \leq \max \left\{ \theta, \frac{\theta - 1}{\theta} \right\} D_{b,X}M_* \sqrt{\frac{2}{\kappa N}}, \tag{5.339}
\]

similar to the Euclidean case. We refer to the SA method based on (5.316),(5.326) and (5.338) as Mirror Descent SA algorithm with constant stepsize policy.

Comparing (5.303) to (5.337), we see that for both the Euclidean and the Mirror Descent SA algorithms, the expected inaccuracy, in terms of the objective values of the approximate solutions is \( O(N^{-1/2}) \). A benefit of the Mirror Descent over the Euclidean algorithm is in potential possibility to reduce the constant factor hidden in \( O(\cdot) \) by adjusting the norm \( \| \cdot \| \) and the distance generating function \( d(\cdot) \) to the geometry of the problem.

Example 5.40 Let \( X := \{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, \ x \geq 0 \} \) be the standard simplex. Consider two setups for the Mirror Descent SA. Namely, the Euclidean setup where the considered norm \( \| \cdot \| := \| \cdot \|_2 \) and \( d(x) := \frac{1}{2}x^T x \), and \( \ell_1 \)-setup where \( \| \cdot \| := \| \cdot \|_1 \) and \( d(\cdot) \) is the entropy function (5.312). The Euclidean setup leads to the Euclidean Robust SA which is easily implementable. Note that the Euclidean diameter of \( X \) is \( \sqrt{2} \) and hence is independent of \( n \). The corresponding efficiency estimate is

\[
E \left[ f(\tilde{x}_{1,N}) - f(\bar{x}) \right] \leq O(1) \max \left\{ \theta, \theta^{-1} \right\} MN^{-1/2}, \tag{5.340}
\]

with \( M^2 = \sup_{x \in X} \mathbb{E} \left[ \| G(x, \xi) \|_2^2 \right] \).

The \( \ell_1 \)-setup corresponds to \( X^* = \{ x \in X : x > 0 \} \), \( D_{b,X} = \sqrt{\ln n} \),

\[
x_1 := \arg\min_{x \in X} d(x) = n^{-1}(1, ..., 1)^T,
\]
5.9. Stochastic Approximation Method

\[ \|x\|_* = \|x\|_\infty \text{ and } \kappa = 1 \] (see (5.313) for verification that \( \kappa = 1 \)). The associated Mirror Descent SA is easily implementable. The prox-function here is

\[ V(x, z) = \sum_{i=1}^{n} z_i \ln \frac{z_i}{x_i}, \]

and the prox-mapping \( P_x(y) \) is given by the explicit formula

\[ [P_x(y)]_i = \frac{x_i e^{-y_i}}{\sum_{k=1}^{n} x_k e^{-y_k}}, \quad i = 1, \ldots, n. \]

The respective efficiency estimate of the \( \ell_1 \)-setup is

\[ \mathbb{E}[f(\tilde{x}_{1,N}) - f(\bar{x})] \leq O(1) \max \{ \theta, \theta^{-1} \} (\ln n)^{1/2} M_* N^{-1/2}, \quad (5.341) \]

with \( M^2 = \sup_{x \in X} \mathbb{E}[\|G(x, \xi)\|_2^2] \), provided that the constant stepsizes (5.338) are used.

To compare (5.341) and (5.340), observe that \( M_* \leq M \), and the ratio \( M_* / M \) can be as small as \( n^{-1/2} \). Thus, the efficiency estimate for the \( \ell_1 \)-setup is never much worse than the estimate for the Euclidean setup, and for large \( n \) can be far better than the latter estimate. That is,

\[ \sqrt{\frac{1}{\ln n}} \leq \frac{M}{\sqrt{\ln n M_*}} \leq \sqrt{\frac{n}{\ln n}}, \]

with both the upper and the lower bounds being achievable. Thus, when \( X \) is a standard simplex of large dimension, we have strong reasons to prefer the \( \ell_1 \)-setup to the usual Euclidean one.

**Comparison with the SAA approach**

Similar to (5.307)–(5.309), by using Chebyshev (Markov) inequality, it is possible to derive from (5.339) an estimate of the sample size \( N \) which guarantees that \( \tilde{x}_{1,N} \) is an \( \varepsilon \)-optimal solution of the true problem with probability at least \( 1 - \alpha \). It is possible, however, to obtain much finer bounds on deviation probabilities when imposing more restrictive assumptions on the distribution of \( G(x, \xi) \). Specifically, assume that there is constant \( M_* > 0 \) such that

\[ \mathbb{E}\left[ \exp \left\{ \left\| G(x, \xi) \right\|_2^2 / M^2 \right\} \right] \leq \exp\{1\}, \quad \forall x \in X. \quad (5.342) \]

Note that condition (5.342) is stronger than (5.333). Indeed, if a random variable \( Y \) satisfies \( \mathbb{E}[\exp\{Y/a]\} \leq \exp\{1\} \) for some \( a > 0 \), then by Jensen inequality

\[ \exp\{\mathbb{E}[Y/a]\} \leq \mathbb{E}[\exp\{Y/a\}] \leq \exp\{1\}. \]

and therefore \( \mathbb{E}[Y] \leq a \). By taking \( Y := \|G(x, \xi)\|_2^2 \) and \( a := M^2 \), we obtain that (5.342) implies (5.333). Of course, condition (5.342) holds if \( \|G(x, \xi)\|_* \leq M_* \) for all \( (x, \xi) \in X \times \Xi \).
Theorem 5.41. Suppose that condition (5.342) is fulfilled. Then for the constant stepsizes (5.338), the following holds for any $\Theta \geq 0$:

$$\Pr \left\{ f(\tilde{x}_{1,N}) - f(\bar{x}) \geq \frac{C(1 + \Theta)}{\sqrt{\kappa N}} \right\} \leq 4 \exp\{-\Theta\},$$

(5.343)

where $C := \left( \max \{\theta, \theta^{-1}\} + 8\sqrt{3} \right) M_d D_{\delta,X} / \sqrt{2}$.

Proof. By (5.331) we have

$$f(\tilde{x}_{1,N}) - f(\bar{x}) \leq A_1 + A_2,$$

(5.344)

where

$$A_1 := \frac{D^2_{\delta,X} + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma^2_t \|G(x_t, \xi^t)\|^2_*}{\sum_{t=1}^{N} \gamma_t} \quad \text{and} \quad A_2 := \sum_{t=1}^{N} \nu_t \Delta_t^T (\bar{x} - x_t).$$

Consider $Y_t := \gamma^2_t \|G(x_t, \xi^t)\|^2_*$ and $c_t := M^2 \gamma^2_t$. Note that by (5.342),

$$\mathbb{E} \left[ \exp\{Y_t/c_t\} \right] \leq \exp\{1\}, \quad i = 1, \ldots, N.$$

(5.345)

Since $\exp\{\cdot\}$ is a convex function we have

$$\exp\left\{ \frac{\sum_{t=1}^{N} Y_t}{\sum_{t=1}^{N} c_t} \right\} = \exp\left\{ \sum_{t=1}^{N} \frac{c_t(Y_t/c_t)}{\sum_{t=1}^{N} c_t} \right\} \leq \sum_{t=1}^{N} \frac{c_t(Y_t/c_t)}{\sum_{t=1}^{N} c_t} \exp\{Y_t/c_t\}.$$

By taking expectation of both sides of the above inequality and using (5.345) we obtain

$$\mathbb{E} \left[ \exp\left\{ \frac{\sum_{t=1}^{N} Y_t}{\sum_{t=1}^{N} c_t} \right\} \right] \leq \exp\{1\}.$$

Consequently by Chebyshev’s inequality we have for any number $\Theta$:

$$\Pr \left[ \exp\left\{ \frac{\sum_{t=1}^{N} Y_t}{\sum_{t=1}^{N} c_t} \right\} \geq \exp\{\Theta\} \right] \leq \frac{\exp\{1\}}{\exp\{\Theta\}} = \exp\{-\Theta\},$$

and hence

$$\Pr \left\{ \sum_{t=1}^{N} Y_t \geq \Theta \sum_{t=1}^{N} c_t \right\} \leq \exp\{-\Theta\} \leq 3 \exp\{-\Theta\}.$$

(5.346)

That is, for any $\Theta$:

$$\Pr \left\{ \sum_{t=1}^{N} \gamma^2_t \|G(x_t, \xi^t)\|^2_* \geq \Theta M^2 \sum_{t=1}^{N} \gamma^2_t \right\} \leq 3 \exp\{-\Theta\}.$$

(5.347)

For the constant stepsize policy (5.338), we obtain by (5.347) that

$$\Pr \left\{ A_1 \geq \max\{\theta, \theta^{-1}\} \frac{M_d D_{\delta,X} (1+\Theta)}{\sqrt{2\kappa N}} \right\} \leq 3 \exp\{-\Theta\}.$$

(5.348)

Consider now the random variable $A_2$. By (5.332) we have that

$$\|\bar{x} - x_t\| \leq \|x_1 - \bar{x}\| + \|x_1 - x_t\| \leq 2\sqrt{2\kappa^{-1/2} D_{\delta,X}},$$

where $D_{\delta,X}$ is the distance between the two sets $\delta$ and $X$.
and hence
\[ |\Delta_t^T (\bar{x} - x_t)|^2 \leq \|\Delta_t\|^2 \|\bar{x} - x_t\|^2 \leq 8\kappa^{-1}D^2_{\delta,X} \|\Delta_t\|^2. \]

We also have that
\[ E\left[ (\bar{x} - x_t)^T \Delta_t | \xi_{[t-1]} \right] = (\bar{x} - x_t)^T E\left[ \Delta_t | \xi_{[t-1]} \right] = 0 \text{ w.p.1,} \]
and by condition (5.342) that
\[ E\left[ \exp\left\{ \frac{\|\Delta_t\|^2}{4M^2} \right\} | \xi_{[t-1]} \right] \leq \exp\{1\} \text{ w.p.1.} \]

Consequently by applying inequality (7.200) of Proposition 7.64 with
\[ \phi_t := \nu_t \Delta_t^T (\bar{x} - x_t) \]
and
\[ \sigma^2_t := 32\kappa^{-1}M^2 D^2_{\delta,X} \nu_t^2 \]
we obtain for any \( \Theta \geq 0 \):
\[ \Pr\left\{ A_2 \geq 4\sqrt{2\kappa^{-1/2}M_* D_{\delta,X} \Theta \sqrt{\sum_{i=1}^N \nu_t^2}} \right\} \leq \exp\{-\Theta^2/3\}. \]

Finally, (5.343) follows from (5.344), (5.348) and (5.350).

By setting \( \varepsilon = \frac{C(1+\Theta)}{\sqrt{\kappa N}} \), we can rewrite the estimate (5.343) in the form\(^{19}\)
\[ \Pr\left\{ f(\tilde{x}_{1,N}) - f(\bar{x}) > \varepsilon \right\} \leq 12 \exp\{-\varepsilon C^{-1} \sqrt{\kappa N}\}. \]

For \( \varepsilon > 0 \) this gives the following estimate of the sample size \( N \) which guarantees that
\( \tilde{x}_{1,N} \) is an \( \varepsilon \)-optimal solution of the true problem with probability at least \( 1 - \alpha \):
\[ N \geq O(1)\varepsilon^{-2} \kappa^{-1}M^2 D^2_{\delta,X} \ln^2 \left( \frac{12}{\alpha} \right). \]

This estimate is similar to the respective estimate (5.123) of the sample size for the SAA method. However, as far as complexity of solving the problem numerically is concerned, it could be mentioned that the SAA method requires a solution of the generated optimization problem, while an SA algorithm is based on computing a single subgradient \( G(x_j, \xi_j) \) at each iteration point. As a result, for the same sample size \( N \) it typically takes considerably less computation time to run an SA algorithm than to solve the corresponding SAA problem.

### 5.9. Accuracy Certificates for Mirror Descent SA Solutions

We discuss now a way to estimate lower and upper bounds for the optimal value of problem (5.1) by employing SA iterates. This will give us an accuracy certificate for obtained solutions. Assume that we run an SA procedure with respective iterates \( x_1, \ldots, x_N \) computed

\(^{19}\)The constant 12 in the right hand side of (5.351) comes from the simple estimate \( 4 \exp\{1\} < 12 \).
according to formula (5.316). As before, set

\[ \nu_t := \frac{\gamma_t}{\sum_{\tau=1}^T \gamma_t}, \quad t = 1, \ldots, N, \quad \text{and } \bar{x}_{1,N} := \sum_{t=1}^N \nu_t x_t. \]

We assume now that the stochastic objective value \( F(x, \xi) \), as well as the stochastic subgradient \( G(x, \xi) \), are computable at a given point \( (x, \xi) \in X \times \Xi \).

Consider

\[ f_N^* := \min_{x \in X} f^N(x) \quad \text{and} \quad f^N := \sum_{t=1}^N \nu_t f(x_t), \quad (5.353) \]

where

\[ f^N(x) := \sum_{t=1}^N \nu_t \left[ f(x_t) + g(x_t)^T (x - x_t) \right]. \quad (5.354) \]

Since \( \nu_t > 0 \) and \( \sum_{t=1}^N \nu_t = 1 \), by convexity of \( f(x) \) we have that the function \( f^N(x) \) underestimates \( f(x) \) everywhere on \( X \), and hence \( f_N^* \leq \vartheta^* \). Since \( \bar{x}_{1,N} \in X \) we also have that \( \vartheta^* \leq f(\bar{x}_{1,N}) \), and by convexity of \( f \) that \( f(\bar{x}_{1,N}) \leq f^N \). It follows that \( \vartheta^* \leq f^N \). That is, for any realization of the random process \( \xi^1, \ldots \), we have that

\[ f_N^* \leq \vartheta^* \leq f^N. \quad (5.355) \]

It follows, of course, that \( \mathbb{E}[f_N^*] \leq \vartheta^* \leq \mathbb{E}[f^N] \) as well.

Along with the “unobservable” bounds \( f_N^*, f^N \), consider their observable (computable) counterparts

\[ f_N := \min_{x \in X} \left\{ \sum_{t=1}^N \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^T (x - x_t)] \right\}, \quad (5.356) \]

which will be referred to as online bounds. The bound \( f_N \) can be easily calculated while running the SA procedure. The bound \( f_N \) involves solving the optimization problem of minimizing a linear in \( x \) objective function over set \( X \). If the set \( X \) is defined by linear constraints, this is a linear programming problem.

Since \( x_t \) is a function of \( \xi^{[t-1]} \) and \( \xi^t \) is independent of \( \xi^{[t-1]} \), we have that

\[ \mathbb{E}[f_N] = \sum_{t=1}^N \nu_t \mathbb{E}\left\{ \mathbb{E}[F(x_t, \xi^t)] | \xi^{[t-1]} \right\} = \sum_{t=1}^N \nu_t \mathbb{E}[f(x_t)] = \mathbb{E}[f^N] \]

and

\[ \mathbb{E}[f_N] = \mathbb{E}\left\{ \min_{x \in X} \left\{ \sum_{t=1}^N \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^T (x - x_t)] \right\} | \xi^{[t-1]} \right\} \]

\[ \leq \mathbb{E}\left\{ \min_{x \in X} \left\{ \mathbb{E}\left[ \sum_{t=1}^N \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^T (x - x_t)] \right] | \xi^{[t-1]} \right\} \right\} \]

\[ = \mathbb{E}\left[ \min_{x \in X} f_N(x) \right] = \mathbb{E}[f^N]. \]

\(^{20}\)Recall that \( \vartheta^* \) denotes the optimal value of the true problem (5.1).
5.9. Stochastic Approximation Method

It follows that
\[ E[f^N] \leq \vartheta^* \leq E[\overline{f}^N]. \] (5.357)

That is, on average $f^N$ and $\overline{f}^N$ give, respectively, a lower and an upper bound for the optimal value $\vartheta^*$ of the optimization problem (5.1).

In order to see how good are the bounds $f^N$ and $\overline{f}^N$ let us estimate expectations of the corresponding errors. We will need the following result.

**Lemma 5.42.** Let $\zeta_t \in \mathbb{R}^n$, $v_1 \in X^*$ and $v_{t+1} = P_{v_t}(\zeta_t)$, $t = 1, \ldots, N$. Then
\[ \sum_{t=1}^{N} \zeta_t^T(v_t - u) \leq V(v_1, u) + (2\kappa)^{-1} \sum_{t=1}^{N} \|\zeta_t\|_2^2, \quad \forall u \in X. \] (5.358)

**Proof.** By the estimate (5.317) of Lemma 5.38 with $x = v_t$ and $y = \zeta_t$ we have that the following inequality holds for any $u \in X$,
\[ V(v_{t+1}, u) \leq V(v_t, u) + \zeta_t^T(u - v_t) + (2\kappa)^{-1} \|\zeta_t\|_2^2. \] (5.359)

Summing this over $t = 1, \ldots, N$, we obtain
\[ V(v_{N+1}, u) \leq V(v_1, u) + \sum_{t=1}^{N} \zeta_t^T(u - v_t) + (2\kappa)^{-1} \sum_{t=1}^{N} \|\zeta_t\|_2^2. \] (5.360)

Since $V(v_{N+1}, u) \geq 0$, (5.358) follows. \qed

Consider again condition (5.333), that is
\[ E[\|G(x, \xi)\|_2^2] \leq M^2, \quad \forall x \in X, \] (5.361)
and the following condition: there is a constant $Q > 0$ such that
\[ \text{Var}[F(x, \xi)] \leq Q^2, \quad \forall x \in X. \] (5.362)

Note that, of course, $\text{Var}[F(x, \xi)] = E[(F(x, \xi) - f(x))^2]$.

**Theorem 5.43.** Suppose that conditions (5.361) and (5.362) hold. Then
\[ E[f^* - f^N] \leq \frac{2D_{\overline{f}, X}^2 + 2\kappa^{-1}M^2 \sum_{t=1}^{N} \gamma_t^2}{\sum_{t=1}^{N} \gamma_t}, \] (5.363)
\[ E[\overline{f}^N - f^*] \leq Q \sqrt{\sum_{t=1}^{N} \nu_t^2}, \] (5.364)
\[ E[f^N - f^*] \leq \left( Q + 4\sqrt{2\kappa^{-1/2}M_{\text{var}}D_{\overline{f}, X}} \right) \sqrt{\sum_{t=1}^{N} \nu_t^2} \]
\[ + \frac{D_{\overline{f}, X}^2 + 2\kappa^{-1}M^2 \sum_{t=1}^{N} \gamma_t^2}{\sum_{t=1}^{N} \gamma_t}. \] (5.365)
Proof. If in Lemma 5.42 we take \( v_1 := x_1 \) and \( \zeta_t := \gamma_t G(x_t, \xi^t) \), then the corresponding iterates \( v_t \) coincide with \( x_t \). Therefore, we have by (5.358) and since \( V(x_t, u) \leq D_{0,X}^2 \) that
\[
\sum_{t=1}^{N} \gamma_t (x_t - u)^T G(x_t, \xi^t) \leq D_{0,X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|G(x_t, \xi^t)\|^2, \quad \forall u \in X. \tag{5.366}
\]

It follows that for any \( u \in X \) (compare with (5.325)),
\[
\sum_{t=1}^{N} \nu_t \left[ -f(x_t) + (x_t - u)^T g(x_t) \right] + \sum_{t=1}^{N} \nu_t f(x_t) \leq D_{0,X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|G(x_t, \xi^t)\|^2 + \sum_{t=1}^{N} \nu_t \Delta_t^T (x_t - u),
\]
where \( \Delta_t := G(x_t, \xi^t) - g(x_t) \). Since
\[
f^{*N} - f^{*N}_u = \sum_{t=1}^{N} \nu_t f(x_t) + \max_{u \in X} \sum_{t=1}^{N} \nu_t \left[ -f(x_t) + (x_t - u)^T g(x_t) \right],
\]
it follows that
\[
f^{*N} - f^{*N}_u \leq D_{0,X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|G(x_t, \xi^t)\|^2 + \max_{u \in X} \sum_{t=1}^{N} \nu_t \Delta_t^T (x_t - u). \tag{5.367}
\]

Let us estimate the second term in the right hand side of (5.367). By using Lemma 5.42 with \( v_1 := x_1 \) and \( \zeta_t := \gamma_t \Delta_t \), and the corresponding iterates \( v_{t+1} = P_{v_t} (\zeta_t) \), we obtain
\[
\sum_{t=1}^{N} \gamma_t \Delta_t^T (v_t - u) \leq D_{0,X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|\Delta_t\|^2, \quad \forall u \in X. \tag{5.368}
\]
Moreover,
\[
\Delta_t^T (v_t - u) = \Delta_t^T (x_t - u) + \Delta_t^T (v_t - x_t),
\]
and hence it follows by (5.368) that
\[
\max_{u \in X} \sum_{t=1}^{N} \nu_t \Delta_t^T (x_t - u) \leq \sum_{t=1}^{N} \nu_t \Delta_t^T (x_t - v_t) + \frac{D_{0,X}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \|\Delta_t\|^2}{\sum_{t=1}^{N} \gamma_t}. \tag{5.369}
\]
Moreover, \( \mathbb{E} [\Delta_t | \xi_{[t-1]}] = 0 \) and \( v_t \) and \( x_t \) are functions of \( \xi_{[t-1]} \), and hence
\[
\mathbb{E} [(x_t - v_t)^T \Delta_t] = \mathbb{E} [(x_t - v_t)^T \mathbb{E} [\Delta_t | \xi_{[t-1]}]] = 0. \tag{5.370}
\]
In view of condition (5.361) we have that \( \mathbb{E} [\|\Delta_t\|^2] \leq 4M^2_t \), and hence it follows from (5.369) and (5.370) that
\[
\mathbb{E} \left[ \max_{u \in X} \sum_{t=1}^{N} \nu_t \Delta_t^T (x_t - u) \right] \leq D_{0,X}^2 + 2\kappa^{-1} M^2_t \sum_{t=1}^{N} \gamma_t^2. \tag{5.371}
\]
Therefore, by taking expectation of both sides of (5.37) and using (5.361) together with (5.371) we obtain (5.363).

In order to proof (5.364) let us observe that

\[ \mathcal{T}^N - f^N = \sum_{t=1}^{N} \nu_t (F(x_t, \xi^t) - f(x_t)), \]

and that for \( 1 \leq s < t \leq N, \)

\[
\mathbb{E} \left[ (F(x_s, \xi^s) - f(x_s))(F(x_t, \xi^t) - f(x_t)) \right] = \mathbb{E} \left\{ \mathbb{E} \left[ (F(x_s, \xi^s) - f(x_s))(F(x_t, \xi^t) - f(x_t)) \middle| \xi_{t-1} \right] \right\} = \mathbb{E} \left\{ (F(x_s, \xi_s) - f(x_s)) \mathbb{E} \left[ (F(x_t, \xi^t) - f(x_t)) \middle| \xi_{t-1} \right] \right\} = 0.
\]

Therefore

\[
\mathbb{E} \left[ (\mathcal{T}^N - f^N)^2 \right] = \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left[ (F(x_t, \xi^t) - f(x_t))^2 \right] = \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left\{ \mathbb{E} \left[ (F(x_t, \xi^t) - f(x_t))^2 \middle| \xi_{t-1} \right] \right\} \leq Q^2 \sum_{t=1}^{N} \nu_t^2,
\]

where the last inequality is implied by condition (5.362). Since for any random variable \( Y \) we have that \( \sqrt{\mathbb{E}[Y]} \geq \mathbb{E}[|Y|] \), the inequality (5.364) follows from (5.372).

Let us now look at (5.365). Denote

\[ \tilde{f}^N(x) := \sum_{t=1}^{N} \nu_t [F(x_t, \xi^t) + G(x_t, \xi^t)^T (x - x_t)]. \]

Then

\[ \left| \mathcal{T}^N - f^N \right| = \left| \min_{x \in \mathcal{X}} \tilde{f}^N(x) - \min_{x \in \mathcal{X}} f^N(x) \right| \leq \max_{x \in \mathcal{X}} \left| \tilde{f}^N(x) - f^N(x) \right| \]

and

\[ \tilde{f}^N(x) - f^N(x) = \mathcal{T}^N - f^N + \sum_{t=1}^{N} \nu_t \Delta_t^N (x_t - x), \]

and hence

\[ \left| \mathcal{T}^N - f^N \right| \leq \left| \mathcal{T}^N - f^N \right| + \max_{x \in \mathcal{X}} \sum_{t=1}^{N} \nu_t \Delta_t^N (x_t - x). \] (5.373)

For \( \mathbb{E} \left[ (\mathcal{T}^N - f^N)^2 \right] \) we already have the estimate (5.364).

By (5.368) we have

\[
\max_{x \in \mathcal{X}} \sum_{t=1}^{N} \nu_t \Delta_t^N (x_t - x) \leq \sum_{t=1}^{N} \nu_t \Delta_t^N (x_t - x_t) + D_{\mathcal{X}}^2 + (2\kappa)^{-1} \sum_{t=1}^{N} \gamma_t^2 \| \Delta_t \|^2 \sum_{t=1}^{N} \gamma_t.
\]

(5.374)

Let us observe that for \( 1 \leq s < t \leq N, \)

\[
\mathbb{E} \left[ (\Delta_s^N (x_s - v_s))(\Delta_t^N (x_t - v_t)) \right] = \mathbb{E} \left\{ \mathbb{E} \left[ (\Delta_s^N (x_s - v_s))(\Delta_t^N (x_t - v_t)) \middle| \xi_{t-1} \right] \right\} = \mathbb{E} \left\{ (\Delta_s^N (x_s - v_s)) \mathbb{E} \left[ (\Delta_t^N (x_t - v_t)) \middle| \xi_{t-1} \right] \right\} = 0.
\]
Therefore, by condition (5.361) we have

$$\mathbb{E} \left[ \left( \sum_{t=1}^{N} \nu_t \Delta_t \tau(x_t - v_t) \right)^2 \right] = \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left[ \| \Delta_t \tau(x_t - v_t) \|^2 \right]$$

$$\leq \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left[ \| \Delta_t \|^2 \| x_t - v_t \|^2 \right] = \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left[ \| x_t - v_t \|^2 \mathbb{E} \| \Delta_t \|^2 | \xi_{t-1} \| \right]$$

$$\leq 4M^2 \sum_{t=1}^{N} \nu_t^2 \mathbb{E} \left[ \| x_t - v_t \|^2 \right] \leq 32\kappa^{-1} M^2 D^2_{\xi,X} \sum_{t=1}^{N} \nu_t^2,$$

where the last inequality follows by (5.332). It follows that

$$\mathbb{E} \left[ \left( \sum_{t=1}^{N} \nu_t \Delta_t \tau(x_t - v_t) \right) \right] \leq 4\sqrt{2} \kappa^{-1/2} M, D, x \sqrt{\sum_{t=1}^{N} \nu_t^2}. \quad (5.375)$$

Putting together (5.373),(5.374),(5.375) and (5.364) we obtain (5.365). \[\square\]

For the constant stepsize policy (5.338), all estimates given in the right hand sides of (5.363),(5.364) and (5.365) are of order \(O(N^{-1/2})\). It follows that under the specified conditions, difference between the upper \(\bar{f}_N\) and lower \(\underline{f}_N\) bounds converge on average to zero, with increase of the sample size \(N\), at a rate of \(O(N^{-1/2})\). It is also possible to derive respective large deviations rates of convergence (Lan, Nemirovski and Shapiro [114]).

**Remark 16.** The lower SA bound \(\underline{f}_N\) can be compared with the respective SAA bound \(\hat{\vartheta}_N\) obtained by solving the corresponding SAA problem (see section 5.6.1). Suppose that the same sample \(\xi^1, \ldots, \xi^N\) is employed for both the SA and SAA methods, that \(F(\cdot, \xi)\) is convex for all \(\xi \in \Xi\), and \(G(x, \xi) \in \partial_x F(x, \xi)\) for all \((x, \xi) \in X \times \Xi\). By convexity of \(F(\cdot, \xi)\) and definition of \(\underline{f}_N\), we have

$$\hat{\vartheta}_N = \min_{x \in X} \left\{ N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right\}$$

$$\geq \min_{x \in X} \left\{ \sum_{t=1}^{N} \nu_t \left[ F(x_t, \xi^t) + G(x_t, \xi^t) \tau(x_t - x_t) \right] \right\} = \underline{f}_N. \quad (5.376)$$

Therefore, for the same sample, the SA lower bound \(\underline{f}_N\) is weaker than the SAA lower bound \(\hat{\vartheta}_N\). However, it should be noted that the SA lower bound can be computed much faster than the respective SAA lower bound.

**Exercises**

5.1. Suppose that set \(X\) is defined by constraints in the form (5.11) with constraint functions given as expectations as in (5.12) and the set \(X_N\) defined in (5.13). Show that if sample average functions \(\hat{g}_{i,N}\) converge uniformly to \(g_i\) w.p.1 on a neighborhood of \(x\) and \(g_i\) are continuous, \(i = 1, \ldots, p\), then condition (a) of Theorem 5.5 holds.

5.2. Specify regularity conditions under which equality (5.28) follows from (5.24).

5.3. Let \(X \subset \mathbb{R}^n\) be a closed convex set. Show that the multifunction \(x \mapsto \mathcal{N}_X(x)\) is closed.

5.4. Prove the following extension of Theorem 5.7. Let \(g : \mathbb{R}^m \rightarrow \mathbb{R}\) be continuously differentiable function, \(F_i(x, \xi), i = 1, \ldots, m\), be random lower semicontinuous
functions, \( f_i(x) := E[F_i(x, \xi)], i = 1, \ldots, m, f(x) = (f_1(x), \ldots, f_m(x)) \), \( X \) be a nonempty compact subset of \( \mathbb{R}^n \) and consider the optimization problem

\[
\min_{x \in X} g(f(x)). \tag{5.377}
\]

Moreover, let \( \xi^1, \ldots, \xi^N \) be an iid random sample, \( \hat{f}_{iN}(x) := \frac{1}{N} \sum_{j=1}^N F_i(x, \xi^j), i = 1, \ldots, m, \hat{f}_N(x) = (\hat{f}_{1N}(x), \ldots, \hat{f}_{mN}(x)) \) be the corresponding sample average functions and

\[
\min_{x \in X} g(\hat{f}_N(x)) \tag{5.378}
\]

be the associated SAA problem. Suppose that conditions (A1) and (A2) (used in Theorem 5.7) hold for every function \( F_i(x, \xi), i = 1, \ldots, m \). Let \( \hat{\vartheta}^* \) and \( \vartheta^* \) be the optimal values of problems (5.377) and (5.378), respectively, and \( S \) be the set of optimal solutions of problem (5.377). Show that

\[
\hat{\vartheta}_N - \vartheta^* = \inf_{x \in S} \left( \sum_{i=1}^m w_i(x) [\hat{f}_{iN}(x) - f_i(x)] \right) + o_p(N^{-1/2}), \tag{5.379}
\]

where

\[
w_i(x) := \left. \frac{\partial g(y_1, \ldots, y_m)}{\partial y_i} \right|_{y = f(x)}, i = 1, \ldots, m.
\]

Moreover, if \( S = \{ \bar{x} \} \) is a singleton, then

\[
N^{1/2} \left( \hat{\vartheta}_N - \vartheta^* \right) \Rightarrow \mathcal{N}(0, \sigma^2), \tag{5.380}
\]

where \( \bar{w}_i := w_i(\bar{x}) \) and

\[
\sigma^2 = \text{Var} \left[ \sum_{i=1}^m \bar{w}_i F_i(\bar{x}, \xi) \right]. \tag{5.381}
\]

Hint: consider function \( V : C(X) \times \cdots \times C(X) \to \mathbb{R} \) defined as \( V(\psi_1, \ldots, \psi_m) := \inf_{x \in X} g(\psi_1(x), \ldots, \psi_m(x)) \), and apply the functional CLT together with Delta and Danskin Theorems.

5.5. Consider matrix \( \begin{bmatrix} H & A \\ A^T & 0 \end{bmatrix} \) defined in (5.43). Assuming that matrix \( H \) is positive definite and matrix \( A \) has full column rank, verify that

\[
\begin{bmatrix} H & A \\ A^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H^{-1} - H^{-1}A(A^TH^{-1}A)^{-1}A^TH^{-1} & H^{-1}A(A^TH^{-1}A)^{-1} \\ (A^TH^{-1}A)^{-1}A^TH^{-1} & -(A^TH^{-1}A)^{-1} \end{bmatrix}.
\]

Using this identity write the asymptotic covariance matrix of \( N^{1/2} \begin{bmatrix} \hat{x}_N - \bar{x} \\ \lambda_N - \bar{\lambda} \end{bmatrix} \), given in (5.44), explicitly.

5.6. Consider the minimax stochastic problem (5.45), the corresponding SAA problem (5.46) and let

\[
\Delta_N := \sup_{x \in X, y \in Y} \left| \hat{f}_N(x, y) - f(x, y) \right|. \tag{5.382}
\]
Consider the SAA problem (5.236) giving approximation of the first stage of the
problem
\[ \overline{\vartheta}_N = \arg \min_{\vartheta} \mathbb{E}[f(x, \vartheta)] \]
and that if \( \hat{x}_N \) is a \( \delta \)-optimal solution of the SAA
problem (5.46), then \( \hat{x}_N \) is a \( (\delta + 2\Delta_N) \)-optimal solution of the minimax
problem (5.45).

(ii) By using Theorem 7.65 conclude that, under appropriate regularity conditions,
for any \( \varepsilon > 0 \) there exist positive constants \( C = C(\varepsilon) \) and \( \beta = \beta(\varepsilon) \) such that
\[
\Pr \left\{ \left| \hat{\vartheta}_N - \vartheta^* \right| \geq \varepsilon \right\} \leq C e^{-N\beta}. \tag{5.383}
\]

(iii) By using bounds (7.222) and (7.223) derive an estimate, similar to (5.113), of
the sample size \( N \) which guarantees with probability at least \( 1 - \alpha \) that a \( \delta \)-optimal
solution \( \hat{x}_N \) of the SAA problem (5.46) is an \( \varepsilon \)-optimal solution of the minimax
problem (5.45). Specify required regularity conditions.

5.8. Consider the following chance constrained problem
\[
\min_{x \in X} f(x) \quad \text{subject to } \Pr \{ T(\xi)x + h(\xi) \in C \} \geq 1 - \alpha, \tag{5.384}
\]
where \( X \subset \mathbb{R}^n \) is a closed convex set, \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is a convex function, \( C \subset \mathbb{R}^m \)
is a convex closed set, \( \alpha \in (0, 1) \), matrix \( T(\xi) \) and vector \( h(\xi) \) are functions of random vector \( \xi \). For example, if
\[
C := \{ z : z = -Wy - w, \; y \in \mathbb{R}^l, \; w \in \mathbb{R}^m \}, \tag{5.385}
\]
then, for a given \( x \in X \), the constraint \( T(\xi)x + h(\xi) \in C \) means that the system
\( Wy + T(\xi)x + h(\xi) \leq 0 \) has a feasible solution. Extend the results of section 5.7
to the setting of problem (5.384).

5.9. Consider the SAA problem (5.236) giving approximation of the first stage of the
corresponding three stage stochastic program. Let
\[
\hat{\vartheta}_{N_1, N_2} := \inf_{x_1 \in A_1} \hat{f}_{N_1, N_2}(x_1)
\]
be the optimal value and \( \bar{x}_{N_1, N_2} \) be an optimal solution of problem (5.236). Consider asymptotics of \( \hat{\vartheta}_{N_1, N_2} \) and \( \bar{x}_{N_1, N_2} \) as \( N_1 \) tends to infinity while \( N_2 \) is fixed.
Let \( \vartheta^*_N \) be the optimal value and \( S_{N_2} \) be the set of optimal solutions of the problem
\[
\min_{x_1 \in A_1} \left\{ f_1(x_1) + \mathbb{E} \left[ Q_{2, N_2}(x_1, \xi_2) \right] \right\}, \tag{5.386}
\]
where the expectation is taken with respect to the distribution of the random vector
\( (\xi_2^1, \xi_2^2, ..., \xi_2^{N_2}) \).

(i) By using results of section 5.1.1 show that \( \hat{\vartheta}_{N_1, N_2} \to \vartheta^*_N \) w.p.1 and distance from \( \bar{x}_{N_1, N_2} \) to \( S_{N_2} \) tends to 0 w.p.1 as \( N_1 \to \infty \). Specify required regularity
Exercises

(ii) Show that, under appropriate regularity conditions,

\[ \tilde{\vartheta}_{N_1, N_2} = \inf_{x_1 \in S_{N_2}} \tilde{f}_{N_1, N_2}(x_1) + o_p(N_1^{-1/2}). \]  

(5.387)

Conclude that if, moreover, \( S_{N_2} = \{ \bar{x}_1 \} \) is a singleton, then

\[ N_1^{1/2} (\tilde{\vartheta}_{N_1, N_2} - \vartheta^*_2) \overset{D}{\to} N(0, \sigma^2(\bar{x}_1)), \]  

(5.388)

where \( \sigma^2(\bar{x}_1) := \text{Var}[\hat{Q}_{2,N_2}(x_1, \xi^i_2)] \). Hint: use Theorem 5.7.
Chapter 6

Risk Averse Optimization

Andrzej Ruszczyński and Alexander Shapiro

6.1 Introduction

So far, we discussed stochastic optimization problems, in which the objective function was defined as the expected value \( f(x) := \mathbb{E}[F(x, \omega)] \). The function \( F : \mathbb{R}^n \times \Omega \to \mathbb{R} \) models the random outcome, for example, the random cost, and is assumed to be sufficiently regular so that the expected value function is well defined. For a feasible set \( X \subset \mathbb{R}^n \), the stochastic optimization model

\[
\min_{x \in X} f(x)
\]  

optimizes the random outcome \( F(x, \omega) \) on average. This is justified when the Law of Large Numbers can be invoked and we are interested in the long term performance, irrespective of the fluctuations of specific outcome realizations. The shortcomings of such an approach can be clearly illustrated by the example of portfolio selection discussed in section 1.4. Consider problem (1.34) of maximizing the expected return rate. Its optimal solution suggests to concentrate the investment in the assets having the highest expected return rate. This is not what we would consider reasonable, because it leaves out all considerations of the involved risk of losing all invested money. In this section we discuss stochastic optimization from a point of view of risk averse optimization.

A classical approach to risk averse preferences is based on the expected utility theory, which has its roots in mathematical economics (we already touched upon this subject in section 1.4). In this theory, in order to compare two random outcomes we consider expected values of some scalar transformations \( u : \mathbb{R} \to \mathbb{R} \) of the realization of these outcomes. In a minimization problem, a random outcome \( Z_1 \) (understood as a scalar random variable) is
preferred over a random outcome $Z_2$, if
\[ \mathbb{E}[u(Z_1)] < \mathbb{E}[u(Z_2)]. \]

The function $u(\cdot)$, called the disutility function, is assumed to be nondecreasing and convex. Following this principle, instead of problem (6.1), we construct the problem
\[
\min_{x \in X} \mathbb{E}[u(F(x, \omega))].
\]
(6.2)

Observe that it is still an expected value problem, but the function $F$ is replaced by the composition $u \circ F$. Since $u(\cdot)$ is convex, we have by Jensen’s inequality that
\[ u(\mathbb{E}[F(x, \omega)]) \leq \mathbb{E}[u(F(x, \omega))]. \]

That is, a sure outcome of $\mathbb{E}[F(x, \omega)]$ is at least as good as the random outcome $F(x, \omega)$. In a maximization problem, we assume that $u(\cdot)$ is concave (and still nondecreasing). We call it a utility function in this case. Again, Jensen’s inequality yields the preference in terms of expected utility:
\[ u(\mathbb{E}[F(x, \omega)]) \geq \mathbb{E}[u(F(x, \omega))]. \]

One of the basic difficulties in using the expected utility approach is specifying the utility or disutility function. They are very difficult to elicit; even the authors of this book cannot specify their utility functions in simple stochastic optimization problems. Moreover, using some arbitrarily selected utility functions may lead to solutions which are difficult to interpret and explain. Modern approach to modeling risk aversion in optimization problems uses the concept of risk measures. These are, generally speaking, functionals which take as their argument the entire collection of realizations $Z(\omega) = F(x, \omega)$, $\omega \in \Omega$, understood as an object in an appropriate vector space. In the following sections we introduce this concept.

6.2 Mean–risk models

6.2.1 Main ideas of mean–risk analysis

The main idea of mean–risk models is to characterize the uncertain outcome $Z_x(\omega) = F(x, \omega)$ by two scalar characteristics: the mean $\mathbb{E}[Z]$ describing the expected outcome, and the risk (dispersion measure) $\mathbb{D}[Z]$ which measures the uncertainty of the outcome. In the mean–risk approach, we select from the set of all possible solutions those that are efficient: for a given value of the mean they minimize the risk and for a given value of risk – maximize the mean. Such an approach has many advantages: it allows one to formulate the problem as a parametric optimization problem, and it facilitates the trade-off analysis between mean and risk.

Let us describe the mean–risk analysis on the example of the minimization problem (6.1). Suppose that the risk functional is defined as the variance $\mathbb{D}[Z] := \text{Var}[Z]$, which is well defined for $Z \in L_2(\Omega, \mathcal{F}, \mathbb{P})$. The variance, although not the best choice, is easiest to start from. It is also important in finance. Later in this chapter we discuss in much detail desirable properties of the risk functionals.
6.2. Mean–risk models

In the mean–risk approach, we aim at finding efficient solutions of the problem with two objectives, namely, $\mathbb{E}[Z_x]$ and $\mathbb{D}[Z_x]$, subject to the feasibility constraint $x \in X$. This can be accomplished by techniques of multi-objective optimization. Most convenient, from our perspective, is the idea of scalarization. For a coefficient $c \geq 0$, we form a composite objective functional

$$\rho[Z] := \mathbb{E}[Z] + c\mathbb{D}[Z].$$

(6.3)

The coefficient $c$ plays the role of the price of risk. We formulate the problem

$$\min_{x \in X} \mathbb{E}[Z_x] + c\mathbb{D}[Z_x].$$

(6.4)

By varying the value of the coefficient $c$, we can generate in this way a large ensemble of efficient solutions. We already discussed this approach for the portfolio selection problem, with $\mathbb{D}[Z] := \text{Var}[Z]$, in section 1.4.

An obvious deficiency of variance as a measure of risk is that it treats the excess over the mean equally as the shortfall. After all, in the minimization case, we are not concerned if a particular realization of $Z$ is significantly below its mean, we do not want it to be too large. Two particular classes of risk functionals, which we discuss next, play an important role in the theory of mean–risk models.

6.2.2 Semideviations

An important group of risk functionals (representing dispersion measures) are central semideviations. The upper semideviation of order $p$ is defined as

$$\sigma^+_p[Z] := \left( \mathbb{E}\left[ (Z - \mathbb{E}[Z])^p \right] \right)^{1/p},$$

(6.5)

where $p \in [1, \infty)$ is a fixed parameter. It is natural to assume here that considered random variables (uncertain outcomes) $Z : \Omega \to \mathbb{R}$ belong to the space $L_p(\Omega, \mathcal{F}, P)$, i.e., that they have finite $p$-th order moments. That is, $\sigma^+_p[Z]$ is well defined and finite valued for all $Z \in L_p(\Omega, \mathcal{F}, P)$. The corresponding mean–risk model has the general form

$$\min_{x \in X} \mathbb{E}[Z_x] + c\sigma^+_p[Z_x].$$

(6.6)

The upper semideviation measure is appropriate for minimization problems, where $Z_x(\omega) = F(x, \omega)$ represents a cost, as a function of $\omega \in \Omega$. It is aimed at penalization of an excess of $Z_x$ over its mean. If we are dealing with a maximization problem, where $Z_x$ represents some reward or profit, the corresponding risk functional is the lower semideviation

$$\sigma^-_p[Z] := \left( \mathbb{E}\left[ (\mathbb{E}[Z] - Z)^p \right] \right)^{1/p},$$

(6.7)

where $Z \in L_p(\Omega, \mathcal{F}, P)$. The resulting mean–risk model has the form

$$\max_{x \in X} \mathbb{E}[Z_x] - c\sigma^-_p[Z_x].$$

(6.8)

In the special case of $p = 1$, both left and right first order semideviations are related to the mean absolute deviation

$$\sigma_1(Z) := \mathbb{E}|Z - \mathbb{E}[Z]|.$$

(6.9)
**Proposition 6.1.** The following identity holds

\[ \sigma^-_1[Z] = \sigma^+_1[Z] = \frac{1}{2} \sigma_1[Z], \quad \forall Z \in L_1(\Omega, \mathcal{F}, P). \quad (6.10) \]

**Proof.** Denote by \( H(\cdot) \) the cumulative distribution function (cdf) of \( Z \) and let \( \mu := \mathbb{E}[Z] \).

We have

\[ \sigma^-_1[Z] = \int_{-\infty}^{\mu} (\mu - z) \, dH(z) = \int_{-\infty}^{\infty} (\mu - z) \, dH(z) - \int_{\mu}^{\infty} (\mu - z) \, dH(z). \]

The first integral on the right hand side is equal to 0, and thus \( \sigma^-_1[Z] = \sigma^+_1[Z] \). The identity \( (6.10) \) follows now from the equation \( \sigma_1[Z] = \sigma^-_1[Z] + \sigma^+_1[Z] \).

We conclude that using the mean absolute deviation instead of the semideviation in mean–risk models has the same effect, just the parameter \( c \) has to be halved. The identity \( (6.10) \) does not extend to semideviations of higher orders, unless the distribution of \( Z \) is symmetric.

### 6.2.3 Weighted mean deviations from quantiles

Let \( H_Z(z) = \Pr(Z \leq z) \) be the cdf of the random variable \( Z \) and \( \alpha \in (0,1) \). Recall that the **left-side \( \alpha \)-quantile** of \( H_Z \) is defined as

\[ H_Z^{-1}(\alpha) := \inf\{t : H_Z(t) \geq \alpha\}, \quad (6.11) \]

and the **right-side \( \alpha \)-quantile** as

\[ \sup\{t : H_Z(t) \leq \alpha\}. \quad (6.12) \]

If \( Z \) represents losses, the (left-side) quantile \( H_Z^{-1}(1-\alpha) \) is also called **Value-at-Risk** and denoted \( \mathbb{V} \mathbb{R}_\alpha(Z) \), i.e.,

\[ \mathbb{V} \mathbb{R}_\alpha(Z) = H_Z^{-1}(1-\alpha) = \inf\{t : \Pr(Z \leq t) \geq 1-\alpha\} = \inf\{t : \Pr(Z > t) \leq \alpha\}. \]

Its meaning is the following: **losses larger than** \( \mathbb{V} \mathbb{R}_\alpha(Z) \) **occur with probability not exceeding** \( \alpha \). Note that

\[ \mathbb{V} \mathbb{R}_\alpha(Z + \tau) = \mathbb{V} \mathbb{R}_\alpha(Z) + \tau, \quad \forall \tau \in \mathbb{R}. \quad (6.13) \]

The weighted mean deviation from a quantile is defined as follows:

\[ q_\alpha[Z] := \mathbb{E} \left[ \max \left\{ (1-\alpha)H_Z^{-1}(\alpha) - Z, \alpha(Z - H_Z^{-1}(\alpha)) \right\} \right]. \quad (6.14) \]

The functional \( q_\alpha[Z] \) is well defined and finite valued for all \( Z \in L_1(\Omega, \mathcal{F}, P) \). It can be easily shown that

\[ q_\alpha[Z] = \min_{\tau \in \mathbb{R}} \left\{ \varphi(t) := \mathbb{E} \left[ \max \left\{ (1-\alpha)(t - Z), \alpha(Z - t) \right\} \right] \right\}. \quad (6.15) \]
6.2. Mean–risk models

Indeed, the right and left side derivatives of the function \( \varphi(\cdot) \) are
\[
\varphi'_+(t) = (1-\alpha)\Pr[Z \leq t] - \alpha \Pr[Z > t], \\
\varphi'_-(t) = (1-\alpha)\Pr[Z < t] - \alpha \Pr[Z \geq t].
\]

At the optimal \( t \) the right derivative is nonnegative and the left derivative nonpositive, and thus
\[
\Pr[Z < t] \leq \alpha \leq \Pr[Z \leq t].
\]
This means that every \( \alpha \)-quantile is a minimizer in (6.15).

The risk functional \( q_\alpha[\cdot] \) can be used in mean–risk models, both in the case of minimization
\[
\min_{x \in X} E[Z_x] + cq_1-\alpha[Z_x], \tag{6.16}
\]
and in the case of maximization
\[
\max_{x \in X} E[Z_x] - cq_\alpha[Z_x]. \tag{6.17}
\]
We use \( 1-\alpha \) in the minimization problem and \( \alpha \) in the maximization problem, because in practical applications we are interested in these quantities for small \( \alpha \).

6.2.4 Average Value-at-Risk

The mean–deviation from quantile model is closely related to the concept of Average Value-at-Risk\(^1\). Suppose that \( Z \) represents losses and we want to satisfy the chance constraint:
\[
\mathcal{VAR}_\alpha[Z_x] \leq 0. \tag{6.18}
\]
Recall that
\[
\mathcal{VAR}_\alpha[Z] = \inf \{ t : \Pr(Z \leq t) \geq 1 - \alpha \},
\]
and hence constraint (6.18) is equivalent to the constraint \( \Pr(Z_x \leq 0) \geq 1 - \alpha \). We have that\(^2\) \( \Pr(Z_x > 0) = E\left[1_{(0,\infty)}(Z_x)\right] \), and hence constraint (6.18) can also be written as the expected value constraint:
\[
E\left[1_{(0,\infty)}(Z_x)\right] \leq \alpha. \tag{6.19}
\]
The source of difficulties with probabilistic (chance) constraints is that the step function \( 1_{(0,\infty)}(\cdot) \) is not convex and, even worse, it is discontinuous at zero. As a result, chance constraints are often nonconvex, even if the function \( x \mapsto Z_x \) is convex almost surely. One possibility is to approach such problems by constructing a convex approximation of the expected value on the left of (6.19).

\(^1\)Average Value-at-Risk is often called Conditional Value-at-Risk. We adopt here the term 'Average' rather than 'Conditional' Value-at-Risk in order to avoid awkward notation and terminology while discussing later conditional risk mappings.

\(^2\)Recall that \( 1_{(0,\infty)}(z) = 0 \) if \( z \leq 0 \), and \( 1_{(0,\infty)}(z) = 1 \) if \( z > 0 \).
Let $\psi : \mathbb{R} \to \mathbb{R}$ be a nonnegative valued, nondecreasing, convex function such that $\psi(z) \geq 1_{(0, \infty)}(z)$ for all $z \in \mathbb{R}$. By noting that $1_{(0, \infty)}(tz) = 1_{(0, \infty)}(z)$ for any $t > 0$ and $z \in \mathbb{R}$, we have that $\psi(tz) \geq 1_{(0, \infty)}(z)$ and hence the following inequality holds,

$$\inf_{t>0} \mathbb{E}[\psi(tZ)] \geq \mathbb{E}[1_{(0, \infty)}(Z)].$$

Consequently, the constraint

$$\inf_{t>0} \mathbb{E}[\psi(tZ_{\theta})] \leq \alpha \quad (6.20)$$

is a conservative approximation of the chance constraint (6.18) in the sense that the feasible set defined by (6.20) is contained in the feasible set defined by (6.18).

Of course, the smaller the function $\psi(\cdot)$ is the better this approximation will be. From this point of view the best choice of $\psi(\cdot)$ is to take piecewise linear function $\psi(z) := [1 + \gamma z]_+$, for some $\gamma > 0$. Since constraint (6.20) is invariant with respect to scale change of $\psi(\gamma z)$, we have that $\psi(z) := [1 + z]_+$ gives the best choice of such a function. For this choice of function $\psi(\cdot)$, we have that constraint (6.20) is equivalent to

$$\inf_{t>0} \{t\mathbb{E}[t^{-1} + Z]_+ - \alpha\} \leq 0,$$

or equivalently

$$\inf_{t>0} \{\alpha^{-1}\mathbb{E}[Z + t^{-1}]_+ - t^{-1}\} \leq 0.$$

Now replacing $t$ with $-t^{-1}$ we get the form:

$$\inf_{t<0} \{t + \alpha^{-1}\mathbb{E}[Z - t]_+\} \leq 0 \quad (6.21)$$

The quantity

$$\text{AV}@\mathcal{R}_\alpha(Z) := \inf_{t \in \mathbb{R}} \left\{t + \alpha^{-1}\mathbb{E}[Z - t]_+\right\} \quad (6.22)$$

is called the Average Value-at-Risk$^3$ of $Z$ (at level $\alpha$). Note that $\text{AV}@\mathcal{R}_\alpha(Z)$ is well defined and finite valued for every $Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$.

The function $\phi(t) := t + \alpha^{-1}\mathbb{E}[Z - t]_+$ is convex. Its derivative at $t$ is equal to $1 + \alpha^{-1}[H_Z(t) - 1]$, provided that the cdf $H_Z(\cdot)$ is continuous at $t$. If $H_Z(\cdot)$ is discontinuous at $t$, then the respective right and left side derivatives of $\phi(\cdot)$ are given by the same formula with $H_Z(t)$ understood as the corresponding right and left side limits. Therefore the minimum of $\phi(t)$, over $t \in \mathbb{R}$, is attained on the interval $[t^*, t^{**}]$, where

$$t^* := \inf\{z : H_Z(z) \geq 1 - \alpha\} \quad \text{and} \quad t^{**} := \sup\{z : H_Z(z) \leq 1 - \alpha\} \quad (6.23)$$

are the respective left and right side quantiles. Recall that the left-side quantile $t^* = \text{VaR}_\alpha(Z)$.

Since the minimum of $\phi(t)$ is attained at $t^* = \text{VaR}_\alpha(Z)$, we have that $\text{AV}@\mathcal{R}_\alpha(Z)$ is bigger than $\text{VaR}_\alpha(Z)$ by the nonnegative amount of $\alpha^{-1}\mathbb{E}[Z - t^*]_+$. Therefore

$$\inf_{t \in \mathbb{R}} \left\{t + \alpha^{-1}\mathbb{E}[Z - t]_+\right\} \leq 0 \quad \text{implies} \quad t^* \leq 0,$$

---

$^3$In some publications the above concept of ‘Average Value-at-Risk’ is called ‘Conditional Value-at-Risk’ and is denoted $\text{CV@R}_\alpha$. 
and hence constraint (6.21) is equivalent to AV@R_\alpha(Z) \leq 0. Therefore, the constraint
\[ AV@R_\alpha[Z_x] \leq 0 \]  
(6.24)
is equivalent to the constraint (6.21) and gives a conservative approximation\(^4\) of the chance constraint (6.18).

The function \( \rho(Z) := AV@R_\alpha(Z) \), defined on a space of random variables, is convex, i.e., if \( Z \) and \( Z' \) are two random variables and \( t \in [0, 1] \), then
\[ \rho(tZ + (1-t)Z') \leq t\rho(Z) + (1-t)\rho(Z'). \]

This follows from the fact that the function \( t + \alpha^{-1}E[Z - t]^+ \) is convex jointly in \( t \) and \( Z \). Also \( \rho(\cdot) \) is monotone, i.e., if \( Z \) and \( Z' \) are two random variables such that with probability one \( Z \geq Z' \), then \( \rho(Z) \geq \rho(Z') \). It follows that if \( G(\cdot, \xi) \) is convex for a.e. \( \xi \in \Xi \), then the function \( \rho[G(\cdot, \xi)] \) is also convex. Indeed, by convexity of \( G(\cdot, \xi) \) and monotonicity of \( \rho(\cdot) \), we have for any \( t \in [0, 1] \) that
\[ \rho[G(tZ + (1-t)Z'), \xi] \leq t\rho[G(Z, \xi)] + (1-t)\rho[G(Z', \xi)], \]
and hence by convexity of \( \rho(\cdot) \) that
\[ \rho[G(tZ + (1-t)Z'), \xi] \leq t\rho[G(Z, \xi)] + (1-t)\rho[G(Z', \xi)]. \]

Consequently, (6.24) is a convex conservative approximation of the chance constraint (6.18). Moreover, from the considered point of view, (6.24) is the best convex conservative approximation of the chance constraint (6.18).

We can now relate the concept of Average Value-at-Risk to mean deviations from quantiles. Recall that (see (6.14))
\[ q_\alpha[Z] := E\left[\max\{(1-\alpha)(H^{-1}_Z(\alpha) - Z), \alpha(Z - H^{-1}_Z(\alpha))\}\right]. \]

**Theorem 6.2.** Let \( Z \in L_1(\Omega, \mathcal{F}, P) \) and \( H(z) \) be its cdf. Then the following identities hold true
\[ AV@R_\alpha(Z) = \frac{1}{\alpha} \int_{1-\alpha}^1 V@R_1^{-\alpha}(Z) d\tau = E[Z] + \frac{1}{\alpha} q_\alpha[Z]. \]  
(6.26)

Moreover, if \( H(z) \) is continuous at \( z = V@R_\alpha(Z) \), then
\[ AV@R_\alpha(Z) = \frac{1}{\alpha} \int_{V@R_\alpha(Z)}^{+\infty} zdH(z) = E[Z | Z \geq V@R_\alpha(Z)]. \]  
(6.27)

**Proof.** As discussed before, the minimum in (6.22) is attained at \( t^* = H^{-1}(1-\alpha) = V@R_\alpha(Z) \). Therefore
\[ AV@R_\alpha(Z) = t^* + \alpha^{-1}E[Z - t^*]_+ = t^* + \alpha^{-1} \int_{t^*}^{+\infty} (z - t^*)dH(z). \]

\(^4\)It is easy to see that for any \( \tau \in \mathbb{R}, AV@R_\alpha(Z + \tau) = AV@R_\alpha(Z) + \tau. \)

Consequently the constraint \( AV@R_\alpha[Z_x] \leq \tau \) gives a conservative approximation of the chance constraint \( V@R_\alpha[Z_x] \leq \tau \).
Moreover,
\[ \int_{t^*}^{+\infty} dH(z) = \Pr(Z \geq t^*) = 1 - \Pr(Z \leq t^*) = \alpha, \]
provided that \( \Pr(Z = t^*) = 0 \), i.e., that \( H(z) \) is continuous at \( z = V@R_\alpha(Z) \). This shows the first equality in (6.27), and then the second equality in (6.27) follows provided that \( \Pr(Z = t^*) = 0 \).

The first equality in (6.26) follows from the first equality in (6.27) by the substitution \( \tau = H(z) \). Finally, we have
\[
AV@R_\alpha(Z) = t^* + \alpha^{-1}E[Z - t^*]_+ = E[Z] + E \{ -Z + t^* + \alpha^{-1}[Z - t^*]_+ \} \\
= E[Z] + E \left[ \max \{ \alpha^{-1}(1 - \alpha)(Z - t^*), t^* - Z \} \right].
\]
This proves the last equality in (6.26).

The first equation in (6.26) motivates the term Average Value-at-Risk. The last equation in (6.27) explains the origin of the alternative term Conditional Value-at-Risk.

Theorem 6.2 allows us to show an important relation between the absolute semideviation \( \sigma^+_1[Z] \) and the mean deviation from quantile \( q_\alpha[Z] \).

**Corollary 6.3.** For every \( Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P) \) we have
\[
\sigma^+_1[Z] = \max_{\alpha \in [0,1]} q_\alpha[Z] = \min_{t \in \mathbb{R}} \max_{\alpha \in [0,1]} E \{ (1 - \alpha)[t - Z]_+ + \alpha[Z - t]_+ \}. \tag{6.28}
\]

**Proof.** From (6.26) we get
\[
q_{1-\alpha}[Z] = \int_{1-\alpha}^{1} H^{-1}_Z(\tau) d\tau - \alpha E[Z].
\]
The right derivative of the right hand side with respect to \( \alpha \) equals \( H^{-1}_Z(1 - \alpha) - \alpha E[Z] \). As it is nonincreasing, the function \( \alpha \mapsto q_{1-\alpha}[Z] \) is concave. Moreover, its maximum is achieved at \( \alpha^*_1 \) for which \( E[Z] \) is the \((1 - \alpha^*_1)\)-quantile of \( Z \). Substituting the minimizer \( t^* = E[Z] \) into (6.22) we conclude that
\[
AV@R_{\alpha^*_1}(Z) = E[Z] + \frac{1}{\alpha^*_1} \sigma^+_1[Z].
\]
Comparison with (6.26) yields the first equality in (6.28). To prove the second equality we recall relation (6.15) and note that
\[
\max \left\{ (1 - \alpha)[t - Z], \alpha(Z - t) \right\} = (1 - \alpha)[t - Z]_+ + \alpha[Z - t]_+.
\]
Thus
\[
\sigma^+_1[Z] = \max_{\alpha \in [0,1]} \min_{t \in \mathbb{R}} E \{ (1 - \alpha)[t - Z]_+ + \alpha[Z - t]_+ \}.
\]
As the function under the “max-min” operation is linear with respect to \( \alpha \in [0,1] \) and convex with respect to \( t \), the “max” and “min” operations can be exchanged. This proves the second equality in (6.28).
It also follows from (6.26) that the minimization problem (6.16) can be equivalently written as follows
\[
\min_{x \in \mathcal{X}} \mathbb{E}[Z_x] + c\mathbb{1}_{1-\alpha}[Z_x] = \min_{x \in \mathcal{X}} (1 - c\alpha)\mathbb{E}[Z_x] + c\alpha\mathbb{V}\otimes \mathcal{R}_\alpha[Z_x]
\]
(6.29)
\[
= \min_{x \in \mathcal{X}, t \in \mathbb{R}} \mathbb{E} \left[ (1 - c\alpha)Z_x + c(\alpha t + [Z_x - t])_+ \right].
\]
Both \(x\) and \(t\) are variables in this problem. We conclude that for this specific mean–risk model, an equivalent expected value formulation has been found. If \(c \in [0, \alpha^{-1}]\) and the function \(x \mapsto Z_x\) is convex, problem (6.29) is convex.

The maximization problem (6.17) can be equivalently written as follows:
\[
\max_{x \in \mathcal{X}} \mathbb{E}[Z_x] - c\mathbb{1}_{\alpha}[Z_x] = -\min_{x \in \mathcal{X}} \mathbb{E}[-Z_x] + c\mathbb{1}_{1-\alpha}[-Z_x]
\]
(6.30)
\[
= -\min_{x \in \mathcal{X}, t \in \mathbb{R}} \mathbb{E} \left[ -(1 - c\alpha)Z_x + c(\alpha t + [-Z_x - t])_+ \right]
\]
\[
= \max_{x \in \mathcal{X}, t \in \mathbb{R}} \mathbb{E} \left[ (1 - c\alpha)Z_x + c(\alpha t - [t - Z_x]_+) \right].
\]
(6.31)
In the last problem we replaced \(t\) by \(-t\) to stress the similarity with (6.29). Again, if \(c \in [0, \alpha^{-1}]\) and the function \(x \mapsto Z_x\) is convex, problem (6.30) is convex.

### 6.3 Coherent Risk Measures

Let \((\Omega, \mathcal{F})\) be a sample space, equipped with the sigma algebra \(\mathcal{F}\), on which considered uncertain outcomes (random functions \(Z = Z(\omega)\)) are defined. By a risk measure we understand a function \(\rho(Z)\) which maps \(Z\) into the extended real line \(\mathbb{R} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}\). In order to make this concept precise we need to define a space \(\mathcal{Z}\) of allowable random functions \(Z(\omega)\) for which \(\rho(Z)\) is defined. It seems that a natural choice of \(\mathcal{Z}\) will be the space of all \(\mathcal{F}\)-measurable functions \(Z : \Omega \rightarrow \mathbb{R}\). However, typically, this space is too large for development of a meaningful theory. Unless stated otherwise, we deal in this chapter with spaces \(\mathcal{Z} := \mathcal{L}_p(\Omega, \mathcal{F}, P)\), where \(p \in [1, +\infty)\) (see section 7.3 for an introduction of these spaces). By assuming that \(Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)\), we assume that random variable \(Z(\omega)\) has a finite \(p\)-th order moment with respect to the reference probability measure \(P\). Also by considering function \(\rho\) to be defined on the space \(\mathcal{L}_p(\Omega, \mathcal{F}, P)\), it is implicitly assumed that actually \(\rho\) is defined on classes of functions which can differ on sets of \(P\)-measure zero, i.e., \(\rho(Z) = \rho(Z')\) if \(P\{\omega : Z(\omega) \neq Z'(\omega)\} = 0\).

We assume throughout this chapter that risk measures \(\rho : \mathcal{Z} \rightarrow \mathbb{R}\) are proper. That is, \(\rho(Z) > -\infty\) for all \(Z \in \mathcal{Z}\) and the domain
\[
\text{dom}(\rho) := \{Z \in \mathcal{Z} : \rho(Z) < +\infty\}
\]
is nonempty. We consider the following axioms associated with a risk measure \(\rho\). For \(Z, Z' \in \mathcal{Z}\) we denote by \(Z \succeq Z'\) the pointwise partial order meaning \(Z(\omega) \geq Z'(\omega)\) for a.e. \(\omega \in \Omega\). We also assume that the smaller the realizations of \(Z\), the better; for example \(Z\) may represent a random cost.

\[\text{This partial order corresponds to the cone } C := \mathcal{L}_p^+(\Omega, \mathcal{F}, P), \text{ see the discussion of section 7.3, page } 407, \text{ following equation (7.251).} \]
Chapter 6. Risk Averse Optimization

(R1) Convexity:
\[ \rho(tZ + (1-t)Z') \leq t\rho(Z) + (1-t)\rho(Z') \]
for all \( Z, Z' \in \mathcal{Z} \) and all \( t \in [0,1] \).

(R2) Monotonicity: If \( Z, Z' \in \mathcal{Z} \) and \( Z \succeq Z' \), then \( \rho(Z) \geq \rho(Z') \).

(R3) Translation Equivariance: If \( a \in \mathbb{R} \) and \( Z \in \mathcal{Z} \), then \( \rho(Z + a) = \rho(Z) + a \).

(R4) Positive Homogeneity: If \( t > 0 \) and \( Z \in \mathcal{Z} \), then \( \rho(tZ) = t\rho(Z) \).

It is said that a risk measure \( \rho \) is coherent if it satisfies the above conditions (R1)–(R4). An example of a coherent risk measure is the Average Value-at-Risk \( \rho(Z) := \text{AV@R}_\alpha(Z) \) (further examples of risk measures will be discussed in section 6.3.2). It is natural to assume in this example that \( Z \) has a finite first order moment, i.e., to use \( Z = L^1(\Omega, \mathcal{F}, P) \). For such space \( Z \) in this example, \( \rho(Z) \) is finite (real valued) for all \( Z \in \mathcal{Z} \).

If the random outcome represents a reward, i.e., larger realizations of \( Z \) are preferred, we can define a risk measure \( \varrho(Z) = \rho(-Z) \), where \( \rho \) satisfies axioms (R1)–(R4). In this case, the function \( \varrho \) also satisfies (R1) and (R4). The axioms (R2) and (R3) change to:

(R2a) Monotonicity: If \( Z, Z' \in \mathcal{Z} \) and \( Z \succeq Z' \), then \( \varrho(Z) \leq \varrho(Z') \).

(R3a) Translation Equivariance: If \( a \in \mathbb{R} \) and \( Z \in \mathcal{Z} \), then \( \varrho(Z + a) = \varrho(Z) - a \).

All our considerations regarding risk measures satisfying (R1)–(R4) have their obvious counterparts for risk measures satisfying (R1)-(R2a)-(R3a)-(R4).

With each space \( \mathcal{Z} := L^p(\Omega, \mathcal{F}, P) \) is associated its dual space \( \mathcal{Z}^* := L^q(\Omega, \mathcal{F}, P) \), where \( q \in (1, +\infty) \) is such that \( \frac{1}{p} + \frac{1}{q} = 1 \). For \( Z \in \mathcal{Z} \) and \( \zeta \in \mathcal{Z}^* \) their scalar product is defined as

\[ \langle \zeta, Z \rangle := \int_\Omega \zeta(\omega)Z(\omega)dP(\omega). \quad (6.32) \]

Recall that the conjugate function \( \rho^* : \mathcal{Z}^* \to \mathbb{R} \) of a risk measure \( \rho \) is defined as

\[ \rho^*(\zeta) := \sup_{Z \in \mathcal{Z}} \{ \langle \zeta, Z \rangle - \rho(Z) \}, \quad (6.33) \]

and the conjugate of \( \rho^* \) (the biconjugate function) as

\[ \rho^{**}(Z) := \sup_{\zeta \in \mathcal{Z}^*} \{ \langle \zeta, Z \rangle - \rho^*(\zeta) \}. \quad (6.34) \]

By the Fenchel-Moreau Theorem (Theorem 7.71) we have that if \( \rho : \mathcal{Z} \to \mathbb{R} \) is convex, proper and lower semicontinuous, then \( \rho^{**} = \rho \), i.e., \( \rho(\cdot) \) has the representation

\[ \rho(Z) = \sup_{\zeta \in \mathcal{Z}^*} \{ \langle \zeta, Z \rangle - \rho^*(\zeta) \}, \quad \forall Z \in \mathcal{Z}. \quad (6.35) \]

Conversely, if the representation (6.35) holds for some proper function \( \rho^* : \mathcal{Z}^* \to \mathbb{R} \), then \( \rho \) is convex, proper and lower semicontinuous. Note that if \( \rho \) is convex, proper and lower
semicontinuous, then its conjugate function \( \rho^* \) is also proper. Clearly, we can write the representation (6.35) in the following equivalent form

\[
\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \left\{ \langle \zeta, Z \rangle - \rho^*(\zeta) \right\}, \quad \forall Z \in \mathcal{Z},
\]

(6.36)

where \( \mathfrak{A} := \text{dom}(\rho^*) \) is the domain of the conjugate function \( \rho^* \).

The following basic duality result for convex risk measures is a direct consequence of the Fenchel-Moreau Theorem.

**Theorem 6.4.** Suppose that \( \rho : \mathcal{Z} \to \mathbb{R} \) is convex, proper and lower semicontinuous. Then the representation (6.36) holds with \( \mathfrak{A} := \text{dom}(\rho^*) \). Moreover, we have that: (i) Condition (R2) holds iff every \( \zeta \in \mathfrak{A} \) is nonnegative, i.e., \( \zeta(\omega) \geq 0 \) for a.e. \( \omega \in \Omega \); (ii) Condition (R3) holds iff \( \int_{\Omega} \zeta dP = 1 \) for every \( \zeta \in \mathfrak{A} \); (iii) Condition (R4) holds iff \( \rho(\cdot) \) is the support function of the set \( \mathfrak{A} \), i.e., can be represented in the form

\[
\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \langle \zeta, Z \rangle, \quad \forall Z \in \mathcal{Z}.
\]

(6.37)

**Proof.** If \( \rho : \mathcal{Z} \to \mathbb{R} \) is convex, proper, and lower semicontinuous, then representation (6.36) is valid by virtue of the Fenchel-Moreau Theorem (Theorem 7.71).

Now suppose that assumption (R2) holds. It follows that \( \rho^*(\zeta) = +\infty \) for every \( \zeta \in \mathcal{Z}^* \) which is not nonnegative. Indeed, if \( \zeta \in \mathcal{Z}^* \) is not nonnegative, then there exists a set \( \Delta \in \mathcal{F} \) of positive measure such that \( \zeta(\omega) < 0 \) for all \( \omega \in \Delta \). Consequently, for \( \tilde{Z} := 1_\Delta \) we have that \( \langle \zeta, \tilde{Z} \rangle < 0 \). Take any \( Z \) in the domain of \( \rho \), i.e., such that \( \rho(Z) \) is finite, and consider \( Z_t := Z - t\tilde{Z} \). Then for \( t \geq 0 \), we have that \( Z \geq Z_t \), and assumption (R2) implies that \( \rho(Z) \geq \rho(Z_t) \). Consequently,

\[
\rho^*(\zeta) \geq \sup_{t \in \mathbb{R}_+} \left\{ \langle \zeta, Z_t \rangle - \rho(Z_t) \right\} \geq \sup_{t \in \mathbb{R}_+} \left\{ \langle \zeta, Z \rangle - t\langle \zeta, \tilde{Z} \rangle - \rho(Z) \right\} = +\infty.
\]

Conversely, suppose that every \( \zeta \in \mathfrak{A} \) is nonnegative. Then for every \( \zeta \in \mathfrak{A} \) and \( Z \geq Z' \), we have that \( \langle \zeta, Z' \rangle \geq \langle \zeta, Z \rangle \). By (6.36), this implies that if \( Z \geq Z' \), then \( \rho(Z) \geq \rho(Z') \). This completes the proof of assertion (i).

Suppose that assumption (R3) holds. Then for every \( Z \in \text{dom}(\rho) \) we have

\[
\rho^*(\zeta) \geq \sup_{a \in \mathbb{R}} \left\{ \langle \zeta, Z + a \rangle - \rho(Z + a) \right\} = \sup_{a \in \mathbb{R}} \left\{ a \int_{\Omega} \zeta dP - a + \langle \zeta, Z \rangle - \rho(Z) \right\}.
\]

It follows that \( \rho^*(\zeta) = +\infty \) for any \( \zeta \in \mathcal{Z}^* \) such that \( \int_{\Omega} \zeta dP \neq 1 \). Conversely, if \( \int_{\Omega} \zeta dP = 1 \), then \( \langle \zeta, Z + a \rangle = \langle \zeta, Z \rangle + a \), and hence condition (R3) follows by (6.36). This completes the proof of (ii).

Clearly, if (6.37) holds, then \( \rho \) is positively homogeneous. Conversely, if \( \rho \) is positively homogeneous, then its conjugate function is the indicator function of a convex subset of \( \mathcal{Z}^* \). Consequently, the representation (6.37) follows by (6.36). \( \square \)

It follows from the above theorem that if \( \rho \) is a risk measure satisfying conditions (R1)–(R3) and is proper and lower semicontinuous, then the representation (6.36) holds.
with $\mathbb{A}$ being a subset of the set of probability density functions,

$$
\mathbb{P} := \left\{ \zeta \in \mathbb{Z}^* : \int_\Omega \zeta(\omega)dP(\omega) = 1, \ \zeta \succeq 0 \right\}.
$$

(6.38)

If, moreover, $\rho$ is positively homogeneous (i.e., condition (R4) holds), then its conjugate $\rho^*$ is the indicator function of a convex set $\mathbb{A} \subset \mathbb{Z}^*$, and $\mathbb{A}$ is equal to the subdifferential $\partial \rho(0)$ of $\rho$ at $0 \in \mathbb{Z}$. Furthermore, $\rho(0) = 0$ and hence by the definition of $\partial \rho(0)$ we have that

$$
\mathbb{A} = \{ \zeta \in \mathbb{P} : \langle \zeta, Z \rangle \leq \rho(Z), \ \forall Z \in \mathbb{Z} \}.
$$

(6.39)

The set $\mathbb{A}$ is weakly$^*$ closed. Recall that if the space $\mathbb{Z}$, and hence $\mathbb{Z}^*$, is reflexive, then a convex subset of $\mathbb{Z}^*$ is closed in the weak$^*$ topology of $\mathbb{Z}^*$ iff it is closed in the strong (norm) topology of $\mathbb{Z}^*$. If $\rho$ is positively homogeneous and continuous, then $\mathbb{A} = \partial \rho(0)$ is a bounded (and weakly$^*$ compact) subset of $\mathbb{Z}^*$ (see Proposition 7.74).

We have that if $\rho$ is a coherent risk measure, then the corresponding set $\mathbb{A}$ is a set of probability density functions. Consequently, for any $\zeta \in \mathbb{A}$ we can view $\langle \zeta, Z \rangle$ as the expectation $\mathbb{E}_\zeta[Z]$ taken with respect to the probability measure $\zeta dP$, defined by the density $\zeta$. Consequently representation (6.37) can be written in the form

$$
\rho(Z) = \sup_{\zeta \in \mathbb{A}} \mathbb{E}_\zeta[Z], \ \forall Z \in \mathbb{Z}.
$$

(6.40)

Definition of a risk measure $\rho$ depends on a particular choice of the corresponding space $\mathbb{Z}$. In many cases there is a natural choice of $\mathbb{Z}$ which ensures that $\rho(Z)$ is finite valued for all $Z \in \mathbb{Z}$. We shall see such examples in section 6.3.2. By Theorem 7.79 we have the following result which shows that for real valued convex and monotone risk measures, the assumption of lower semicontinuity in Theorem 6.4 holds automatically.

**Proposition 6.5.** Let $\mathbb{Z} := L_p(\Omega, \mathcal{F}, P)$, with $p \in [1, +\infty]$, and $\rho : \mathbb{Z} \rightarrow \mathbb{R}$ be a real valued risk measure satisfying conditions (R1) and (R2). Then $\rho$ is continuous and subdifferentiable on $\mathbb{Z}$.

Theorem 6.4 together with Proposition 6.5 imply the following basic duality result.

**Theorem 6.6.** Let $\rho : \mathbb{Z} \rightarrow \mathbb{R}$, where $\mathbb{Z} := L_p(\Omega, \mathcal{F}, P)$ with $p \in [1, +\infty]$. Then $\rho$ is a real valued coherent risk measure iff there exists a convex bounded and weakly$^*$ closed set $\mathbb{A} \subset \mathbb{P}$ such that the representation (6.37) holds.

**Proof.** If $\rho : \mathbb{Z} \rightarrow \mathbb{R}$ is a real valued coherent risk measure, then by Proposition 6.5 it is continuous, and hence by Theorem 6.4 the representation (6.37) holds with $\mathbb{A} = \partial \rho(0)$. Moreover, the subdifferential of a convex continuous function is bounded and weakly$^*$ closed (and hence is weakly$^*$ compact).

Conversely, if the representation (6.37) holds with the set $\mathbb{A}$ being a convex subset of $\mathbb{P}$ and weakly$^*$ compact, then $\rho$ is real valued and satisfies conditions (R1)–(R4).

Of course, the analysis simplifies considerably if the space $\Omega$ is finite, say $\Omega := \{\omega_1, \ldots, \omega_K\}$ equipped with sigma algebra of all subsets of $\Omega$ and respective (positive)
6.3. Coherent Risk Measures

probabilities \( p_1, \ldots, p_K \). Then every function \( Z : \Omega \to \mathbb{R} \) is measurable and the space \( Z \) of all such functions can be identified with \( \mathbb{R}^K \) by identifying \( Z \in Z \) with the vector \( (Z(\omega_1), \ldots, Z(\omega_K)) \in \mathbb{R}^K \). The dual of the space \( \mathbb{R}^K \) can be identified with itself by using the standard scalar product in \( \mathbb{R}^K \), and the set \( \mathcal{P} \) becomes

\[
\mathcal{P} = \left\{ \zeta \in \mathbb{R}^K : \sum_{k=1}^K p_k \zeta_k = 1, \zeta \geq 0 \right\}. \tag{6.41}
\]

The above set \( \mathcal{P} \) forms a convex bounded subset of \( \mathbb{R}^K \), and hence the set \( \mathfrak{A} \) is also bounded.

6.3.1 Differentiability Properties of Risk Measures

Let \( \rho : Z \to \mathbb{R} \) be a convex proper lower semicontinuous risk measure. By convexity and lower semicontinuity of \( \rho \) we have that \( \rho^{**} = \rho \), and hence by Proposition 7.73 that

\[
\partial \rho(Z) = \arg \sup_{\zeta \in \mathfrak{A}} \{ \langle \zeta, Z \rangle - \rho^*(\zeta) \}, \tag{6.42}
\]

provided that \( \rho(Z) \) is finite. If, moreover, \( \rho \) is positively homogeneous, then \( \mathfrak{A} = \partial \rho(0) \) and

\[
\partial \rho(Z) = \arg \max_{\zeta \in \mathfrak{A}} \{ \langle \zeta, Z \rangle \}. \tag{6.43}
\]

As we know conditions (R1)–(R3) imply that \( \mathfrak{A} \) is a subset of the set \( \mathcal{P} \) of probability density functions. Consequently, under conditions (R1)–(R3), \( \partial \rho(Z) \) is a subset of \( \mathcal{P} \) as well.

We also have that if \( \rho \) is finite valued and continuous at \( Z \), then \( \partial \rho(Z) \) is a nonempty bounded and weakly* compact subset of \( Z^* \), \( \rho \) is Hadamard directionally differentiable and subdifferentiable at \( Z \) and

\[
\rho'(Z, H) = \sup_{\zeta \in \partial \rho(Z)} \langle \zeta, H \rangle, \quad \forall H \in Z. \tag{6.44}
\]

In particular, if \( \rho \) is continuous at \( Z \) and \( \partial \rho(Z) \) is a singleton, i.e., \( \partial \rho(Z) \) consists of unique point denoted \( \nabla \rho(Z) \), then \( \rho \) is Hadamard differentiable at \( Z \) and

\[
\rho'(Z, \cdot) = \langle \nabla \rho(Z), \cdot \rangle. \tag{6.45}
\]

We often have to deal with composite functions \( \rho \circ F : \mathbb{R}^n \to \mathbb{R} \), where \( F : \mathbb{R}^n \to Z \) is a mapping. We write \( f(x, \omega) \), or \( f_\omega(x) \), for \( [F(x)](\omega) \), and view \( f(x, \omega) \) as a random function defined on the measurable space \( (\Omega, \mathcal{F}) \). We say that the mapping \( F \) is convex if the function \( f(\cdot, \omega) \) is convex for every \( \omega \in \Omega \).

**Proposition 6.7.** If the mapping \( F : \mathbb{R}^n \to Z \) is convex and \( \rho : Z \to \mathbb{R} \) satisfies conditions (R1)–(R2), then the composite function \( \phi(\cdot) := \rho(F(\cdot)) \) is convex.

**Proof.** For any \( x, x' \in \mathbb{R}^n \) and \( t \in [0, 1] \), we have by convexity of \( F(\cdot) \) and monotonicity of \( \rho(\cdot) \) that

\[
\rho(F(tx + (1-t)x')) \leq t \rho(F(x)) + (1-t) \rho(F(x')).
\]
Hence convexity of $\rho(\cdot)$ implies that
\[
\rho(F(tx + (1-t)x')) \leq t\rho(F(x)) + (1-t)\rho(F(x')).
\]
This proves convexity of $\rho(F(\cdot))$. $\square$

It should be noted that the monotonicity condition (R2) was essential in the above derivation of convexity of the composite function.

Let us discuss differentiability properties of the composite function $\phi = \rho \circ F$ at a point $\bar{x} \in \mathbb{R}^n$. As before we assume that $Z := \mathcal{L}_p(\Omega, \mathcal{F}, P)$. The mapping $F : \mathbb{R}^n \to Z$ maps a point $x \in \mathbb{R}^n$ into a real valued function (or rather a class of functions which may differ on sets of $P$-measure zero) $[F(x)](\cdot)$ on $\Omega$, also denoted $f(x, \cdot)$, which is an element of $\mathcal{L}_p(\Omega, \mathcal{F}, P)$. If $F$ is convex, then $f(\cdot, \omega)$ is convex valued and hence is continuous and has (finite valued) directional derivatives at $\bar{x}$, denoted $f'_x(\bar{x}, h)$. These properties are inherited by the mapping $F$.

**Lemma 6.8.** Let $Z := \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and $F : \mathbb{R}^n \to Z$ be a convex mapping. Then $F$ is continuous, directionally differentiable and
\[
[F'(\bar{x}, h)](\omega) = f'_x(\bar{x}, h), \quad \omega \in \Omega, \ h \in \mathbb{R}^n.
\] (6.46)

**Proof.** In order to show continuity of $F$ we need to verify that, for an arbitrary point $\bar{x} \in \mathbb{R}^n$, $\|F(x) - F(\bar{x})\|_p$ tends to zero as $x \to \bar{x}$. By the Lebesgue Dominated Convergence Theorem and continuity of $f(\cdot, \omega)$ we can write that
\[
\lim_{\bar{x} \to x} \int_{\Omega} \|f(x, \omega) - f(\bar{x}, \omega)\|^p dP(\omega) = \int_{\Omega} \lim_{\bar{x} \to x} \|f(x, \omega) - f(\bar{x}, \omega)\|^p dP(\omega) = 0, \quad (6.47)
\]
provided that there exists a neighborhood $U \subset \mathbb{R}^n$ of $\bar{x}$ such that the family $\{\|f(x, \omega) - f(\bar{x}, \omega)\|^p\}_{x \in U}$ is dominated by a $P$-integrable function, or equivalently that $\{\|f(x, \omega) - f(\bar{x}, \omega)\|\}_{x \in U}$ is dominated by a function from the space $\mathcal{L}_p(\Omega, \mathcal{F}, P)$. Since $f(\bar{x}, \cdot)$ belongs to $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, it suffices to verify this dominance condition for $\{\|f(x, \omega)\|\}_{x \in U}$.

Now let $x_1, \ldots, x_{n+1} \in \mathbb{R}^n$ be such points that the set $U := \text{conv}\{x_1, \ldots, x_{n+1}\}$ forms a neighborhood of the point $\bar{x}$, and let $g(\omega) := \max\{f(x_1, \omega), \ldots, f(x_{n+1}, \omega)\}$. By convexity of $f(\cdot, \omega)$ we have that $f(x, \cdot) \leq g(\cdot)$ for all $x \in U$. Also since every $f(x_i, \cdot)$, $i = 1, \ldots, n + 1$, is an element of $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, we have that $g \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ as well. That is, $g(\omega)$ gives an upper bound for $\{f(x, \omega)\}_{x \in U}$. Also by convexity of $f(\cdot, \omega)$ we have that
\[
f(x, \omega) \geq 2f(\bar{x}, \omega) - f(2\bar{x} - x, \omega).
\]
By shrinking the neighborhood $U$ if necessary, we can assume that $U$ is symmetrical around $\bar{x}$, i.e., if $x \in U$, then $2\bar{x} - x \in U$. Consequently, we have that $\tilde{g}(\omega) := 2f(\bar{x}, \omega) - g(\omega)$ gives a lower bound for $\{f(x, \omega)\}_{x \in U}$, and $\tilde{g} \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$. This shows that the required domination condition holds and hence $F$ is continuous at $\bar{x}$ by (6.47).

Now for $h \in \mathbb{R}^n$ and $t > 0$ denote
\[
R_t(\omega) := t^{-1} |f(\bar{x} + th, \omega) - f(\bar{x}, \omega)| \quad \text{and} \quad Z(\omega) := f'_x(\bar{x}, h), \quad \omega \in \Omega.
\]
Note that $f(\bar{x} + th, \cdot)$ and $f(\bar{x}, \cdot)$ are elements of the space $L_p(\Omega, \mathcal{F}, P)$, and hence $R_t(\cdot)$ is also an element of $L_p(\Omega, \mathcal{F}, P)$ for any $t > 0$. Since for a.e. $\omega \in \Omega$, $f(\cdot, \omega)$ is convex real valued, we have that $R_t(\omega)$ is monotonically nonincreasing and converges to $Z(\omega)$ as $t \downarrow 0$. Therefore we have that $R_t(\cdot) \geq Z(\cdot)$ for any $t > 0$. Again by convexity of $f(\cdot, \omega)$, we have that for $t > 0$,

$$Z(\cdot) \geq t^{-1} \left[ f(\bar{x}, \cdot) - f(\bar{x} - th, \cdot) \right].$$

We obtain that $Z(\cdot)$ is bounded from above and below by functions which are elements of the space $L_p(\Omega, \mathcal{F}, P)$, and hence $Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ as well.

We have that $R_t(\cdot) - Z(\cdot)$, and hence $|R_t(\cdot) - Z(\cdot)|^p$, are monotonically decreasing to zero as $t \downarrow 0$ and for any $t > 0$, $E \| R_t - Z \|^p$ is finite. It follows by the Monotone Convergence Theorem that $E \| R_t - Z \|^p$ tends to zero as $t \downarrow 0$. That is, $R_t$ converges to $Z$ in the norm topology of $Z$. Since $R_t = t^{-1}[F(\bar{x} + th) - F(\bar{x})]$, this shows that $F$ is directionally differentiable at $\bar{x}$ and formula (6.46) follows.

The following theorem can be viewed as an extension of Theorem 7.4.6 where a similar result is derived for $\rho(\cdot) := E[\cdot]$.

**Theorem 6.9.** Let $Z := L_p(\Omega, \mathcal{F}, P)$ and $F : \mathbb{R}^n \rightarrow Z$ be a convex mapping. Suppose that $\rho$ is convex, finite valued and continuous at $Z := F(\bar{x})$. Then the composite function $\phi = \rho \circ F$ is directionally differentiable at $\bar{x}$, $\phi'(\bar{x}, h)$ is finite valued for every $h \in \mathbb{R}^n$ and

$$\phi'(\bar{x}, h) = \sup_{\zeta \in \partial \rho(Z)} \int_{\Omega} f'_\omega(\bar{x}, h) \zeta(\omega) dP(\omega). \quad (6.48)$$

**Proof.** Since $\rho$ is continuous at $\bar{Z}$, it follows that $\rho$ subdifferentiable and Hadamard directionally differentiable at $\bar{Z}$ and formula (6.44) holds. Also by Lemma 6.8, mapping $F$ is directionally differentiable. Consequently, we can apply the chain rule (see Proposition 7.5.8) to conclude that $\phi(\cdot)$ is directionally differentiable at $\bar{x}$, $\phi'(\bar{x}, h)$ is finite valued and

$$\phi'(\bar{x}, h) = \rho'(\bar{Z}, F'(\bar{x}, h)). \quad (6.49)$$

Together with (6.44) and (6.46), the above formula (6.49) implies (6.48).

It is also possible to write formula (6.48) in terms of the corresponding subdifferentials.

**Theorem 6.10.** Let $Z := L_p(\Omega, \mathcal{F}, P)$ and $F : \mathbb{R}^n \rightarrow Z$ be a convex mapping. Suppose that $\rho$ satisfies conditions (R1) and (R2) and is finite valued and continuous at $Z := F(\bar{x})$. Then the composite function $\phi = \rho \circ F$ is subdifferentiable at $\bar{x}$ and

$$\partial \phi(\bar{x}) = \text{cl} \left( \bigcup_{\zeta \in \partial \rho(Z)} \int_{\Omega} \partial f(\bar{x}) \zeta(\omega) dP(\omega) \right). \quad (6.50)$$

**Proof.** Since, by Lemma 6.8, $F$ is continuous at $\bar{x}$ and $\rho$ is continuous at $F(\bar{x})$, we have that $\phi$ is continuous at $\bar{x}$, and hence $\phi(x)$ is finite valued for all $x$ in a neighborhood of $\bar{x}$.
Moreover, by Proposition 6.7, \( \phi(\cdot) \) is convex, and hence is continuous in a neighborhood of \( \bar{x} \) and is subdifferentiable at \( \bar{x} \). Also formula (6.48) holds. It follows that \( \phi'(\bar{x}, \cdot) \) is convex, continuous, positively homogeneous and

\[
\phi'(\bar{x}, \cdot) = \sup_{\zeta \in \partial \rho(\bar{Z})} \eta_\zeta(\cdot),
\]

where

\[
\eta_\zeta(\cdot) := \int_\Omega f'_\omega(\bar{x}, \cdot) \zeta(\omega) dP(\omega).
\]

Because of the condition (R2) we have that every \( \zeta \in \partial \rho(\bar{Z}) \) is nonnegative. Consequently the corresponding function \( \eta_\zeta \) is convex and positively homogeneous, and hence is the support function of the set \( \partial \eta_\zeta(0) \). The supremum of these functions, given by the right hand side of (6.51), is the support function of the set \( \bigcup_{\zeta \in \partial \rho(\bar{Z})} \partial \eta_\zeta(0) \). Applying Theorem 7.47 and using the fact that the subdifferential of \( f' \cup \eta \) given by the right hand side of (6.51), is the support function of the set \( \partial \eta \) and hence is the support function of the set \( \eta \).

Consequently the corresponding function \( \eta_\zeta \) is convex continuous and positively homogeneous, and hence is the support function of the set \( \partial \eta_\zeta(0) \). The supremum of these functions, given by the right hand side of (6.51), is the support function of the set \( \bigcup_{\zeta \in \partial \rho(\bar{Z})} \partial \eta_\zeta(0) \).

Applying Theorem 7.47 and using the fact that the subdifferential of \( f'(\bar{x}, \cdot) \) at 0 \( \in \mathbb{R}^n \) coincides with \( \partial f_\omega(\bar{x}) \), we obtain

\[
\partial \eta_\zeta(0) = \int_\Omega \partial f_\omega(\bar{x}) \zeta(\omega) dP(\omega).
\]

Since \( \partial \rho(\bar{Z}) \) is convex, it is straightforward to verify that the set \( \bigcup_{\zeta \in \partial \rho(\bar{Z})} \partial \eta_\zeta(0) \) is also convex. Consequently it follows by (6.51) and (6.53) that the subdifferential of \( \phi'(\bar{x}, \cdot) \) at 0 \( \in \mathbb{R}^n \) is equal to the topological closure of the set \( \bigcup_{\zeta \in \partial \rho(\bar{Z})} \partial \eta_\zeta(0) \), i.e., is given by the right hand side of (6.50). It remains to note that the subdifferential of \( \phi'(\bar{x}, \cdot) \) at 0 \( \in \mathbb{R}^n \) coincides with \( \partial \phi(\bar{x}) \).

Under the assumptions of the above theorem we have that the composite function \( \phi \) is convex and is continuous (in fact even Lipschitz continuous) in a neighborhood of \( \bar{x} \). It follows that \( \phi \) is differentiable\(^6\) at \( \bar{x} \) iff \( \partial \phi(\bar{x}) \) is a singleton. This leads to the following result, where for \( \zeta \geq 0 \), we say that a property holds for \( \zeta \)-a.e. \( \omega \in \Omega \) if the set of points \( A \in \mathcal{F} \) where it does not hold has \( \zeta dP \) measure zero, i.e., \( \int_A \zeta(\omega) dP(\omega) = 0 \). Of course, if \( P(A) = 0 \), then \( \int_A \zeta(\omega) dP(\omega) = 0 \). That is, if a property holds for a.e. \( \omega \in \Omega \) with respect to \( P \), then it holds for \( \zeta \)-a.e. \( \omega \in \Omega \).

**Corollary 6.11.** Let \( Z := \mathcal{L}_p(\Omega, \mathcal{F}, P) \) and \( F : \mathbb{R}^n \to Z \) be a convex mapping. Suppose that \( \rho \) satisfies conditions (R1) and (R2) and is finite valued and continuous at \( Z := F(\bar{x}) \).

Then the composite function \( \phi = \rho \circ F \) is differentiable at \( \bar{x} \) if and only if the following two properties hold:

1. For every \( \zeta \in \partial \rho(Z) \) the function \( f(\cdot, \omega) \) is differentiable at \( \bar{x} \) for \( \zeta \)-a.e. \( \omega \in \Omega \).
2. \( \int_\Omega \nabla f_\omega(\bar{x}) \zeta(\omega) dP(\omega) \) is the same for every \( \zeta \in \partial \rho(Z) \).

In particular, if \( \partial \rho(Z) = \{ \zeta \} \) is a singleton, then \( \phi \) is differentiable at \( \bar{x} \) if and only if \( f(\cdot, \omega) \) is differentiable at \( \bar{x} \) for \( \zeta \)-a.e. \( \omega \in \Omega \), in which case

\[
\nabla \phi(\bar{x}) = \int_\Omega \nabla f_\omega(\bar{x}) \zeta(\omega) dP(\omega).
\]

\(^6\)Note that since \( \phi(\cdot) \) is Lipschitz continuous near \( \bar{x} \), the notions of Gâteaux and Fréchet differentiability at \( \bar{x} \) are equivalent here.
Proof. By Theorem 6.10 we have that \( \phi \) is differentiable at \( \bar{x} \) iff the set on the right hand side of (6.50) is a singleton. Clearly this set is a singleton iff the set \( \int_{\Omega} \partial f_{a}(\bar{x}) \zeta(\omega) dP(\omega) \) is a singleton and is the same for every \( \zeta \in \partial \rho(\bar{Z}) \). Since \( \partial f_{a}(\bar{x}) \) is a singleton iff \( f_{a}(\cdot) \) is differentiable at \( \bar{x} \), in which case \( \partial f_{a}(\bar{x}) = \{ \nabla f_{a}(\bar{x}) \} \), we obtain that \( \phi \) is differentiable at \( \bar{x} \) iff conditions (i) and (ii) hold. The second assertion then follows. \( \square \)

Of course, if the set inside the parentheses on the right hand side of (6.50) is closed, then there is no need to take its topological closure. This holds true in the following case.

Corollary 6.12. Suppose that the assumptions of Theorem 6.10 are satisfied and for every \( \zeta \in \partial \rho(\bar{Z}) \) the function \( f_{\omega}(\cdot) \) is differentiable at \( \bar{x} \) for \( \zeta \text{-a.e. } \omega \in \Omega \). Then

\[
\partial \phi(\bar{x}) = \bigcup_{\zeta \in \partial \rho(\bar{Z})} \int_{\Omega} \nabla f_{a}(\bar{x}) \zeta(\omega) dP(\omega). \tag{6.55}
\]

Proof. In view of Theorem 6.10 we only need to show that the set on the right hand side of (6.55) is closed. As \( \rho \) is continuous at \( \bar{Z} \), the set \( \partial \rho(\bar{Z}) \) is weakly* compact. Also the mapping \( \zeta \mapsto \int_{\Omega} \nabla f_{a}(\bar{x}) \zeta(\omega) dP(\omega) \), from \( Z^{*} \) to \( R^{n} \), is continuous with respect to the weak* topology of \( Z^{*} \) and the standard topology of \( R^{n} \). It follows that the image of the set \( \partial \rho(\bar{Z}) \) by this mapping is compact and hence is closed, i.e., the set at the right hand side of (6.55) is closed. \( \square \)

### 6.3.2 Examples of Risk Measures

In this section we discuss several examples of risk measures which are commonly used in applications. In each of the following examples it is natural to use the space \( Z := L_{p}(\Omega, F, P) \) for an appropriate \( p \in [1, +\infty) \). Note that if a random variable \( Z \) has a \( p \)-th order finite moment, then it has finite moments of any order \( p' \) smaller than \( p \), i.e., if \( 1 \leq p' < p \) and \( Z \in L_{p'}(\Omega, F, P) \), then \( Z \in L_{p'}(\Omega, F, P) \). This gives a natural embedding of \( L_{p'}(\Omega, F, P) \) into \( L_{p'}(\Omega, F, P) \) for \( p' < p \). Note, however, that this embedding is not continuous. Unless stated otherwise, all expectations and probabilistic statements will be made with respect to the probability measure \( P \).

Before proceeding to particular examples let us consider the following construction. Let \( \rho : Z \rightarrow R \) and define

\[
\tilde{\rho}(Z) := E[Z] + \inf_{t \in R} \rho(Z - t). \tag{6.56}
\]

Clearly we have that for any \( a \in R \),

\[
\tilde{\rho}(Z + a) = E[Z + a] + \inf_{t \in R} \rho(Z + a - t) = E[Z] + a + \inf_{t \in R} \rho(Z - t) = \tilde{\rho}(Z) + a.
\]

That is, \( \tilde{\rho} \) satisfies condition (R3) irrespectively whether \( \rho \) does. It is not difficult to see that if \( \rho \) satisfies conditions (R1) and (R2), then \( \tilde{\rho} \) satisfies these conditions as well. Also if \( \rho \) is
positively homogeneous, then so is \( \tilde{\rho} \). Let us calculate the conjugate of \( \tilde{\rho} \). We have

\[
\tilde{\rho}^*(\zeta) = \sup_{Z \in \mathcal{Z}} \{\langle \zeta, Z \rangle - \tilde{\rho}(Z)\} = \sup_{Z \in \mathcal{Z}} \{\langle \zeta, Z \rangle - \mathbb{E}[Z] - \inf_{t \in \mathbb{R}} \rho(Z - t)\}
\]

\[
= \sup_{Z \in \mathcal{Z}, t \in \mathbb{R}} \{\langle \zeta, Z \rangle - \mathbb{E}[Z] - \rho(Z - t)\}
\]

\[
= \sup_{Z \in \mathcal{Z}, t \in \mathbb{R}} \{\langle \zeta - 1, Z \rangle + t(\mathbb{E}[\zeta] - 1) - \rho(Z)\}.
\]

It follows that

\[
\tilde{\rho}^*(\zeta) = \begin{cases} 
\rho^*(\zeta - 1), & \text{if } \mathbb{E}[\zeta] = 1, \\
+\infty, & \text{if } \mathbb{E}[\zeta] \neq 1.
\end{cases}
\]

The construction below can be viewed as a homogenization of a risk measure \( \rho : \mathcal{Z} \to \mathbb{R} \). Define

\[
\hat{\rho}(Z) := \inf_{\tau > 0} \tau \rho(\tau^{-1}Z).
\]

For any \( t > 0 \), by making change of variables \( \tau \mapsto \tau t \), we obtain that \( \hat{\rho}(tZ) = t\hat{\rho}(Z) \). That is, \( \hat{\rho} \) is positively homogeneous whether \( \rho \) is or isn’t. Clearly, if \( \rho \) is positively homogeneous to start with, then \( \rho = \hat{\rho} \).

If \( \rho \) is convex, then so is \( \hat{\rho} \). Indeed, observe that if \( \rho \) is convex, then function \( \varphi(\tau, Z) := \tau \rho(\tau^{-1}Z) \) is convex jointly in \( Z \) and \( \tau > 0 \). This can be verified directly as follows. For \( t \in [0, 1], \tau_1, \tau_2 > 0 \) and \( Z_1, Z_2 \in \mathcal{Z} \), and setting \( \tau := t\tau_1 + (1-t)\tau_2 \) and \( Z := tZ_1 + (1-t)Z_2 \), we have

\[
t[\tau_1 \rho(\tau_1^{-1}Z_1)] + (1-t)[\tau_2 \rho(\tau_2^{-1}Z_2)] = \tau \left[ \frac{t\tau_1}{\tau} \rho(\tau_1^{-1}Z_1) + \frac{(1-t)\tau_2}{\tau} \rho(\tau_2^{-1}Z_2) \right]
\]

\[
\geq \varphi \left( \frac{t}{\tau} Z_1 + \frac{(1-t)}{\tau} Z_2 \right) = \tau \rho(\tau^{-1}Z).
\]

Minimizing convex function \( \varphi(\tau, Z) \) over \( \tau > 0 \), we obtain a convex function. It is also not difficult to see that if \( \rho \) satisfies conditions (R2) and (R3), then so is \( \hat{\rho} \).

Let us calculate the conjugate of \( \hat{\rho} \). We have

\[
\hat{\rho}^*(\zeta) = \sup_{Z \in \mathcal{Z}} \{\langle \zeta, Z \rangle - \hat{\rho}(Z)\} = \sup_{Z \in \mathcal{Z}, \tau > 0} \{\langle \zeta, Z \rangle - \tau \rho(\tau^{-1}Z)\}
\]

\[
= \sup_{Z \in \mathcal{Z}, \tau > 0} \{\tau[\langle \zeta, Z \rangle - \rho(Z)\}.
\]

It follows that \( \hat{\rho}^* \) is the indicator function of the set

\[
\mathcal{A} := \{\zeta \in \mathcal{Z}^* : \langle \zeta, Z \rangle \leq \rho(Z), \forall Z \in \mathcal{Z} \}.
\] \hspace{1cm} (6.58)

If, moreover, \( \hat{\rho} \) is lower semicontinuous, then \( \hat{\rho} \) is equal to the conjugate of \( \hat{\rho}^* \), and hence \( \hat{\rho} \) is the support function of the above set \( \mathcal{A} \).

**Example 6.13 (Utility model)** It is possible to relate the theory of convex risk measures with the utility model. Let \( g : \mathbb{R} \to \mathbb{R} \) be a proper convex nondecreasing lower semicontinuous function such that the expectation \( \mathbb{E}[g(Z)] \) is well defined for all \( Z \in \mathcal{Z} \) (it is allowed
6.3. Coherent Risk Measures

here for $\mathbb{E}[g(Z)]$ to take value $+\infty$, but not $-\infty$ since the corresponding risk measure is required to be proper. We can view the function $g$ as a disutility function\(^7\).

**Proposition 6.14.** Let $g : \mathbb{R} \to \overline{\mathbb{R}}$ be a proper convex nondecreasing lower semicontinuous function. Suppose that the risk measure

$$ \rho(Z) := \mathbb{E}[g(Z)] \quad (6.59) $$

is well defined and proper. Then $\rho$ is convex, lower semicontinuous, satisfies the monotonicity condition (R2) and the representation (6.35) holds with

$$ \rho^*(\zeta) = \mathbb{E}[g^*(\zeta)]. \quad (6.60) $$

Moreover, if $\rho(Z)$ is finite, then

$$ \partial \rho(Z) = \{ \zeta \in \mathbb{Z}^* : \zeta(\omega) \in \partial g(Z(\omega)) \text{ a.e. } \omega \in \Omega \}. \quad (6.61) $$

**Proof.** Since $g$ is lower semicontinuous and convex, we have by Fenchel-Moreau Theorem that

$$ g(z) = \sup_{\alpha \in \mathbb{R}} \{ \alpha z - g^*(\alpha) \}, $$

where $g^*$ is the conjugate of $g$. As $g$ is proper, the conjugate function $g^*$ is also proper. It follows that

$$ \rho(Z) = \mathbb{E} \left[ \sup_{\alpha \in \mathbb{R}} \{ \alpha Z - g^*(\alpha) \} \right]. \quad (6.62) $$

By the interchangeability principle (Theorem 7.80) for the space $\mathcal{M} := \mathbb{Z}^* = \mathcal{L}_q(\Omega, \mathcal{F}, P)$, which is decomposable, we obtain

$$ \rho(Z) = \sup_{\zeta \in \mathbb{Z}^*} \{ \langle \zeta, Z \rangle - \mathbb{E}[g^*(\zeta)] \}. \quad (6.63) $$

It follows that $\rho$ is convex and lower semicontinuous, and representation (6.35) holds with the conjugate function given in (6.60). Moreover, since the function $g$ is nondecreasing, it follows that $\rho$ satisfies the monotonicity condition (R2).

Since $\rho$ is convex proper and lower semicontinuous, and hence $\rho^{**} = \rho$, we have by Proposition 7.73 that

$$ \partial \rho(Z) = \arg \max_{\zeta \in \mathbb{R}} \{ \mathbb{E}[\zeta Z - g^*(\zeta)] \}, \quad (6.64) $$

assuming that $\rho(Z)$ is finite. Together with formula (7.253) of the interchangeability principle (Theorem 7.80) this implies (6.61).

The above risk measure $\rho$, defined in (6.59), does not satisfy condition (R3) unless $g(z) \equiv z$. We can consider the corresponding risk measure $\tilde{\rho}$, defined in (6.56), which in the present case can be written as

$$ \tilde{\rho}(Z) = \inf_{t \in \mathbb{R}} \mathbb{E}[Z + g(Z - t)]. \quad (6.65) $$

\(^7\)We consider here minimization problems, and that is why we speak about disutility. Any disutility function $g$ corresponds to a utility function $u : \mathbb{R} \to \mathbb{R}$ defined by $u(z) = -g(-z)$. Note that the function $u$ is concave and nondecreasing since the function $g$ is convex and nondecreasing.
We have that $\tilde{\rho}$ is convex since $g$ is convex, that $\tilde{\rho}$ is monotone (i.e., condition (R2) holds) if $z + g(z)$ is monotonically nondecreasing, and that $\tilde{\rho}$ satisfies condition (R3). If, moreover, $\tilde{\rho}$ is lower semicontinuous, then the dual representation

$$\tilde{\rho}(Z) = \sup_{z \in Z^*} \left\{ \langle \zeta - 1, Z \rangle - E[g^*(\zeta)] \right\}$$  \hspace{1cm} (6.66)$$

holds. \hfill \blacksquare

**Example 6.15 (Average Value-at-Risk)** The risk measure $\rho$ associated with disutility function $g$, defined in (6.59), is positively homogeneous only if $g$ is positively homogeneous. Suppose now that $g(z) := \max\{az, bz\}$, where $b \geq a$. Then $g(\cdot)$ is positively homogeneous and convex. It is natural here to use the space $Z := L_1(\Omega, \mathcal{F}, P)$, since $E[g(Z)]$ is finite for every $Z \in L_1(\Omega, \mathcal{F}, P)$. The conjugate function of $g$ is the indicator function $g^* = \mathbb{I}_{[a,b]}$. Therefore it follows by Proposition 6.14 that the representation (6.37) holds with

$$\mathfrak{A} = \{ \zeta \in L_\infty(\Omega, \mathcal{F}, P) : \zeta(\omega) \in [a, b] \text{ a.e. } \omega \in \Omega \}. $$

Note that the dual space $Z^* = L_\infty(\Omega, \mathcal{F}, P)$, of the space $Z := L_1(\Omega, \mathcal{F}, P)$, appears naturally in the corresponding representation (6.37) since, of course, the condition that \text{"$\zeta(\omega) \in [a, b]$ for a.e. $\omega \in \Omega$" implies that $\zeta$ is essentially bounded.}

Consider now the risk measure

$$\tilde{\rho}(Z) := E[Z] + \inf_{t \in \mathbb{R}} E\{ \beta_1[t - Z]_+ + \beta_2[Z - t]_+ \}, \quad Z \in L_1(\Omega, \mathcal{F}, P),$$  \hspace{1cm} (6.67)$$

where $\beta_1 \in [0, 1]$ and $\beta_2 \geq 0$. This risk measure can be recognized as risk measure defined in (6.65), associated with function $g(z) := \beta_1 [-z]_+ + \beta_2 [z]_+$. For specified $\beta_1$ and $\beta_2$, the function $z + g(z)$ is convex and nondecreasing, and $\tilde{\rho}$ is a continuous coherent risk measure. For $\beta_1 \in (0, 1]$ and $\beta_2 > 0$, the above risk measure $\tilde{\rho}(Z)$ can be written in the form

$$\tilde{\rho}(Z) = (1 - \beta_1)E[Z] + \beta_1 AV@\mathbb{R}_+(Z),$$  \hspace{1cm} (6.68)$$

where $\alpha := \beta_1/(\beta_1 + \beta_2)$. Note that the right hand side of (6.67) attains its minimum at $t^* = V@\mathbb{R}_+(Z)$. Therefore the second term on the right hand side of (6.67) is the weighted measure of deviation from the quantile $V@\mathbb{R}_+(Z)$, discussed in section 6.2.3.

The respective conjugate function is the indicator function of the set $\mathfrak{A} := \text{dom}(\tilde{\rho}^*)$, and $\tilde{\rho}$ can be represented in the dual form (6.37) with

$$\mathfrak{A} = \{ \zeta \in L_\infty(\Omega, \mathcal{F}, P) : \zeta(\omega) \in [1 - \beta_1, 1 + \beta_2] \text{ a.e. } \omega \in \Omega, \ E[\zeta] = 1 \}. $$  \hspace{1cm} (6.69)$$

In particular, for $\beta_1 = 1$ we have that $\tilde{\rho}(\cdot) = AV@\mathbb{R}_+(\cdot)$, and hence the dual representation (6.37) of AV@\mathbb{R}_+ holds with the set

$$\mathfrak{A} = \{ \zeta \in L_\infty(\Omega, \mathcal{F}, P) : \zeta(\omega) \in [0, 1] \text{ a.e. } \omega \in \Omega, \ E[\zeta] = 1 \}. $$  \hspace{1cm} (6.70)$$

Since AV@\mathbb{R}_+(\cdot) is convex and continuous, it is subdifferentiable and its subdifferentials can be calculated using formula (6.43). That is,

$$\partial(\text{AV@\mathbb{R}_+})(Z) = \arg \max_{\zeta \in Z^*} \left\{ \langle \zeta, Z \rangle : \zeta(\omega) \in [0, 1] \text{ a.e. } \omega \in \Omega, \ E[\zeta] = 1 \right\}. $$  \hspace{1cm} (6.71)$$
Consider the maximization problem on the right hand side of (6.71). The Lagrangian of that problem is

\[ L(\zeta, \lambda) = \langle \zeta, Z \rangle + \lambda(1 - \mathbb{E}[\zeta]) = \langle \zeta, Z - \lambda \rangle + \lambda, \]

and its (Lagrangian) dual is the problem:

\[
\min_{\lambda \in \mathbb{R}} \max_{\zeta(\cdot) \in [0, \alpha^{-1}]} \{ \langle \zeta, Z - \lambda \rangle + \lambda \}. \tag{6.72}
\]

We have that

\[
\max_{\zeta(\cdot) \in [0, \alpha^{-1}]} \langle \zeta, Z - \lambda \rangle = \alpha^{-1} \mathbb{E}([Z - \lambda]_+),
\]

and hence the dual problem (6.72) can be written as

\[
\min_{\lambda \in \mathbb{R}} \alpha^{-1} \mathbb{E}([Z - \lambda]_+) + \lambda. \tag{6.73}
\]

The set of optimal solutions of problem (6.73) is the interval with the end points given by the left and right side \((1 - \alpha)\)-quantiles of the cdf \(H_Z(z) = \Pr(Z \leq z)\) of \(Z(\omega)\). Since the set of optimal solutions of the dual problem (6.72) is a compact subset of \(\mathbb{R}\), there is no duality gap between the maximization problem on the right hand side of (6.71) and its dual (6.72) (see Theorem 7.10). It follows that the set of optimal solutions of the right hand side of (6.71), and hence the subdifferential \(\partial (\mathbb{V} \mathbb{A} \mathbb{V} @ \mathbb{R}_\alpha)(Z)\), is given by such feasible \(\zeta\) that \((\zeta, \lambda)\) is a saddle point of the Lagrangian \(L(\zeta, \lambda)\) for any \((1 - \alpha)\)-quantile \(\lambda\). Recall that the left-side \((1 - \alpha)\)-quantile of the cdf \(H_Z(z)\) is called Value-at-Risk and denoted \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\).

Suppose for the moment that the set of \((1 - \alpha)\)-quantiles of \(H_Z\) is a singleton, i.e., consists of one point \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\). Then we have

\[
\partial (\mathbb{V} \mathbb{A} \mathbb{V} @ \mathbb{R}_\alpha)(Z) = \left\{ \zeta : \mathbb{E}[\zeta] = 1, \begin{array}{ll}
\zeta(\omega) = \alpha^{-1}, & \text{if } Z(\omega) > \mathbb{V} @ \mathbb{R}_\alpha(Z), \\
\zeta(\omega) = 0, & \text{if } Z(\omega) < \mathbb{V} @ \mathbb{R}_\alpha(Z), \\
\zeta(\omega) \in [0, \alpha^{-1}], & \text{if } Z(\omega) = \mathbb{V} @ \mathbb{R}_\alpha(Z).
\end{array} \right\} \tag{6.74}
\]

If the set of \((1 - \alpha)\)-quantiles of \(H_Z\) is not a singleton, then the probability that \(Z(\omega)\) belongs to that set is zero. Consequently, formula (6.74) still holds with the left-side quantile \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\) can be replaced by any \((1 - \alpha)\)-quantile of \(H_Z\).

It follows that \(\partial (\mathbb{V} \mathbb{A} \mathbb{V} @ \mathbb{R}_\alpha)(Z)\) is a singleton, and hence \(\mathbb{V} \mathbb{A} \mathbb{V} @ \mathbb{R}_\alpha(\cdot)\) is Hadamard differentiable at \(Z\), iff the following condition holds:

\[
\Pr(Z < \mathbb{V} @ \mathbb{R}_\alpha(Z)) = 1 - \alpha \quad \text{or} \quad \Pr(Z > \mathbb{V} @ \mathbb{R}_\alpha(Z)) = \alpha. \tag{6.75}
\]

Again if the set of \((1 - \alpha)\)-quantiles is not a singleton, then the left-side quantile \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\) in the above condition (6.75) can be replaced by any \((1 - \alpha)\)-quantile of \(H_Z\). Note that condition (6.75) is always satisfied if the cdf \(H_Z(\cdot)\) is continuous at \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\), but may also hold even if \(H_Z(\cdot)\) is discontinuous at \(\mathbb{V} @ \mathbb{R}_\alpha(Z)\).

**Example 6.16 (Exponential utility function risk measure)** Consider utility risk measure \(\rho\), defined in (6.59), associated with the exponential disutility function \(g(z) := e^z\). That is, \(\rho(Z) := \mathbb{E}[e^Z]\). A natural question is what space \(Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)\) to use here. Let
us observe that unless the sigma algebra $\mathcal{F}$ has a finite number of elements, in which case $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ is finite dimensional, there exist such $Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ that $\mathbb{E}[e^Z] = +\infty$.

In fact for any $p \in [1, +\infty)$ the domain of $\rho$ forms a dense subset of $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ and $\rho(\cdot)$ is discontinuous at every $Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ unless $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ is finite dimensional. Nevertheless, for any $p \in [1, +\infty)$ the risk measure $\rho$ is proper and, by Proposition 6.14, is convex and lower semicontinuous. Note that if $Z : \Omega \to \mathbb{R}$ is an $\mathcal{F}$-measurable function such that $\mathbb{E}[e^Z]$ is finite, then $Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ for any $p \geq 1$. Therefore, by formula (6.61) of Proposition 6.14, we have that if $\mathbb{E}[e^Z]$ is finite, then $\partial \rho(Z) = \{e^Z\}$ is a singleton. It could be mentioned that although $\rho(\cdot)$ is subdifferentiable at every $Z \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ where it is finite and has unique subgradient $e^Z$, it is discontinuous and nondifferentiable at $Z$ unless $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ is finite dimensional.

The above risk measure associated with the exponential disutility function is not positively homogeneous and does not satisfy condition (R3). Let us consider instead the following risk measure

$$\rho_e(Z) := \ln \mathbb{E}[e^Z],$$

(6.76)
defined on $Z = \mathcal{L}_p(\Omega, \mathcal{F}, P)$ for some $p \in [1, +\infty)$. Since $\ln(\cdot)$ is continuous on the positive half of the real line and $\mathbb{E}[e^Z] > 0$, it follows from the above that $\rho_e$ has the same domain as $\rho(Z) = \mathbb{E}[e^Z]$, and is lower semicontinuous and proper. It is also can be verified that $\rho_e$ is convex (see derivations of section 7.2.8 following equation (7.181)). Moreover, for any $a \in \mathbb{R}$,

$$\ln \mathbb{E}[e^{Z + a}] = \ln (e^a \mathbb{E}[e^Z]) = \ln a + \mathbb{E}[e^Z],$$

i.e., $\rho_e$ satisfies condition (R3).

Let us calculate the conjugate of $\rho_e$. We have that

$$\rho_e^*(\zeta) = \sup_{Z \in \mathcal{Z}} \{ \mathbb{E}[\zeta Z] - \ln \mathbb{E}[e^Z] \}.$$

(6.77)

Since $\rho_e$ satisfies conditions (R2) and (R3), it follows that $\text{dom}(\rho_e) \subset \mathcal{P}$, where $\mathcal{P}$ is the set of density functions (see (6.38)). By writing (first-order) optimality conditions for the optimization problem on the right hand side of (6.77), it is straightforward to verify that for $\zeta \in \mathcal{P}$ such that $\zeta(\omega) > 0$ for a.e. $\omega \in \Omega$, a point $\tilde{Z}$ is an optimal solution of that problem if $\tilde{Z} = \ln \zeta + a$ for some $a \in \mathbb{R}$. Substituting this into the right hand side of (6.77), and noting that the obtained expression does not depend on $a$, we obtain

$$\rho_e^*(\zeta) = \begin{cases} \mathbb{E}[\zeta \ln \zeta], & \text{if } \zeta \in \mathcal{P}, \\ +\infty, & \text{if } \zeta \not\in \mathcal{P}. \end{cases}$$

(6.78)

Note that $x \ln x$ tends to zero as $x \downarrow 0$. Therefore, we set $0 \ln 0 = 0$ in the above formula (6.78). Note also that $x \ln x$ is bounded for $x \in [0, 1]$. Therefore, $\text{dom}(\rho_e^*) = \mathcal{P}$ for any $p \in [1, +\infty)$.

Furthermore, we can apply the homogenization procedure to $\rho_e$ (see (6.57)). That is, consider the following risk measure

$$\tilde{\rho}_e(Z) := \inf_{\tau > 0} \tau \ln \mathbb{E}[e^{-\tau Z}].$$

(6.79)

Risk measure $\tilde{\rho}_e$ satisfies conditions (R1)–(R4), i.e., it is a coherent risk measure. Its conjugate $\tilde{\rho}_e^*$ is the indicator function of the set (see (6.58)):

$$\mathcal{A} := \{ \zeta \in \mathcal{Z}^* : \mathbb{E}[\zeta Z] \leq \ln \mathbb{E}[e^Z], \forall Z \in \mathcal{Z} \}.$$

(6.80)
6.3. Coherent Risk Measures

Note that since \(e^z\) is a convex function it follows by Jensen inequality that \(\mathbb{E}[\mathbb{E}[Z]] \leq \ln \mathbb{E}[e^Z]\). Consequently \(\zeta(\cdot) = 1\) is an element of the above set \(\mathfrak{A}\).

Example 6.17 (Mean-variance risk measure) Consider

\[
\rho(Z) := \mathbb{E}[Z] + c \mathbb{V}ar[Z],
\]

where \(c \geq 0\) is a given constant. It is natural to use here the space \(\mathcal{Z} := L_2(\Omega, \mathcal{F}, P)\) since for any \(Z \in L_2(\Omega, \mathcal{F}, P)\) the expectation \(\mathbb{E}[Z]\) and variance \(\mathbb{V}ar[Z]\) are well defined and finite. We have here that \(\mathcal{Z}^* = \mathcal{Z}\) (i.e., \(\mathcal{Z}\) is a Hilbert space) and for \(Z \in \mathcal{Z}\) its norm is given by \(\|Z\|_2 = \sqrt{\mathbb{E}[Z^2]}\). We also have that

\[
\|Z\|_2^2 = \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta, Z \rangle - \frac{1}{4}\|\zeta\|^2 \}.
\]

Indeed, it is not difficult to verify that the maximum on the right hand side of (6.82) is attained at \(\zeta = 2Z\).

We have that \(\mathbb{V}ar[Z] = \|Z - \mathbb{E}[Z]\|^2_2\), and since \(\|\cdot\|_2^2\) is a convex and continuous function on the Hilbert space \(\mathcal{Z}\), it follows that \(\rho(\cdot)\) is convex and continuous. Also because of (6.82), we can write

\[
\mathbb{V}ar[Z] = \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta, Z - \mathbb{E}[Z] \rangle - \frac{1}{4}\|\zeta\|^2 \}.
\]

Since

\[
\langle \zeta, Z - \mathbb{E}[Z] \rangle = \langle \zeta, Z \rangle - \mathbb{E}[\zeta] - \mathbb{E}[\zeta],
\]

we can rewrite the last expression as follows:

\[
\mathbb{V}ar[Z] = \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta - \mathbb{E}[\zeta], Z \rangle - \frac{1}{4}\|\zeta\|^2 \}
\]

\[
= \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta - \mathbb{E}[\zeta], Z \rangle - \frac{1}{4}\mathbb{V}ar[\zeta] - \frac{1}{4}(\mathbb{E}[\zeta])^2 \}.
\]

Since \(\zeta - \mathbb{E}[\zeta]\) and \(\mathbb{V}ar[\zeta]\) are invariant under transformations of \(\zeta\) to \(\zeta + a\), where \(a \in \mathbb{R}\), the above maximization can be restricted to such \(\zeta \in \mathcal{Z}\) that \(\mathbb{E}[\zeta] = 0\). Consequently

\[
\mathbb{V}ar[Z] = \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta, Z \rangle - \frac{1}{2}\mathbb{V}ar[\zeta] \}.
\]

Therefore the risk measure \(\rho\), defined in (6.81), can be expressed as

\[
\rho(Z) = \mathbb{E}[Z] + c \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta, Z \rangle - \frac{1}{2}\mathbb{V}ar[\zeta] \},
\]

and hence for \(c > 0\) (by making change of variables \(\zeta' = c\zeta + 1\)) as

\[
\rho(Z) = \sup_{\zeta \in \mathcal{Z}} \{ \langle \zeta, Z \rangle - \frac{1}{2c}\mathbb{V}ar[\zeta] \}.
\]
It follows that for any \( c > 0 \) the function \( \rho \) is convex, continuous and

\[
\rho^*(\zeta) = \begin{cases} \frac{\sqrt{\text{Var}[\zeta]}}{c}, & \text{if } \mathbb{E}[\zeta] = 1, \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.85}
\]

The function \( \rho \) satisfies the translation equivariance condition (R3), e.g., because the domain of its conjugate contains only \( \zeta \) such that \( \mathbb{E}[\zeta] = 1 \). However, for any \( c > 0 \) the function \( \rho \) is not positively homogeneous and it does not satisfy the monotonicity condition (R2), because the domain of \( \rho^* \) contains density functions which are not nonnegative.

Since \( \text{Var}[Z] = \langle Z, Z \rangle - (\mathbb{E}[Z])^2 \), it is straightforward to verify that \( \rho(\cdot) \) is (Fréchet) differentiable and

\[
\nabla \rho(Z) = 2cz - 2c\mathbb{E}[Z] + 1. \tag{6.86}
\]

Example 6.18 (Mean-deviation risk measures of order \( p \)) For \( Z := \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P}) \) and \( Z^* := \mathcal{L}_q(\Omega, \mathcal{F}, \mathbb{P}) \), with \( p \in [1, +\infty) \) and \( c \geq 0 \), consider

\[
\rho(Z) := \mathbb{E}|Z| + c\left(\mathbb{E}[|Z - \mathbb{E}[|Z|^p]|^p]\right)^{1/p}. \tag{6.87}
\]

We have that \( (\mathbb{E}[|Z|^p])^{1/p} = \|Z\|_p \), where \( \|\cdot\|_p \) denotes the norm of the space \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P}) \). The function \( \rho \) is convex continuous and positively homogeneous. Also

\[
\|Z\|_p = \sup_{\|\zeta\|_q \leq 1} \langle \zeta, Z \rangle, \tag{6.88}
\]

and hence

\[
(\mathbb{E}[|Z - \mathbb{E}[|Z|^p]|^p])^{1/p} = \sup_{\|\zeta\|_q \leq 1} \langle \zeta, Z - \mathbb{E}[Z] \rangle = \sup_{\|\zeta\|_q \leq 1} \langle \zeta - \mathbb{E}[\zeta], Z \rangle. \tag{6.89}
\]

It follows that representation (6.37) holds with the set \( \mathfrak{A} \) given by

\[
\mathfrak{A} = \{\zeta' \in Z^* : \zeta' = 1 + \zeta - \mathbb{E}[\zeta], \|\zeta\|_q \leq c\}. \tag{6.90}
\]

We obtain here that \( \rho \) satisfies conditions (R1), (R3) and (R4).

The monotonicity condition (R2) is more involved. Suppose that \( p = 1 \). Then \( q = +\infty \) and hence for any \( \zeta' \in \mathfrak{A} \) and a.e. \( \omega \in \Omega \) we have

\[
\zeta'(\omega) = 1 + \zeta(\omega) - \mathbb{E}[\zeta] \geq 1 - |\zeta(\omega)| - \mathbb{E}[|\zeta|] \geq 1 - 2c.
\]

It follows that if \( c \in [0, 1/2] \), then \( \zeta'(\omega) \geq 0 \) for a.e. \( \omega \in \Omega \), and hence condition (R2) follows. Conversely, take \( \zeta := c(-1_A + 1_{\Omega \setminus A}) \), for some \( A \in \mathcal{F} \), and \( \zeta' = 1 + \zeta - \mathbb{E}[\zeta] \). We have that \( \|\zeta\|_\infty = c \) and \( \zeta'(\omega) = 1 - 2c + 2cP(A) \) for all \( \omega \in A \). It follows that if \( c > 1/2 \), then \( \zeta'(\omega) < 0 \) for all \( \omega \in A \), provided that \( P(A) \) is small enough. We obtain that for \( c > 1/2 \) the monotonicity property (R2) does not hold if the following condition is satisfied:

For any \( \varepsilon > 0 \) there exists \( A \in \mathcal{F} \) such that \( \varepsilon > P(A) > 0. \tag{6.91} \)
That is, for $p = 1$ the mean-deviation measure $\rho$ satisfies (R2) if, and provided that condition (6.91) holds, only if $c \in [0, 1/2]$. (The above condition (6.91) holds, in particular, if the measure $P$ is nonatomic.)

Suppose now that $p > 1$. For a set $A \in \mathcal{F}$ and $\alpha > 0$ let us take $\zeta := -\alpha 1_A$ and $\zeta' = 1 + \zeta - E[\zeta]$. Then $\|\zeta\|_q = \alpha P(A)^{1/q}$ and $\zeta'(\omega) = 1 - \alpha + \alpha P(A)$ for all $\omega \in A$. It follows that if $p > 1$, then for any $c > 0$ the mean-deviation measure $\rho$ does not satisfy (R2) provided that condition (6.91) holds.

Since $\rho$ is convex continuous, it is subdifferentiable. By (6.43) and because of (6.90) and (6.83) we have here that $\partial \rho(Z)$ is formed by vectors $\zeta' = 1 + \zeta - E[\zeta]$ such that $\zeta \in \text{arg max}_{\|\zeta\|_q \leq c} \{\zeta, Z - E[Z]\}$. That is,

$$\partial \rho(Z) = \{\zeta' = 1 + c \zeta - c E[\zeta] : \zeta \in \mathcal{S}_Y\},$$

(6.92)

where $Y(\omega) = Z(\omega) - E[Z]$ and $\mathcal{S}_Y$ is the set of contact points of $Y$. If $p \in (1, +\infty)$, then the set $\mathcal{S}_Y$ is a singleton, i.e., there is unique contact point $\zeta^*_Y$, provided that $Y(\omega)$ is not zero for a.e. $\omega \in \Omega$. In that case $\rho(\cdot)$ is Hadamard differentiable at $Z$ and

$$\nabla \rho(Z) = 1 + c \zeta^*_Y - c E[\zeta^*_Y]$$

(6.93)

(an explicit form of the contact point $\zeta^*_Y$ is given in equation (7.238)). If $Y(\omega)$ is zero for a.e. $\omega \in \Omega$, i.e., $Z(\omega)$ is constant w.p.1, then $\mathcal{S}_Y = \{\zeta \in \mathcal{Z}^* : \|\zeta\|_q \leq 1\}$.

For $p = 1$ the set $\mathcal{S}_Y$ is described in equation (7.239). It follows that if $p = 1$, and hence $q = +\infty$, then the subdifferential $\partial \rho(Z)$ is a singleton iff $Z(\omega) \neq E[Z]$ for a.e. $\omega \in \Omega$, in which case

$$\nabla \rho(Z) = \left\{ \zeta : \begin{array}{ll} \zeta(\omega) = 1 + 2c(1 - \Pr(Z > E[Z])), & \text{if } Z(\omega) > E[Z], \\
\zeta(\omega) = 1 - 2c \Pr(Z > E[Z]), & \text{if } Z(\omega) < E[Z]. \end{array} \right.$$  

(6.94)

\[\]

**Example 6.19 (Mean-upper-semideviation of order $p$)** Let $Z := L_p(\Omega, \mathcal{F}, P)$ and for $c \geq 0$ consider

$$\rho(Z) := E[Z] + c \left( E \left[ \left| Z - E[Z] \right|^p \right] \right)^{1/p}. $$

(6.95)

For any $c \geq 0$ this function satisfies conditions (R1), (R3) and (R4), and similarly to the derivations of Example 6.18 it can be shown that representation (6.37) holds with the set $\mathfrak{A}$ given by

$$\mathfrak{A} = \{\zeta' \in \mathcal{Z}^* : \zeta' = 1 + \zeta - E[\zeta], \|\zeta\|_q \leq c, \zeta \geq 0\}. $$

(6.96)

Since $|E[\zeta]| \leq E|\zeta| \leq \|\zeta\|_q$ for any $\zeta \in L_q(\Omega, \mathcal{F}, P)$, we have that every element of the above set $\mathfrak{A}$ is nonnegative and has its expected value equal to 1. This means that the monotonicity condition (R2) holds, if and, provided that condition (6.91) holds, only if $c \in [0, 1]$. That is, $\rho$ is a coherent risk measure if $c \in [0, 1]$.

Since $\rho$ is convex continuous, it is subdifferentiable. Its subdifferential can be calculated in a way similar to the derivations of Example 6.18. That is, $\partial \rho(Z)$ is formed by vectors $\zeta' = 1 + \zeta - E[\zeta]$ such that

$$\zeta \in \text{arg max} \{\zeta, Y : \|\zeta\|_q \leq c, \zeta \geq 0\}.$$  

(6.97)

\[8\text{We denote } [a]^+_p := (\max\{0, a\})^p.\]
where \( Y := Z - \mathbb{E}[Z] \). Suppose that \( p \in (1, +\infty) \). Then the set of maximizers on the right hand side of (6.97) is not changed if \( Y \) is replaced by \( Y_+ \), where \( Y_+(\cdot) := \lfloor Y(\cdot) \rfloor_+ \). Consequently, if \( Z(\omega) \) is not constant for a.e. \( \omega \in \Omega \), and hence \( Y_+ \neq 0 \), then \( \partial \rho(Z) \) is a singleton and

\[
\nabla \rho(Z) = 1 + c \, \zeta^*_Y - c \, \mathbb{E}[\zeta^*_Y],
\]

(6.98)

where \( \zeta^*_Y \) is the contact point of \( Y_+ \) (note that the contact point of \( Y_+ \) is nonnegative since \( Y_+ \geq 0 \)).

Suppose now that \( p = 1 \) and hence \( q = +\infty \). Then the set on the right hand side of (6.97) is formed by \( \zeta(\cdot) \) such that: \( \zeta(\omega) = c \) if \( Y(\omega) > 0 \), \( \zeta(\omega) = 0 \), if \( Y(\omega) < 0 \), and \( \zeta(\omega) \in [0, c] \) if \( Y(\omega) = 0 \). It follows that \( \partial \rho(Z) \) is a singleton iff \( Z(\omega) \neq \mathbb{E}[Z] \) for a.e. \( \omega \in \Omega \), in which case

\[
\nabla \rho(Z) = \begin{cases} 
\zeta : 
\zeta(\omega) = 1 + c(1 - \mathbb{P}(Z > \mathbb{E}[Z])), & \text{if } Z(\omega) > \mathbb{E}[Z], \\
\zeta(\omega) = 1 - c \mathbb{P}(Z > \mathbb{E}[Z]), & \text{if } Z(\omega) < \mathbb{E}[Z].
\end{cases}
\]

(6.99)

It can be noted that by Lemma 6.1

\[
\mathbb{E}(\lfloor Z - \mathbb{E}[Z] \rfloor) = 2\mathbb{E}(\lfloor Z - \mathbb{E}[Z] \rfloor_+).
\]

(6.100)

Consequently, formula (6.99) can be derived directly from (6.94).

**Example 6.20 (Mean-upper-semivariance from a target)** Let \( Z := \mathcal{L}_2(\Omega, \mathcal{F}, P) \) and for a weight \( c \geq 0 \) and a target \( \tau \in \mathbb{R} \) consider

\[
\rho(Z) := \mathbb{E}[Z] + c \mathbb{E}\left[ (Z - \tau)_+^2 \right].
\]

(6.101)

This is a convex and continuous risk measure. We can now use (6.63) with \( g(z) := z + c[z - \tau]^2 \). Since

\[
g^*(\alpha) = \begin{cases} 
(\alpha - 1)^2/4c + \tau(\alpha - 1), & \text{if } \alpha \geq 1, \\
+\infty, & \text{otherwise},
\end{cases}
\]

we obtain that

\[
\rho(Z) = \sup_{\zeta \in Z, \zeta(\cdot) \geq 1} \left\{ \mathbb{E}[\zeta Z - \tau \mathbb{E}[\zeta - 1] - \frac{1}{4c} \mathbb{E}[\zeta - 1]^2] \right\}.
\]

(6.102)

Consequently, representation (6.36) holds with \( \mathfrak{A} = \{ \zeta \in Z : \zeta - 1 \geq 0 \} \) and

\[
\rho^*(\zeta) = \tau \mathbb{E}[\zeta - 1] + \frac{1}{4c} \mathbb{E}[\zeta - 1]^2, \quad \zeta \in \mathfrak{A}.
\]

If \( c > 0 \) then conditions (R3) and (R4) are not satisfied by this risk measure.

Since \( \rho \) is convex continuous, it is subdifferentiable. Moreover, by using (6.61) we obtain that its subdifferentials are singletons and hence \( \rho(\cdot) \) is differentiable at every \( Z \in Z \), and

\[
\nabla \rho(Z) = \begin{cases} 
\zeta : 
\zeta(\omega) = 1 + 2c(Z(\omega) - \tau), & \text{if } Z(\omega) \geq \tau, \\
\zeta(\omega) = 1, & \text{if } Z(\omega) < \tau.
\end{cases}
\]

(6.103)

The above formula can be also derived directly and it can be shown that \( \rho \) is differentiable in the sense of Fréchet.
6.3. Coherent Risk Measures

Example 6.21 (Mean-upper-semideviation of order $p$ from a target) Let $\mathcal{Z}$ be the space $L_p(\Omega, \mathcal{F}, P)$, and for $c \geq 0$ and $\tau \in \mathbb{R}$ consider

$$
\rho(Z) := \mathbb{E}[Z] + c \left( \mathbb{E} \left[ \left( Z - \tau \right)_+^p \right] \right)^{1/p}.
$$

(6.104)

For any $c \geq 0$ and $\tau$ this risk measure satisfies conditions (R1) and (R2), but not (R3) and (R4) if $c > 0$. We have

$$
\left( \mathbb{E} \left[ (Z - \tau)_+^p \right] \right)^{1/p} = \sup_{\|z\| \leq 1} \mathbb{E}(\zeta[Z - \tau]_+) = \sup_{\|\zeta\| \leq 1, \zeta(\cdot) \geq 0} \mathbb{E}(\zeta[Z - \tau]_+) = \sup_{\|\zeta\| \leq 1, \zeta(\cdot) \geq 0} \mathbb{E}([\zeta Z - \tau \zeta]_+).
$$

We obtain that representation (6.36) holds with

$$
\mathcal{A} = \{ \zeta \in \mathcal{Z}^* : \|\zeta\|_q \leq c, \zeta \geq 0 \}
$$

and $\rho^*(\zeta) = \tau \mathbb{E}[\zeta]$ for $\zeta \in \mathcal{A}$. 

6.3.3 Law Invariant Risk Measures and Stochastic Orders

As in the previous sections, unless stated otherwise we assume here that $\mathcal{Z} = L_p(\Omega, \mathcal{F}, P)$, $p \in [1, +\infty)$. We say that random outcomes $Z_1 \in \mathcal{Z}$ and $Z_2 \in \mathcal{Z}$ have the same distribution, with respect to the reference probability measure $P$, if $P(Z_1 \leq z) = P(Z_2 \leq z)$ for all $z \in \mathbb{R}$. We write this relation as $Z_1 \overset{D}{=} Z_2$. In all examples considered in section 6.3.2, the risk measures $\rho(Z)$ discussed there were dependent only on the distribution of $Z$. That is, each risk measure $\rho(Z)$, considered in section 6.3.2, could be formulated in terms of the cumulative distribution function (cdf) $H_Z(t) := P(Z \leq t)$ associated with $Z \in \mathcal{Z}$. We call such risk measures law invariant (or law based, or version independent).

Definition 6.22. A risk measure $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ is law invariant, with respect to the reference probability measure $P$, if for all $Z_1, Z_2 \in \mathcal{Z}$ we have the implication

$$
\{ Z_1 \overset{D}{=} Z_2 \} \Rightarrow \{ \rho(Z_1) = \rho(Z_2) \}.
$$

Suppose for the moment that the set $\Omega = \{ \omega_1, ..., \omega_K \}$ is finite with respective probabilities $p_1, ..., p_K$ such that any two partial sums of $p_k$ are different, i.e., $\sum_{k \in A} p_k = \sum_{k \in B} p_k$, for $A, B \subset \{ 1, ..., K \}$, iff $A = B$. Then $Z_1, Z_2 : \Omega \rightarrow \mathbb{R}$ have the same distribution only if $Z_1 = Z_2$. In that case any risk measure, defined on the space of random variables $Z : \Omega \rightarrow \mathbb{R}$, is law invariant. Therefore, for a meaningful discussion of law invariant risk measures it is natural to consider nonatomic probability spaces.

A particular example of law invariant coherent risk measure is the Average Value-at-Risk measure $AV@R_\alpha$. Clearly, a convex combination $\sum_{i=1}^{m} \mu_i AV@R_{\alpha_i}$, with $\alpha_i \in (0, 1]$, $\mu_i \geq 0$, $\sum_{i=1}^{m} \mu_i = 1$, of Average Value-at-Risk measures is also a law invariant coherent risk measure. Moreover, maximum of several law invariant coherent risk measures is again
Chapter 6. Risk Averse Optimization

a law invariant coherent risk measure. It turns out that any law invariant coherent risk measure can be constructed by the operations of taking convex combinations and maximum from the class of Average Value-at-Risk measures.

**Theorem 6.23 (Kusuoka).** Suppose that the probability space $(\Omega, \mathcal{F}, P)$ is nonatomic and let $\rho : \mathcal{Z} \to \mathbb{R}$ be a law invariant lower semicontinuous coherent risk measure. Then there exists a set $\mathcal{M}$ of probability measures on the interval $[0, 1]$ (equipped with its Borel sigma algebra) such that

$$\rho(Z) = \sup_{\mu \in \mathcal{M}} \int_0^1 \text{AV}^\alpha\text{R}_\alpha(Z) d\mu(\alpha), \quad \forall Z \in \mathcal{Z}. \quad (6.105)$$

In order to prove this we will need the following result.

**Lemma 6.24.** Let $(\Omega, \mathcal{F}, P)$ be a nonatomic probability space and $Z := \mathcal{L}_p(\Omega, \mathcal{F}, P)$. Then for $Z \in \mathcal{Z}$ and $\zeta \in \mathcal{Z}^*$ we have

$$\sup_{Y \sim Z} \int_{\Omega} \zeta(\omega) Y(\omega) dP(\omega) = \int_0^1 H_\zeta^{-1}(t) H_Z^{-1}(t) dt, \quad (6.106)$$

where $H_\zeta$ and $H_Z$ are the cdf’s of $\zeta$ and $Z$, respectively.

**Proof.** First we prove formula (6.106) for finite set $\Omega = \{\omega_1, \ldots, \omega_n\}$ with equal probabilities $P(\{\omega_i\}) = 1/n$, $i = 1, \ldots, n$. For a function $Y : \Omega \to \mathbb{R}$ denote $Y_i := Y(\omega_i)$, $i = 1, \ldots, n$. We have here that $Y \sim Z$ if $Y_i = Z_{\pi(i)}$ for some permutation $\pi$ of the set $\{1, \ldots, n\}$, and $\int_{\Omega} \zeta Y dP = n^{-1} \sum_{i=1}^n \zeta_i Y_i$. Moreover,

$$\sum_{i=1}^n \zeta_i Y_i \leq \sum_{i=1}^n \zeta_{[i]} Y_{[i]}, \quad (6.107)$$

where $\zeta_{[1]} \leq \cdots \leq \zeta_{[n]}$ are numbers $\zeta_1, \ldots, \zeta_n$ arranged in the increasing order, and $Y_{[1]} \leq \cdots Y_{[n]}$ are numbers $Y_1, \ldots, Y_n$ arranged in the increasing order. It follows that

$$\sup_{Y \sim Z} \int_{\Omega} \zeta(\omega) Y(\omega) dP(\omega) = n^{-1} \sum_{i=1}^n \zeta_{[i]} Z_{[i]}. \quad (6.108)$$

It remains to note that in the considered case the right hand side of (6.108) coincides with the right hand side of (6.106).

Now if the space $(\Omega, \mathcal{F}, P)$ is nonatomic, we can partition $\Omega$ into $n$ disjoint subsets, each of the same $P$-measure $1/n$, and it suffices to verify formula (6.106) for functions which are piecewise constant on such partitions. This reduces the problem to the case considered above. □

---

Footnote: Inequality (6.107) is called the Hardy–Littlewood–Polya inequality (compare with the proof of Theorem 4.50).
6.3. Coherent Risk Measures

**Proof.** [Proof of Theorem 6.23] By the dual representation (6.37) of Theorem 6.4, we have that for $Z \in \mathcal{Z}$,

$$\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \int_{\Omega} \zeta(\omega)Z(\omega)dP(\omega),$$

(6.109)

where $\mathfrak{A}$ is a set of probability density functions in $\mathcal{Z}^*$. Since $\rho$ is law invariant, we have that

$$\rho(Z) = \sup_{Y \in \mathcal{D}(Z)} \rho(Y),$$

where $\mathcal{D}(Z) := \{Y \in \mathcal{Z} : Y \sim Z\}$. Consequently

$$\rho(Z) = \sup_{Y \in \mathcal{D}(Z)} \left[ \sup_{\zeta \in \mathfrak{A}} \int_{\Omega} \zeta(\omega)Y(\omega)dP(\omega) \right] = \sup_{Y \in \mathcal{D}(Z)} \int_{0}^{1} \zeta(\omega)Y(\omega)dP(\omega).$$

(6.110)

Moreover, by Lemma 6.24 we have

$$\sup_{Y \in \mathcal{D}(Z)} \int_{\Omega} \zeta(\omega)Y(\omega)dP(\omega) = \int_{0}^{1} H_{\zeta}^{-1}(t)H_{Z}^{-1}(t)dt,$$

(6.111)

where $H_{\zeta}$ and $H_{Z}$ are the cdf’s of $\zeta(\omega)$ and $Z(\omega)$, respectively. Recalling that $H_{Z}^{-1}(t) = V@R_{1-t}(Z)$, we can write (6.111) in the form

$$\sup_{Y \in \mathcal{D}(Z)} \int_{\Omega} \zeta(\omega)Y(\omega)dP(\omega) = \int_{0}^{1} H_{\zeta}^{-1}(t)V@R_{1-t}(Z)dt,$$

(6.112)

which together with (6.110) imply that

$$\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \int_{0}^{1} H_{\zeta}^{-1}(t)V@R_{1-t}(Z)dt.$$  

(6.113)

For $\zeta \in \mathfrak{A}$, the function $H_{\zeta}^{-1}(t)$ is monotonically nondecreasing on $[0,1]$ and can be represented in the form

$$H_{\zeta}^{-1}(t) = \int_{1-t}^{1} \alpha^{-1}d\mu(\alpha)$$

(6.114)

for some measure $\mu$ on $[0,1]$. Moreover, for $\zeta \in \mathfrak{A}$ we have that $\int \zeta dP = 1$, and hence

$$\int_{0}^{1} H_{\zeta}^{-1}(t)dt = \int \zeta dP = 1,$$

and therefore

$$1 = \int_{0}^{1} \int_{1-t}^{1} \alpha^{-1}d\mu(\alpha)dt = \int_{0}^{1} \int_{1-\alpha}^{1} \alpha^{-1}dtd\mu(\alpha) = \int_{0}^{1} d\mu(\alpha).$$

Consequently $\mu$ is a probability measure on $[0,1]$. Also (see Theorem 6.2) we have

$$AV@R_{\alpha}(Z) = \frac{1}{\alpha} \int_{1-\alpha}^{1} V@R_{1-t}(Z)dt,$$
and hence
\[
\int_0^1 \text{AV@} R_t(Z)d\mu(\alpha) = \int_0^1 \int_{1-\alpha}^1 \text{AV@} R_{1-t}(Z)d\mu(\alpha) dt \\
= \int_0^1 V@ R_{1-t}(Z) \left( \int_{1-\alpha}^1 d\mu(\alpha) \right) dt \\
= \int_0^1 V@ R_{1-t}(Z) H^{-1}_\zeta(t) dt.
\]

By (6.113) this completes the proof, with the correspondence between \(\zeta \in \mathcal{A}\) and \(\mu \in \mathcal{M}\) given by equation (6.114).

\[\square\]

**Example 6.25** Consider \(\rho := \text{AV@} R_t\), risk measure for some \(\gamma \in (0, 1)\). Assume that the corresponding probability space is \(\Omega = [0,1]\) equipped with its Borel sigma algebra and uniform probability measure \(P\). We have here (see (6.70))
\[
\mathcal{A} = \left\{ \zeta : 0 \leq \zeta(\omega) \leq \gamma^{-1}, \ \omega \in [0,1], \ \int_0^1 \zeta(\omega)d\omega = 1 \right\}.
\]

Consequently the family of cumulative distribution functions \(H^{-1}_\zeta, \zeta \in \mathcal{A}\), is formed by left side continuous monotonically nondecreasing on \([0,1]\) functions with \(\int_0^1 H^{-1}_\zeta(t) dt = 1\) and range values \(0 \leq H^{-1}_\zeta(t) \leq \gamma^{-1}\), \(t \in [0,1]\). Since \(V@ R_{1-t}(Z)\) is monotonically nondecreasing in \(t\) function, it follows that the maximum in the right hand side of (6.113) is attained at \(\zeta \in \mathcal{A}\) such that \(H^{-1}_\zeta(t) = 0\) for \(t \in [0,1-\gamma]\), and \(H^{-1}_\zeta(t) = \gamma^{-1}\) for \(t \in (1-\gamma, 1]\). The corresponding measure \(\mu\), defined by equation (6.114), is given by function \(\mu(\alpha) = 1\) for \(\alpha \in [0, \gamma]\), and \(\mu(\alpha) = 1\) for \(\alpha \in (\gamma, 1]\), i.e., \(\mu\) is the measure of mass 1 at the point \(\gamma\). By the above proof of Theorem 6.23, this \(\mu\) is the maximizer of the right hand side of (6.105). It follows that the representation (6.105) recovers the measure \(\text{AV@} R_{1-}\), as it should be.

For law invariant risk measures it makes sense to discuss their monotonicity properties with respect to various stochastic orders defined for (real valued) random variables. Many stochastic orders can be characterized by a class \(\mathcal{U}\) of functions \(u : \mathbb{R} \to \mathbb{R}\) as follows. For (real valued) random variables \(Z_1\) and \(Z_2\) it is said that \(Z_2\) dominates \(Z_1\), denoted \(Z_2 \succeq_u Z_1\), if \(E[u(Z_2)] \geq E[u(Z_1)]\) for all \(u \in \mathcal{U}\) for which the corresponding expectations do exist. This stochastic order is called the **integral stochastic order** with generator \(\mathcal{U}\). In particular, the **usual stochastic order**, written \(Z_2 \succeq_{(1)} Z_1\), corresponds to the generator \(\mathcal{U}\) formed by all non-decreasing functions \(u : \mathbb{R} \to \mathbb{R}\). Equivalently, \(Z_2 \succeq_{(1)} Z_1\) iff \(H_{Z_2}(t) \leq H_{Z_1}(t)\) for all \(t \in \mathbb{R}\). The relation \(\succeq_{(1)}\) is also frequently called the **first order stochastic dominance** (see Definition 4.3). We say that the integral stochastic order is **increasing** if all functions in the set \(\mathcal{U}\) are non-decreasing. The usual stochastic order is an example of increasing integral stochastic order.

**Definition 6.26.** A law invariant risk measure \(\rho : Z \to \mathbb{R}\) is consistent (monotone) with the integral stochastic order \(\succeq_u\), if for all \(Z_1, Z_2 \in Z\) we have the implication
\[
\{ Z_2 \succeq_u Z_1 \} \Rightarrow \{ \rho(Z_2) \geq \rho(Z_1) \}.
\]

For an increasing integral stochastic order we have that if \(Z_2(\omega) \geq Z_1(\omega)\) for a.e. \(\omega \in \Omega\), then \(u(Z_2(\omega)) \geq u(Z_1(\omega))\) for any \(u \in \mathcal{U}\) and a.e. \(\omega \in \Omega\), and hence \(\mathbb{E}[u(Z_2)] \geq \mathbb{E}[u(Z_1)]\).
6.3. Coherent Risk Measures

\[ \mathbb{E}[u(Z_1)] \]. That is, if \( Z_2 \succeq Z_1 \) in the almost sure sense, then \( Z_2 \succeq \rho \, Z_1 \). It follows that if \( \rho \) is law invariant and consistent with respect to an increasing integral stochastic order, then it satisfies the monotonicity condition (R2). In other words if \( \rho \) does not satisfy condition (R2), then it cannot be consistent with any increasing integral stochastic order. In particular, for \( c > 1 \) the mean-semideviation risk measure, defined in Example 6.19, is not consistent with any increasing integral stochastic order, provided that condition (6.91) holds.

A general way of proving consistency of law invariant risk measures with stochastic orders can be obtained via the following construction. For a given pair of random variables \( Z_1 \) and \( Z_2 \) in \( \mathcal{Z} \), consider another pair of random variables, \( \hat{Z}_1 \) and \( \hat{Z}_2 \), which have distributions identical to the original pair, i.e., \( \hat{Z}_1 \sim Z_1 \) and \( \hat{Z}_2 \sim Z_2 \). The construction is such that the postulated consistency result becomes evident. For this method to be applicable, it is convenient to assume that the probability space \( (\Omega, \mathcal{F}, P) \) is nonatomic. Then there exists a measurable function \( U : \Omega \to \mathbb{R} \) (uniform random variable) such that \( P(U \leq t) = t \) for all \( t \in [0, 1] \).

**Theorem 6.27.** Suppose that the probability space \( (\Omega, \mathcal{F}, P) \) is nonatomic. Then the following holds: if a risk measure \( \rho : \mathcal{Z} \to \mathbb{R} \) is law invariant, then it is consistent with the usual stochastic order iff it satisfies the monotonicity condition (R2).

**Proof.** By the discussion preceding the theorem, it is sufficient to prove that (R2) implies consistency with the usual stochastic order.

For a uniform random variable \( U(\omega) \) consider the random variables \( \hat{Z}_1 := H_{Z_1}^{-1}(U) \) and \( \hat{Z}_2 := H_{Z_2}^{-1}(U) \). We obtain that if \( Z_2 \succeq (1) Z_1 \), then \( \hat{Z}_2(\omega) \geq \hat{Z}_1(\omega) \) for all \( \omega \in \Omega \), and hence by virtue of (R2), \( \rho(\hat{Z}_2) \geq \rho(\hat{Z}_1) \). By construction, \( \hat{Z}_1 \sim Z_1 \) and \( \hat{Z}_2 \sim Z_2 \). Since the risk measure is law invariant, we conclude that \( \rho(Z_2) \geq \rho(Z_1) \). Consequently, the risk measure \( \rho \) is consistent with the usual stochastic order. \( \square \)

It is said that \( Z_1 \) is smaller than \( Z_2 \) in the **increasing convex order**, written \( Z_1 \preceq_{\text{icx}} Z_2 \), if \( \mathbb{E}[u(Z_1)] \leq \mathbb{E}[u(Z_2)] \) for all increasing convex functions \( u : \mathbb{R} \to \mathbb{R} \) such that the expectations exist. Clearly this is an integral stochastic order with the corresponding generator given by the set of increasing convex functions. It is equivalent to the **second order stochastic dominance** relation for the negative variables: \( -Z_1 \preceq (2) -Z_2 \) (recall that we are dealing here with minimization rather than maximization problems). Indeed, applying Definition 4.4 to \( -Z_1 \) and \( -Z_2 \) for \( k = 2 \) and using identity (4.7) we see that

\[
\mathbb{E}\{ |Z_1 - \eta|_+ \} \leq \mathbb{E}\{ |Z_2 - \eta|_+ \}, \quad \forall \eta \in \mathbb{R}.
\]  

(6.115)

Since any convex nondecreasing function \( u(z) \) can be arbitrarily close approximated by a positive combination of functions \( u_k(z) = \beta_k + [z - \eta_k]^+ \), inequality (6.115) implies that \( \mathbb{E}[u(Z_1)] \leq \mathbb{E}[u(Z_2)] \), as claimed (compare with the statement (4.8)).

**Theorem 6.28.** Suppose that the probability space \( (\Omega, \mathcal{F}, P) \) is nonatomic. Then any law invariant lower semicontinuous coherent risk measure \( \rho : \mathcal{Z} \to \mathbb{R} \) is consistent with the increasing convex order.

**Proof.** By using definition (6.22) of \( \mathcal{A}/\mathcal{D}_a \), and the property that \( Z_1 \preceq_{\text{icx}} Z_2 \) iff condition
(6.115) holds, it is straightforward to verify that $AV@R_\alpha$ is consistent with the increasing convex order. Now by using the representation (6.105) of Theorem 6.23 and noting that the operations of taking convex combinations and maximum preserve consistency with the increasing convex order, we can complete the proof.

Remark 17. For convex risk measures (without the positive homogeneity property) Theorem 6.28 in the space $L_1(\Omega, \mathcal{F}, P)$ can be derived from Theorem 4.52, which for the increasing convex order can be written as follows:

$$\{Z \in L_1(\Omega, \mathcal{F}, P) : Z \preceq_{icx} Y\} = \text{cl conv} \{Z \in L_1(\Omega, \mathcal{F}, P) : Z \preceq_{(1)} Y\}.$$  

If $Z$ is an element of the set on the left hand side, then there exists a sequence of random variables $Z^k \to Z$, which are convex combinations of some elements of the set on the right hand side, that is,

$$Z^k = \sum_{j=1}^{N_k} \alpha_j^k Z_j^k, \quad \sum_{j=1}^{N_k} \alpha_j^k = 1, \quad \alpha_j^k \geq 0, \quad Z_j^k \preceq_{(1)} Y.$$  

By convexity of $\rho$ and by Theorem 6.27, we obtain

$$\rho(Z^k) \leq \sum_{j=1}^{N_k} \alpha_j^k \rho(Z_j^k) \leq \sum_{j=1}^{N_k} \alpha_j^k \rho(Y) = \rho(Y).$$  

Passing to the limit with $k \to \infty$ and using lower semicontinuity of $\rho$, we obtain $\rho(Z) \leq \rho(Y)$, as required.

If $p > 1$ the domain of $\rho$ can be extended to $L_p(\Omega, \mathcal{F}, P)$, while preserving its lower semicontinuity (cf., Filipović and Svindland [66]).

Remark 18. For some measures of risk, in particular, for the mean-semideviation measures, defined in Example 6.19, and for the Average Value-at-Risk, defined in Example 6.15, consistency with the increasing convex order can be proved without the assumption that the probability space $(\Omega, \mathcal{F}, P)$ is nonatomic by using the following construction. Let $(\Omega, \mathcal{F}, P)$ be a nonatomic probability space, for example we can take $\Omega$ as the interval $[0, 1]$ equipped with its Borel sigma algebra and uniform probability measure $P$. Then for any finite set of probabilities $p_k > 0$, $k = 1, \ldots, K$, $\sum_{k=1}^{K} p_k = 1$, we can construct a partition of the set $\Omega = \bigcup_{k=1}^{K} A_k$ such that $P(A_k) = p_k$, $k = 1, \ldots, K$. Consider the linear subspace of the respective space $L_p(\Omega, \mathcal{F}, P)$ formed by piecewise constant on the sets $A_k$ functions $Z : \Omega \to \mathbb{R}$. We can identify this subspace with the space of random variables defined on a finite probability space of cardinality $K$ with the respective probabilities $p_k$, $k = 1, \ldots, K$.

By the above theorem the mean-upper-semideviation risk measure (of order $p$) defined on $(\Omega, \mathcal{F}, P)$ is consistent with the increasing convex order. This property is preserved by restricting it to the constructed subspace. This shows that the mean-upper-semideviation risk measures are consistent with the increasing convex order on any finite probability space. This can be extended to the general probability spaces by continuity arguments.
6.3. Coherent Risk Measures

**Corollary 6.29.** Suppose that the probability space \((\Omega, \mathcal{F}, P)\) is nonatomic. Let \(\rho : \mathcal{Z} \to \mathbb{R}\) be a law invariant lower semicontinuous coherent risk measure and \(\mathcal{G}\) be a sigma subalgebra of the sigma algebra \(\mathcal{F}\). Then

\[
\rho \left( E[Z|\mathcal{G}] \right) \leq \rho(Z), \quad \forall Z \in \mathcal{Z},
\]

and

\[
E[Z] \leq \rho(Z), \quad \forall Z \in \mathcal{Z}.
\]

**Proof.** Consider \(Z \in \mathcal{Z}\) and \(Z' := E[Z|\mathcal{G}]\). For every convex function \(u : \mathbb{R} \to \mathbb{R}\) we have

\[
E[u(Z')] = E[u(E[Z|\mathcal{G}])] \leq \mathbb{E}(E[u(Z)|\mathcal{G}]) = \mathbb{E}[u(Z)],
\]

where the inequality is implied by Jensen’s inequality. This shows that \(Z' \preceq_{icx} Z\), and hence (6.116) follows by Theorem 6.28.

In particular, for \(\mathcal{G} := \{\Omega, \emptyset\}\), it follows by (6.116) that \(\rho(Z) \geq \rho(E[Z])\), and since \(\rho(E[Z]) = E[Z]\) this completes the proof. □

An intuitive interpretation of property (6.116) is that if we reduce variability of a random variable \(Z\) by employing conditional averaging \(Z' = E[Z|\mathcal{G}]\), then the risk measure \(\rho(Z')\) becomes smaller, while \(E[Z'] = E[Z]\).

6.3.4 Relation to Ambiguous Chance Constraints

Owing to the dual representation (6.36), measures of risk are related to robust and ambiguous models. Consider a chance constraint of the form

\[
P\{C(x, \omega) \leq 0\} \geq 1 - \alpha.
\]

(6.118)

Here \(P\) is a probability measure on a measurable space \((\Omega, \mathcal{F})\) and \(C : \mathbb{R}^n \times \Omega \to \mathbb{R}\) is a random function. It is assumed in this formulation of chance constraint that the probability measure (distribution), with respect to which the corresponding probabilities are calculated, is known. Suppose now that the underlying probability distribution is not known exactly, but rather is assumed to belong to a specified family of probability distributions. Problems involving such constrains are called ambiguous chance constrained problems. For a specified uncertainty set \(\mathfrak{A}\) of probability measures on \((\Omega, \mathcal{F})\), the corresponding ambiguous chance constraint defines a feasible set \(X \subset \mathbb{R}^n\) which can be written as

\[
X := \{x : \mu\{C(x, \omega) \leq 0\} \geq 1 - \alpha, \forall \mu \in \mathfrak{A}\}. \tag{6.119}
\]

The set \(X\) can be written in the following equivalent form

\[
X = \left\{ x \in \mathbb{R}^n : \sup_{\mu \in \mathfrak{A}} \mathbb{E}_\mu [1_{A_x}] \leq \alpha \right\}, \tag{6.120}
\]

where \(A_x := \{\omega \in \Omega : C(x, \omega) > 0\}\). Recall that, by the duality representation (6.37), with the set \(\mathfrak{A}\) is associated a coherent risk measure \(\rho\), and hence (6.120) can be written as

\[
X = \{x \in \mathbb{R}^n : \rho(1_{A_x}) \leq \alpha\}. \tag{6.121}
\]
We discuss now constraints of the form (6.121) where the respective risk measure is defined in a direct way. As before, we use spaces \( Z = \mathcal{L}_p(\Omega, \mathcal{F}, P) \), where \( P \) is viewed as a reference probability measure.

It is not difficult to see that if \( \rho \) is a law invariant risk measure, then for any \( A \in \mathcal{F} \) the quantity \( \rho(1_A) \) depends only on \( P(A) \). Indeed, if \( Z := 1_A \) for some \( A \in \mathcal{F} \), then its cdf \( H_Z(z) := P(Z \leq z) \) is

\[
H_Z(z) = \begin{cases} 
0, & \text{if } z < 0, \\
1 - P(A), & \text{if } 0 \leq z < 1, \\
1, & \text{if } 1 \leq z,
\end{cases}
\]

which clearly depends only on \( P(A) \).

- With every law invariant real valued risk measure \( \rho : Z \to \mathbb{R} \) we associate function \( \varphi_\rho \) defined as \( \varphi_\rho(t) := \rho(1_A) \), where \( A \in \mathcal{F} \) is any event such that \( P(A) = t \), and \( t \in T := \{ P(A) : A \in \mathcal{F} \} \).

The function \( \varphi_\rho \) is well defined because for law invariant risk measure \( \rho \) the quantity \( \rho(1_A) \) depends only on the probability \( P(A) \) and hence \( \rho(1_A) \) is the same for any \( A \in \mathcal{F} \) such that \( P(A) = t \) for a given \( t \in T \). Clearly \( T \) is a subset of the interval \([0,1]\), and \( 0 \in T \) (since \( \emptyset \in \mathcal{F} \) and \( 1 \in T \) (since \( \Omega \in \mathcal{F} \)). If \( P \) is a nonatomic measure, then for any \( A \in \mathcal{F} \) the set \( \{ P(B) : B \subset A, B \in \mathcal{F} \} \) coincides with the interval \([0, P(A)]\). In particular, if \( P \) is nonatomic, then the set \( T = \{ P(A) : A \in \mathcal{F} \} \), on which \( \varphi_\rho \) is defined, coincides with the interval \([0,1]\).

**Proposition 6.30.** Let \( \rho : Z \to \mathbb{R} \) be a (real valued) law invariant coherent risk measure. Suppose that the reference probability measure \( P \) is nonatomic. Then \( \varphi_\rho(\cdot) \) is a continuous nondecreasing function defined on the interval \([0,1]\) such that \( \varphi_\rho(0) = 0 \) and \( \varphi_\rho(1) = 1 \), and \( \varphi_\rho(t) \geq t \) for all \( t \in [0,1] \).

**Proof.** Since the coherent risk measure \( \rho \) is real valued, it is continuous. Because \( \rho \) is continuous and positively homogeneous, \( \rho(0) = 0 \) and hence \( \varphi_\rho(0) = 0 \). Also by (R3), we have that \( \rho(1_A) = 1 \) and hence \( \varphi_\rho(1) = 1 \). By Corollary 6.29 we have that \( \rho(1_A) \geq P(A) \) for any \( A \in \mathcal{F} \), and hence \( \varphi_\rho(t) \geq t \) for all \( t \in [0,1] \).

Let \( t_k \in [0,1] \) be a monotonically increasing sequence tending to \( t^* \). Since \( P \) is a nonatomic, there exists a sequence \( A_1 \subset A_2 \subset \ldots \) of \( \mathcal{F} \)-measurable sets such that \( P(A_k) = t_k \) for all \( k \in \mathbb{N} \). It follows that the set \( A := \bigcup_{k=1}^\infty A_k \) is \( \mathcal{F} \)-measurable and \( P(A) = t^* \). Since \( 1_A \) converges (in the norm topology of \( Z \)) to \( 1_A \), it follows by continuity of \( \rho \) that \( \rho(1_A) \) tends to \( \rho(1_A) \), and hence \( \varphi_\rho(t_k) \) tends to \( \varphi_\rho(t^*) \). In a similar way we have that \( \varphi_\rho(t_k) \to \varphi_\rho(t^*) \) for a monotonically decreasing sequence \( t_k \) tending to \( t^* \). This shows that \( \varphi_\rho \) is continuous.

For any \( 0 \leq t_1 < t_2 \leq 1 \) there exist sets \( A, B \in \mathcal{F} \) such that \( B \subset A \) and \( P(B) = t_1 \), \( P(A) = t_2 \). Since \( 1_A \geq 1_B \), it follows by monotonicity of \( \rho \) that \( \rho(1_A) \geq \rho(1_B) \). This implies that \( \varphi_\rho(t_2) \geq \varphi_\rho(t_1) \), i.e., \( \varphi_\rho \) is nondecreasing.

Now consider again the set \( X \) of the form (6.119). Assuming conditions of Proposition 6.30, we obtain that this set \( X \) can be written in the following equivalent form

\[
X = \{ x : P(C(x, \omega) \leq 0) \geq 1 - \alpha^* \},
\]

(6.122)
6.3. Coherent Risk Measures 289

where \( \alpha^* := \varphi^{-1}_p(\alpha) \). That is, \( X \) can be defined by a chance constraint with respect to the reference distribution \( P \) and with the respective significance level \( \alpha^* \). Since \( \varphi_p(t) \geq t \), for any \( t \in [0, 1] \), it follows that \( \alpha^* \leq \alpha \). Let us consider some examples.

Consider Average Value-at-Risk measure \( \rho := AV\oplus R_\gamma \), \( \gamma \in (0, 1] \). By direct calculations it is straightforward to verify that for any \( A \in \mathcal{F} \),

\[
AV\oplus R_\gamma (1_A) = \begin{cases} 
\gamma^{-1} P(A), & \text{if } P(A) \leq \gamma, \\
1, & \text{if } P(A) > \gamma.
\end{cases}
\]

Consequently the corresponding function \( \varphi_p(t) = \gamma^{-1} t \) for \( t \in [0, \gamma] \), and \( \varphi_p(t) = 1 \) for \( t \in [\gamma, 1] \). Now let \( p \) be a convex combination of Average Value-at-Risk measures, i.e., \( p := \sum_{i=1}^m \lambda_i p_i \), with \( p_i := AV\oplus R_{\gamma_i} \) and positive weights \( \lambda_i \) summing up to one. By the definition of the function \( \varphi_p \) we have then that \( \varphi_p = \sum_{i=1}^m \lambda_i \varphi_{p_i} \). It follows that \( \varphi_p : [0, 1] \to [0, 1] \) is a piecewise linear nondecreasing concave function with \( \varphi_p(0) = 0 \) and \( \varphi_p(1) = 1 \). More generally, let \( \lambda \) be a probability measure on \( (0, 1] \) and \( p := \int_{0}^{1} AV\oplus R_{\gamma} d\lambda(\gamma) \). In that case the corresponding function \( \varphi_p \) becomes a nondecreasing concave function with \( \varphi_p(0) = 0 \) and \( \varphi_p(1) = 1 \). We also can consider measures \( p \) given by the maximum of such integral functions over some set \( \mathcal{M} \) of probability measures on \( (0, 1] \). In that case the respective function \( \varphi_p \) becomes the maximum of the corresponding nondecreasing concave functions. By Theorem 6.23 this actually gives the most general form of the function \( \varphi_p \).

For instance, let \( Z := L^1(\Omega, \mathcal{F}, P) \) and \( \rho(Z) := (1 - \beta) \mathbb{E}[Z] + \beta AV\oplus R_\gamma(Z) \), where \( \beta, \gamma \in (0, 1) \) and the expectations are taken with respect to the reference distribution \( P \). This risk measure was discussed in example 6.15. Then

\[
\varphi_p(t) = \begin{cases} 
(1 - \beta + \gamma^{-1} \beta) t, & \text{if } t \in [0, \gamma], \\
\beta + (1 - \beta) t, & \text{if } t \in (\gamma, 1].
\end{cases}
\]  

(6.123)

It follows that for this risk measure and for \( \alpha \leq \beta + (1 - \beta) \gamma \),

\[
\alpha^* = \frac{\alpha}{1 + \beta (\gamma^{-1} - 1)}.
\]

(6.124)

In particular, for \( \beta = 1 \), i.e., for \( \rho = AV\oplus R_\gamma \), we have that \( \alpha^* = \gamma \alpha \).

As another example consider the mean-upper-semideviation risk measure of order \( p \). That is, \( Z := L^p(\Omega, \mathcal{F}, P) \) and

\[
\rho(Z) := \mathbb{E}[Z] + c \left( \mathbb{E} \left[ \left( Z - \mathbb{E}[Z] \right)^p \right] \right)^{1/p}
\]

(see example 6.19). We have here that \( \rho(1_A) = P(A) + c [P(A)(1 - P(A))]^{1/p} \), and hence

\[
\varphi_p(t) = t + ct^{1/p}(1 - t), \quad t \in [0, 1].
\]

(6.125)

In particular, for \( p = 1 \) we have that \( \varphi_p(t) = (1 + c) t - ct^2 \), and hence

\[
\alpha^* = 1 + c - \sqrt{(1 + c)^2 - 4ac}.
\]

(6.126)

Note that for \( c > 1 \) the above function \( \varphi_p(t) \) is not monotonically nondecreasing on the interval \( [0, 1] \). This should be not surprising since for \( c > 1 \) and nonatomic \( P \), the corresponding mean-upper-semideviation risk measure is not monotone.
6.4 Optimization of Risk Measures

As before, we use spaces $\mathcal{Z} = \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and $\mathcal{Z}^* = \mathcal{L}_q(\Omega, \mathcal{F}, P)$. Consider the composite function $\phi(\cdot) := \rho(F(\cdot))$, also denoted $\phi = \rho \circ F$, associated with a mapping $F : \mathbb{R}^n \to \mathcal{Z}$ and a risk measure $\rho : \mathcal{Z} \to \mathbb{R}$. We already studied properties of such composite functions in section 6.3.1. Again we write $f(x, \omega)$ or $f_n(x)$ for $[F(x)](\omega)$, and view $f(x, \omega)$ as a random function defined on the measurable space $(\Omega, \mathcal{F})$. Note that $F(x)$ is an element of space $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ and hence $f(x, \cdot)$ is $\mathcal{F}$-measurable and finite valued. If, moreover, $f(\cdot, \omega)$ is continuous for a.e. $\omega \in \Omega$, then $f(x, \omega)$ is a Carathéodory function, and hence is random lower semicontinuous.

In this section we discuss optimization problems of the form

$$\min_{x \in X} \{ \phi(x) := \rho(F(x)) \}.$$  \hfill (6.127)

Unless stated otherwise, we assume that the feasible set $X$ is a nonempty convex closed subset of $\mathbb{R}^n$. Of course, if we use $\rho(\cdot) := \mathbb{E}[\cdot]$, then problem (6.127) becomes a standard stochastic problem of optimizing (minimizing) the expected value of the random function $f(x, \omega)$. In that case we can view the corresponding optimization problem as risk neutral. However, a particular realization of $f(x, \omega)$ could be quite different from its expectation $\mathbb{E}[f(x, \omega)]$. This motivates an introduction, in the corresponding optimization procedure, of some type of risk control. In the analysis of Portfolio Selection (see section 1.4) we discussed an approach of using variance as a measure of risk. There is, however, a problem with such approach since the corresponding mean-variance risk measure is not monotone (see Example 6.17). We shall discuss this later.

Unless stated otherwise we assume that the risk measure $\rho$ is proper, lower semicontinuous and satisfies conditions (R1)–(R2). By Theorem 6.4 we can use representation (6.36) to write problem (6.127) in the form

$$\min_{x \in X} \sup_{\zeta \in \mathfrak{A}} \Phi(x, \zeta),$$  \hfill (6.128)

where $\mathfrak{A} := \text{dom}(\rho^*)$ and the function $\Phi : \mathbb{R}^n \times \mathcal{Z}^* \to \mathbb{R}$ is defined by

$$\Phi(x, \zeta) := \int_{\Omega} f(x, \omega) \zeta(\omega) dP(\omega) - \rho^*(\zeta).$$  \hfill (6.129)

If, moreover, $\rho$ is positively homogeneous, then $\rho^*$ is the indicator function of the set $\mathfrak{A}$ and hence $\rho^*(\cdot)$ is identically zero on $\mathfrak{A}$. That is, if $\rho$ is a proper lower semicontinuous coherent risk measure, then problem (6.127) can be written as the minimax problem

$$\min_{x \in X} \sup_{\zeta \in \mathfrak{A}} \mathbb{E}_\zeta[f(x, \omega)],$$  \hfill (6.130)

where

$$\mathbb{E}_\zeta[f(x, \omega)] := \int_{\Omega} f(x, \omega) \zeta(\omega) dP(\omega)$$

denotes the expectation with respect to $\zeta dP$. Note that, by the definition, $F(x) \in \mathcal{Z}$ and $\zeta \in \mathcal{Z}^*$, and hence

$$\mathbb{E}_\zeta[f(x, \omega)] = (F(x), \zeta)$$
is finite valued.

Suppose that the mapping $F : \mathbb{R}^n \to Z$ is convex, i.e., for a.e. $\omega \in \Omega$ the function $f(\cdot, \omega)$ is convex. This implies that for every $\zeta \succeq 0$ the function $\Phi(\cdot, \zeta)$ is convex and if, moreover, $\zeta \in \mathfrak{A}$, then $\Phi(\cdot, \zeta)$ is real valued and hence continuous. We also have that $\langle F(x), \zeta \rangle$ is linear and $\rho^*(\zeta)$ is convex in $\zeta \in Z^*$, and hence for every $x \in X$ the function $\Phi(x, \cdot)$ is concave. Therefore, under various regularity conditions, there is no duality gap between problem (6.127) and its dual

$$\max_{\zeta \in \mathfrak{A}} \inf_{x \in X} \left\{ \int_{\Omega} f(x, \omega)\zeta(\omega)dP(\omega) - \rho^*(\zeta) \right\},$$

(6.131)

which is obtained by interchanging the “min” and “max” operators in (6.128) (recall that the set $X$ is assumed to be nonempty closed and convex). In particular, if there exists a saddle point $(\bar{x}, \bar{\zeta}) \in X \times \mathfrak{A}$ of the minimax problem (6.128), then there is no duality gap between problems (6.128) and (6.131), and $\bar{x}$ and $\bar{\zeta}$ are optimal solutions of (6.128) and (6.131), respectively.

**Proposition 6.31.** Suppose that mapping $F : \mathbb{R}^n \to Z$ is convex and risk measure $\rho : Z \to \mathbb{R}$ is proper, lower semicontinuous and satisfies conditions (R1)–(R2). Then $(\bar{x}, \bar{\zeta}) \in X \times \mathfrak{A}$ is a saddle point of $\Phi(x, \zeta)$ if and only if $\bar{\zeta} \in \partial \rho(\bar{Z})$ and

$$0 \in N_X(\bar{x}) + E_{\bar{\zeta}}[\partial f_\omega(\bar{x})],$$

(6.132)

where $\bar{Z} := F(\bar{x})$.

**Proof.** By the definition, $(\bar{x}, \bar{\zeta})$ is a saddle point of $\Phi(x, \zeta)$ iff

$$\bar{x} \in \arg\min_{x \in X} \Phi(x, \zeta) \text{ and } \bar{\zeta} \in \arg\max_{\zeta \in \mathfrak{A}} \Phi(\bar{x}, \zeta).$$

(6.133)

The first of the above conditions means that $\bar{x} \in \arg\min_{x \in X} \psi(x)$, where

$$\psi(x) := \int_{\Omega} f(x, \omega)\zeta(\omega)dP(\omega).$$

Since $X$ is convex and $\psi(\cdot)$ is convex real valued, by the standard optimality conditions this holds iff $0 \in N_X(\bar{x}) + \partial \psi(\bar{x})$. Moreover, by Theorem 7.47 we have $\partial \psi(\bar{x}) = E_{\bar{\zeta}}[\partial f_\omega(\bar{x})]$. Therefore, condition (6.132) and the first condition in (6.133) are equivalent. The second condition (6.133) and the condition $\bar{\zeta} \in \partial \rho(\bar{Z})$ are equivalent by equation (6.42).

Under the assumptions of Proposition 6.31, existence of $\bar{\zeta} \in \partial \rho(\bar{Z})$ in equation (6.132) can be viewed as an optimality condition for problem (6.127). Sufficiency of that condition follows directly from the fact that it implies that $(\bar{x}, \bar{\zeta})$ is a saddle point of the min-max problem (6.128). In order for that condition to be necessary we need to verify existence of a saddle point for problem (6.128).

**Proposition 6.32.** Let $\bar{x}$ be an optimal solution of the problem (6.127). Suppose that the mapping $F : \mathbb{R}^n \to Z$ is convex and risk measure $\rho : Z \to \mathbb{R}$ is proper, lower
semicontinuous and satisfies conditions (R1)–(R2), and is continuous at \( \bar{Z} := F(\bar{x}) \). Then there exists \( \bar{\zeta} \in \partial \rho(\bar{Z}) \) such that \((\bar{x}, \bar{\zeta})\) is a saddle point of \( \Phi(x, \zeta) \).

**Proof.** By monotonicity of \( \rho \) (condition (R2)) it follows from the optimality of \( \bar{x} \) that \((\bar{x}, \bar{Z})\) is an optimal solution of the problem

\[
\min_{(x, Z) \in S} \rho(Z),
\]

where \( S := \{(x, Z) \in X \times \mathcal{Z} : F(x) \succeq Z\} \). Since \( F \) is convex, the set \( S \) is convex, and since \( F \) is continuous (see Lemma 6.8) the set \( S \) is closed. Also because \( \rho \) is convex and continuous at \( \bar{Z} \), the following (first order) optimality condition holds at \((\bar{x}, \bar{Z})\) (see Remark 31, page 406):

\[
0 \in \partial \rho(\bar{Z}) \times \{0\} + \mathcal{N}_S(\bar{x}, \bar{Z}).
\]

This means that there exists \( \bar{\zeta} \in \partial \rho(\bar{Z}) \) such that \((-\bar{\zeta}, 0) \in \mathcal{N}_S(\bar{x}, \bar{Z})\). This in turn implies that

\[
\langle \bar{\zeta}, Z - \bar{Z} \rangle \geq 0, \ \forall (x, Z) \in S.
\]

Setting \( Z := F(x) \) we obtain that

\[
\langle \bar{\zeta}, F(x) - F(\bar{x}) \rangle \geq 0, \ \forall x \in X.
\]

It follows that \( \bar{x} \) is a minimizer of \( \langle \bar{\zeta}, F(x) \rangle \) over \( x \in X \), and hence \( \bar{x} \) is a minimizer of \( \Phi(x, \zeta) \) over \( x \in X \). That is, \( \bar{x} \) satisfies first of the two conditions in (6.133). Moreover, as it was shown in the proof of Proposition 6.31, this implies condition (6.132), and hence \((\bar{x}, \bar{\zeta})\) is a saddle point by Proposition 6.31. \( \square \)

**Corollary 6.33.** Suppose that problem (6.127) has optimal solution \( \bar{x} \), the mapping \( F : \mathbb{R}^n \to \mathcal{Z} \) is convex and risk measure \( \rho : \mathcal{Z} \to \mathbb{R} \) is proper, lower semicontinuous and satisfies conditions (R1)–(R2), and is continuous at \( \bar{Z} := F(\bar{x}) \). Then there is no duality gap between problems (6.128) and (6.131), and problem (6.131) has an optimal solution.

Propositions 6.31 and 6.32 imply the following optimality conditions.

**Theorem 6.34.** Suppose that mapping \( F : \mathbb{R}^n \to \mathcal{Z} \) is convex and risk measure \( \rho : \mathcal{Z} \to \mathbb{R} \) is proper, lower semicontinuous and satisfies conditions (R1)–(R2). Consider a point \( \bar{x} \in X \) and let \( \bar{Z} := F(\bar{x}) \). Then a sufficient condition for \( \bar{x} \) to be an optimal solution of the problem (6.127) is existence of \( \bar{\zeta} \in \partial \rho(\bar{Z}) \) such that equation (6.132) holds. This condition is also necessary if \( \rho \) is continuous at \( \bar{Z} \).

It could be noted that if \( \rho(\cdot) := \mathbb{E}[^{\cdot}()] \), then its subdifferential consists of unique sub-gradient \( \zeta(\cdot) \equiv 1 \). In that case condition (6.132) takes the form

\[
0 \in \mathcal{N}_X(\bar{x}) + \mathbb{E}[\partial f_\omega(\bar{x})].
\]

Note that since it is assumed that \( F(x) \in L^p_\rho(\Omega, \mathcal{F}, P) \), the expectation \( \mathbb{E}[f_\omega(x)] \) is well defined and finite valued for all \( x \), and hence \( \partial \mathbb{E}[f_\omega(x)] = \mathbb{E}[\partial f_\omega(x)] \) (see Theorem 7.47).
6.4. Optimization of Risk Measures

6.4.1 Dualization of Nonanticipativity Constraints

We assume again that \( Z = \mathcal{L}_p(\Omega, F, P) \) and \( Z^* = \mathcal{L}_q(\Omega, F, P) \), that \( F : \mathbb{R}^n \to Z \) is convex and \( \rho : Z \to \mathbb{R} \) is proper lower semicontinuous and satisfies conditions (R1) and (R2). A way of representing problem (6.127) is to consider the decision vector \( x \) as a function of the elementary event \( \omega \in \Omega \) and then to impose an appropriate nonanticipativity constraint. That is, let \( \mathfrak{M} \) be a linear space of \( F \)-measurable mappings \( \chi : \Omega \to \mathbb{R}^n \). Define \( F^*_\chi(\omega) := f(\chi(\omega), \omega) \) and

\[
\mathfrak{M}_\chi := \{ \chi \in \mathfrak{M} : \chi(\omega) \in X, \text{ a.e. } \omega \in \Omega \}. \tag{6.139}
\]

We assume that the space \( \mathfrak{M} \) is chosen in such a way that \( F^*_\chi \in Z \) for every \( \chi \in \mathfrak{M} \) and for every \( x \in \mathbb{R}^n \) the constant mapping \( \chi(\omega) \equiv x \) belongs to \( \mathfrak{M} \). Then we can write problem (6.127) in the following equivalent form

\[
\begin{aligned}
\min_{(\chi, x) \in \mathfrak{M}_\chi \times \mathbb{R}^n} & \rho(F^*_\chi) \\
\text{subject to } & \chi(\omega) = x, \text{ a.e. } \omega \in \Omega. \tag{6.140}
\end{aligned}
\]

Formulation (6.140) allows developing a duality framework associated with the nonanticipativity constraint \( \chi(\cdot) = x \). In order to formulate such duality we need to specify the space \( \mathfrak{M} \) and its dual. It looks natural to use \( \mathfrak{M} := \mathcal{L}_p(\Omega, F, P; \mathbb{R}^n) \), for some \( p' \in [1, +\infty) \), and its dual \( \mathfrak{M}^* := \mathcal{L}_{q'}(\Omega, F, P; \mathbb{R}^n) \), \( q' \in (1, +\infty) \). It is also possible to employ \( \mathfrak{M} := \mathcal{L}_\infty(\Omega, F, P; \mathbb{R}^n) \). Unfortunately this Banach space is not reflexive. Nevertheless, it can be paired with the space \( \mathcal{L}_1(\Omega, F, P; \mathbb{R}^n) \) by defining the corresponding scalar product in the usual way. As long as the risk measure is lower semicontinuous and subdifferentiable in the corresponding weak topology, we can use this setting as well.

The (Lagrangian) dual of problem (6.140) can be written in the form

\[
\max_{\lambda \in \mathfrak{M}^*} \left\{ \inf_{(\chi, x) \in \mathfrak{M}_\chi \times \mathbb{R}^n} L(\chi, x, \lambda) \right\}, \tag{6.141}
\]

where

\[
L(\chi, x, \lambda) := \rho(F^*_\chi) + \mathbb{E}[\lambda^T(\chi - x)], \quad (\chi, x, \lambda) \in \mathfrak{M} \times \mathbb{R}^n \times \mathfrak{M}^*. \tag{6.142}
\]

Note that

\[
\inf_{x \in \mathbb{R}^n} L(\chi, x, \lambda) = \begin{cases} L(\chi, 0, \lambda), & \text{if } \mathbb{E}[\lambda] = 0, \\ -\infty, & \text{if } \mathbb{E}[\lambda] \neq 0. \end{cases}
\]

Therefore the dual problem (6.142) can be rewritten in the form

\[
\max_{\lambda \in \mathfrak{M}^*} \left\{ \inf_{\chi \in \mathfrak{M}_\chi} L_0(\chi, \lambda) \right\} \quad \text{subject to } \mathbb{E}[\lambda] = 0, \tag{6.143}
\]

where \( L_0(\chi, \lambda) := L(\chi, 0, \lambda) = \rho(F^*_\chi) + \mathbb{E}[\lambda^T\chi] \).

We have that the optimal value of problem (6.140) (which is the same as the optimal value of problem (6.127)) is greater than or equal to the optimal value of its dual (6.143). Moreover, under some regularity conditions, their optimal values are equal to each other. In particular, if Lagrangian \( L(\chi, x, \lambda) \) has a saddle point \( (\bar{\chi}, \bar{x}, \bar{\lambda}) \), then there is no duality gap between problems (6.140) and (6.143), and \( (\bar{\chi}, \bar{x}) \) and \( \bar{\lambda} \) are optimal solutions of...
problems (6.140) and (6.143), respectively. Noting that \( L(\chi, 0, \lambda) \) is linear in \( x \) and in \( \lambda \), we have that \( ((\hat{\chi}, \hat{x}), \hat{\lambda}) \) is a saddle point of \( L(\chi, x, \lambda) \) iff the following conditions hold:

\[
\begin{align*}
\hat{\chi}(\omega) &= \hat{x}, \text{ a.e. } \omega \in \Omega, \quad \text{and} \quad E[\lambda] = 0, \\
\hat{\chi} &\in \arg \min_{\chi \in M} L_0(\chi, \hat{\lambda}).
\end{align*}
\]

Unfortunately, it may not be easy to verify existence of such saddle point.

We can approach the duality analysis by conjugate duality techniques. For a perturbation vector \( y \in \mathfrak{M} \) consider the problem

\[
\min_{(\chi, x) \in M \times \mathbb{R}^n} \rho(F_\chi) \quad \text{subject to} \quad \chi(\omega) = x + y(\omega),
\]

and let \( \vartheta(y) \) be its optimal value. Note that a perturbation in the vector \( x \), in the constraints of problem (6.140), can be absorbed into \( y(\omega) \). Clearly for \( y = 0 \), problem (6.145) coincides with the unperturbed problem (6.140), and \( \vartheta(0) \) is the optimal value of the unperturbed problem (6.140). Assume that \( \vartheta(0) \) is finite. Then there is no duality gap between problem (6.140) and its dual (6.141) iff \( \vartheta(y) \) is lower semicontinuous at \( y = 0 \). Again it may be not easy to verify lower semicontinuity of the optimal value function \( \vartheta : \mathfrak{M} \to \mathbb{R} \).

By the general theory of conjugate duality we have the following result.

**Proposition 6.35.** Suppose that \( F : \mathbb{R}^n \to \mathbb{Z} \) is convex, \( \rho : \mathbb{Z} \to \mathbb{R} \) satisfies conditions (R1)–(R2) and the function \( \rho(F_\chi) \), from \( \mathfrak{M} \) to \( \mathbb{R} \), is lower semicontinuous. Suppose, further, that \( \vartheta(0) \) is finite and \( \vartheta(y) < +\infty \) for all \( y \) in a neighborhood (in the norm topology) of \( 0 \in \mathfrak{M} \). Then there is no duality gap between problems (6.140) (6.141), and the dual problem (6.141) has an optimal solution.

**Proof.** Since \( \rho \) satisfies conditions (R1) and (R2) and \( F \) is convex, we have that the function \( \rho(F_\chi) \) is convex, and by the assumption it is lower semicontinuous. The assertion then follows by a general result of conjugate duality for Banach spaces (see Theorem 7.77).

\[ \square \]

In order to apply the above result we need to verify lower semicontinuity of the function \( \rho(F_\chi) \). This function is lower semicontinuous if \( \rho(\cdot) \) is lower semicontinuous and the mapping \( \chi \mapsto F_\chi \), from \( \mathfrak{M} \) to \( \mathbb{Z} \), is continuous. If the set \( \Omega \) is finite, and hence the spaces \( \mathbb{Z} \) and \( \mathfrak{M} \) are finite dimensional, then continuity of \( \chi \mapsto F_\chi \) follows from the continuity of \( F \). In the infinite dimensional setting this should be verified by specialized methods. The assumption that \( \vartheta(0) \) is finite means that the optimal value of the problem (6.140) is finite, and the assumption that \( \vartheta(y) < +\infty \) means that the corresponding problem (6.145) has a feasible solution.

**Interchangeability Principle for Risk Measures**

By removing the nonanticipativity constraint \( \chi(\cdot) = x \) we obtain the following relaxation of the problem (6.140):

\[
\min_{\chi \in \mathfrak{M} \chi} \rho(F_\chi),
\]

(6.146)
where $M_X$ is defined in (6.139). Similarly to the interchangeability principle for the expec-
tation operator (Theorem 7.80), we have the following result for monotone risk measures. By $\inf_{x \in X} F(x)$ we denote the pointwise minimum, i.e.,

$$\left[ \inf_{x \in X} F(x) \right](\omega) := \inf_{x \in X} f(x, \omega), \quad \omega \in \Omega. \quad (6.147)$$

**Proposition 6.36.** Let $Z := L_p(\Omega, F, P)$ and $M := L_p(\Omega, F, P; \mathbb{R}^n)$, where $p, p' \in [1, +\infty]$, $M_X$ be defined in (6.139), $\rho : Z \to \mathbb{R}$ be a proper risk measure satisfying monotonicity condition (R2), and $F : \mathbb{R}^n \to Z$ be such that $\inf_{x \in X} F(x) \in Z$. Suppose that $\rho$ is continuous at $\Psi := \inf_{x \in X} F(x)$. Then

$$\inf_{\chi \in M_X} \rho(F_\chi) = \rho \left( \inf_{x \in X} F(x) \right). \quad (6.148)$$

**Proof.** For any $\chi \in M_X$ we have that $\chi(\cdot) \in X$, and hence the following inequality holds

$$\left[ \inf_{x \in X} F(x) \right](\omega) \leq F_\chi(\omega), \quad \text{a.e. } \omega \in \Omega. \quad (6.149)$$

By monotonicity of $\rho$ this implies that $\rho(\Psi) \leq \rho(F_\chi)$, and hence

$$\rho(\Psi) \leq \inf_{\chi \in M_X} \rho(F_\chi). \quad (6.150)$$

Since $\rho$ is proper we have that $\rho(\Psi) > -\infty$. If $\rho(\Psi) = +\infty$, then by (6.149) the left hand side of (6.148) is also $+\infty$ and hence (6.148) holds. Therefore we can assume that $\rho(\Psi)$ is finite.

Let us derive now the converse of (6.149) inequality. Since it is assumed that $\Psi \in Z$, we have that $\Psi(\omega)$ is finite valued for a.e. $\omega \in \Omega$ and measurable. Therefore for a sequence $\varepsilon_k \downarrow 0$ and a.e. $\omega \in \Omega$ and all $k \in \mathbb{N}$, we can choose $\chi_k(\omega) \in X$ such that $|f(\chi_k(\omega), \omega) - \Psi(\omega)| \leq \varepsilon_k$ and $\chi_k(\cdot)$ are measurable. We also can truncate $\chi_k(\cdot)$, if necessary, in such a way that each $\chi_k$ belongs to $M_X$, and $f(\chi_k(\omega), \omega)$ monotonically converges to $\Psi(\omega)$ for a.e. $\omega \in \Omega$. We have then that $f(\chi_k(\cdot), \cdot) - \Psi(\cdot)$ is nonnegative valued and is dominated by a function from the space $Z$. It follows by the Lebesgue Dominated Convergence Theorem that $F_{\chi_k}$ converges to $\Psi$ in the norm topology of $Z$. Since $\rho$ is continuous at $\Psi$, it follows that $\rho(F_{\chi_k})$ tends to $\rho(\Psi)$. Also $\inf_{\chi \in M_X} \rho(F_\chi) \leq \rho(F_{\chi_k})$, and hence the required converse inequality

$$\inf_{\chi \in M_X} \rho(F_\chi) \leq \rho(\Psi) \quad (6.151)$$

follows. \( \Box \)

**Remark 19.** It follows from (6.148) that if

$$\bar{\chi} \in \arg \min_{\chi \in M_X} \rho(F_\chi), \quad (6.151)$$
then
\[ \bar{\chi}(\omega) \in \arg \min_{x \in X} f(x, \omega), \quad \text{a.e. } \omega \in \Omega. \] (6.152)

Conversely, suppose that the function \( f(x, \omega) \) is random lower semicontinuous. Then the multifunction \( \omega \mapsto \arg \min_{x \in X} f(x, \omega) \) is measurable. Therefore \( \bar{\chi}(\omega) \) in the left hand side of (6.152) can be chosen to be measurable. If, moreover, \( \bar{\chi} \in \mathcal{M} \) (this holds, in particular, if the set \( X \) is bounded and hence \( \bar{\chi}(\cdot) \) is bounded), then the inclusion (6.151) follows.

Consider now a setting of two stage programming. That is, suppose that the function \( [F(x)](\omega) = f(x, \omega) \), of the first stage problem
\[ \min_{x \in X} \rho(F(x)), \] (6.153)
is given by the optimal value of the following second stage problem
\[ \min_{y \in G(x, \omega)} g(x, y, \omega), \] (6.154)
where \( g : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R} \) and \( G : \mathbb{R}^n \times \Omega \to \mathbb{R}^m \). Under appropriate regularity conditions, from which the most important is the monotonicity condition (R2), we can apply the interchangeability principle to the optimization problem (6.154) to obtain
\[ \rho(F(x)) = \inf_{y(\cdot) \in G(x, \cdot)} \rho(g(x, y(\omega), \omega)), \] (6.155)
where now \( y(\cdot) \) is an element of an appropriate functional space and the notation \( y(\cdot) \in G(x, \cdot) \) means that \( y(\omega) \in G(x, \omega) \) w.p.1. If the interchangeability principle (6.155) holds, then the two stage problem (6.153)–(6.154) can be written as one large optimization problem:
\[ \min_{x \in X, y(\cdot) \in G(x, \cdot)} \rho(g(x, y(\omega), \omega)). \] (6.156)

In particular, suppose that the set \( \Omega \) is finite, say \( \Omega = \{\omega_1, \ldots, \omega_K\} \), i.e., there is a finite number \( K \) of scenarios. In that case we can view function \( Z : \Omega \to \mathbb{R} \) as vector \( (Z(\omega_1), \ldots, Z(\omega_K)) \in \mathbb{R}^K \), and hence to identify the space \( \mathcal{Z} \) with \( \mathbb{R}^K \). Then problem (6.156) takes the form
\[ \min_{x \in X, y_k \in G(x, \omega_k), k = 1, \ldots, K} \rho \left[ (g(x, y_1, \omega_1), \ldots, g(x, y_K, \omega_K)) \right]. \] (6.157)

Moreover, consider the linear case where \( X := \{x : Ax = b, \ x \geq 0\} \), \( g(x, y, \omega) := c^T x + q(\omega)^T y \) and
\[ G(x, \omega) := \{y : T(\omega)x + W(\omega)y = h(\omega), \ y \geq 0\}. \]
Assume that \( \rho \) satisfies conditions (R1)–(R3) and the set \( \Omega = \{\omega_1, \ldots, \omega_K\} \) is finite. Then problem (6.157) takes the form:
\[ \min_{x, y_1, \ldots, y_K} c^T x + \rho \left[ (q_1^T y_1, \ldots, q_K^T y_K) \right] \] (6.158)
s.t. \( Ax = b, \ x \geq 0, \ T_k x + W_k y_k = h_k, \ y_k \geq 0, \ k = 1, \ldots, K, \)
where \( (q_k, T_k, W_k, h_k) := (q(\omega_k), T(\omega_k), W(\omega_k), h(\omega_k)), k = 1, \ldots, K. \)
6.4. Optimization of Risk Measures

6.4.2 Examples

Let \( Z := \mathcal{L}_1(\Omega, \mathcal{F}, P) \) and consider

\[
\rho(Z) := \mathbb{E}[Z] + \inf_{t \in \mathbb{R}} \mathbb{E}\{\beta_1[t - Z]_+ + \beta_2[Z - t]_+\}, \quad Z \in Z,
\]

where \( \beta_1 \in [0, 1] \) and \( \beta_2 \geq 0 \) are some constants. Properties of this risk measure were studied in Example 6.15 (see equations (6.67) and (6.68) in particular). We can write the corresponding optimization problem (6.127) in the following equivalent form

\[
\min_{(x, t) \in X \times \mathbb{R}} \mathbb{E}\{f_\omega(x) + \beta_1[t - f_\omega(x)]_+ + \beta_2[f_\omega(x) - t]_+\}.
\]

That is, by adding one extra variable we can formulate the corresponding optimization problem as an expectation minimization problem.

Risk Averse Optimization of an Inventory Model

Let us consider again the inventory model analyzed in section 1.2. Recall that the objective of that model is to minimize the total cost

\[
F(x, d) = cx + b[d - x]_+ + h[x - d]_+,
\]

where \( c, b \) and \( h \) are nonnegative constants representing costs of ordering, back ordering and holding, respectively. Again we assume that \( b > c > 0 \), i.e., the back order cost is bigger than the ordering cost. A risk averse extension of the corresponding (expected value) problem (1.4) can be formulated in the form

\[
\min_{x \geq 0} \{f(x) := \rho[F(x, D)]\},
\]

where \( \rho \) is a specified risk measure.

Assume that the risk measure \( \rho \) is coherent, i.e., satisfies conditions (R1)–(R4), and that demand \( D = D(\omega) \) belongs to an appropriate space \( Z = \mathcal{L}_p(\Omega, \mathcal{F}, P) \). Assume, further, that \( \rho : Z \to \mathbb{R} \) is real valued. It follows that there exists a convex set \( \mathfrak{A} \subset \mathfrak{P} \), where \( \mathfrak{P} \subset \mathbb{R}^* \) is the set of probability density functions, such that

\[
\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \int_{\Omega} Z(\omega)\zeta(\omega)dP(\omega), \quad Z \in Z.
\]

Consequently we have that

\[
\rho[F(x, D)] = \sup_{\zeta \in \mathfrak{A}} \int_{\Omega} F(x, D(\omega))\zeta(\omega)dP(\omega).
\]

To each \( \zeta \in \mathfrak{P} \) corresponds the cumulative distribution function \( H \) of \( D \) with respect to the measure \( Q := \zeta dP \), that is,

\[
H(z) = Q(D \leq z) = \mathbb{E}_\zeta[1_{D \leq z}] = \int_{\{\omega: D(\omega) \leq z\}} \zeta(\omega)dP(\omega).
\]
We have then that
\[ \int_{\Omega} F(x, D(\omega)) \zeta(\omega) dP(\omega) = \int F(x, z) dH(z). \]

Denote by \( \mathfrak{H} \) the set of cumulative distribution functions \( H \) associated with densities \( \zeta \in \mathfrak{X} \). The correspondence between \( \zeta \in \mathfrak{X} \) and \( H \in \mathfrak{H} \) is given by formula (6.163), and depends on \( D(\cdot) \) and the reference probability measure \( P \). Then we can rewrite (6.162) in the form
\[ \rho[F(x, D)] = \sup_{H \in \mathfrak{H}} \int F(x, z) dH(z) = \sup_{H \in \mathfrak{H}} \mathbb{E}_H[F(x, D)]. \]

This leads to the following minimax formulation of the risk averse optimization problem (6.161):
\[ \min_{x \geq 0} \sup_{H \in \mathfrak{H}} \mathbb{E}_H[F(x, D)]. \]

Note that we also have that \( \rho(D) = \sup_{H \in \mathfrak{H}} \mathbb{E}_H[D] \).

In the subsequent analysis we deal with the minimax formulation (6.165), rather than the risk averse formulation (6.161), viewing \( \mathfrak{H} \) as a given set of cumulative distribution functions. We show next that the minimax problem (6.165), and hence the risk averse problem (6.161), structurally is similar to the corresponding (expected value) problem (1.4). We assume that every \( H \in \mathfrak{H} \) is such that \( H(z) = 0 \) for any \( z < 0 \) (recall that the demand cannot be negative). We also assume that \( \sup_{H \in \mathfrak{H}} \mathbb{E}_H[D] < +\infty \), which follows from the assumption that \( \rho(\cdot) \) is real valued.

**Proposition 6.37.** Let \( \mathfrak{H} \) be a set of cumulative distribution functions such that \( H(z) = 0 \) for any \( H \in \mathfrak{H} \) and \( z < 0 \), and \( \sup_{H \in \mathfrak{H}} \mathbb{E}_H[D] < +\infty \). Consider function \( f(x) := \sup_{H \in \mathfrak{H}} \mathbb{E}_H[F(x, D)] \). Then there exists a cdf \( \bar{H} \), depending on the set \( \mathfrak{H} \) and \( \eta := b/(b + h) \), such that \( \bar{H}(z) = 0 \) for any \( z < 0 \), and the function \( f(x) \) can be written in the form
\[ f(x) = b \sup_{H \in \mathfrak{H}} \mathbb{E}_H[D] + (c - b)x + (b + h) \int_{-\infty}^{x} \bar{H}(z) dz. \]

**Proof.** We have (see formula (1.5)) that for \( H \in \mathfrak{H} \),
\[ \mathbb{E}_H[F(x, D)] = b \mathbb{E}_H[D] + (c - b)x + (b + h) \int_{0}^{x} H(z) dz. \]

Therefore we can write \( f(x) = (c - b)x + (b + h)g(x) \), where
\[ g(x) := \sup_{H \in \mathfrak{H}} \left\{ \eta \mathbb{E}_H[D] + \int_{-\infty}^{x} H(z) dz \right\}. \]

Since every \( H \in \mathfrak{H} \) is a monotonically nondecreasing function, we have that \( x \mapsto \int_{-\infty}^{x} H(z) dz \) is a convex function. It follows that the function \( g(x) \) is given by the maximum of convex functions and hence is convex. Moreover, \( g(x) \geq 0 \) and
\[ g(x) \leq \eta \sup_{H \in \mathfrak{H}} \mathbb{E}_H[D] + [x]_. \]
and hence \( g(x) \) is finite valued for any \( x \in \mathbb{R} \). Also for any \( H \in \mathcal{M} \) and \( z < 0 \) we have that \( H(z) = 0 \), and hence \( g(x) = \eta \sup_{H \in \mathcal{M}} E_H[D] \) for any \( x < 0 \).

Consider the right hand side derivative of \( g(x) \):

\[
g^+(x) := \lim_{t \downarrow 0} \frac{g(x + t) - g(x)}{t},
\]

and define \( \bar{H}(\cdot) := g^+(\cdot) \). Since \( g(x) \) is real valued convex, its right hand side derivative \( g^+(x) \) exists, is finite and for any \( x \geq 0 \) and \( a < 0 \),

\[
g(x) = g(a) + \int_a^x g^+(z)dz = \eta \sup_{H \in \mathcal{M}} E_H[D] + \int_{-\infty}^x \bar{H}(z)dz. \tag{6.169}
\]

Note that definition of the function \( g(\cdot) \), and hence \( \bar{H}(\cdot) \), involves the constant \( \eta \) and set \( \mathcal{M} \) only. Let us also observe that the right hand side derivative \( g^+(x) \), of a real valued convex function, is monotonically nondecreasing and right side continuous. Moreover, \( g^+(x) = 0 \) for \( x < 0 \) since \( g(x) \) is constant for \( x < 0 \). We also have that \( g^+(x) \) tends to one as \( x \to +\infty \). Indeed, since \( g^+(x) \) is monotonically nondecreasing it tends to a limit, denoted \( r \), as \( x \to +\infty \). We have then that \( g(x)/x \to r \) as \( x \to +\infty \). It follows from (6.168) that \( r \leq 1 \), and by (6.167) that for any \( H \in \mathcal{M} \),

\[
\lim_{x \to +\infty} \frac{g(x)}{x} \geq \lim_{x \to +\infty} \frac{1}{x} \int_{-\infty}^x \bar{H}(z)dz \geq 1,
\]

and hence \( r \geq 1 \). It follows that \( r = 1 \).

We obtain that \( \bar{H}(\cdot) = g^+(\cdot) \) is a cumulative distribution function of some probability distribution and the representation (6.166) holds.

It follows from the representation (6.166) that the set of optimal solutions of the risk measure problem (6.161) is an interval given by the set of \( \kappa \)-quantiles of the cdf \( \bar{H}(\cdot) \), where \( \kappa := \frac{b-a}{b+a} \) (compare with Remark 1, page 3).

In some specific cases it is possible to calculate the corresponding cdf \( \bar{H} \) in a closed form. Consider the risk measure \( \rho \) defined in (6.159):

\[
\rho(Z) := E[Z] + \inf_{t \in \mathbb{R}} E\{\beta_1|t - Z|_+ + \beta_2[Z - t]_+\},
\]

where the expectations are taken with respect to some reference cdf \( H^*(\cdot) \). The corresponding set \( \mathcal{M} \) is formed by cumulative distribution functions \( H(\cdot) \) such that

\[
(1 - \beta_1) \int_S dH^* \leq \int_S dH \leq (1 + \beta_2) \int_S dH^* \tag{6.170}
\]

for any Borel set \( S \subset \mathbb{R} \) (compare with formula (6.69)). Recall that for \( \beta_1 = 1 \) this risk measure is \( \rho(Z) = AV(R_\alpha(Z)) \) with \( \alpha = 1/(1 + \beta_2) \). Suppose that the reference distribution of the demand is uniform on the interval \([0, 1]\), i.e., \( H^*(z) = z \) for \( z \in [0, 1] \). It follows that any \( H \in \mathcal{M} \) is continuous, \( H(0) = 0 \) and \( H(1) = 1 \), and

\[
E_H[D] = \int_0^1 zdH(z) = zH(z)|_0^1 - \int_0^1 H(z)dz = 1 - \int_0^1 H(z)dz.
\]
Consequently we can write function $g(x)$, defined in (6.167), for $x \in [0, 1]$ in the form

$$g(x) = \eta + \sup_{H \in \mathcal{M}} \left\{ (1 - \eta) \int_0^x H(z)dz - \eta \int_x^1 H(z)dz \right\}. \quad (6.171)$$

Suppose, further, that $h = 0$ (i.e., there are no holding costs) and hence $\eta = 1$. In that case

$$g(x) = 1 - \inf_{H \in \mathcal{M}} \int_x^1 H(z)dz, \text{ for } x \in [0, 1]. \quad (6.172)$$

By using the first inequality of (6.170) with $S := [0, z]$ we obtain that $H(z) \geq (1 - \beta_1)z$ for any $H \in \mathcal{M}$ and $z \in [0, 1]$. Similarly, by the second inequality of (6.170) with $S := [z, 1]$ we have that $H(z) \geq 1 + (1 + \beta_2)(z - 1)$ for any $H \in \mathcal{M}$ and $z \in [0, 1]$. Consequently, the cdf

$$\bar{H}(z) := \max\{(1 - \beta_1)z, (1 + \beta_2)z - \beta_2\}, \quad z \in [0, 1], \quad (6.173)$$

is dominated by any other cdf $H \in \mathcal{M}$, and it can be verified that $\bar{H} \in \mathcal{M}$. Therefore the minimum on the right hand side of (6.172) is attained at $\bar{H}$ for any $x \in [0, 1]$, and hence this cdf $\bar{H}$ fulfills equation (6.166).

Note that for any $\beta_1 \in (0, 1)$ and $\beta_2 > 0$, the cdf $\bar{H}(\cdot)$ defined in (6.173) is strictly less than the reference cdf $H^*(\cdot)$ on the interval $(0, 1)$. Consequently the corresponding risk averse optimal solution $\bar{H}^{-1}(\kappa)$ is bigger than the risk neutral optimal solution $H^*^{-1}(\kappa)$. It should be not surprising that in the absence of holding costs it will be safer to order a larger quantity of the product.

**Risk Averse Portfolio Selection**

Consider the portfolio selection problem introduced in section 1.4. A risk averse formulation of the corresponding optimization problem can be written in the form

$$\min_{x \in X} \rho\left(-\sum_{i=1}^n \xi_i x_i\right), \quad (6.174)$$

where $\rho$ is a chosen risk measure and $X := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = W_0, \ x \geq x\}$. We use the negative of the return as an argument of the risk measure, because we developed our theory for the minimization, rather than maximization framework. An example below shows what could be a problem of using risk measures with dispersions measured by variance or standard deviation.

**Example 6.38** Let $n = 2$, $W_0 = 1$ and the risk measure $\rho$ be of the form

$$\rho(Z) := \mathbb{E}[Z] + c \mathbb{D}[Z], \quad (6.175)$$

where $c > 0$ and $\mathbb{D}[\cdot]$ is a dispersion measure. Let the dispersion measure be either $\mathbb{D}[Z] := \sqrt{\text{Var}[Z]}$ or $\mathbb{D}[Z] := \text{Var}[Z]$. Suppose, further, that the space $\Omega := \{\omega_1, \omega_2\}$ consists of two points with associated probabilities $p$ and $1 - p$, for some $p \in (0, 1)$. Define (random) return rates $\xi_1, \xi_2 : \Omega \to \mathbb{R}$ as follows: $\xi_1(\omega_1) = a$ and $\xi_1(\omega_2) = 0$, where $a$ is some positive number, and $\xi_2(\omega_1) = \xi_2(\omega_2) = 0$. Obviously, is better to invest...
in asset 1 than asset 2. Now, for $\mathbb{E}[Z] := \sqrt{\operatorname{Var}[Z]}$, we have that $\rho(-\xi_2) = 0$ and $\rho(-\xi_1) = -pa + ca\sqrt{p(1-p)}$. It follows that $\rho(-\xi_1) > \rho(-\xi_2)$ for any $c > 0$ and $p < (1+c)^{-2}$. Similarly, for $\mathbb{E}[Z] := \sqrt{\operatorname{Var}[Z]}$ we have that $\rho(-\xi_1) = -pa + ca^2 p(1-p)$, $\rho(-\xi_2) = 0$, and hence $\rho(-\xi_1) > \rho(-\xi_2)$ again, provided $p < 1 - (ca)^{-1}$. That is, although $\xi_2$ dominates $\xi_1$ in the sense that $\xi_1(\omega) \geq \xi_2(\omega)$ for every possible realization of $(\xi_1(\omega), \xi_2(\omega))$, we have that $\rho(\xi_1) > \rho(\xi_2)$.

Here $[F(x)](\omega) := -\xi_1(\omega)x_1 - \xi_2(\omega)x_2$. Let $\bar{x} := (1, 0)$ and $x^* := (0, 1)$. Note that the feasible set $X$ is formed by vectors $t\bar{x} + (1-t)x^*$, $t \in [0, 1]$. We have that $[F(x)](\omega) = -\xi_1(\omega)x_1$, and hence $[F(x)](\omega)$ is dominated by $[F(x)](\omega)$ for any $x \in X$ and $\omega \in \Omega$. And yet, under the specified conditions, we have that $\theta[F(x)](\omega) = \rho(-\xi_1)$ is greater than $\rho[F(x^*)] = \rho(-\xi_2)$, and hence $\bar{x}$ is not an optimal solution of the corresponding optimization (minimization) problem. This should be not surprising, because the chosen risk measure is not monotone, i.e., it does not satisfy the condition (R2), for $c > 0$ (see Examples 6.17 and 6.18).

Suppose now that $\rho$ is a real valued coherent risk measure. We can write then problem (6.174) in the corresponding min-max form (6.130), that is

$$\min \sup_{x \in X} \sum_{i=1}^{n} \left(-\mathbb{E}_{\xi}[\xi_i]\right) x_i.$$

Equivalently,

$$\max \inf_{x \in X} \sum_{i=1}^{n} \left(\mathbb{E}_{\xi}[\xi_i]\right) x_i. \quad (6.167)$$

Since the feasible set $X$ is compact, problem (6.174) always has an optimal solution $\bar{x}$. Also (see Proposition 6.32) the min-max problem (6.174) has a saddle point, and $(\bar{x}, \bar{\xi})$ is a saddle point iff

$$\bar{\xi} \in \partial \rho(\bar{Z}) \quad \text{and} \quad \bar{x} \in \arg \max_{x \in X} \sum_{i=1}^{n} \hat{\mu}_i x_i, \quad (6.177)$$

where $\bar{Z}(\omega) := -\sum_{i=1}^{n} \xi_i(\omega) \bar{x}_i$ and $\hat{\mu}_i := \mathbb{E}_{\xi}[\xi_i]$.

An interesting insight into the risk adverse solution is provided by its game-theoretical interpretation. For $W_0 = 1$ the portfolio allocations $x$ can be interpreted as a mixed strategy of the investor (for another $W_0$ the fractions $x_i/W_0$ are the mixed strategy). The measure $\xi$ represents the mixed strategy of the opponent (the market). It is not chosen from the set of all possible mixed strategies, but rather from the set $\mathfrak{F}$. The risk adverse solution (6.177) corresponds to the equilibrium of the game.

It is not difficult to see that the set $\arg \max_{x \in X} \sum_{i=1}^{n} \hat{\mu}_i x_i$ is formed by all convex combinations of vectors $W_0 e_i$, $i \in \mathcal{I}$, where $e_i \in \mathbb{R}^n$ denotes $i$-th coordinate vector (with zero entries except the $i$-th entry equal 1), and

$$\mathcal{I} := \{ i' : \hat{\mu}_{i'} = \max_{1 \leq i \leq n} \hat{\mu}_i, \ i' = 1, \ldots, n \}.$$

Also $\partial \rho(\bar{Z}) \subset \mathfrak{F}$ (see formula (6.43) for the subdifferential $\partial \rho(\bar{Z})$).
6.5 Statistical Properties of Risk Measures

All examples of risk measures discussed in section 6.3.2 were constructed with respect to a reference probability measure (distribution) \( P \). Suppose now that the “true” probability distribution \( P \) is estimated by an empirical measure (distribution) \( P_N \) based on a sample of size \( N \). In this section we discuss statistical properties of the respective estimates of the “true values” of the corresponding risk measures.

6.5.1 Average Value-at-Risk

Recall that the Average Value-at-Risk, \( \text{AV}_{\alpha}(Z) \), at a level \( \alpha \in (0,1) \) of a random variable \( Z \), is given by the optimal value of the minimization problem

\[
\min_{t \in \mathbb{R}} E \{ t + \alpha^{-1}[Z - t]_+ \},
\]

where the expectation is taken with respect to the probability distribution \( P \) of \( Z \). We assume that \( E|Z| < +\infty \), which implies that \( \text{AV}_{\alpha}(Z) \) is finite. Suppose now that we have an iid random sample \( Z^1, ..., Z^N \) of \( N \) realizations of \( Z \). Then we can estimate \( \theta^* := \text{AV}_{\alpha}(Z) \) by replacing distribution \( P \) with its empirical estimate \( \frac{1}{N} \sum_{j=1}^{N} \Delta(Z^j) \). This leads to the sample estimate \( \hat{\theta}_N \), of \( \theta^* = \text{AV}_{\alpha}(Z) \), given by the optimal value of the following problem

\[
\min_{t \in \mathbb{R}} \left\{ t + \frac{1}{\alpha N} \sum_{j=1}^{N} [Z^j - t]_+ \right\}.
\]

Let us observe that problem (6.178) can be viewed as a stochastic programming problem and problem (6.179) as its sample average approximation. That is,

\[
\theta^* = \inf_{t \in \mathbb{R}} f(t) \quad \text{and} \quad \hat{\theta}_N = \inf_{t \in \mathbb{R}} \hat{f}_N(t),
\]

where

\[
f(t) = t + \alpha^{-1}E[Z - t]_+ \quad \text{and} \quad \hat{f}_N(t) = t + \frac{1}{\alpha N} \sum_{j=1}^{N} [Z^j - t]_+.
\]

Therefore, results of section 5.1 can be applied here in a straightforward way. Recall that the set of optimal solutions of problem (6.178) is the interval \([t^*, t^{**}]\), where

\[
t^* = \inf \{ z : H_Z(z) \geq 1 - \alpha \} = \text{AV}_{\alpha}(Z) \quad \text{and} \quad t^{**} = \sup \{ z : H_Z(z) \leq 1 - \alpha \}
\]

are the respective left and right side \((1 - \alpha)\)-quantiles of the distribution of \( Z \) (see page 260). Since for any \( \alpha \in (0,1) \) the interval \([t^*, t^{**}]\) is finite and problem (6.178) is convex, we have by Theorem 5.4 that

\[
\hat{\theta}_N \to \theta^* \quad \text{w.p.1 as } N \to \infty.
\]

That is, \( \hat{\theta}_N \) is a consistent estimator of \( \theta^* = \text{AV}_{\alpha}(Z) \).

\footnote{Recall that \( \Delta(z) \) denotes measure of mass one at point \( z \).}
Assume now that $E[Z^2] < +\infty$. Then the assumptions (A1) and (A2) of Theorem 5.7 hold, and hence

$$\hat{\theta}_N = \inf_{t \in [t^*, t^*]} \check{f}_N(t) + o_p(N^{-1/2}). \tag{6.181}$$

Moreover, if $t^* = t^{**}$, i.e., the left and right side $(1 - \alpha)$-quantiles of the distribution of $Z$ are the same, then

$$N^{1/2} \left( \hat{\theta}_N - \theta^* \right) \overset{D}{\to} N(0, \sigma^2), \tag{6.182}$$

where $\sigma^2 = \alpha^{-2} \text{Var} ([Z - t^*]_+)$. The estimator $\hat{\theta}_N$ has a negative bias, i.e., $E[\hat{\theta}_N] - \theta^* \leq 0$, and (see Proposition 5.6)

$$E[\hat{\theta}_N] \leq E[\hat{\theta}_{N+1}], \quad N = 1, \ldots, \tag{6.183}$$

i.e., the bias is monotonically decreasing with increase of the sample size $N$. If $t^* = t^{**}$, then this bias is of order $O(N^{-1})$ and can be estimated using results of section 5.13. The first and second order derivatives of the expectation function $f(t)$ here are $f'(t) = 1 + \alpha^{-1}(H_Z(t) - 1)$, provided that the cumulative distribution function $H_Z(\cdot)$ is continuous at $t$, and $f''(t) = \alpha^{-1} h_Z(t)$, provided that the density $h_Z(t) = \partial H_Z(t)/\partial t$ exists. We obtain (see Theorem 5.8 and the discussion on page 169), under appropriate regularity conditions, in particular if $t^* = t^{**} = V@R_{\alpha}(Z)$ and the density $h_Z(t^*) = \partial H_Z(t^*)/\partial t$ exists and $h_Z(t^*) \neq 0$, that

$$\hat{\theta}_N - \check{f}_N(t^*) = N^{-1} \inf_{t \in \mathbb{R}} \left\{ \tau Z + \frac{1}{2} \tau^2 f''(t^*) \right\} + o_p(N^{-1})$$

$$= - \frac{\alpha^2}{2h_Z(t^*)} + o_p(N^{-1}), \tag{6.184}$$

where $Z \sim N(0, \gamma^2)$ with

$$\gamma^2 = \text{Var} \left( \alpha^{-1} \frac{\partial [Z - t^*]}{\partial t} \right) = H_Z(t^*) \left( 1 - H_Z(t^*) \right) \alpha^2 = 1 - \frac{\alpha}{\alpha}. \tag{6.185}$$

Consequently, under appropriate regularity conditions,

$$N \left[ \hat{\theta}_N - \check{f}_N(t^*) \right] \overset{D}{\to} \left[ \frac{1 - \alpha}{2h_Z(t^*)} \right] \chi^2_1 \tag{6.185}$$

and (see Remark 29 on page 385)

$$E[\hat{\theta}_N] - \theta^* = - \frac{1 - \alpha}{2N h_Z(t^*)} + o(N^{-1}). \tag{6.186}$$

### 6.5.2 Absolute Semideviation Risk Measure

Consider the mean absolute semideviation risk measure

$$\rho_c(Z) := E \left\{ Z + c[Z - E(Z)]_+ \right\}, \tag{6.187}$$

where $c \in [0, 1]$ and the expectation is taken with respect to the probability distribution $P$ of $Z$. We assume that $E|Z| < +\infty$, and hence $\rho_c(Z)$ is finite. For a random sample
Z^1, ..., Z^N of Z, the corresponding estimator of \( \theta^* := \rho_c(Z) \) is

\[
\hat{\theta}_N = N^{-1} \sum_{j=1}^{N} (Z^j + c|Z^j - \bar{Z}|_+), \tag{6.188}
\]

where \( \bar{Z} = N^{-1} \sum_{j=1}^{N} Z^j \).

We have that \( \rho_c(Z) \) is equal to the optimal value of the following convex-concave minimax problem

\[
\min_{t \in \mathbb{R}} \max_{\gamma \in [0,1]} \mathbb{E}[F(t, \gamma, Z)], \tag{6.189}
\]

where

\[
F(t, \gamma, z) := z + c\gamma[z - t]_+ + c(1 - \gamma)[t - z]_+ = z + c|z - t|_+ + c(1 - \gamma)(z - t). \tag{6.190}
\]

This follows by virtue of Corollary 6.3. More directly we can argue as follows. Denote \( \mu := \mathbb{E}[Z] \). We have that

\[
\sup_{\gamma \in [0,1]} \mathbb{E}\{Z + c\gamma[Z - t]_+ + c(1 - \gamma)[t - Z]_+\} = \mathbb{E}[Z] + c \max \{\mathbb{E}([Z - t]_+), \mathbb{E}([t - Z]_+)\}.
\]

Moreover, \( \mathbb{E}([Z - t]_+) = \mathbb{E}([t - Z]_+) \) if \( t = \mu \), and either \( \mathbb{E}([Z - t]_+) \) or \( \mathbb{E}([t - Z]_+) \) is bigger than \( \mathbb{E}([Z - \mu]_+) \) if \( t \neq \mu \). This implies the assertion and also shows that the minimum in (6.189) is attained at unique point \( t^* = \mu \). It also follows that the set of saddle points of the minimax problem (6.189) is given by \( \{\mu\} \times [\gamma^*, \gamma^{**}] \), where

\[
\gamma^* = \text{Pr}(Z < \mu) \quad \text{and} \quad \gamma^{**} = \text{Pr}(Z \leq \mu) = H_Z(\mu). \tag{6.191}
\]

In particular, if the cdf \( H_Z(\cdot) \) is continuous at \( \mu = \mathbb{E}[Z] \), then there is unique saddle point \((\mu, H_Z(\mu))\).

Consequently, \( \hat{\theta}_N \) is equal to the optimal value of the corresponding SAA problem

\[
\min_{t \in \mathbb{R}} \max_{\gamma \in [0,1]} N^{-1} \sum_{j=1}^{N} F(t, \gamma, Z^j). \tag{6.192}
\]

Therefore we can apply results of section 5.1.4 in a straightforward way. We obtain that \( \hat{\theta}_N \) converges w.p.1 to \( \theta^* \) as \( N \to \infty \). Moreover, assuming that \( \mathbb{E}[Z^2] < +\infty \) we have by Theorem 5.10 that

\[
\hat{\theta}_N = \max_{\gamma \in [\gamma^*, \gamma^{**}]} N^{-1} \sum_{j=1}^{N} F(\mu, \gamma, Z^j) + o_p(N^{-1/2}) = \bar{Z} + cN^{-1} \sum_{j=1}^{N} [Z^j - \mu]_+ + c\Psi(\bar{Z} - \mu) + o_p(N^{-1/2}), \tag{6.193}
\]

where \( \bar{Z} = N^{-1} \sum_{j=1}^{N} Z^j \) and function \( \Psi(\cdot) \) is defined as

\[
\Psi(z) := \begin{cases} 
(1 - \gamma^*)z, & \text{if } z > 0, \\
(1 - \gamma^{**})z, & \text{if } z \leq 0.
\end{cases}
\]
6.5. Statistical Properties of Risk Measures

If, moreover, the cdf \( H_Z(\cdot) \) is continuous at \( \mu \), and hence \( \gamma^* = \gamma^{**} = H_Z(\mu) \), then

\[
N^{1/2}(\hat{\theta}_N - \theta^*) \xrightarrow{D} N(0, \sigma^2),
\]

(6.194)

where \( \sigma^2 = \operatorname{Var}[F(\mu, H_Z(\mu), Z)] \).

This analysis can be extended to risk averse optimization problems of the form (6.127). That is, consider problem

\[
\min_{x \in X} \{ \rho_t[G(x, \xi)] = \mathbb{E} \{ G(x, \xi) + c[G(x, \xi) - \mathbb{E}(G(x, \xi))]_+ \} \},
\]

(6.195)

where \( X \subset \mathbb{R}^n \) and \( G : X \times \Omega \to \mathbb{R} \). Its sample average approximation (SAA) is obtained by replacing the true distribution of the random vector \( \xi \) with the empirical distribution associated with a random sample \( \xi^1, \ldots, \xi^N \), that is

\[
\min_{x \in X} \frac{1}{N} \sum_{j=1}^N \left\{ G(x, \xi^j) + c \left[ G(x, \xi^j) - \frac{1}{N} \sum_{j=1}^N G(x, \xi^j) \right]_+ \right\}.
\]

(6.196)

Assume that the set \( X \) is convex compact and function \( G(\cdot, \xi) \) is convex for a.e. \( \xi \). Then, for \( c \in [0, 1] \), problems (6.195) and (6.196) are convex. By using the min-max representation (6.189), problem (6.195) can be written as the minimax problem

\[
\min_{(x, t) \in X \times \mathbb{R}} \max_{\gamma \in [0, 1]} \mathbb{E} \{ F(t, \gamma, G(x, \xi)) \},
\]

(6.197)

where function \( F(t, \gamma, z) \) is defined in (6.190). The function \( F(t, \gamma, z) \) is convex and monotonically increasing in \( z \). Therefore, by convexity of \( G(\cdot, \xi) \), the function \( F(t, \gamma, G(x, \xi)) \) is convex in \( x \in X \), and hence (6.197) is a convex-concave minimax problem. Consequently results of section 5.1.4 can be applied.

Let \( \theta^* \) and \( \hat{\theta}_N \) be the optimal values of the true problem (6.195) and the SAA problem (6.196), respectively, and \( \hat{S} \) be the set of optimal solutions of the true problem (6.195). By Theorem 5.10 and the above analysis we obtain, assuming that conditions specified in Theorem 5.10 are satisfied, that

\[
\hat{\theta}_N = N^{-1} \inf_{c \in \mathbb{R}^{[G(x, \xi)]}} \max_{\gamma \in [\gamma^*, \gamma^{**}]} \left\{ \sum_{j=1}^N F(t, \gamma, G(x, \xi^j)) \right\} + o_p(N^{-1/2}),
\]

(6.198)

where

\[
\gamma^* := \operatorname{Pr}\{ G(x, \xi) < \mathbb{E}[G(x, \xi)] \} \quad \text{and} \quad \gamma^{**} := \operatorname{Pr}\{ G(x, \xi) \leq \mathbb{E}[G(x, \xi)] \}, \quad x \in S.
\]

Note that the points \((x, \mathbb{E}[G(x, \xi)]), \gamma)\), where \( x \in S \) and \( \gamma \in [\gamma^*, \gamma^{**}] \), form the set of saddle points of the convex-concave minimax problem (6.197), and hence the interval \([\gamma^*, \gamma^{**}] \) is the same for any \( x \in S \).

Moreover, assume that \( \hat{S} = \{ \hat{x} \} \) is a singleton, i.e., problem (6.195) has unique optimal solution \( \hat{x} \), and the cdf of the random variable \( Z = G(\hat{x}, \xi) \) is continuous at \( \mu := \mathbb{E}[G(\hat{x}, \xi)] \), and hence \( \gamma^* = \gamma^{**} \). Then it follows that \( N^{1/2}(\hat{\theta}_N - \theta^*) \) converges in distribution to normal with zero mean and variance

\[
\operatorname{Var} \{ G(\hat{x}, \xi) + c[G(\hat{x}, \xi) - \mu]_+ + c(1 - \gamma^*)(G(\hat{x}, \xi) - \mu) \}.
\]
6.5.3 Von Mises Statistical Functionals

In the two examples, of AV@R\(_{\alpha}\) and absolute semideviation, of risk measures considered in the above sections it was possible to use their variational representations in order to apply results and methods developed in section 5.1. A possible approach to deriving large sample asymptotics of law invariant coherent risk measures is to use the Kusuoka representation described in Theorem 6.23 (such approach was developed in [147]). In this section we discuss an alternative approach of Von Mises statistical functionals borrowed from Statistics.

We view now a (law invariant) risk measure \(\rho(Z)\) as a function \(\mathfrak{F}(P)\) of the corresponding probability measure \(P\). For example, with the (upper) semideviation risk measure \(\sigma^+_{\rho}[Z]\), defined in (6.5), we associate the functional

\[
\mathfrak{F}(P) := \left(\mathbb{E}_P \left( Z - \mathbb{E}_P[Z]\right)^p \right)^{1/p}.
\]  

(6.199)

The sample estimate of \(\mathfrak{F}(P)\) is obtained by replacing probability measure \(P\) with the empirical measure \(P_N\). That is, we estimate \(\theta^* = \mathfrak{F}(P)\) by \(\hat{\theta}_N = \mathfrak{F}(P_N)\).

Let \(Q\) be an arbitrary probability measure, defined on the same probability space as \(P\), and consider the convex combination \((1 - t)P + tQ = P + t(Q - P)\), with \(t \in [0, 1]\), of \(P\) and \(Q\). Suppose that the following limit exists

\[
\mathfrak{F}'(P, Q - P) := \lim_{t \to 0} \frac{\mathfrak{F}(P + t(Q - P)) - \mathfrak{F}(P)}{t}.
\]  

(6.200)

The above limit is just the directional derivative of \(\mathfrak{F}(\cdot)\) at \(P\) in the direction \(Q - P\). If, moreover, the directional derivative \(\mathfrak{F}'(P, \cdot)\) is linear, then \(\mathfrak{F}(\cdot)\) is Gâteaux differentiable at \(P\). Consider now the approximation

\[
\mathfrak{F}(P_N) - \mathfrak{F}(P) \approx \mathfrak{F}'(P, P_N - P).
\]  

(6.201)

By this approximation

\[
N^{1/2}(\hat{\theta}_N - \theta^*) \approx \mathfrak{F}'(P, N^{1/2}(P_N - P)),
\]  

(6.202)

and we can use \(\mathfrak{F}'(P, N^{1/2}(P_N - P))\) to derive asymptotics of \(N^{1/2}(\hat{\theta}_N - \theta^*)\).

Suppose, further, that \(\mathfrak{F}'(P, \cdot)\) is linear, i.e., \(\mathfrak{F}(\cdot)\) is Gâteaux differentiable at \(P\). Then, since \(P_N = N^{-1} \sum_{j=1}^N \Delta(Z^j)\), we have that

\[
\mathfrak{F}'(P, P_N - P) = \frac{1}{N} \sum_{j=1}^N IF_{\mathfrak{F}}(Z^j),
\]  

(6.203)

where

\[
IF_{\mathfrak{F}}(z) := \sum_{j=1}^N \mathfrak{F}'(P, \Delta(z) - P)
\]  

(6.204)

is the so-called influence function (also called influence curve) of \(\mathfrak{F}\).

It follows from the linearity of \(\mathfrak{F}'(P, \cdot)\) that \(\mathbb{E}_P[IF_{\mathfrak{F}}(Z)] = 0\). Indeed, linearity of \(\mathfrak{F}'(P, \cdot)\) means that it is a linear functional and hence can be represented as

\[
\mathfrak{F}'(P, Q - P) = \int g d(Q - P) = \int g dQ - \mathbb{E}_P[g(Z)]
\]
for some function $g$ in an appropriate functional space. Consequently $IF_{\tilde{g}}(z) = g(z) - E_P[g(Z)]$, and hence

$$E_P[IF_{\tilde{g}}(Z)] = E_P[g(Z) - E_P[g(Z)] = 0.$$  

Then by the Central Limit Theorem we have that $N^{-1/2} \sum_{j=1}^{N} IF_{\tilde{g}}(Z_j)$ converges in distribution to normal with zero mean and variance $E_P[IF_{\tilde{g}}(Z)^2]$. This suggests the following asymptotics

$$N^{1/2}(\theta_N - \theta^*) \xrightarrow{D} N(0, E_P[IF_{\tilde{g}}(Z)^2]).$$  

(6.205)

It should be mentioned at this point that the above derivations do not prove in a rigorous way validity of the asymptotics (6.205). The main technical difficulty is to give a rigorous justification for the approximation (6.202) leading to the corresponding convergence in distribution. This can be compared with the Delta method, discussed in section 7.2.7 and applied in section 5.1, where first (and second) order approximations were derived in functional spaces rather than spaces of measures. Anyway, formula (6.205) gives correct asymptotics and is routinely used in statistical applications.

Let us consider, for example, the statistical functional

$$\tilde{g}(P) := E_P[Z - E_P[Z]]_+,$$  

(6.206)

associated with $\sigma^+_1[Z]$. Denote $\mu := E_P[Z]$. Then

$$\tilde{g}(P + t(Q - P)) - \tilde{g}(P) = t \left( E_Q[Z - \mu]_+ - E_P[Z - \mu]_+ \right) + E_P[Z - \mu - t(E_Q[Z] - \mu)]_+ + o(t).$$

Moreover, the right side derivative at $t = 0$ of the second term in the right hand side of the above equation is $(1 - H_Z(\mu))(E_Q[Z] - \mu)$, provided that the cdf $H_Z(z)$ is continuous at $z = \mu$. It follows that if the cdf $H_Z(z)$ is continuous at $z = \mu$, then

$$\tilde{g}'(P, Q - P) = E_Q[Z - \mu]_+ - E_P[Z - \mu]_+ + (1 - H_Z(\mu))(E_Q[Z] - \mu),$$

and hence

$$IF_{\tilde{g}}(z) = [z - \mu]_+ - E_P[Z - \mu]_+ + (1 - H_Z(\mu))(z - \mu).$$  

(6.207)

It can be seen now that $E_P[IF_{\tilde{g}}(Z)] = 0$ and

$$E_P[IF_{\tilde{g}}(Z)^2] = \text{Var}\{[Z - \mu]_+ + (1 - H_Z(\mu))(Z - \mu)\}.$$  

That is, the asymptotics (6.205) here are exactly the same as the ones derived in the previous section 6.5.2 (compare with (6.194)).

In a similar way it is possible to compute the influence function of the statistical functional defined in (6.199), associated with $\sigma^+_p[Z]$, for $p > 1$. For example, for $p = 2$ the corresponding influence function can be computed, provided that the cdf $H_Z(z)$ is continuous at $z = \mu$, as

$$IF_{\tilde{g}}(z) = \frac{1}{2\kappa^2} \left( [z - \mu]^2_+ - \theta^*^2 + 2\kappa(1 - H_Z(\mu))(z - \mu) \right),$$  

(6.208)

where $\theta^* := \tilde{g}(P) = (E_P[Z - \mu]^2_+)^{1/2}$ and $\kappa := E_P[Z - \mu]_+ = \frac{1}{2}E_P[Z - \mu]$. 

6.6 The Problem of Moments

Due to the duality representation (6.37) of a coherent risk measure, the corresponding risk averse optimization problem (6.127) can be written as the minimax problem (6.130). So far, risk measures were defined on an appropriate functional space which in turn was dependent on a reference probability distribution. One can take an opposite point of view by defining a min-max problem of the form

$$\min_{x \in X} \sup_{P \in \mathcal{M}} \mathbb{E}_P[f(x, \omega)]$$

in a direct way, for a specified set $\mathcal{M}$ of probability measures on a measurable space $(\Omega, \mathcal{F})$. Note that we do not assume in this section existence of a reference measure $P$ and do not work in a functional space of corresponding density functions. In fact, it will be essential here to consider discrete measures on the space $(\Omega, \mathcal{F})$. We denote by $\bar{\mathcal{P}}$ the set of probability measures on $(\Omega, \mathcal{F})$.

The set $\mathcal{M}$ can be viewed as an uncertainty set for the underlying probability distribution. Of course, there are various ways to define the uncertainty set $\mathcal{M}$. In some situations, it is reasonable to assume that we have a knowledge about certain moments of the corresponding probability distribution. That is, the set $\mathcal{M}$ is defined by moment constraints as follows

$$\mathcal{M} := \left\{ P \in \mathcal{P} : \begin{array}{l} E_P[\psi_i(\omega)] = b_i, \; i = 1, \ldots, p, \\ E_P[\psi_i(\omega)] \leq b_i, \; i = p + 1, \ldots, q \end{array} \right\},$$

where $\psi_i : \Omega \to \mathbb{R}, \; i = 1, \ldots, q$, are measurable functions. Note that the condition $P \in \mathcal{P}$, i.e., that $P$ is a probability measure, can be formulated explicitly as the constraint $\int_\Omega dP = 1, P \succeq 0$.

We assume that every finite subset of $\Omega$ is $\mathcal{F}$-measurable. This is a mild assumption. For example, if $\Omega$ is a metric space equipped with its Borel sigma algebra, then this is certainly holds true. We denote by $\mathcal{P}_m$ the set of probability measures on $(\Omega, \mathcal{F})$ having a finite support of at most $m$ points. That is, every measure $P \in \mathcal{P}_m$ can be represented in the form $P = \sum_{i=1}^m \alpha_i \Delta(\omega_i)$, where $\alpha_i$ are nonnegative numbers summing up to one and $\Delta(\omega)$ denotes measure of mass one at the point $\omega \in \Omega$. Similarly, we denote $\mathcal{M}_m := \mathcal{M} \cap \mathcal{P}_m$. Note that the set $\mathcal{M}$ is convex while, for a fixed $m$, the set $\mathcal{M}_m$ is not necessarily convex. By Theorem 7.32, to any $P \in \mathcal{M}$ corresponds a probability measure $Q \in \mathcal{P}$ with a finite support of at most $q + 1$ points such that $E_P[\psi_i(\omega)] = E_Q[\psi_i(\omega)], \; i = 1, \ldots, q$. That is, if the set $\mathcal{M}$ is nonempty, then its subset $\mathcal{M}_q+1$ is also nonempty. Consider the function

$$g(x) := \sup_{P \in \mathcal{M}} \mathbb{E}_P[f(x, \omega)].$$

---

11 The set $\mathcal{P}$ of probability measures should be distinguished from the set $\mathcal{P}$ of probability density functions used before.

12 Recall that the notation “$P \succeq 0$” means that $P$ is a nonnegative (not necessarily probability) measure on $(\Omega, \mathcal{F})$. 

---
### 6.6. The Problem of Moments

**Proposition 6.39.** For any $x \in X$ we have that

$$g(x) = \sup_{P \in M^*_{q+1}} \mathbb{E}_P[f(x, \omega)].$$

**Proof.** If the set $\mathcal{M}$ is empty, then its subset $\mathcal{M}^*_{q+1}$ is also empty, and hence $g(x)$ as well as the optimal value of the right hand side of (6.212) are equal to $+\infty$. So suppose that $\mathcal{M}$ is nonempty. Consider a point $x \in X$ and $P \in \mathcal{M}$. By Theorem 7.32 there exists $Q \in \mathcal{M}^*_{q+2}$ such that $\mathbb{E}_P[f(x, \omega)] = \mathbb{E}_Q[f(x, \omega)]$. It follows that $g(x)$ is equal to the maximum of $\mathbb{E}_P[f(x, \omega)]$ over $P \in M^*_{q+2}$, which in turn is equal to the optimal value of the problem

$$\max_{\omega_1, \ldots, \omega_m \in \Omega} \sum_{j=1}^m \alpha_j f(x, \omega_j)$$

subject to

$$\sum_{j=1}^m \alpha_j \psi_i(\omega_j) = b_i, \ i = 1, \ldots, p,$$

$$\sum_{j=1}^m \alpha_j \psi_i(\omega_j) \leq b_i, \ i = p + 1, \ldots, q,$$

$$\sum_{j=1}^m \alpha_j = 1,$$

where $m := q + 2$. For fixed $\omega_1, \ldots, \omega_m \in \Omega$, the above is a linear programming problem. Its feasible set is bounded and its optimum is attained at an extreme point of its feasible set which has at most $q + 1$ nonzero components of $\alpha$. Therefore it suffices to take the maximum over $P \in M^*_{q+1}$. \(\square\)

For a given $x \in X$, the (Lagrangian) dual of the problem

$$\max_{P \in \mathcal{M}} \mathbb{E}_P[f(x, \omega)]$$

is the problem

$$\min_{\lambda \in \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^q} \sup_{P \geq 0} L_x(P, \lambda),$$

where

$$L_x(P, \lambda) := \int_{\Omega} f(x, \omega)dP(\omega) + \lambda_0(1 - \int_{\Omega} dP(\omega)) + \sum_{i=1}^{q} \lambda_i (b_i - \int_{\Omega} \psi_i(\omega)dP(\omega)).$$

It is straightforward to verify that

$$\sup_{P \geq 0} L_x(P, \lambda) = \begin{cases} \lambda_0 + \sum_{i=1}^{q} b_i \lambda_i, & \text{if } f(x, \omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \psi_i(\omega) \leq 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

The last assertion follows since for any $\tilde{\omega} \in \Omega$ and $\alpha > 0$ we can take $P := \alpha \Delta(\tilde{\omega})$, in which case

$$\mathbb{E}_P \left[ f(x, \omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \psi_i(\omega) \right] = \alpha \left[ f(x, \tilde{\omega}) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \psi_i(\tilde{\omega}) \right].$$
Consequently, we can write the dual problem (6.215) in the form

$$\begin{align*}
\text{Min} & \quad \lambda_0 + \sum_{i=1}^{q} b_i \lambda_i \\
\text{subject to} & \quad \lambda_0 + \sum_{i=1}^{q} \lambda_i \psi_i(\omega) \geq f(x, \omega), \; \omega \in \Omega.
\end{align*}$$

(6.216)

If the set $\Omega$ is finite, then problem (6.214) and its dual (6.216) are linear programming problems. In that case there is no duality gap between these problems unless both are infeasible. If the set $\Omega$ is infinite, then the dual problem (6.216) becomes a linear semi-infinite programming problem. In that case one needs to verify some regularity conditions in order to ensure the “no duality gap” property. One such regularity condition will be: “the dual problem (6.216) has a nonempty and bounded set of optimal solutions” (see Theorem 7.8). Another regularity condition ensuring the “no duality gap” property is: “the set $\Omega$ is a compact metric space equipped with its Borel sigma algebra and functions $\psi_i(\cdot)$, $i = 1, \ldots, q$, and $f(x, \cdot)$ are continuous on $\Omega$”.

If for every $x \in X$ there is no duality gap between problems (6.214) and (6.216), then the corresponding min-max problem (6.209) is equivalent to the following semi-infinite programming problem

$$\begin{align*}
\text{Min} & \quad x \in X, \lambda \in \mathbb{R} \times \mathbb{R}^{p \times q} \\
\text{subject to} & \quad \lambda_0 + \sum_{i=1}^{q} \lambda_i \psi_i(\omega) \geq f(x, \omega), \; \omega \in \Omega.
\end{align*}$$

(6.217)

Remark 20. Let $\Omega$ be a nonempty measurable subset of $\mathbb{R}^d$, equipped with its Borel sigma algebra, and $\mathcal{M}$ be the set of all probability measures supported on $\Omega$. Then by the above analysis we have that it suffices in problem (6.209) to take the maximum over measures of mass one, and hence problem (6.209) is equivalent to the following (deterministic) minimax problem

$$\begin{align*}
\text{Min} & \quad \sup_{\omega \in \Omega} f(x, \omega).
\end{align*}$$

(6.218)

### 6.7 Multistage Risk Averse Optimization

In this section we discuss an extension of risk averse optimization to a multistage setting. In order to simplify the presentation we start our analysis in the following section with a discrete process where evolution of the state of the system is represented by a scenario tree.

#### 6.7.1 Scenario Tree Formulation

Consider a scenario tree representation of evolution of the corresponding data process (see section 3.1.3). The basic idea of multistage stochastic programming is that if we are currently at a state of the system at stage $t$, represented by a node of the scenario tree, then our
decision at that node is based on our knowledge about the next possible realizations of the data process, which are represented by its children nodes at stage \( t+1 \). In the risk neutral approach we optimize the corresponding conditional expectation of the objective function. This allows to write the associated dynamic programming equations. This idea can be extended to optimization of a risk measure conditional on a current state of the system. We are going to discuss now such construction in detail.

As in section 3.1.3, we denote by \( \Omega \) the set of all nodes at stage \( t = 1, \ldots, T \), by \( K_t := |\Omega_t| \) the cardinality of \( \Omega_t \) and by \( C_a \) the set of children nodes of a node \( a \) of the tree. Note that \( \{C_a\}_{a \in \Omega_t} \) forms a partition of the set \( \Omega_{t+1} \), i.e., \( C_a \cap C_{a'} = \emptyset \) if \( a \neq a' \) and \( \Omega_{t+1} = \bigcup_{a \in \Omega_t} C_a, t = 1, \ldots, T-1 \). With the set \( \Omega_T \) we associate sigma algebra \( \mathcal{F}_T \) of all its subsets. Let \( \mathcal{F}_{T-1} \) be the subalgebra of \( \mathcal{F}_T \) generated by sets \( C_a, a \in \Omega_{T-1} \), i.e., these sets form the set of elementary events of \( \mathcal{F}_{T-1} \) (recall that \( \{C_a\}_{a \in \Omega_{T-1}} \) forms a partition of \( \Omega_T \)). By this construction there is a one-to-one correspondence between elementary events of \( \mathcal{F}_{T-1} \) and the set \( \Omega_{T-1} \) of nodes at stage \( T-1 \). By continuing this process we construct a sequence of sigma algebras \( \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T \) (such a sequence of nested sigma algebras is called filtration). Note that \( \mathcal{F}_1 \) corresponds to the unique root node and hence \( \mathcal{F}_1 = \{\emptyset, \Omega_T\} \). In this construction there is a one-to-one correspondence between nodes of \( \Omega_T \) and elementary events of the sigma algebra \( \mathcal{F}_t \), and hence we can identify every node \( a \in \Omega_t \) with an elementary event of \( \mathcal{F}_t \). By taking all children of every node of \( C_a \) at later stages, we eventually can identify with \( C_a \) a subset of \( \Omega_T \).

Suppose, further, that there is a probability distribution defined on the scenario tree. As it was discussed in section 3.1.3, such probability distribution can be defined by introducing conditional probabilities of going from a node of the tree to its children nodes. That is, with a node \( a \in \Omega_t \) is associated a probability vector\(^{13} \) \( \rho^a \in \mathbb{R}^{|C_a|} \) of conditional probabilities of moving from \( a \) to nodes of \( C_a \). Equipped with probability vector \( \rho^a \), the set \( C_a \) becomes a probability space, with the corresponding sigma algebra of all subsets of \( C_a \), and any function \( Z : C_a \to \mathbb{R} \) can be viewed as a random variable. Since the space of functions \( Z : C_a \to \mathbb{R} \) can be identified with the space \( \mathbb{R}^{|C_a|} \), we identify such random variable \( Z \) with an element of the vector space \( \mathbb{R}^{|C_a|} \). With every \( Z \in \mathbb{R}^{|C_a|} \) is associated the expectation \( E_{\rho^a}[Z] \), which can be considered as a conditional expectation given that we are currently at node \( a \).

Now with every node \( a \) at stage \( t = 1, \ldots, T-1 \) we associate a risk measure \( \rho^a(Z) \) defined on the space of functions \( Z : C_a \to \mathbb{R} \), that is we choose a family of risk measures

\[
\rho^a : \mathbb{R}^{|C_a|} \to \mathbb{R}, \quad a \in \Omega_t, \quad t = 1, \ldots, T-1.
\]  

(6.219)

Of course, there are many ways how such risk measures can be defined. For instance, for a given probability distribution on the scenario tree, we can use conditional expectations

\[
\rho^a(Z) := E_{\rho^a}[Z], \quad a \in \Omega_t, \quad t = 1, \ldots, T-1.
\]  

(6.220)

Such choice of risk measures \( \rho^a \) leads to the risk neutral formulation of a corresponding multistage stochastic program. For a risk averse approach we can use any class of coherent

---

\(^{13}\) A vector \( p = (p_1, \ldots, p_n) \in \mathbb{R}^n \) is said to be a probability vector if all its components \( p_i \) are nonnegative and \( \sum_{i=1}^n p_i = 1 \). If \( Z = (Z_1, \ldots, Z_n) \in \mathbb{R}^n \) is viewed as a random variable, then its expectation with respect to \( p \) is \( E_p[Z] = \sum_{i=1}^n p_i Z_i \).
risk measures discussed in section 6.3.2, as, for example,

\[ \rho^a[Z] := \inf_{t \in \mathbb{R}} \{ t + \lambda_a^{-1} \mathbb{E}_{\rho^a}[Z - t]_+ \}, \quad \lambda_a \in (0, 1), \]  

(6.221)
corresponding to AV@R risk measure and

\[ \rho^c[Z] := \mathbb{E}_{\rho^c}[Z] + c_a \mathbb{E}_{\rho^c}[Z - \mathbb{E}_{\rho^c}[Z]]_+, \quad c_a \in [0, 1], \]  

(6.222)
corresponding to the absolute semideviation risk measure.

Since \( \Omega_{t+1} \) is the union of the disjoint sets \( C_a, a \in \Omega_t \), we can write \( \mathbb{R}^{K_{t+1}} \) as the Cartesian product of the spaces \( \mathbb{R}^{|C_a|}, a \in \Omega_t \). That is, \( \mathbb{R}^{K_{t+1}} = \mathbb{R}^{|C_{a_1}|} \times \cdots \times \mathbb{R}^{|C_{a_{K_t}}|} \),

where \( \{a_1, \ldots, a_{K_t}\} = \Omega_t \). Define the mappings

\[ \rho_{t+1} := (\rho^{a_1}, \ldots, \rho^{a_{K_t}}) : \mathbb{R}^{K_{t+1}} \rightarrow \mathbb{R}^{K_t}, \quad t = 1, \ldots, T - 1, \]  

(6.223)
associated with risk measures \( \rho^a \). Recall that the set \( \Omega_{t+1} \) of nodes at stage \( t+1 \) is identified with the set of elementary events of sigma algebra \( \mathcal{F}_{t+1} \), and its sigma subalgebra \( \mathcal{F}_t \) is generated by sets \( C_a, a \in \Omega_t \).

We denote by \( \mathcal{Z}_T \) the space of all functions \( \Omega_T \rightarrow \mathbb{R} \). As it was mentioned before we can identify every such function with a vector of the space \( \mathbb{R}^{K_T} \), i.e., the space \( \mathcal{Z}_T \) can be identified with the space \( \mathbb{R}^{K_T} \). We have that a function \( \Omega_T \rightarrow \mathbb{R} \) is \( \mathcal{F}_{T-1} \)-measurable iff it is constant on every set \( C_a, a \in \Omega_{T-1} \). We denote by \( \mathcal{Z}_{T-1} \) the subspace of \( \mathcal{Z}_T \) formed by \( \mathcal{F}_{T-1} \)-measurable functions. The space \( \mathcal{Z}_{T-1} \) can be identified with \( \mathbb{R}^{K_{T-1}} \). And so on, we can construct a sequence \( \mathcal{Z}_t, t = 1, \ldots, T, \) of spaces of \( \mathcal{F}_t \)-measurable functions \( \Omega_T \rightarrow \mathbb{R} \) such that \( \mathcal{Z}_1 \subset \cdots \subset \mathcal{Z}_T \) and each \( \mathcal{Z}_t \) can be identified with the space \( \mathbb{R}^{K_t} \). Recall that \( K_1 = 1 \), and hence \( \mathcal{Z}_1 \) can be identified with \( \mathbb{R} \). We view the mapping \( \rho_{t+1} \), defined in (6.223), as a mapping from the space \( \mathcal{Z}_{t+1} \) into the space \( \mathcal{Z}_t \). Conversely, with any mapping \( \rho_{t+1} : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t \) we can associate a family of risk measures of the form (6.219).

We say that a mapping \( \rho_{t+1} : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t \) is a conditional risk mapping if it satisfies the following conditions\(^\dagger\):

**(R’1) Convexity:**

\[ \rho_{t+1}(\alpha Z + (1 - \alpha) Z') \preceq \alpha \rho_{t+1}(Z) + (1 - \alpha) \rho_{t+1}(Z'), \]

for any \( Z, Z' \in \mathcal{Z}_{t+1} \) and \( \alpha \in [0, 1] \).

**(R’2) Monotonicity:** If \( Z, Z' \in \mathcal{Z}_{t+1} \) and \( Z \succeq Z' \), then \( \rho_{t+1}(Z) \succeq \rho_{t+1}(Z') \).

**(R’3) Translation Equivariance:** If \( Y \in \mathcal{Z}_t \) and \( Z \in \mathcal{Z}_{t+1} \), then \( \rho_{t+1}(Z + Y) = \rho_{t+1}(Z) + Y \).

**(R’4) Positive homogeneity:** If \( \alpha \geq 0 \) and \( Z \in \mathcal{Z}_{t+1} \), then \( \rho_{t+1}(\alpha Z) = \alpha \rho_{t+1}(Z) \).

\(^\dagger\)For \( Z_1, Z_2 \in \mathcal{Z}_t \), the inequality \( Z_2 \succeq Z_1 \) is understood componentwise.

It is straightforward to see that conditions (R’1), (R’2) and (R’4) hold iff the corresponding conditions (R1), (R2) and (R4), defined in section 6.3, hold for every risk measure \( \rho^a \)
associated with \( \rho_{t+1} \). Also by construction of \( \rho_{t+1} \), we have that condition (R'3) holds iff condition (R3) holds for all \( \rho^a \). That is, \( \rho_{t+1} \) is a conditional risk mapping iff every corresponding risk measure \( \rho^a \) is a coherent risk measure.

By Theorem 6.4 with each coherent risk measure \( \rho^a, a \in \Omega_t \), is associated a set \( \mathfrak{A}(a) \) of probability measures (vectors) such that

\[
\rho^a(Z) = \max_{p \in \mathfrak{A}(a)} E_p[Z]. \tag{6.224}
\]

Here \( Z \in \mathbb{R}^{K_{t+1}} \) is a vector corresponding to function \( Z : \Omega_{t+1} \rightarrow \mathbb{R} \), and \( \mathfrak{A}(a) = \mathfrak{A}_{t+1}(a) \) is a closed convex set of probability vectors \( p \in \mathbb{R}^{K_{t+1}} \) such that \( p_k = 0 \) if \( k \in \Omega_{t+1} \setminus C_a \), i.e., all probability measures of \( \mathfrak{A}_{t+1}(a) \) are supported on the set \( C_a \).

We can now represent the corresponding conditional risk mapping \( \rho_{t+1} \) as a maximum of conditional expectations as follows. Let \( \nu = (\nu_a)_{a \in \Omega_t} \) be a probability distribution on \( \Omega_t \), assigning positive probability \( \nu_a \) to every \( a \in \Omega_t \), and define

\[
\mathcal{C}_{t+1} := \left\{ \mu = \sum_{a \in \Omega_t} \nu_a p^a : p^a \in \mathfrak{A}_{t+1}(a) \right\}. \tag{6.225}
\]

It is not difficult to see that \( \mathcal{C}_{t+1} \subseteq \mathbb{R}^{K_{t+1}} \) is a convex set of probability vectors. Moreover, since each \( \mathfrak{A}_{t+1}(a) \) is compact, the set \( \mathcal{C}_{t+1} \) is also compact and hence is closed. Consider a probability distribution (measure) \( \mu = \sum_{a \in \Omega_t} \nu_a p^a \in \mathcal{C}_{t+1} \). We have that, for \( a \in \Omega_t \), the corresponding conditional distribution given the event \( C_a \) is \( p^a \), and if

\[
E_{\mu}[Z|\mathcal{F}_t](a) = \nu_a E_{p^a}[Z], \quad Z \in \mathcal{Z}_{t+1}. \tag{6.226}
\]

It follows then by (6.224) that

\[
\rho_{t+1}(Z) = \max_{\mu \in \mathcal{C}_{t+1}} E_{\mu}[Z|\mathcal{F}_t]. \tag{6.227}
\]

where the maximum on the right hand side of (6.227) is taken pointwise in \( a \in \Omega_t \). That is, formula (6.227) means that

\[
[\rho_{t+1}(Z)](a) = \max_{p \in \mathfrak{A}_{t+1}(a)} E_p[Z], \quad Z \in \mathcal{Z}_{t+1}, \quad a \in \Omega_t. \tag{6.228}
\]

Note that, in this construction, choice of the distribution \( \nu \) is arbitrary and any distribution of \( \mathcal{C}_{t+1} \) agrees with the distribution \( \nu \) on \( \Omega_t \).

We are ready now to give a formulation of risk averse multistage programs. For a sequence \( \rho_{t+1} : \mathcal{Z}_{t+1} \rightarrow \mathcal{Z}_t, t = 1, \ldots, T - 1 \), of conditional risk mappings consider the following risk averse formulation analogous to the nested risk neutral formulation (3.1):

\[
\begin{align*}
\min_{x_1 \in \mathcal{X}_1} & \quad f_1(x_1) + p_2 \left\{ \inf_{x_2 \in \mathcal{X}_2(x_1, \omega)} f_2(x_2, \omega) + \ldots \right. \\
& \quad + \rho_{t-1} \left[ \inf_{x_{t-1} \in \mathcal{X}_{t-1}(x_{t-2}, \omega)} f_{t-1}(x_{t-1}, \omega) \right] \\
& \quad + \rho_T \left[ \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \omega)} f_T(x_T, \omega) \right]. \tag{6.229}
\end{align*}
\]

\(^{15}\)Recall that the conditional expectation \( E_{\mu} \cdot |\mathcal{F}_t \) is a mapping from \( \mathcal{Z}_{t+1} \) into \( \mathcal{Z}_t \).
Here $\omega$ is an element of $\Omega := \Omega_T$, the objective functions $f_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}$ are real valued functions and $X_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}$, $t = 2, \ldots, T$, are multifunctions such that $f_t(x_{t} \cdot \cdot \cdot)$ and $X_t(x_{t-1} \cdot \cdot \cdot)$ are $F_t$-measurable for all $x_t$ and $x_{t-1}$. Note that if the corresponding risk measures $\rho^a$ are defined as conditional expectations (6.220), then the above multistage problem (6.229) coincides with the risk neutral multistage problem (3.1).

There are several ways how the above nested formulation (6.229) can be formalized. Similarly to (3.3), we can write problem (6.229) in the form

$$
\text{Min}_{x_1, x_2, \ldots, x_T} f_1(x_1) + \rho_2 \left[ f_2(x_2(\omega), \omega) + \ldots + \rho_{T-1} \left[ f_{T-1} (x_{T-1}(\omega), \omega) + \rho_T[f_T(x_T(\omega), \omega)] \right] \right]
$$

(6.230)

s.t. $x_1 \in X_1, \ x_t(\omega) \in X_t(x_{t-1}(\omega), \omega), \ t = 2, \ldots, T$.

Optimization in (6.230) is performed over functions $x_t : \Omega \to \mathbb{R}$, $t = 1, \ldots, T$, satisfying the corresponding constraints, which imply that each $x_t(\omega)$ is $F_t$-measurable and hence each $f_t(x_t(\omega), \omega)$ is $F_t$-measurable. The requirement for $x_t(\omega)$ to be $F_t$-measurable is another way of formulating the nonanticipativity constraints. Therefore, it can be viewed that the optimization in (6.230) is performed over feasible policies.

Consider the function $\rho : Z_1 \times \cdots \times Z_T \to \mathbb{R}$ defined as

$$
\rho(Z_1, \ldots, Z_T) := Z_1 + \rho_2 \left[ Z_2 + \cdots + \rho_{T-1} \left[ Z_{T-1} + \rho_T[Z_T] \right] \right].
$$

(6.231)

By condition (R’3) we have that

$$
\rho_{T-1} \left[ Z_{T-1} + \rho_T[Z_T] \right] = \rho_{T-1} \circ \rho_T[Z_{T-1} + Z_T].
$$

By continuing this process we obtain that

$$
\bar{\rho}(Z_1, \ldots, Z_T) = \bar{\rho}(Z_1 + \cdots + Z_T),
$$

(6.232)

where $\bar{\rho} := \rho_2 \circ \cdots \circ \rho_T$. We refer to $\bar{\rho}$ as the composite risk measure. That is,

$$
\bar{\rho}(Z_1 + \cdots + Z_T) = Z_1 + \rho_2 \left[ Z_2 + \cdots + \rho_{T-1} \left[ Z_{T-1} + \rho_T[Z_T] \right] \right],
$$

(6.233)

defined for $Z_t \in Z_t$, $t = 1, \ldots, T$. Recall that $Z_1$ is identified with $\mathbb{R}$, and hence $Z_1$ is a real number and $\bar{\rho} : Z_T \to \mathbb{R}$ is a real valued function. Conditions (R’1)–(R’4) imply that $\bar{\rho}$ is a coherent risk measure.

As above we have that since $f_{T-1}(x_{T-1}(\omega), \omega)$ is $F_{T-1}$-measurable, it follows by condition (R’3) that

$$
f_{T-1}(x_{T-1}(\omega), \omega) + \rho_T[f_T(x_T(\omega), \omega)] = \rho_T \left[ f_{T-1}(x_{T-1}(\omega), \omega) + f_T(x_T(\omega), \omega) \right].
$$

Continuing this process backwards we obtain that the objective function of (6.230) can be formulated using the composite risk measure. That is, problem (6.230) can be written in the form

$$
\text{Min}_{x_1, x_2, \ldots, x_T} \bar{\rho} \left[ f_1(x_1) + f_2(x_2(\omega), \omega) + \cdots + f_T(x_T(\omega), \omega) \right]
$$

s.t. $x_1 \in X_1, \ x_t(\omega) \in X_t(x_{t-1}(\omega), \omega), \ t = 2, \ldots, T$.

(6.234)
6.7. Multistage Risk Averse Optimization

If the conditional risk mappings are defined as the respective conditional expectations, then the composite risk measure \( \bar{\rho} \) becomes the corresponding expectation operator, and (6.234) coincides with the multistage program written in the form (3.3). Unfortunately, it is not easy to write the composite risk measure \( \bar{\rho} \) in a closed form even for relatively simple conditional risk mappings other than conditional expectations.

An alternative approach to formalize the nested formulation (6.229) is to write dynamic programming equations. That is, for the last period \( T \) we have

\[
Q_T(x_{T-1}, \omega) := \inf_{x_T \in X_T(x_{T-1}, \omega)} f_T(x_T, \omega),
\]

(6.235)

\[
Q_T(x_{T-1}, \omega) := \rho_T(Q_T(x_{T-1}, \omega)),
\]

(6.236)

and for \( t = T-1, \ldots, 2 \), we recursively apply the conditional risk measures

\[
Q_t(x_{t-1}, \omega) := \rho_t [Q_t(x_{t-1}, \omega)],
\]

(6.237)

where

\[
Q_t(x_{t-1}, \omega) := \inf_{x_t \in X_t(x_{t-1}, \omega)} \left\{ f_t(x_t, \omega) + Q_{t+1}(x_t, \omega) \right\}.
\]

(6.238)

Of course, equations (6.237) and (6.238) can be combined into one equation\(^{16}\):

\[
Q_t(x_{t-1}, \omega) = \inf_{x_t \in X_t(x_{t-1}, \omega)} \left\{ f_t(x_t, \omega) + \rho_{t+1} [Q_{t+1}(x_t, \omega)] \right\}.
\]

(6.239)

Finally, at the first stage we solve the problem

\[
\min_{x_1 \in X_1} f_1(x_1) + \rho_2 [Q_2(x_1, \omega)].
\]

(6.240)

It is important to emphasize that conditional risk mappings \( \rho_t(Z) \) are defined on real valued functions \( Z(\omega) \). Therefore it is implicitly assumed in the above equations that the cost-to-go (value) functions \( Q_t(x_{t-1}, \omega) \) are real valued. In particular, this implies that the considered problem should have relatively complete recourse. Also in the above development of the dynamic programming equations the monotonicity condition (R'2) plays a crucial role, because only then we can move the optimization under the risk operation.

**Remark 21.** By using representation (6.227), we can write the dynamic programming equations (6.239) in the form

\[
Q_t(x_{t-1}, \omega) = \inf_{x_t \in X_t(x_{t-1}, \omega)} \left\{ f_t(x_t, \omega) + \sup_{\mu \in \mathcal{E}_{t+1}} \mathbb{E}_\mu \left[ Q_{t+1}(x_t) | \mathcal{F}_t \right] (\omega) \right\}.
\]

(6.241)

Note that the left and right hand side functions in (6.241) are \( \mathcal{F}_t \)-measurable, and hence this equation can be written in terms of \( a \in \Omega_t \) instead of \( \omega \in \Omega \). Recall that every \( \mu \in \mathcal{E}_{t+1} \) is representable in the form

\[
\mu = \sum_{a \in \Omega_t} \nu_a \delta^a \quad \text{(see (6.225)) and that}
\]

\[
\mathbb{E}_\mu \left[ Q_{t+1}(x_t) | \mathcal{F}_t \right] (a) = \mathbb{E}_{\mu^a} [Q_{t+1}(x_t)], \quad a \in \Omega_t.
\]

(6.242)

\(^{16}\)With some abuse of the notation we write \( Q_{t+1}(x_t, \omega) \) for the value of \( Q_{t+1}(x_t) \) at \( \omega \in \Omega \), and \( \rho_{t+1} [Q_{t+1}(x_t, \omega)] \) for \( \rho_{t+1} [Q_{t+1}(x_t)] (\omega) \).
We say that the problem is convex if the functions $f_t(\cdot, \omega)$, $Q_t(\cdot, \omega)$ and the sets $X_t(x_{t-1}, \omega)$ are convex for every $\omega \in \Omega$ and $t = 1, \ldots, T$. If the problem is convex, then (since the set $C_{t+1}$ is convex compact) the ‘inf’ and ‘sup’ operators on the right hand side of (6.241) can be interchanged to obtain a dual problem, and for a given $x_{t-1}$ and every $\omega \in \Omega_t$ the dual problem has an optimal solution $\bar{\mu}^* \in \mathfrak{A}_{t+1}(a)$. Consequently, for $\bar{\mu}_{t+1} := \sum_{a \in \Omega_t} \nu_a \bar{\mu}^a$ an optimal solution of the original problem and the corresponding cost-to-go functions satisfy the following dynamic programming equations:

$$Q_t(x_{t-1}, \omega) = \inf_{x_t \in X_t(x_{t-1}, \omega)} \left\{ f_t(x_t, \omega) + \mathbb{E}_{\bar{\mu}_{t+1}} [Q_{t+1}(x_t)|\mathcal{F}_t](\omega) \right\}. \quad (6.243)$$

Moreover, it is possible to choose the “worst case” distributions $\bar{\mu}_{t+1}$ in a consistent way, i.e., such that each $\bar{\mu}_{t+1}$ coincides with $\bar{\mu}_1$ on $\mathcal{F}_t$. That is, consider the first-stage problem (6.240). We have that (recall that at the first stage there is only one node, $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{C}_2 = \mathfrak{A}_2$)

$$\rho[Q_2(x_1)] = \sup_{\mu \in \mathcal{C}_2} \mathbb{E}_{\mu}[Q_2(x_1)|\mathcal{F}_1] = \sup_{\mu \in \mathcal{C}_2} \mathbb{E}_{\mu}[Q_2(x_1)]. \quad (6.244)$$

By convexity and since $\mathcal{C}_2$ is compact, we have that there is $\bar{\rho}_2 \in \mathcal{C}_2$ (an optimal solution of the dual problem) such that the optimal value of the first stage problem is equal to the optimal value and the set of optimal solutions of the first stage problem is contained in the set of optimal solutions of the problem

$$\min_{x_1 \in \mathcal{X}_1} \mathbb{E}_{\bar{\rho}_2}[Q_2(x_1)]. \quad (6.245)$$

Let $\bar{x}_1$ be an optimal solution of the first stage problem. Then we can choose $\bar{\mu}_2 \in \mathcal{C}_2$, of the form $\bar{\mu}_2 := \sum_{a \in \mathcal{A}_2} \nu_a \bar{\rho}^a$, such that equation (6.243) holds with $t = 2$ and $x_1 = \bar{x}_1$. Moreover, we can take the probability measure $\nu = (\nu_a)_{a \in \mathcal{A}_2}$ to be the same as $\bar{\rho}_2$, and hence to ensure that $\bar{\mu}_2$ coincides with $\bar{\mu}_2$ on $\mathcal{F}_2$. Next, for every node $a \in \Omega_2$ choose a corresponding (second-stage) optimal solution and repeat the construction to produce an appropriate $\bar{\mu}_3 \in \mathcal{C}_3$, and so on for later stages.

In that way, assuming existence of optimal solutions, we can construct a probability distribution $\bar{\mu}_2, \ldots, \bar{\mu}_T$ on the considered scenario tree such that the obtained multistage problem, of the standard form (3.1), has the same cost-to-go (value) functions as the original problem (6.229) and has an optimal solution which also is an optimal solution of the problem (6.249) (in that sense the obtained multistage problem, driven by dynamic programming equations (6.243), is “almost equivalent” to the original problem).

**Remark 22.** Let us define, for every node $a \in A_t$, $t = 1, \ldots, T - 1$, the corresponding set $\mathfrak{A}(a) = \mathfrak{A}_{t+1}(a)$ to be the set of all probability measures (vectors) on the set $C_a$ (recall that $C_a \subset \Omega_{t+1}$ is the set of children nodes of $a$, and that all probability measures of $\mathfrak{A}_{t+1}(a)$ are supported on $C_a$). Then the maximum on the right hand side of (6.224) is attained at a measure of mass one at a point of the set $C_a$. Consequently, by (6.242), for such choice of the sets $\mathfrak{A}_{t+1}(a)$ the dynamic programming equations (6.241) can be written as

$$Q_t(x_{t-1}, a) = \inf_{x_t \in X_t(x_{t-1}, a)} \left\{ f_t(x_t, a) + \max_{\omega \in C_a} Q_{t+1}(x_t, \omega) \right\}, \quad a \in \Omega_t. \quad (6.246)$$

It is interesting to note (see Remark 21, page 315) that if the problem is convex, then it is possible to construct a probability distribution (on the considered scenario tree),
defined by a sequence $\mu_t$, $t = 2, \ldots, T$, of consistent probability distributions, such that the obtained (risk neutral) multistage program is “almost equivalent” to the “min-max” formulation (6.246).

6.7.2 Conditional Risk Mappings

In this section we discuss a general concept of conditional risk mappings which can be applied to a risk averse formulation of multistage programs. The material of this section can be considered as an extension to an infinite dimensional setting the developments presented in the previous section. Similarly to the presentation of coherent risk measures, given in section 6.3, we use the framework of $L_p$ spaces, $p \in [1, +\infty)$. That is, let $\Omega$ be a sample space equipped with sigma algebras $\mathcal{F}_1 \subset \mathcal{F}_2$ (i.e., $\mathcal{F}_1$ is subalgebra of $\mathcal{F}_2$) and a probability measure $P$ on $(\Omega, \mathcal{F}_2)$. Consider the spaces $\mathcal{Z}_1 := L_p(\Omega, \mathcal{F}_1, P)$ and $\mathcal{Z}_2 := L_p(\Omega, \mathcal{F}_2, P)$. Since $\mathcal{F}_1$ is a subalgebra of $\mathcal{F}_2$, it follows that $\mathcal{Z}_1 \subset \mathcal{Z}_2$.

We say that a mapping $\rho : \mathcal{Z}_2 \to \mathcal{Z}_1$ is a conditional risk mapping if it satisfies the following conditions:

(R1) **Convexity:**

$$\rho(\alpha Z + (1 - \alpha)Z') \leq \alpha \rho(Z) + (1 - \alpha)\rho(Z'),$$

for any $Z, Z' \in \mathcal{Z}_2$ and $\alpha \in [0, 1]$.

(R2) **Monotonicity:** If $Z, Z' \in \mathcal{Z}_2$ and $Z \geq Z'$, then $\rho(Z) \geq \rho(Z')$.

(R3) **Translation Equivariance:** If $Y \in \mathcal{Z}_1$ and $Z \in \mathcal{Z}_2$, then

$$\rho(Y + Z) = \rho(Z) + Y.$$

(R4) **Positive homogeneity:** If $\alpha \geq 0$ and $Z \in \mathcal{Z}_2$, then $\rho(\alpha Z) = \alpha \rho(Z)$.

The above conditions coincide with the respective conditions of the previous section which were defined in a finite dimensional setting. If the sigma algebra $\mathcal{F}_1$ is trivial, i.e., $\mathcal{F}_1 = \{\emptyset, \Omega\}$, then the space $\mathcal{Z}_1$ can be identified with $\mathbb{R}$, and conditions (R1)–(R4) define a coherent risk measure. Examples of coherent risk measures, discussed in section 6.3.2, have conditional risk mapping analogues which are obtained by replacing the expectation operator with the corresponding conditional expectation $\mathbb{E}[-|\mathcal{F}_1]$ operator. Let us look at some examples.

**Conditional expectation.** In itself, the conditional expectation mapping $\mathbb{E}[-|\mathcal{F}_1] : \mathcal{Z}_2 \to \mathcal{Z}_1$ is a conditional risk mapping. Indeed, for any $p \geq 1$ and $Z \in L_p(\Omega, \mathcal{F}_2, P)$ we have by Jensen inequality that $\mathbb{E}[-|\mathcal{F}_1] \geq |\mathbb{E}[Z|\mathcal{F}_1]|^p$, and hence

$$\int_{\Omega} |\mathbb{E}[Z|\mathcal{F}_1]|^p dP \leq \int_{\Omega} \mathbb{E}[-|\mathcal{F}_1]| dP = \mathbb{E}[Z] < +\infty. \quad (6.247)$$

This shows that, indeed, $\mathbb{E}[-|\mathcal{F}_1]$ maps $\mathcal{Z}_2$ into $\mathcal{Z}_1$. The conditional expectation is a linear operator, and hence conditions (R1) and (R4) follow. The monotonicity condition (R2) also clearly holds, and condition (R3) is a property of conditional expectation.
Conditional AV@R. An analogue of AV@R risk measure can be defined as follows. Let \( Z_i := L_1(\Omega, F_i, P), i = 1, 2 \). For \( \alpha \in (0, 1) \) define mapping \( AV@R_\alpha(\cdot | F_1) : Z_2 \rightarrow Z_1 \) as

\[
[AV@R_\alpha(Z|F_1)](\omega) := \inf_{Y \in Z_1} \left\{ Y(\omega) + \alpha^{-1}E[[Z - Y]_+]|F_1](\omega) \right\}, \quad \omega \in \Omega. \tag{6.248}
\]

It is not difficult to verify that, indeed, this mapping satisfies conditions (R’1)–(R’4). Similarly to (6.69), for \( \beta \in [0, 1] \) and \( \alpha \in (0, 1) \), we can also consider the following conditional risk mapping

\[
\rho_{\alpha, \beta|F_1}(Z) := (1 - \beta)E[Z|F_1] + \beta AV@R_\alpha(Z|F_1). \tag{6.249}
\]

Of course, the above conditional risk mapping \( \rho_{\alpha, \beta|F_1} \) corresponds to the coherent risk measure \( \rho_{\alpha, \beta}(Z) := (1 - \beta)E[Z] + \beta AV@R_\alpha(Z) \).

**Conditional mean-upper-semideviation.** An analogue of the mean-upper-semideviation risk measure (of order \( p \)) can be constructed as follows. Let \( Z_i := L_p(\Omega, F_i, P), i = 1, 2 \). For \( c \in [0, 1] \) define

\[
\rho_{c|F_1}(Z) := E[Z|F_1] + c \left( E \left[ \left[ Z - E[Z|F_1] \right]^p F_1 \right] \right)^{1/p}. \tag{6.250}
\]

In particular, for \( p = 1 \) this gives an analogue of the absolute semideviation risk measure.

In the discrete case of scenario tree formulation (discussed in the previous section) the above examples correspond to taking the same respective risk measure at every node of the considered tree at stage \( t = 1, \ldots, T \). □

Consider a conditional risk mapping \( \rho : Z_2 \rightarrow Z_1 \). With a set \( A \in F_1 \), such that \( P(A) \neq 0 \), we associate the function

\[
\rho_A(Z) := E[\rho(Z)|A], \quad Z \in Z_2, \tag{6.251}
\]

where \( E[Y|A] := \frac{1}{P(A)} \int_A Y dP \) denotes the conditional expectation of random variable \( Y \in Z_1 \) given event \( A \in F_1 \). Clearly conditions (R’1)–(R’4) imply that the corresponding conditions (R1)–(R4) hold for \( \rho_A \), and hence \( \rho_A \) is a coherent risk measure defined on the space \( Z_2 = L_p(\Omega, F_2, P) \). Moreover, for any \( B \in F_1 \) we have by (R’3) that

\[
\rho_A(Z + \alpha 1_B) := E[\rho(Z) + \alpha 1_B|A] = \rho_A(Z) + \alpha P(B|A), \quad \forall \alpha \in \mathbb{R}, \tag{6.252}
\]

where \( P(B|A) = P(B \cap A)/P(A) \).

Since \( \rho_A \) is a coherent risk measure, by Theorem 6.4 it can be represented in the following form

\[
\rho_A(Z) = \sup_{\zeta \in \mathfrak{A}(A)} \int_{\Omega} \zeta(\omega)Z(\omega)dP(\omega), \tag{6.253}
\]

for some set \( \mathfrak{A}(A) \subset L_q(\Omega, F_2, P) \) of probability density functions. Let us make the following observation.

- Each density \( \zeta \in \mathfrak{A}(A) \) is supported on the set \( A \).
Indeed, for any \( B \in \mathcal{F}_1 \), such that \( P(B \cap A) = 0 \), and any \( \alpha \in \mathbb{R} \), we have by (6.252) that \( \rho_A(Z + \alpha 1_B) = \rho_A(Z) \). On the other hand, if there exists \( \zeta \in \mathfrak{A}(A) \) such that \( \int_B \zeta dP > 0 \), then it follows from (6.253) that \( \rho_A(Z + \alpha 1_B) \) tends to \( +\infty \) as \( \alpha \to +\infty \).

Similarly to (6.227), we show now that a conditional risk mapping can be represented as a maximum of a family of conditional expectations. We consider a situation where the subalgebra \( \mathcal{F}_1 \) has a countable number of elementary events. That is, there is a (countable) partition \( \{A_i\}_{i \in \mathbb{N}} \), of the sample space \( \Omega \) which generates \( \mathcal{F}_1 \), i.e., \( \bigcup_{i \in \mathbb{N}} A_i = \Omega \), the sets \( A_i \), \( i \in \mathbb{N} \), are disjoint and form the family of elementary events of sigma algebra \( \mathcal{F}_1 \). Since \( \mathcal{F}_1 \) is a subalgebra of \( \mathcal{F}_2 \), we have of course that \( A_i \in \mathcal{F}_2 \), \( i \in \mathbb{N} \). We also have that a function \( Z : \Omega \to \mathbb{R} \) is \( \mathcal{F}_1 \)-measurable iff it is constant on every set \( A_i \), \( i \in \mathbb{N} \).

Consider a conditional risk mapping \( \rho : Z_\Omega \to Z_\Omega \). Let

\[ \mathfrak{N} := \{i \in \mathbb{N} : P(A_i) \neq 0\} \]

and \( \rho_{A_i}, i \in \mathfrak{N} \), be the corresponding coherent risk measures defined in (6.251). By (6.253) with every \( \rho_{A_i}, i \in \mathfrak{N} \), is associated set \( \mathfrak{A}(A_i) \) of probability density functions, supported on the set \( A_i \), such that

\[ \rho_{A_i}(Z) = \sup_{\zeta \in \mathfrak{A}(A_i)} \int_{\Omega} \zeta(\omega)Z(\omega)dP(\omega). \]  

(6.254)

Now let \( \nu = (\nu_i)_{i \in \mathbb{N}} \) be a probability distribution (measure) on \( (\Omega, \mathcal{F}_1) \), assigning probability \( \nu_i \) to the event \( A_i \), \( i \in \mathbb{N} \). Assume that \( \nu \) is such that \( \nu(A_i) = 0 \) iff \( P(A_i) = 0 \) (i.e., \( \mu \) is absolutely continuous with respect to \( P \) and \( P \) is absolutely continuous with respect to \( \nu \) on \( (\Omega, \mathcal{F}_1) \)), otherwise the probability measure \( \nu \) is arbitrary. Define the following family of probability measures on \( (\Omega, \mathcal{F}_2) \):

\[ \mathcal{C} := \left\{ \mu = \sum_{i \in \mathfrak{N}} \nu_i \mu_i : d\mu_i = \zeta_i dP, \ z_i \in \mathfrak{A}(A_i), \ i \in \mathfrak{N} \right\}. \]

(6.255)

Note that since \( \sum_{i \in \mathfrak{N}} \nu_i = 1 \), every \( \mu \in \mathcal{C} \) is a probability measure. For \( \mu \in \mathcal{C} \), with respective densities \( \zeta_i \in \mathfrak{A}(A_i) \) and \( d\mu_i = \zeta_i dP \), and \( Z \in Z_2 \) we have that

\[ \mathbb{E}_\mu[Z|\mathcal{F}_1] = \sum_{i \in \mathfrak{N}} \mathbb{E}_{\mu_i}[Z|\mathcal{F}_1]. \]

(6.256)

Moreover, since \( \zeta_i \) is supported on \( A_i \),

\[ \mathbb{E}_{\mu_i}[Z|\mathcal{F}_1](\omega) = \begin{cases} \int_{A_i} Z\zeta_i dP, \quad &\text{if } \omega \in A_i, \\ 0, \quad &\text{otherwise}. \end{cases} \]

(6.257)

By the max-representations (6.254) it follows that for \( Z \in Z_2 \) and \( \omega \in A_i \),

\[ \sup_{\mu \in \mathcal{C}} \mathbb{E}_\mu[Z|\mathcal{F}_1](\omega) = \sup_{\zeta_i \in \mathfrak{A}(A_i)} \int_{A_i} Z\zeta_i dP = \rho_{A_i}(Z). \]

(6.258)

Also since \( [\rho(Z)](\cdot) \) is \( \mathcal{F}_1 \)-measurable, and hence is constant on every set \( A_i \), we have that \( [\rho(Z)](\omega) = \rho_{A_i}(Z) \) for every \( \omega \in A_i \), \( i \in \mathfrak{N} \). We obtain the following result.
**Proposition 6.40.** Let $Z_i := L_p(\Omega, \mathcal{F}_i, P)$, $i = 1, 2$, with $\mathcal{F}_1 \subset \mathcal{F}_2$, and let $\rho : Z_2 \to Z_1$ be a conditional risk mapping. Suppose that $\mathcal{F}_1$ has a countable number of elementary events. Then

$$\rho(Z) = \sup_{\mu \in \mathcal{E}} \mathbb{E}_\mu[Z|\mathcal{F}_1], \quad \forall Z \in Z_2,$$

where $\mathcal{E}$ is a family of probability measures on $(\Omega, \mathcal{F}_2)$, specified in (6.255), corresponding to a probability distribution $\nu$ on $(\Omega, \mathcal{F}_1)$.

### 6.7.3 Risk Averse Multistage Stochastic Programming

There are several ways how risk averse stochastic programming can be formulated in a multistage setting. We discuss now a nested formulation similar to the derivations of section 6.7.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T$ be a sequence of nested sigma algebras with $\mathcal{F}_1 = \{\emptyset, \Omega\}$ being trivial sigma algebra and $\mathcal{F}_T = \mathcal{F}$ (such sequence of sigma algebras is called a filtration). For $p \in [1, +\infty)$ let $\mathcal{Z}_i := L_p(\Omega, \mathcal{F}_i, P)$, $t = 1, \ldots, T$, be the corresponding sequence of spaces of $\mathcal{F}_t$-measurable and $p$-integrable functions, and $\rho_{t+1|\mathcal{F}_t} : \mathcal{Z}_{t+1} \to \mathcal{Z}_t$, $t = 1, \ldots, T-1$, be a selected family of conditional risk mappings. It is straightforward to verify that the composition

$$\rho_{t+1|\mathcal{F}_{t+1}} \circ \cdots \circ \rho_{T|\mathcal{F}_{T-1}} : \mathcal{Z}_T \xrightarrow{\rho_{T|\mathcal{F}_{T-2}}} \mathcal{Z}_{T-1} \xrightarrow{\rho_{T-1|\mathcal{F}_{T-2}}} \cdots \xrightarrow{\rho_{2|\mathcal{F}_1}} \mathcal{Z}_1,$$

is a family of probability measures on $(\Omega, \mathcal{F}_2)$, specified in (6.255), corresponding to a probability distribution $\nu$ on $(\Omega, \mathcal{F}_1)$.

Here $f_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}$ and $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}^{n_t}$, $t = 2, \ldots, T$, are such that $f_t(x_t, \cdot) \in \mathcal{Z}_t$ and $\mathcal{X}_t(x_{t-1}, \cdot)$ are $\mathcal{F}_t$-measurable for all $x_t$ and $x_{t-1}$.

As it was discussed in section 6.7.1, the above nested formulation (6.261) has two equivalent interpretations. Namely, it can be formulated as

$$\min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \rho_{2|\mathcal{F}_1} \left[ f_2(x_2, \omega) + \cdots + \rho_{T-1|\mathcal{F}_{T-2}} \left[ f_{T-1}(x_{T-1}, \omega) + \rho_{T|\mathcal{F}_{T-1}} \left[ f_T(x_T, \omega) \right] \right] \right].$$

Here $f_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}$ and $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \Omega \to \mathbb{R}^{n_t}$, $t = 2, \ldots, T$, are such that $f_t(x_t, \cdot) \in \mathcal{Z}_t$ and $\mathcal{X}_t(x_{t-1}, \cdot)$ are $\mathcal{F}_t$-measurable for all $x_t$ and $x_{t-1}$.
where the optimization is performed over \( F_t \)-measurable \( x_t : \Omega \rightarrow \mathbb{R} \), \( t = 1, \ldots, T \), satisfying the corresponding constraints, and such that \( f_t(x_t(\cdot)) \in Z_t \). Recall that the nonanticipativity is enforced here by the \( F_t \)-measurability of \( x_t(\cdot) \). By using the composite risk measure \( \bar{\rho} := \rho_2 \circ \cdots \circ \rho_T \), we can also write (6.262) in the form

\[
\begin{align*}
\min_{x_1, x_2, \ldots, x_T} & \quad \bar{\rho}[f_1(x_1) + f_2(x_2(\omega), \omega) + \cdots + f_T(x_T(\omega), \omega)] \\
\text{s.t.} & \quad x_1 \in \mathcal{X}_1, \quad x_t(\omega) \in \mathcal{X}_t(x_{t-1}(\omega), \omega), \quad t = 2, \ldots, T.
\end{align*}
\]  

(6.263)

Recall that for \( Z_t \in Z_t, t = 1, \ldots, T \),

\[
\bar{\rho}(Z_1 + \cdots + Z_T) = Z_1 + \rho_2[Z_1 \cdots + \rho_T[Z_T - 1 + \rho_T[Z_T - 1 + Z_T]]],
\]  

(6.264)

and that conditions (R’1)–(R’4) imply that \( \bar{\rho} : Z_T \rightarrow \mathbb{R} \) is a coherent risk measure.

Alternatively we can write the corresponding dynamic programming equations (compare with (6.235)–(6.240)):

\[
Q_T(x_{T-1}, \omega) = \inf_{x_T \in \mathcal{X}_T(x_{T-1}, \omega)} f_T(x_T, \omega),
\]  

(6.265)

\[
Q_t(x_{t-1}, \omega) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \omega)} \left\{ f_t(x_t, \omega) + Q_{t+1}(x_t, \omega) \right\}, \quad t = T - 1, \ldots, 2,
\]  

(6.266)

where

\[
Q_t(x_{t-1}, \omega) = \rho_t[x_{t-1}] Q_t(x_{t-1}, \omega), \quad t = T, \ldots, 2.
\]  

(6.267)

Finally, at the first stage we solve the problem

\[
\begin{align*}
\min_{x_1 \in \mathcal{X}_1} & \quad f_1(x_1) + \rho_2[x_1] Q_2(x_1, \omega),
\end{align*}
\]  

(6.268)

We need to ensure here that the cost-to-go functions are \( \rho \)-integrable, i.e., \( Q_t(x_{t-1}, \cdot) \in Z_t \) for \( t = 1, \ldots, T - 1 \) and all feasible \( x_{t-1} \).

In applications we are often deal with a data process represented by a sequence of random vectors \( \xi_1, \ldots, \xi_T \), say defined on a probability space \( (\Omega, \mathcal{F}, P) \). We can associate, with this data process filtration \( \mathcal{F}_t := \sigma(\xi_1, \ldots, \xi_t), t = 1, \ldots, T \), where \( \sigma(\xi_1, \ldots, \xi_t) \) denotes the smallest sigma algebra with respect to which \( \xi_t = (\xi_1, \ldots, \xi_t) \) is measurable. However, it is more convenient to deal with conditional risk mappings defined directly in terms of the data process rather that the respective sequence of sigma algebras. For example, consider

\[
\rho_t[\xi_{t-1}](Z) := (1 - \beta_t) \mathbb{E}[Z|\xi_{t-1}] + \beta_t \text{AV@R}_{\alpha_t}(Z|\xi_{t-1}), \quad t = 2, \ldots, T,
\]  

(6.269)

where

\[
\text{AV@R}_{\alpha_t}(Z|\xi_{t-1}) := \inf_{Y \in \mathcal{F}_{t-1}} \{ Y + \alpha_t^{-1} \mathbb{E}[|Z - Y| |\xi_{t-1}] \}.
\]  

(6.270)

Here \( \beta_t \in [0, 1] \) and \( \alpha_t \in (0, 1) \) are chosen constants, \( Z_t := L_1(\Omega, \mathcal{F}_t, P) \), where \( \mathcal{F}_t \) is the smallest filtration associated with the process \( \xi_t \), and the minimum on the right hand side of (6.270) is taken pointwise in \( \omega \in \Omega \). Compared with (6.248), the conditional AV@R is
defined in (6.270) in terms of the conditional expectation with respect to the history \(\xi_{t-1}\) of the data process rather than the corresponding sigma algebra \(\mathcal{F}_{t-1}\). We can also consider conditional mean-upper-semideviation risk mappings of the form:

\[
\rho_t(\xi_{t-1}) := E[Z|\xi_{t-1}] + \epsilon_t \left( E\left[\left( Z - E[Z|\xi_{t-1}] \right)_+^p | \xi_{t-1} \right] \right)^{1/p},
\]

(6.271) defined in terms of the data process. Note that with \(\rho_t(\xi_{t-1})\), defined in (6.269) or (6.271), is associated coherent risk measure \(\rho_t\) which is obtained by replacing the conditional expectations with respective (unconditional) expectations. Note also that if random variable \(Z \in Z_t\) is independent of \(\xi_{t-1}\), then the conditional expectations on the right hand sides of (6.269)–(6.271) coincide with the respective unconditional expectations, and hence \(\rho_t(\xi_{t-1})\) does not depend on \(\xi_{t-1}\) and coincides with \(\rho_t(Z)\).

Let us also assume that the objective functions \(f_t(x_t, \xi_t)\) and feasible sets \(\mathcal{X}_t(x_{t-1}, \xi_t)\) are given in terms of the data process. Then formulation (6.262) takes the form

\[
\min_{x_1, x_2, \ldots, x_T} f_1(x_1) + \rho_{2|\xi_1} \left( f_2(x_2(\xi_2), \xi_2) + \ldots + \rho_{T-1|\xi_{T-2}} \left[ f_{T-1}(x_{T-1}(\xi_{T-1}), \xi_{T-1}) + \rho_{T|\xi_{T-1}} \left[ f_T(x_T(\xi_T), \xi_T) \right] \right] \right)
\]

s.t. \(x_1 \in \mathcal{X}_1, x_t(\xi_t) \in \mathcal{X}_t(x_{t-1}(\xi_{t-1}), \xi_t), t = 2, \ldots, T,
\]

(6.272)

where the optimization is performed over feasible policies.

The corresponding dynamic programming equations (6.266)–(6.267) take the form:

\[
Q_t(x_{t-1}, \xi_t) = \min_{x_t \in \mathcal{X}_t(x_{t-1}, \xi_t)} \left[ f_t(x_t, \xi_t) + Q_{t+1}(x_t, \xi_t) \right],
\]

(6.273)

where

\[
Q_{t+1}(x_t, \xi_t) = \rho_{t+1|\xi_t} \left[ Q_{t+1}(x_t, \xi_{t+1}) \right].
\]

(6.274)

Note that if the process \(\xi_t\) is stagewise independent, then the conditional expectations coincide with the respective unconditional expectations, and hence (similar to the risk neutral case) functions \(Q_{t+1}(x_t, \xi_t)\) do not depend on \(\xi_t\), and the cost-to-go functions \(Q_t(x_{t-1}, \xi_t)\) depend only on \(\xi_t\) rather than \(\xi_{t-1}\).

Of course, if we set \(\rho_{t|\xi_{t-1}}(\cdot) := E[\cdot|\xi_{t-1}]\), then the above equations (6.273) coincide with the corresponding risk neutral dynamic programming equations. Also in that case the composite measure \(\bar{\rho}\) becomes the corresponding expectation operator and hence formulation (6.263) coincides with the respective risk neutral formulation (3.3). Unfortunately, in the general case it is quite difficult to write the composite measure \(\bar{\rho}\) in an explicit form.

**Multi-period Coherent Risk Measures**

It is possible to approach risk averse multistage stochastic programming in the following framework. As before, let \(\mathcal{F}_t\) be a filtration and \(Z_t := L_p(\Omega, \mathcal{F}_t, P), t = 1, \ldots, T\). Consider the space \(Z := Z_1 \times \cdots \times Z_T\). Recall that since \(\mathcal{F}_1 = \{\emptyset, \Omega\}\), the space \(Z_1\) can be
identified with \( \mathbb{R} \). With space \( Z \) we can associate its dual space \( Z^* := Z_1^* \times \cdots \times Z_T^* \), where \( Z_t^* = L_q(\Omega, \mathcal{F}_t, P) \) is the dual of \( Z_t \). For \( Z = (Z_1, \ldots, Z_T) \in Z \) and \( \zeta = (\zeta_1, \ldots, \zeta_T) \in Z^* \) scalar product is defined in the natural way:

\[
(\zeta, Z) := \sum_{t=1}^{T} \int_{\Omega} \zeta_t(\omega) Z_t(\omega) dP(\omega).
\]

(6.275)

Note that \( Z \) can be equipped with a norm, consistent with \( \| \cdot \|_p \) norms of its components, which makes it a Banach space. For example, we can use \( \| Z \| := \sum_{t=1}^{T} \| Z_t \|_p \). This norm induces the dual norm \( \| \zeta \|^* = \max \{ \| \zeta_1 \|_q, \ldots, \| \zeta_T \|_q \} \) on the space \( Z^* \).

Consider a function \( \varrho : Z \to \mathbb{R} \). For such a function it makes sense to talk about conditions (R1),(R2) and (R4) defined in section 6.3, with \( Z \cong Z' \) understood componentwise. We say that \( \varrho(\cdot) \) is a \textit{multiperiod risk measure} if it satisfies the respective conditions (R1),(R2) and (R4). Similarly to the analysis of section 6.3, we have the following results. By Theorem 7.79 it follows from convexity (condition (R1)) and monotonicity (condition (R2)), and since \( \varrho(\cdot) \) is real valued, that \( \varrho(\cdot) \) is continuous. By the Fenchel-Moreau Theorem, we have that convexity, continuity and positive homogeneity (condition (R4)) imply the dual representation:

\[
\varrho(Z) = \sup_{\zeta \in \mathfrak{A}} (\zeta, Z), \quad \forall Z \in Z,
\]

(6.276)

where \( \mathfrak{A} \) is a convex, bounded and weakly* closed subset of \( Z^* \) (and hence, by the Banach-Alaoglu Theorem, \( \mathfrak{A} \) is weakly* compact). Moreover, it is possible to show, exactly in the same way as in the proof of Theorem 6.4, that condition (R2) holds iff \( \zeta \succeq 0 \) for every \( \zeta \in \mathfrak{A} \). Conversely, if \( \varrho \) is given in the form (6.276) with \( \mathfrak{A} \) being a convex weakly* compact subset of \( Z^* \) such that \( \zeta \succeq 0 \) for every \( \zeta \in \mathfrak{A} \), then \( \varrho \) is a (real valued) multiperiod risk measure. An analogue of the condition (R3) (translation equivariance) is more involved, we will discuss this later.

For any multiperiod risk measure \( \varrho \) we can formulate the following risk averse multistage program

\[
\begin{align*}
\text{Min} & \quad \varrho \left( f_1(x_1), f_2(x_2(\omega), \omega), \ldots, f_T(x_T(\omega), \omega) \right) \\
\text{s.t.} & \quad x_1 \in \mathcal{X}_1, \ x_t(\omega) \in \mathcal{X}_t(x_{t-1}(\omega), \omega), \ t = 2, \ldots, T,
\end{align*}
\]

(6.277)

where optimization is performed over \( \mathcal{F}_t \)-measurable \( x_t : \Omega \to \mathbb{R}, t = 1, \ldots, T \), satisfying the corresponding constraints, and such that \( f_t(x_t(\cdot), \cdot) \in Z_t \). The nonanticipativity is enforced here by the \( \mathcal{F}_t \)-measurability of \( x_t(\omega) \).

Let us make the following observation. If we are currently at a certain stage of the system, then we know the past and hence it is reasonable to require that our decisions should be based on that information alone and should not involve unknown data. This is the nonanticipativity constraint, which was discussed in the previous sections. However, if we believe in the considered model, we also have an idea what can and what cannot happen in the future. Think, for example, about a scenario tree representing evolution of the data process. If we are currently at a certain node of that tree, representing current state of the system, we already know that only scenarios passing through this node can happen in the future. Therefore, apart form the nonanticipativity constraint, it is also reasonable to think about the following concept, which we refer to as the \textit{time consistency} principle:
Chapter 6. Risk Averse Optimization

• At every state of the system, optimality of our decisions should not depend on scenarios which we already know cannot happen in the future.

In order to formalize this concept of time consistency we need to say, of course, what we optimize (say minimize) at every state of the process, i.e., to formulate a respective optimality criterion associated with every state of the system. The risk neutral formulation (3.3) of multistage stochastic programming, discussed in Chapter 3, automatically satisfies the time consistency requirement (see below). The risk averse case is more involved and needs a discussion. We say that multiperiod risk measure \( \varrho \) is time consistent if the corresponding multistage problem (6.277) satisfies the above principle of time consistency.

Consider the class of functionals \( \varrho : \mathcal{Z} \to \mathbb{R} \) of the form (6.231), i.e., functionals representable as

\[
\varrho(Z_1, \ldots, Z_T) = Z_1 + \rho_2[Z_2 + \cdots + \rho_{T-1}[Z_{T-1} + \rho_T[Z_T]]], \tag{6.278}
\]

where \( \rho_{t+1}|_{\mathcal{F}_t} : \mathcal{Z}_{t+1} \to \mathcal{Z}_t, t = 1, \ldots, T-1, \) is a sequence of conditional risk mappings.

It is not difficult to see that conditions (R’1),(R’2) and (R’4) (defined in section 6.7.2), applied to every conditional risk mapping \( \rho_{t+1}|_{\mathcal{F}_t} \), imply respective conditions (R1),(R2) and (R4) for the functional \( \varrho \) of the form (6.278). That is, equation (6.278) defines a particular class of multiperiod risk measures.

Of course, for \( \varrho \) of the form (6.278), optimization problem (6.277) coincides with the nested formulation (6.262). Recall that if the set \( \Omega \) is finite, then we can formulate multistage risk averse optimization in the framework of scenario trees. As it was discussed in section 6.7.1, nested formulation (6.262) is implied by the approach where with every node of the scenario tree is associated a coherent risk measure applied to the next stage of the scenario tree. In particular, this allows to write the corresponding dynamic programming equations and implies that an associated optimal policy has the decomposition property. That is, if the process reached a certain node at stage \( t \), then the remaining decisions of the optimal policy are also optimal with respect to this node considered as the starting point of the process. It follows that the multiperiod risk measure of the form (6.278) is time consistent and the corresponding approach to risk averse optimization satisfies the time consistency principle.

It is interesting and important to give an intrinsic characterization of the nested approach to multiperiod risk measures. Unfortunately, this seems to be too difficult and we will give only a partial answer to this question. Let observe first that for any \( Z = (Z_1, \ldots, Z_T) \in \mathcal{Z} \),

\[
\mathbb{E}[Z_1 + \cdots + Z_T] = Z_1 + \mathbb{E}|_{\mathcal{F}_1}[Z_2 + \cdots + \mathbb{E}|_{\mathcal{F}_{T-1}}[Z_{T-1} + \mathbb{E}|_{\mathcal{F}_T}[Z_T]]], \tag{6.279}
\]

where \( \mathbb{E}|_{\mathcal{F}_t}[\cdot] = \mathbb{E}[\cdot | \mathcal{F}_t] \) are the corresponding conditional expectation operators. That is, the expectation risk measure \( \varrho(Z_1, \ldots, Z_T) := \mathbb{E}[Z_1 + \cdots + Z_T] \) is time consistent and the risk neutral formulation (3.3) of multistage stochastic programming satisfies the time consistency principle.

Consider the following condition.

(R3-d) For any \( Z = (Z_1, \ldots, Z_T) \in \mathcal{Z}, Y_t \in \mathcal{Z}_t, t = 1, \ldots, T-1, \) and \( a \in \mathbb{R} \) it holds that

\[
\varrho(Z_1, \ldots, Z_t, Z_{t+1} + Y_t, \ldots, Z_T) = \varrho(Z_1, \ldots, Z_t + Y_t, Z_{t+1}, \ldots, Z_T), \tag{6.280}
\]
\( \varrho(Z_1 + a, \ldots, Z_T) = a + \varrho(Z_1, \ldots, Z_T). \) \hspace{1cm} (6.281)

**Proposition 6.41.** Let \( \varrho : \mathcal{Z} \rightarrow \mathbb{R} \) be a multiperiod risk measure. Then the following conditions (i)-(iii) are equivalent:

(i) There exists a coherent risk measure \( \bar{\varrho} : \mathcal{Z}_T \rightarrow \mathbb{R} \) such that

\[
\varrho(Z_1, \ldots, Z_T) = \bar{\varrho}(Z_1 + \ldots + Z_T), \quad \forall (Z_1, \ldots, Z_T) \in \mathcal{Z}.
\] 

(ii) Condition (R3-d) is fulfilled.

(iii) There exists a nonempty, convex, bounded and weakly* closed subset \( \mathfrak{A}_T \) of probability density functions \( \mathfrak{P}_T \subset \mathcal{Z}_T^* \) such that the dual representation (6.276) holds with the corresponding set \( \mathfrak{A} \) of the form

\[
\mathfrak{A} = \{(\zeta_1, \ldots, \zeta_T) : \zeta_T \in \mathfrak{A}_T, \; \zeta_t = \mathbb{E}[\zeta_T | \mathcal{F}_t], \; t = 1, \ldots, T - 1 \}.
\] 

**Proof.** If condition (i) is satisfied, then for any \( Z = (Z_1, \ldots, Z_T) \in \mathcal{Z} \) and \( Y_t \in \mathfrak{A}_t \),

\[
\varrho(Z_1, \ldots, Z_t, Z_{t+1} + Y_t, \ldots, Z_T) = \bar{\varrho}(Z_1 + \ldots + Z_T + Y_t)
\]

\[
= \varrho(Z_1, \ldots, Z_t + Y_t, Z_{t+1}, \ldots, Z_T).
\]

Property (6.281) also follows by condition (R3) of \( \bar{\varrho} \). That is, condition (i) implies condition (R3-d).

Conversely, suppose that condition (R3-d) holds. Then for \( Z = (Z_1, Z_2, \ldots, Z_T) \) we have that \( \varrho(Z_1, Z_2, \ldots, Z_T) = \varrho(0, Z_1 + Z_2, \ldots, Z_T) \). Continuing in this way, we obtain that

\[
\varrho(Z_1, \ldots, Z_T) = \varrho(0, \ldots, 0, Z_1 + \ldots + Z_T).
\]

Define

\[
\bar{\varrho}(W_T) := \varrho(0, \ldots, 0, W_T), \quad W_T \in \mathcal{Z}_T.
\]

Conditions (R1),(R2),(R4) for \( \varrho \) imply the respective conditions for \( \bar{\varrho} \). Moreover, for \( a \in \mathbb{R} \) we have

\[
\bar{\varrho}(W_T + a) = \varrho(0, \ldots, 0, W_T + a) = \varrho(0, \ldots, a, W_T) = \ldots = \varrho(a, \ldots, 0, W_T)
\]

\[
= a + \varrho(0, \ldots, 0, W_T) = \bar{\varrho}(W_T) + a.
\]

That is, \( \bar{\varrho} \) is a coherent risk measure, and hence (ii) implies (i).

Now suppose that condition (i) holds. By the dual representation (see Theorem 6.4 and Proposition 6.5) there exists a convex, bounded and weakly* closed set \( \mathfrak{A}_T \subset \mathfrak{P}_T \) such that

\[
\bar{\varrho}(W_T) = \sup_{\zeta_T \in \mathfrak{A}_T} \langle \zeta_T, W_T \rangle, \quad W_T \in \mathcal{Z}_T.
\] 

(6.284)

Moreover, for \( W_T = Z_1 + \ldots + Z_T \) we have

\[\mathbb{E}[\zeta_T W_T] = \sum_{t=1}^T \mathbb{E}[\zeta_T Z_t], \quad \text{and since } Z_t \text{ is } F_t\text{-measurable,}
\]

\[\mathbb{E}[\zeta_T Z_t] = \mathbb{E}[\mathbb{E}[\zeta_T Z_t | F_t]], \quad \mathbb{E}[Z_t \mathbb{E}[\zeta_T | F_t]].
\] 

(6.285)

That is, (i) implies (iii). Conversely, suppose that (iii) holds. Then equation (6.284) defines a coherent risk measure \( \bar{\varrho} \). The dual representation (6.276) together with (6.283) imply (6.282). This shows that conditions (i) and (iii) are equivalent. \( \square \)
As we know, condition (i) of the above proposition is necessary for the multiperiod risk measure \( \rho \) to be representable in the nested form (6.278) (see section 6.7.3 and equation (6.264) in particular). This condition, however, is not sufficient. It seems to be quite difficult to give a complete characterization of coherent risk measures \( \tilde{\rho} \) representable in the form

\[
\tilde{\rho}(Z_1 + \ldots + Z_T) = Z_1 + \rho_{Z_i|\mathcal{F}_1} \left[ Z_2 + \ldots + \rho_{\tau-1|\mathcal{F}_{\tau-1}} \left[ Z_T + \rho_{\tau|\mathcal{F}_{\tau-1}} (Z_T) \right] \right],
\]

(6.286)

for all \( Z = (Z_1, \ldots, Z_T) \in \mathcal{Z} \), and some sequence \( \rho_{t+1|\mathcal{F}_t} : \mathcal{Z}_{t+1} \to \mathcal{Z}_t \), \( t = 1, \ldots, T-1 \), of conditional risk mappings.

**Remark 23.** Of course, condition \( \zeta_t = \mathbb{E}[\zeta_T | \mathcal{F}_t] \), \( t = 1, \ldots, T-1 \), of equation (6.283), can be written as

\[
\zeta_t = \mathbb{E}[\zeta_{t+1} | \mathcal{F}_t], \quad t = 1, \ldots, T-1.
\]

(6.287)

That is, if representation (6.282) holds for some coherent risk measure \( \rho(\cdot) \), then any element \( (\zeta_1, \ldots, \zeta_T) \) of the dual set \( \mathcal{A} \), in the representation (6.276) of \( \rho(\cdot) \), forms a martingale sequence.

**Example 6.42** Let \( \rho_{t \backslash \mathcal{F}_{t-1}} : \mathcal{Z}_T \to \mathcal{Z}_{\tau-1} \) be a conditional risk mapping for some \( 2 \leq \tau \leq T \), and let \( \rho_1(Z_1) := Z_1, Z_1 \in \mathbb{R} \), and \( \rho_{1 \backslash \mathcal{F}_{t-1}} := \mathbb{E}_{\mathcal{F}_{t-1}}, t = 2, \ldots, T, t \neq \tau \). That is, we take here all conditional risk mappings to be the respective conditional expectations except (an arbitrary) conditional risk mapping \( \rho_{t \backslash \mathcal{F}_{t-1}} \) at the period \( t = \tau \). It follows that

\[
\rho(Z_1, \ldots, Z_T) = \mathbb{E} \left[ Z_1 + \ldots + Z_{\tau-1} + \rho_{\tau \backslash \mathcal{F}_{\tau-1}} \left[ \mathbb{E}_{\mathcal{F}_{\tau}} [Z_\tau + \ldots + Z_T] \right] \right]
\]

(6.288)

That is,

\[
\tilde{\rho}(W_T) = \mathbb{E} \left[ \rho_{\tau \backslash \mathcal{F}_{\tau-1}} \left[ \mathbb{E}_{\mathcal{F}_{\tau}} [W_T] \right] \right], \quad W_T \in \mathcal{Z}_T,
\]

(6.289)

is the corresponding (composite) coherent risk measure.

Coherent risk measures of the form (6.289) have the following property:

\[
\tilde{\rho}(W_T + Y_{\tau-1}) = \rho(W_T) + \mathbb{E}[W_T], \quad \forall W_T \in \mathcal{Z}_T, \forall Y_{\tau-1} \in \mathcal{Z}_{\tau-1}.
\]

(6.290)

By (6.283) the above condition (6.290) means that the corresponding set \( \mathcal{A} \), defined in (6.283), has the additional property that \( \zeta_t = \mathbb{E}[\zeta_T] = 1 \), \( t = 1, \ldots, \tau-1 \), i.e., these components of \( \zeta \in \mathcal{A} \) are constants (equal to one).

In particular, for \( \tau = T \) the composite risk measure (6.289) becomes

\[
\tilde{\rho}(W_T) = \mathbb{E} \left[ \rho_{\tau \backslash \mathcal{F}_{\tau-1}} [W_T] \right], \quad W_T \in \mathcal{Z}_T.
\]

(6.291)

Let, further, \( \rho_{\tau \backslash \mathcal{F}_{\tau-1}} : \mathcal{Z}_T \to \mathcal{Z}_{\tau-1} \) be the conditional mean absolute deviation, i.e.,

\[
\rho_{\tau \backslash \mathcal{F}_{\tau-1}}[Z_T] := \mathbb{E}_{\mathcal{F}_{\tau-1}} \left[ Z_T + c |Z_T - \mathbb{E}_{\mathcal{F}_{\tau-1}}[Z_T]| \right],
\]

(6.292)

\( c \in [0, 1/2] \). The corresponding composite coherent risk measure here is

\[
\tilde{\rho}(W_T) = \mathbb{E}[W_T] + c \mathbb{E} \left[ W_T - \mathbb{E}_{\mathcal{F}_{\tau-1}}[W_T] \right], \quad W_T \in \mathcal{Z}_T.
\]

(6.293)
For $T > 2$ the risk measure (6.293) is different from the mean absolute deviation measure
\[ \bar{\rho}(W_T) := \mathbb{E}[W_T] + c \mathbb{E}[W_T - \mathbb{E}[W_T]], \quad W_T \in Z_T, \] (6.294)
and that the multiperiod risk measure
\[ \rho(Z_1, \ldots, Z_T) := \bar{\rho}(Z_1 + \ldots + Z_T) = \mathbb{E}[Z_1 + \ldots + Z_T] + c \mathbb{E}[Z_1 + \ldots + Z_T - \mathbb{E}[Z_1 + \ldots + Z_T]], \]
corresponding to (6.294), is not time consistent.

**Risk Averse Multistage Portfolio Selection**

We discuss now the example of portfolio selection. A nested formulation of multistage portfolio selection can be written as
\[
\begin{align*}
\text{Min } & \left\{ \bar{\rho}(-W_T) := \rho_1 \left[ \cdots \rho_{T-1} | W_{T-2} \left[ \rho_T | W_{T-1} (-W_T) \right] \right] \right\} \\
\text{s.t. } & W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}, \quad \sum_{i=1}^{n} x_{i,t} = W_t, \quad x_t \geq 0, \quad t = 0, \ldots, T - 1.
\end{align*}
\] (6.295)

We use here conditional risk mappings formulated in terms of the respective conditional expectations, like the conditional $\mathbb{A} \mathbb{V} \mathbb{R}$ (see (6.269)) and conditional mean semideviations (see (6.271)), and the notation $\rho_t(W_{t-1})$ stands for a conditional risk mapping defined in terms of the respective conditional expectations given $W_{t-1}$. By $\rho_t(\cdot)$ we denote the corresponding (unconditional) risk measures. For example, to the conditional $\mathbb{A} \mathbb{V} \mathbb{R}_n(\cdot | \xi_{[t-1]})$ corresponds the respective (unconditional) $\mathbb{A} \mathbb{V} \mathbb{R}_n(\cdot)$. If we set $\rho_t(W_{t-1}) := \mathbb{E}[W_{t-1}, t = 1, \ldots, T$ then since
\[ \mathbb{E}\left[ \cdots \mathbb{E}\left[ \mathbb{E}[-W_T | W_{T-1}] | W_{T-2} \right] \right] = \mathbb{E}[-W_T], \]
we obtain the risk neutral formulation. Note also that in order to formulate this as a minimization, rather than a maximization, problem we changed the sign of $\xi_{it}$.

Suppose that the random process $\xi_{i}$ is **stagewise independent**. Let us write dynamic programming equations. At the last stage we have to solve problem
\[
\begin{align*}
\text{Min } & \rho_T | W_{T-1} (-W_T) \\
\text{s.t. } & W_T = \sum_{i=1}^{n} \xi_{i,T} x_{i,T-1}, \quad \sum_{i=1}^{n} x_{i,T-1} = W_{T-1}.
\end{align*}
\] (6.296)

Since $W_{T-1}$ is a function of $\xi_{[T-1]}$, by the stagewise independence we have that $\xi_T$, and hence $W_T$, are independent of $W_{T-1}$. It follows by positive homogeneity of $\rho_T$ that the optimal value of (6.296) is $Q_{T-1}(W_{T-1}) = W_{T-1} \nu_{T-1}$, where $\nu_{T-1}$ is the optimal value of
\[
\begin{align*}
\text{Min } & \rho_T [-W_T] \\
\text{s.t. } & W_T = \sum_{i=1}^{n} \xi_{i,T} x_{i,T-1}, \quad \sum_{i=1}^{n} x_{i,T-1} = 1,
\end{align*}
\] (6.297)
and an optimal solution of (6.296) is \( \bar{x}_{T-1}(W_{T-1}) = W_{T-1}x^*_T \), where \( x^*_T \) is an optimal solution of (6.297). And so on we obtain that the optimal policy \( \bar{x}_t(W_t) \) here is myopic. That is, \( \bar{x}_t(W_t) = W_tx^*_t \), where \( x^*_t \) is an optimal solution of (6.298). (compare with remark 1.4.3, 21). Note that the composite risk measure \( \bar{\rho} \) can be quite complicated here.

An alternative, multiperiod risk averse approach can be formulated as

\[
\min_{x_t \geq 0, W_{t+1}} \rho_{t+1}[-W_{t+1}],
\]
\[
\text{s.t. } W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1}x_{it}, \quad \sum_{i=1}^{n} x_{it} = 1
\]

(6.298)

for an explicitly defined risk measure \( \rho \). Let, for example,

\[
\rho(\cdot) := (1 - \beta)\mathbb{E}[^{\cdot}] + \beta AVaR(\cdot), \quad \beta \in [0, 1], \quad \alpha \in (0, 1).
\]

(6.300)

Then problem (6.299) becomes

\[
\min \ (1 - \beta)\mathbb{E}[-W_T] + \beta (-r + \alpha^{-1}\mathbb{E}[r - W_T])
\]
\[
\text{s.t. } W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1}x_{it}, \quad \sum_{i=1}^{n} x_{it} = W_t, \quad x_t \geq 0, \quad t = 0, \ldots, T-1,
\]

(6.301)

where \( r \in \mathbb{R} \) is the (additional) first stage decision variable. After \( r \) is decided, at the first stage, the problem comes to minimizing \( \mathbb{E}[U(W_T)] \) at the last stage, where \( U(W) := (1 - \beta)W + \beta \alpha^{-1}[W - r]_+ \) can be viewed as a disutility function.

The respective dynamic programming equations become as follows. The last stage value function \( Q_{T-1}(W_{T-1}, r) \) is given by the optimal value of the problem

\[
\min_{x_{T-1} \geq 0, W_T} \mathbb{E}[-(1 - \beta)W_T + \beta \alpha^{-1}[r - W_T]_+]
\]
\[
\text{s.t. } W_T = \sum_{i=1}^{n} \xi_{i,T}x_{i,T-1}, \quad \sum_{i=1}^{n} x_{i,T-1} = W_{T-1}.
\]

(6.302)

Proceeding in this way, at stages \( t = T - 2, \ldots, 1 \) we consider the problems

\[
\min_{x_t \geq 0, W_{t+1}} \mathbb{E} \{ Q_{t+1}(W_{t+1}, r) \}
\]
\[
\text{s.t. } W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1}x_{it}, \quad \sum_{i=1}^{n} x_{it} = W_t.
\]

(6.303)
whose optimal value is denoted $Q_t(W_t, r)$. Finally, at stage $t = 0$ we solve the problem
\[
\min_{x_0 \geq 0, r, W_1} \beta r + \mathbb{E}[Q_1(W_1, r)]
\]
subject to $W_1 = \sum_{i=1}^{n} \xi_i x_{i0}, \sum_{i=1}^{n} x_{i0} = W_0$. \hfill (6.304)

In the above multiperiod risk averse approach the optimal policy is not myopic and the property of time consistency is not satisfied.

**Risk Averse Multistage Inventory Model**

Consider the multistage inventory problem (1.17). The nested risk averse formulation of that problem can be written as
\[
\min_{x_1 \geq y_1} \left[ c_1(x_1 - y_1) + \rho_1 \left[ \psi_1(x_1, D_1) + c_2(x_2 - y_2) + \rho_2 \psi_2(x_2, D_2) + \cdots \right. \right.
\]
\[
\left. \left. + c_{T-1}(x_{T-1} - y_{T-1}) + \rho_{T-1} \psi_{T-1}(x_{T-1}, D_{T-1}) \right. \right.
\]
\[
\left. \left. + c_T(x_T - y_T) + \rho_T \psi_T(x_T, D_T) \right] \right] \right] \right]
\]
subject to $y_{t+1} = x_t - D_t, t = 1, \ldots, T - 1,$

where $y_1$ is a given initial inventory level, $\psi_t(x_t, d_t) := b_t[d_t - x_t]_+ + h_t[x_t - d_t]_+$ and $\rho_{t|D_{t-1}}(\cdot)$, $t = 2, \ldots, T$, are chosen conditional risk mappings. Recall that the notation $\rho_{t|D_{t-1}}(\cdot)$ stands for a conditional risk mapping obtained by using conditional expectations, conditional on $D_{t-1}$, and note that $\rho_1(\cdot)$ is real valued and is a coherent risk measure.

As it was discussed before, there are two equivalent interpretations of problem (6.305). We can write it as an optimization problem with respect to feasible policies $x_t(d_{t-1})$ (compare with (6.272)):
\[
\min_{x_1, x_2, \ldots, x_T} \left[ c_1(x_1 - y_1) + \rho_1 \left[ \psi_1(x_1, D_1) + c_2(x_2 - x_1 + D_1) \right. \right.
\]
\[
\left. \left. + \rho_2 \psi_2(x_2, D_2) + \cdots \right. \right.
\]
\[
\left. \left. + c_{T-1}(x_{T-1} - x_{T-2} + D_{T-2}) + \rho_{T-1} \psi_{T-1}(x_{T-1}, D_{T-1}) \right. \right.
\]
\[
\left. \left. + c_T(x_T - x_{T-1} + D_{T-2}) + \rho_T \psi_T(x_T, D_T) \right] \right]
\]
subject to $x_1 \geq y_1, x_2(D_1) \geq x_1 - D_1, x_t(D_{t-1}) \geq x_{t-1}(D_{t-2}) - D_{t-1}, t = 3, \ldots, T.$

Alternatively we can write dynamic programming equations. At the last stage $t = T$, for observed inventory level $y_T$, we need to solve the problem:
\[
\min_{x_T \geq y_T} c_T(x_T - y_T) + \rho_T \psi_T(x_T, D_T). \hfill (6.307)
\]
The optimal value of problem (6.307) is denoted $Q_T(y_T, D_{[T-1]})$. Continuing in this way, we write for $t = T-1, \ldots, 2$ the following dynamic programming equations

$$
Q_t(y_t, D_{[t-1]}) = \min_{x_t \geq y_t} c_t(x_t - y_t) + \rho_t|D_{[t-1]}| \left[ \psi(x_t, D_t) + Q_{t+1}(x_t - D_t, D_{[t]}) \right].
$$

Finally, at the first stage we need to solve problem

$$
\min_{x_1} c_1(x_1 - y_1) + \rho_1 \left[ \psi(x_1, D_1) + Q_2(x_1 - D_1, D_1) \right].
$$

(6.308)

Suppose now that the process $D_t$ is stagewise independent. Then, by exactly the same argument as in section 1.2.3, the cost-to-go (value) function $Q_t(y_t, d_{[t-1]}) = Q_t(y_t)$, $t = 2, \ldots, T$, is independent of $d_{[t-1]}$, and by convexity arguments the optimal policy $\bar{x}_t = \bar{x}_t(d_{t-1})$ is a basestock policy. That is, $\bar{x}_t = \max\{y_t, x^*_t\}$, where $x^*_t$ is an optimal solution of

$$
\min_{x_t} c_t x_t + \rho_t \left[ \psi(x_t, D_t) + Q_{t+1}(x_t - D_t) \right].
$$

(6.310)

Recall that $\rho_t$ denotes the coherent risk measure corresponding to the conditional risk mapping $\rho_t|D_{[t-1]}$.

---

**Exercises**

6.1. Let $Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ be a random variable with cdf $H(z) := P\{Z \leq z\}$. Note that $\lim_{t\uparrow 1} H(z) = H(t)$ and denote $H^-(t) := \lim_{z\downarrow t} H(z)$. Consider functions $\phi_1(t) := \mathbb{E}[t - Z]_+, \phi_2(t) := \mathbb{E}[Z - t]_+$ and $\phi(t) := \beta_1 \phi_1(t) + \beta_2 \phi_2(t)$, where $\beta_1, \beta_2$ are positive constants. Show that $\phi_1, \phi_2$ and $\phi$ are real valued convex functions with subdifferentials

$$
\partial \phi_1(t) = [H^-(t), H(t)] \quad \text{and} \quad \partial \phi_2(t) = [-1 + H^-(t), -1 + H(t)],
$$

$$
\partial \phi(t) = [(\beta_1 + \beta_2)H^-(t) - \beta_2, (\beta_1 + \beta_2)H(t) - \beta_2].
$$

Conclude that the set of minimizers of $\phi(t)$ over $t \in \mathbb{R}$ is the (closed) interval of $[\beta_2/(\beta_1 + \beta_2)]$-quantiles of $H(\cdot)$.

6.2. (i) Let $Y \sim \mathcal{N}(\mu, \sigma^2)$. Show that

$$
\text{AV@R}_\alpha(Y) = \mu + z_\alpha \sigma,
$$

(6.311)

where $z_\alpha := \Phi^{-1}(1 - \alpha)$, and

$$
\text{AV@R}_\alpha(Y) = \mu + \frac{\sigma}{\alpha \sqrt{2\pi}} e^{-z_\alpha^2/2}.
$$

(6.312)

(ii) Let $Y^1, \ldots, Y^N$ be an iid sample of $Y \sim \mathcal{N}(\mu, \sigma^2)$. Compute the asymptotic variance and asymptotic bias of the sample estimator $\hat{\theta}_N$, of $\theta^* = \text{AV@R}_\alpha(Y)$, defined on page 302.
6.3. Consider the chance constraint

\[ \Pr \left\{ \sum_{i=1}^{n} \xi_i x_i \geq b \right\} \geq 1 - \alpha, \]  

(6.313)

where \( \xi \sim \mathcal{N}(\mu, \Sigma) \) (see problem (1.43)). Note that this constraint can be written as

\[ \mathbb{V} \circ \mathbb{R}_\alpha \left( b - \sum_{i=1}^{n} \xi_i x_i \right) \leq 0. \]  

(6.314)

Consider the following constraint

\[ \mathbb{A} \mathbb{V} \circ \mathbb{R}_\gamma \left( b - \sum_{i=1}^{n} \xi_i x_i \right) \leq 0. \]  

(6.315)

Show that constraints (6.313) and (6.315) are equivalent if

\[ z_\alpha = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \]

(6.316)

6.4. Consider the function \( \phi(x) := \mathbb{A} \mathbb{V} \circ \mathbb{R}_\alpha (F_x) \), where \( F_x = F_x(\omega) = F(x, \omega) \) is a real valued random variable, on a probability space \((\Omega, \mathcal{F}, P)\), depending on \( x \in \mathbb{R}^n \). Assume that: (i) for a.e. \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is continuously differentiable on a neighborhood \( V \) of a point \( x_0 \in \mathbb{R}^n \), (ii) the families \( \|F(x, \omega)\|, x \in V \), and \( \|\nabla_x F(x, \omega)\|, x \in V \), are dominated by a \( P \)-integrable function, (iii) the random variable \( F_{x_0} \) has continuous distribution for all \( x \in V \). Show that, under these conditions, \( \phi(x) \) is directionally differentiable at \( x_0 \) and

\[ \nabla \phi(x_0) = \alpha^{-1} \mathbb{E} \left[ \frac{1}{\gamma_\alpha}(\omega) \nabla_x F(x_0, \omega) \right]. \]  

(6.317)

Hint: use Theorem 7.44 together with Danskin Theorem 7.21.

6.5. Show that the set of saddle points of the minimax problem (6.189) is given by \( \{\mu\} \times [\gamma^*, \gamma^{**}] \), where \( \gamma^* \) and \( \gamma^{**} \) are defined in (6.191).

6.6. Consider the absolute semideviation risk measure

\[ \rho_c(Z) := \mathbb{E} \{ Z + c[Z - \mathbb{E}(Z)]_+ \}, \quad Z \in L_1(\Omega, \mathcal{F}, P), \]

where \( c \in [0, 1] \), and the following risk averse optimization problem:

\[ \min_{x \in X} \mathbb{E} \{ G(x, \xi) + c[G(x, \xi) - \mathbb{E}(G(x, \xi))]_+ \}. \]  

(6.318)

Viewing the optimal value of problem (6.318) as Von Mises statistical functional of the probability measure \( P \), compute its influence function.

Hint: use derivations of section 6.5.3 together with Danskin Theorem.
Consider the newsvendor problem with cost function

\[ F(x, d) = cx + b|d - x|_+ + h|x - d|_+, \quad \text{where } b > c \geq 0, \, h > 0, \]

and the following minimax problem

\[
\min_{x \geq 0} \sup_{H \in \mathcal{H}} \mathbb{E}_H[F(x, D)], \tag{6.319}
\]

where \( \mathcal{H} \) is the set of cumulative distribution functions (probability measures) supported on (final) interval \([l, u] \subseteq \mathbb{R}_+\) and having a given mean \(d \in [l, u] \). Show that for any \( x \in [l, u] \) the maximum of \( \mathbb{E}_H[F(x, D)] \) over \( H \in \mathcal{H} \) is attained at the probability measure \( \bar{H} = p\Delta(l) + (1 - p)\Delta(u) \), where \( p = (d - u)/(l - u) \), i.e., the cdf \( \bar{H}(\cdot) \) is the step function

\[
\bar{H}(z) = \begin{cases} 
0, & \text{if } z < l, \\
p, & \text{if } l \leq z < u, \\
1, & \text{if } u \leq z.
\end{cases}
\]

Conclude that \( \bar{H} \) is the cdf specified in Proposition 6.37, and that \( \bar{x} = \bar{H}^{-1}(\kappa) \), where \( \kappa = (b - c)/(b - h) \), is the optimal solution of problem (6.319). That is, \( \bar{x} = l \) if \( \kappa < p \), and \( \bar{x} = u \) if \( \kappa > p \).

Consider the following version of the newsvendor problem. A newsvendor has to decide about quantity \( x \) of a product to purchase at the cost of \( c \) per unit. He can sell this product at the price \( s \) per unit and unsold products can be returned to the vendor at the price of \( r \) per unit. It is assumed that \( 0 \leq r < s \). If the demand \( D \) turns out to be greater than or equal to the order quantity \( x \), then he makes the profit \( sx - cx = (s - c)x \), while if \( D \) is less than \( x \), his profit is \( sD + r(x - D) - cx \). Thus the profit is a function of \( x \) and \( D \) and is given by

\[
F(x, D) = \begin{cases} 
(s - c)x, & \text{if } x \leq D, \\
(r - c)x + (s - r)D, & \text{if } x > D.
\end{cases} \tag{6.320}
\]

(a) Assuming that demand \( D \geq 0 \) is a random variable with cdf \( H(\cdot) \), show that the expectation function \( f(x) := \mathbb{E}_H[F(x, D)] \) can be represented in the form

\[
f(x) = (s - c)x - (s - r) \int_0^x H(z)dz. \tag{6.321}
\]
Exercises

Conclude that the set of optimal solutions of the problem
\[
\text{Max} \{ f(x) := E[H(F(x, D))] \} \tag{6.322}
\]
is an interval given by the set of \( \kappa \)-quantiles of the cdf \( H(\cdot) \) with \( \kappa := (s-c)/(s-r) \).

(b) Consider the following risk averse version of the Newsvendor problem
\[
\text{Min} \{ \phi(x) := \rho[-F(x, D)] \}. \tag{6.323}
\]
Here \( \rho \) is a real valued coherent risk measure representable in the form (6.164) and \( H^* \) is the corresponding reference cdf.

Show that: (i) The function \( \phi(x) = \rho[-F(x, D)] \) can be represented in the form
\[
\phi(x) = (c-s)x + (s-r) \int_0^x \bar{H}(z)dz \tag{6.324}
\]
for some cdf \( \bar{H} \).

(ii) If \( \rho(\cdot) := \text{AV}_@ R_{\alpha}(\cdot) \), then \( \bar{H}(z) = \max \{ \alpha^{-1}H^*(z), 1 \} \). Conclude that in that case optimal solutions of the risk averse problem (6.323) are smaller than the risk neutral problem (6.322).

6.10. Let \( Z_i := L_p(\Omega, F_i, P), i = 1, 2, \) with \( F_1 \subset F_2 \), and let \( \rho : Z_2 \rightarrow Z_1 \).

(a) Show that if \( \rho \) is a conditional risk mapping, \( Y \in Z_1 \) and \( Y \succeq 0 \), then \( \rho(YZ) = Y\rho(Z) \) for any \( Z \in Z_2 \).

(b) Suppose that the mapping \( \rho \) satisfies conditions (R’1)–(R’3), but not necessarily the positive homogeneity condition (R’4). Show that it can be represented in the form
\[
[\rho(Z)](\omega) = \sup_{\mu \in \mathcal{C}} \{ E[\mu[Z|F_2]](\omega) - [\rho^* signature](\omega) \}, \tag{6.325}
\]
where \( \mathcal{C} \) is a set of probability measures on \( (\Omega, F_2) \) and
\[
[\rho^* signature](\omega) = \sup_{Z \in Z_2} \{ E[\mu[Z|F_1]](\omega) - [\rho(Z)](\omega) \}. \tag{6.326}
\]
You may assume that \( F_1 \) has a countable number of elementary events.

6.11. Consider the following risk averse approach to multistage portfolio selection. Let \( \xi_1, \ldots, \xi_T \) be the respective data process (of random returns) and consider the following chance constrained nested formulation
\[
\text{Max} \ E[W_T] \\
\text{s.t.} W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}, \sum_{i=1}^{n} x_{i,t} = W_t, \ x_t \geq 0, \tag{6.327}
\]
\[
\Pr \{ W_{t+1} \geq \kappa W_t | \xi_{[t]} \} \geq 1 - \alpha, \ t = 0, \ldots, T - 1,
\]
where $\kappa \in (0, 1)$ and $\alpha \in (0, 1)$ are given constants. Dynamic programming equations for this problem can be written as follows. At the last stage $t = T - 1$ the cost-to-go function $Q_{T−1}(W_{T−1}, \xi_{[T−1]})$ is given by the optimal value of the problem

$$
\max_{x_{T-1} \geq 0, W_T} E[W_T | \xi_{[T-1]}]
$$

subject to

$$
W_T = \sum_{i=1}^{n} \xi_{i,T} x_{i,T-1}, \sum_{i=1}^{n} x_{i,T-1} = W_{T-1},
$$

and

$$
\Pr \left\{ W_T \geq \kappa W_{T-1} \mid \xi_{[T-1]} \right\},
$$

and at stage $t = T - 2, \ldots, 1$, the cost-to-go function $Q_t(W_t, \xi_{[t]})$ is given by the optimal value of the problem

$$
\max_{x_t \geq 0, W_{t+1}} E[Q_{t+1}(W_{t+1}, \xi_{[t+1]} | \xi_{[t]})]
$$

subject to

$$
W_{t+1} = \sum_{i=1}^{n} \xi_{i,t+1} x_{i,t}, \sum_{i=1}^{n} x_{i,t} = W_t,
$$

and

$$
\Pr \left\{ W_{t+1} \geq \kappa W_t \mid \xi_{[t]} \right\}.
$$

Assuming that the process $\xi_t$ is stagewise independent show that the optimal policy is myopic and is given by $\bar{x}_t(W_t) = W_t x_t^*$, where $x_t^*$ is an optimal solution of the problem

$$
\max_{x_t \geq 0} \sum_{i=1}^{n} E[\xi_{i,t+1}] x_{i,t}
$$

subject to

$$
\sum_{i=1}^{n} x_{i,t} = 1, \Pr \left\{ \sum_{i=1}^{n} \xi_{i,t+1} x_{i,t} \geq \kappa \right\} \geq 1 - \alpha.
$$
Chapter 7

Background Material

Alexander Shapiro

In this chapter we discuss some concepts and results from convex analysis, probability, functional analysis and optimization theories, needed for a development of the material in this book. Of course, a careful derivation of the required material goes far beyond the scope of this book. We give or outline proofs of some results while others are referred to the literature. Of course, this choice is somewhat subjective.

We denote by $\mathbb{R}^n$ the standard $n$-dimensional vector space, of (column) vectors $x = (x_1, \ldots, x_n)^T$, equipped with the scalar product $x^T y = \sum_{i=1}^n x_i y_i$. Unless stated otherwise we denote by $\| \cdot \|$ the Euclidean norm $\|x\| = \sqrt{x^T x}$. The notation $A^T$ stands for the transpose of matrix (vector) $A$, and " := " stands for "equal by definition" to distinguish it from the usual equality sign. By $\mathbb{R} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ we denote the set of extended real numbers. The domain of an extended real valued function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\text{dom } f := \{ x \in \mathbb{R}^n : f(x) < +\infty \}.$$ 

It is said that $f$ is proper if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and its domain, $\text{dom } f$, is nonempty. The function $f$ is said to be lower semicontinuous (lsc) at a point $x_0 \in \mathbb{R}^n$ if $f(x_0) \leq \lim \inf_{x \to x_0} f(x)$. It is said that $f$ is lower semicontinuous if it is lsc at every point of $\mathbb{R}^n$. The largest lower semicontinuous function which is less than or equal to $f$ is denoted $\text{lsc } f$.

It is not difficult to show that $f$ is lower semicontinuous if and only if (iff) its epigraph $\text{epi } f := \{(x, \alpha) \in \mathbb{R}^{n+1} : f(x) \leq \alpha \}$, is a closed subset of $\mathbb{R}^{n+1}$. We often have to deal with polyhedral functions.

**Definition 7.1.** An extended real valued function $f : \mathbb{R}^n \to \mathbb{R}$ is called polyhedral if it is proper convex and lower semicontinuous, its domain is a convex closed polyhedron and $f(\cdot)$ is piecewise linear on its domain.
By $1_A(\cdot)$ we denote the characteristic\footnote{Function $1_A(\cdot)$ is often also called the indicator function of the set $A$. We call it here characteristic function in order to distinguish it from the indicator function $I_A(\cdot)$.} function

$$1_A(x) := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A, \end{cases} \quad (7.1)$$

and by $\mathbb{I}_A(\cdot)$ the indicator function

$$\mathbb{I}_A(x) := \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \notin A, \end{cases} \quad (7.2)$$

of set $A$.

By $\text{cl}(A)$ we denote the topological closure of set $A \subset \mathbb{R}^n$. For sets $A, B \subset \mathbb{R}^n$ we denote by

$$\text{dist}(x, A) := \inf_{x' \in A} \|x - x'\| \quad (7.3)$$

the distance from $x \in \mathbb{R}^n$ to $A$, and by

$$\mathbb{D}(A, B) := \sup_{x \in A} \text{dist}(x, B) \quad \text{and} \quad \mathbb{H}(A, B) := \max \{ \mathbb{D}(A, B), \mathbb{D}(B, A) \} \quad (7.4)$$

the deviation of the set $A$ from the set $B$ and the Hausdorff distance between the sets $A$ and $B$, respectively. By the definition, $\text{dist}(x, A) = +\infty$ if $A$ is empty, and $\mathbb{H}(A, B) = +\infty$ if $A$ or $B$ is empty.

### 7.1 Optimization and Convex Analysis

#### 7.1.1 Directional Differentiability

Consider a mapping $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$. It is said that $g$ is directionally differentiable at a point $x_0 \in \mathbb{R}^n$ in a direction $h \in \mathbb{R}^n$ if the limit

$$g'(x_0, h) := \lim_{t \downarrow 0} \frac{g(x_0 + th) - g(x_0)}{t} \quad (7.5)$$

exists, in which case $g'(x_0, h)$ is called the directional derivative of $g(x)$ at $x_0$ in the direction $h$. If $g$ is directionally differentiable at $x_0$ in every direction $h \in \mathbb{R}^n$, then it is said that $g$ is directionally differentiable at $x_0$. Note that whenever exists, $g'(x_0, h)$ is positively homogeneous in $h$, i.e., $g'(x_0, th) = tg'(x_0, h)$ for any $t \geq 0$. If $g(x)$ is directionally differentiable at $x_0$ and $g'(x_0, h)$ is linear in $h$, then it is said that $g(x)$ is Gâteaux differentiable at $x_0$. Equation (7.5) can be also written in the form

$$g(x_0 + h) = g(x_0) + g'(x_0, h) + r(h), \quad (7.6)$$

where the remainder term $r(h)$ is such that $r(th)/t \rightarrow 0$, as $t \downarrow 0$, for any fixed $h \in \mathbb{R}^n$. If, moreover, $g'(x_0, h)$ is linear in $h$ and the remainder term $r(h)$ is “uniformly small” in the sense that $r(h)/\|h\| \rightarrow 0$ as $h \rightarrow 0$, i.e., $r(h) = o(h)$, then it is said that $g(x)$ is differentiable at $x_0$ in the sense of Fréchet, or simply differentiable at $x_0$. 
Clearly Fréchet differentiability implies Gâteaux differentiability. The converse of that is not necessarily true. However, the following theorem shows that for locally Lipschitz continuous mappings both concepts do coincide. Recall that a mapping (function) \( g : \mathbb{R}^n \to \mathbb{R}^m \) is said to be Lipschitz continuous on a set \( X \subset \mathbb{R}^n \), if there is a constant \( c \geq 0 \) such that
\[
\|g(x_1) - g(x_2)\| \leq c \|x_1 - x_2\|, \quad \text{for all } x_1, x_2 \in X.
\]
If \( g \) is Lipschitz continuous on a neighborhood of every point of \( X \) (probably with different Lipschitz constants), then it is said that \( g \) is locally Lipschitz continuous on \( X \).

**Theorem 7.2.** Suppose that mapping \( g : \mathbb{R}^n \to \mathbb{R}^m \) is Lipschitz continuous in a neighborhood of a point \( x_0 \in \mathbb{R}^n \) and directionally differentiable at \( x_0 \). Then \( g'(x_0, \cdot) \) is Lipschitz continuous on \( \mathbb{R}^n \) and
\[
\lim_{t \to 0} \frac{g(x_0 + t h_1) - g(x_0) - g'(x_0, h_1) \cdot h_1}{t} = 0. \tag{7.7}
\]

**Proof.** For \( h_1, h_2 \in \mathbb{R}^n \) we have
\[
\|g'(x_0, h_1) - g'(x_0, h_2)\| = \lim_{t \to 0} \frac{\|g(x_0 + t h_1) - g(x_0 + t h_2)\|}{t}.
\]
Also, since \( g \) is Lipschitz continuous near \( x_0 \) with Lipschitz constant \( c \), we have that for \( t > 0 \) small enough
\[
\|g(x_0 + t h_1) - g(x_0 + t h_2)\| \leq ct \|h_1 - h_2\|.
\]
It follows that \( \|g'(x_0, h_1) - g'(x_0, h_2)\| \leq c \|h_1 - h_2\| \) for any \( h_1, h_2 \in \mathbb{R}^n \), i.e., \( g'(x_0, \cdot) \) is Lipschitz continuous on \( \mathbb{R}^n \).

Consider now a sequence \( t_k \downarrow 0 \) and a sequence \( \{h_k\} \) converging to a point \( h \in \mathbb{R}^n \). We have that
\[
g(x_0 + t_k h_k) - g(x_0) = (g(x_0 + t_k h) - g(x_0)) + (g(x_0 + t_k h_k) - g(x_0 + t_k h))
\]
and
\[
\|g(x_0 + t_k h_k) - g(x_0 + t_k h)\| \leq c t_k \|h_k - h\|
\]
for all \( k \) large enough. It follows that
\[
g'(x_0, h) = \lim_{k \to \infty} \frac{g(x_0 + t_k h_k) - g(x_0)}{t_k}. \tag{7.8}
\]
The proof of (7.7) can be completed now by arguing by a contradiction and using the fact that every bounded sequence in \( \mathbb{R}^n \) has a convergent subsequence. \( \square \)

We have that \( g \) is differentiable at a point \( x \in \mathbb{R}^n \), iff
\[
g(x + h) - g(x) = [\nabla g(x)] h + o(h), \tag{7.9}
\]
where \( \nabla g(x) \) is the so-called \( m \times n \) Jacobian matrix of partial derivatives \( [\partial g_i(x)/\partial x_j], \)
\( i = 1, ..., m, j = 1, ..., n \). If \( m = 1 \), i.e., \( g(x) \) is real valued, we call \( \nabla g(x) \) the gradient of \( g \) at \( x \). In that case equation (7.9) takes the form
\[
g(x + h) - g(x) = h^T \nabla g(x) + o(h). \tag{7.10}
\]
Note that when \( g(\cdot) \) is real valued, we write its gradient \( \nabla g(x) \) as a column vector. This is why there is a slight discrepancy between the notation of (7.10) and notation of (7.9) where the Jacobian matrix is of order \( m \times n \). If \( g(x, y) \) is a function (mapping) of two vector variables \( x \) and \( y \) and we consider derivatives of \( g(\cdot, y) \) while keeping \( y \) constant, we write the corresponding gradient (Jacobian matrix) as \( \nabla_x g(x, y) \).

**Clarke Generalized Gradient**

Consider now a locally Lipschitz continuous function \( f : U \to \mathbb{R} \) defined on an open set \( U \subset \mathbb{R}^n \). By Rademacher’s theorem we have that \( f(x) \) is differentiable on \( U \) almost everywhere. That is, the subset of \( U \) where \( f \) is not differentiable has Lebesgue measure zero. At a point \( \bar{x} \in U \) consider the set of all limits of the form \( \lim_{t_k \to 0} \nabla f(x_k) \) such that \( x_k \to \bar{x} \) and \( f \) is differentiable at \( x_k \). This set is nonempty and compact, its convex hull is called Clarke generalized gradient of \( f \) at \( \bar{x} \) and denoted \( \partial^c f(\bar{x}) \). The generalized directional derivative of \( f \) at \( \bar{x} \) is defined as
\[
f^c(\bar{x}, d) := \limsup_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}. \tag{7.11}
\]
It is possible to show that \( f^c(\bar{x}, \cdot) \) is the support function of the set \( \partial^c f(\bar{x}) \). That is
\[
f^c(\bar{x}, d) = \sup_{z \in \partial^c f(\bar{x})} z^T d, \quad \forall d \in \mathbb{R}^n. \tag{7.12}
\]

Function \( f \) is called regular in the sense of Clarke, or Clarke regular, at \( \bar{x} \in \mathbb{R}^n \) if \( f(\cdot) \) is directionally differentiable at \( \bar{x} \) and \( f'(\bar{x}, \cdot) = f^c(\bar{x}, \cdot) \). Any convex function \( f \) is Clarke regular and its Clarke generalized gradient \( \partial^c f(\bar{x}) \) coincides with the respective subdifferential in the sense of convex analysis. For a concave function \( f \), the function \(-f\) is Clarke regular, and we shall call it Clarke regular with the understanding that we modify the regularity requirement above to apply to \(-f\). In this case we have also \( \partial^c (-f)(\bar{x}) = -\partial^c f(\bar{x}) \).

We say that \( f \) is continuously differentiable at a point \( \bar{x} \in U \) if \( \partial^c f(\bar{x}) \) is a singleton. In other words, \( f \) is continuously differentiable at \( \bar{x} \) if \( f \) is differentiable at \( \bar{x} \) and \( \nabla f(\bar{x}) \) is continuous at \( \bar{x} \) on the set where \( f \) is differentiable. Note that continuous differentiability of \( f \) at a point \( \bar{x} \) does not imply differentiability of \( f \) at every point of any neighborhood of the point \( \bar{x} \).

Consider a composite real valued function \( f(x) := g(h(x)) \) with \( h : \mathbb{R}^m \to \mathbb{R}^n \) and \( g : \mathbb{R}^n \to \mathbb{R} \), and assume that \( g \) and \( h \) are locally Lipschitz continuous. Then
\[
\partial^c f(x) \subset \text{cl} \left\{ \text{conv} \left( \sum_{i=1}^n \alpha_i v_i : \alpha \in \partial^c g(y), \ v_i \in \partial h_i(x), \ i = 1, \ldots, n \right) \right\}, \tag{7.13}
\]
where \( \alpha = (\alpha_1, ..., \alpha_n), \ y = h(x) \) and \( h_1, ..., h_n \) are components of \( h \). The equality in (7.13) holds true, if any one of the following conditions is satisfied: (i) \( g \) and \( h_i, \ i = 1, \ldots, n \), are Clarke-regular and every element in \( \partial^c g(y) \) has non-negative components, (ii) \( g \) is differentiable and \( n = 1 \), (iii) \( g \) is Clarke-regular and \( h \) is differentiable.
7.1.2 Elements of Convex Analysis

Let $C$ be a subset of $\mathbb{R}^n$. It is said that $x \in \mathbb{R}^n$ is an interior point of $C$ if there is a neighborhood $N$ of $x$ such that $N \subset C$. The set of interior points of $C$ is denoted $\text{int}(C)$. The convex hull of $C$, denoted $\text{conv}(C)$, is the smallest convex set including $C$. It is said that $C$ is a cone if for any $x \in C$ and $t \geq 0$ it follows that $tx \in C$. The polar cone of a cone $C \subset \mathbb{R}^n$ is defined as

$$C^* := \{z \in \mathbb{R}^n : z^T x \leq 0, \forall x \in C\}. \quad (7.14)$$

We have that the polar of the polar cone $C^{**} = (C^*)^*$ is equal to the topological closure of the convex hull of $C$, and that $C^{**} = C$ iff the cone $C$ is convex and closed.

Let $C$ be a nonempty convex subset of $\mathbb{R}^n$. The affine space generated by $C$ is the space of points in $\mathbb{R}^n$ of the form $tx + (1-t)y$, where $x, y \in C$ and $t \in \mathbb{R}$. It is said that a point $x \in \mathbb{R}^n$ belongs to the relative interior of the set $C$ if $x$ is an interior point of $C$ relative to the affine space generated by $C$, i.e., there exists a neighborhood of $x$ such that its intersection with the affine space generated by $C$ is included in $C$. The relative interior set of $C$ is denoted $\text{ri}(C)$. Note that if the interior of $C$ is nonempty, then the affine space generated by $C$ coincides with $\mathbb{R}^n$, and hence in that case $\text{ri}(C) = \text{int}(C)$. Note also that the relative interior of any convex set $C \subset \mathbb{R}^n$ is nonempty. The recession cone of the set $C$ is formed by vectors $h \in \mathbb{R}^n$ such that for any $x \in C$ and any $t > 0$ it follows that $x + th \in C$. The recession cone of the convex set $C$ is convex and is closed if the set $C$ is closed. Also the convex set $C$ is bounded iff its recession cone is $\{0\}$.

**Theorem 7.3 (Helly).** Let $A_i$, $i \in \mathcal{I}$, be a family of convex subsets of $\mathbb{R}^n$. Suppose that the intersection of any $n+1$ sets of this family is nonempty and either the index set $\mathcal{I}$ is finite or the sets $A_i$, $i \in \mathcal{I}$, are closed and there exists no common nonzero recession direction to the sets $A_i$, $i \in \mathcal{I}$. Then the intersection of all sets $A_i$, $i \in \mathcal{I}$, is nonempty.

The support function $s(\cdot) = s_C(\cdot)$ of a (nonempty) set $C \subset \mathbb{R}^n$ is defined as

$$s(h) := \sup_{z \in C} z^T h. \quad (7.15)$$

The support function $s(\cdot)$ is convex, positively homogeneous and lower semicontinuous. The support function of a set $C$ coincides with the support function of the set $\text{cl}(\text{conv} C)$. If $s_1(\cdot)$ and $s_2(\cdot)$ are support functions of convex closed sets $C_1$ and $C_2$, respectively, then $s_1(\cdot) \leq s_2(\cdot)$ iff $C_1 \subset C_2$, and $s_1(\cdot) = s_2(\cdot)$ iff $C_1 = C_2$.

Let $C \subset \mathbb{R}^n$ be a convex closed set. The normal cone to $C$ at a point $x_0 \in C$ is defined as

$$\mathcal{N}_C(x_0) := \{z : z^T(x - x_0) \leq 0, \forall x \in C\}. \quad (7.16)$$

By definition $\mathcal{N}_C(x_0) := \emptyset$ if $x_0 \notin C$. The topological closure of the radial cone $\mathcal{R}_C(x_0) := \cup_{t>0} \{t(C - x_0)\}$ is called the tangent cone to $C$ at $x_0 \in C$, and denoted $\mathcal{T}_C(x_0)$. Both cones $\mathcal{T}_C(x_0)$ and $\mathcal{N}_C(x_0)$ are closed and convex, and each one is the polar cone of the other.

Consider an extended real valued function $f : \mathbb{R}^n \to \mathbb{R}$. It is not difficult to show that $f$ is convex iff its epigraph $\text{epi} f$ is a convex subset of $\mathbb{R}^{n+1}$. Suppose that $f$ is a convex function and $x_0 \in \mathbb{R}^n$ is a point such that $f(x_0)$ is finite. Then $f(x)$ is directionally
differentiable at $x_0$, its directional derivative $f'(x_0, \cdot)$ is an extended real valued convex positively homogeneous function and can be written in the form
\[
f'(x_0, h) = \inf_{t > 0} \frac{f(x_0 + th) - f(x_0)}{t}.
\] (7.17)

Moreover, if $x_0$ is in the interior of the domain of $f(\cdot)$, then $f(x)$ is Lipschitz continuous in a neighborhood of $x_0$, the directional derivative $f'(x_0, h)$ is finite valued for any $h \in \mathbb{R}^n$, and $f(x)$ is differentiable at $x_0$ iff $f'(x_0, h)$ is linear in $h$.

It is said that a vector $z \in \mathbb{R}^n$ is a subgradient of $f(x)$ at $x_0$ if
\[
f(x) - f(x_0) \geq z^T(x - x_0), \quad \forall x \in \mathbb{R}^n.
\] (7.18)

The set of all subgradients of $f(x)$, at $x_0$, is called the subdifferential and denoted $\partial f(x_0)$. The subdifferential $\partial f(x_0)$ is a closed convex subset of $\mathbb{R}^n$. It is said that $f$ is subdifferentiable at $x_0$ if $\partial f(x_0)$ is nonempty. If $f$ is subdifferentiable at $x_0$, then the normal cone $N_{\text{dom} f(x_0)}$ to the domain of $f$ at $x_0$, forms the recession cone of the set $\partial f(x_0)$. It is also clear that if $f$ is subdifferentiable at $x_0$, then $f(x) \geq -\infty$ for any $x$ and hence $f$ is proper.

By duality theory of convex analysis we have that if the directional derivative $f'(x_0, \cdot)$ is lower semicontinuous, then
\[
f'(x_0, h) = \sup_{z \in \partial f(x_0)} z^T h, \quad \forall h \in \mathbb{R}^n,
\] (7.19)
i.e., $f'(x_0, \cdot)$ is the support function of the set $\partial f(x_0)$. In particular, if $x_0$ is an interior point of the domain of $f(x)$, then $f'(x_0, \cdot)$ is continuous, $\partial f(x_0)$ is nonempty and compact and (7.19) holds. Conversely, if $\partial f(x_0)$ is nonempty and compact, then $x_0$ is an interior point of the domain of $f(x)$. Also $f(x)$ is differentiable at $x_0$ iff $\partial f(x_0)$ is a singleton, i.e., contains only one element, which then coincides with the gradient $\nabla f(x_0)$.

**Theorem 7.4 (Moreau–Rockafellar).** Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, be proper convex functions, $f(\cdot) := f_1(\cdot) + \ldots + f_m(\cdot)$ and $x_0$ be a point such that $f_i(x_0)$ are finite, i.e., $x_0 \in \cap_{i=1}^m \text{dom } f_i$. Then
\[
\partial f_1(x_0) + \ldots + \partial f_m(x_0) \subset \partial f(x_0).
\] (7.20)

Moreover,
\[
\partial f_1(x_0) + \ldots + \partial f_m(x_0) = \partial f(x_0).
\] (7.21)

if any one of the following conditions holds: (i) the set $\cap_{i=1}^m \text{ri}(\text{dom } f_i)$ is nonempty, (ii) the functions $f_1, \ldots, f_k$, $k \leq m$, are polyhedral and the intersection of the sets $\cap_{i=1}^k \text{dom } f_i$ and $\cap_{i=k+1}^m \text{ri}(\text{dom } f_i)$ is nonempty, (iii) there exists a point $\bar{x} \in \text{dom } f_m$ such that $\bar{x} \in \text{int}(\text{dom } f_i), i = 1, \ldots, m - 1$.

In particular, if all functions $f_1, \ldots, f_m$ in the above theorem are polyhedral, then the equation (7.21) holds without an additional regularity condition.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be an extended real valued function. The conjugate function of $f$ is
\[
f^*(z) := \sup_{x \in \mathbb{R}^n} \{z^T x - f(x)\}.
\] (7.22)
The conjugate function $f^* : \mathbb{R}^n \to \mathbb{R}$ is always convex and lower semicontinuous. The conjugate of $f^*$ is denoted $f^{**}$. Note that if $f(x) = -\infty$ at some $x \in \mathbb{R}^n$, then $f^*(\cdot) \equiv +\infty$ and $f^{**}(\cdot) \equiv -\infty$.

**Theorem 7.5 (Fenchel-Moreau).** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a proper extended real valued convex function. Then

$$f^{**} = \text{lsc } f.$$  \hspace{1cm} (7.23)

It follows from (7.23) that if $f$ is proper and convex, then $f^{**} = f$ iff $f$ is lower semicontinuous. Also it immediately follows from the definitions that

$$z \in \partial f(x) \iff f^*(z) + f(x) = z^T x.$$  

By applying that to the function $f^{**}$, instead of $f$, we obtain that $z \in \partial f^{**}(x)$ iff $f^{***}(z) + f^{**}(x) = z^T x$. Now by the Fenchel–Moreau Theorem we have that $f^{***} = f^*$, and hence $z \in \partial f^{**}(x)$ iff $f^*(z) + f^{**}(x) = z^T x$. Consequently we obtain

$$\partial f^{**}(x) = \arg \max_{z \in \mathbb{R}^n} \{ z^T x - f^*(z) \}, \hspace{1cm} (7.24)$$

and if $f^{**}(x) = f(x)$ and is finite, then $\partial f^{**}(x) = \partial f(x)$.

**Strong convexity.** Let $X \subset \mathbb{R}^n$ be a nonempty closed convex set. It is said that a function $f : X \to \mathbb{R}$ is **strongly convex**, with parameter $c > 0$, if\(^2\)

$$tf(x') + (1 - t)f(x) \geq f(tx' + (1 - t)x) + \frac{1}{2}c(1 - t)||x' - x||^2,$$  \hspace{1cm} (7.25)

for all $x, x' \in X$ and $t \in [0, 1]$. It is not difficult to verify that $f$ is strongly convex iff the function $\psi(x) := f(x) - \frac{1}{2}c||x||^2$ is convex on $X$.

Indeed, convexity of $\psi$ means that the inequality

$$tf(x') - \frac{1}{2}c||x'||^2 + (1 - t)f(x) - \frac{1}{2}c(1 - t)||x||^2 \geq f(tx' + (1 - t)x) - \frac{1}{2}c||tx' + (1 - t)x||^2,$$

holds for all $t \in [0, 1]$ and $x, x' \in X$. By the identity

$$t||x'||^2 + (1 - t)||x||^2 - ||tx' + (1 - t)x||^2 = t(1 - t)||x' - x||^2,$$

this is equivalent to (7.25).

If the set $X$ has a nonempty interior and $f : X \to \mathbb{R}$ is continuous and differentiable at every point $x \in \text{int}(X)$, then $f$ is strongly convex iff

$$f(x') \geq f(x) + (x' - x)^T \nabla f(x) + \frac{1}{2}c||x' - x||^2, \hspace{1cm} \forall x, x' \in \text{int}(X), \hspace{1cm} (7.26)$$

or equivalently iff

$$(x' - x)^T (\nabla f(x') - \nabla f(x)) \geq c||x' - x||^2, \hspace{1cm} \forall x, x' \in \text{int}(X). \hspace{1cm} (7.27)$$

\(^2\)Unless stated otherwise we denote by $|| \cdot ||$ the Euclidean norm on $\mathbb{R}^n$. 

---

7.1. Optimization and Convex Analysis

341
7.1.3 Optimization and Duality

Consider a real valued function $L : X \times Y \to \mathbb{R}$, where $X$ and $Y$ are arbitrary sets. We can associate with the function $L(x, y)$ the following two optimization problems:

\[
\begin{align*}
\min_{x \in X} \{ f(x) := \sup_{y \in Y} L(x, y) \}, & \quad (7.28) \\
\max_{y \in Y} \{ g(y) := \inf_{x \in X} L(x, y) \}, & \quad (7.29)
\end{align*}
\]

viewed as dual to each other. We have that for any $x \in X$ and $y \in Y$,

\[
g(y) = \inf_{x' \in X} L(x', y) \leq L(x, y) \leq \sup_{y' \in Y} L(x, y') = f(x),
\]

and hence the optimal value of problem (7.28) is greater than or equal to the optimal value of problem (7.29). It is said that a point $(\bar{x}, \bar{y}) \in X \times Y$ is a saddle point of $L(x, y)$ if

\[
L(\bar{x}, y) \leq L(\bar{x}, \bar{y}) \leq L(x, \bar{y}), \quad \forall (x, y) \in X \times Y. \quad (7.30)
\]

Theorem 7.6. The following holds: (i) The optimal value of problem (7.28) is greater than or equal to the optimal value of problem (7.29). (ii) Problems (7.28) and (7.29) have the same optimal value and each has an optimal solution if and only if there exists a saddle point $(\bar{x}, \bar{y})$. In that case $\bar{x}$ and $\bar{y}$ are optimal solutions of problems (7.28) and (7.29), respectively. (iii) If problems (7.28) and (7.29) have the same optimal value, then the set of saddle points coincides with the Cartesian product of the sets of optimal solutions of (7.28) and (7.29).

Suppose that there is no duality gap between problems (7.28) and (7.29), i.e., their optimal values are equal to each other, and let $\bar{y}$ be an optimal solution of problem (7.29). By the above we have that the set of optimal solutions of problem (7.28) is contained in the set of optimal solutions of the problem

\[
\min_{x \in X} L(x, \bar{y}), \quad (7.31)
\]

and the common optimal value of problems (7.28) and (7.29) is equal to the optimal value of (7.31). In applications of the above results to optimization problems with constraints, the function $L(x, y)$ usually is the Lagrangian and $y$ is a vector of Lagrange multipliers. The inclusion of the set of optimal solutions of (7.28) into the set of optimal solutions of (7.31) can be strict (see the following example).

Example 7.7 Consider the linear problem

\[
\min_{x \in \mathbb{R}} x \quad \text{subject to } x \geq 0. \quad (7.32)
\]

This problem has unique optimal solution $\bar{x} = 0$ and can be written in the minimax form (7.28) with $L(x, y) := x - yx$, $Y := \mathbb{R}_+$, and $X := \mathbb{R}$. The objective function $g(y)$ of its dual (of the form (7.29)) is equal to $-\infty$ for all $y$ except $y = 1$ for which $g(y) = 0$. There is no duality gap here between the primal and dual problems and the dual problem has unique feasible point $\bar{y} = 1$, which is also its optimal solution. The corresponding problem (7.31) takes here the form of minimizing $L(x, 1) \equiv 0$ over $x \in \mathbb{R}$, with the set of optimal solutions equal to $\mathbb{R}$. That is, in this example the set of optimal solutions of (7.28) is a strict subset of the set of optimal solutions of (7.31).
Conjugate Duality

An alternative approach to duality, referred to as conjugate duality, is the following. Consider an extended real valued function \( \psi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \). Let \( \vartheta(y) \) be the optimal value of the parameterized problem
\[
\min_{x \in \mathbb{R}^n} \psi(x, y),
\]
ie., \( \vartheta(y) := \inf_{x \in \mathbb{R}^n} \psi(x, y) \). Note that implicitly the optimization in the above problem is performed over the domain of the function \( \psi(\cdot, y) \), i.e., \( \text{dom} \psi(\cdot, y) \) can be viewed as the feasible set of problem (7.33).

The conjugate of the function \( \vartheta(y) \) can be expressed in terms of the conjugate of \( \psi(x, y) \). That is, the conjugate of \( \psi \) is
\[
\psi^*(x^*, y^*) := \sup_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m} \left\{ (x^*)^T x + (y^*)^T y - \psi(x, y) \right\},
\]
and hence the conjugate of \( \vartheta \) can be written as
\[
\vartheta^*(y^*) := \sup_{y \in \mathbb{R}^m} \left\{ (y^*)^T y - \vartheta(y) \right\} = \sup_{y \in \mathbb{R}^m} \left\{ (y^*)^T y - \inf_{x \in \mathbb{R}^n} \psi(x, y) \right\} = \sup_{y \in \mathbb{R}^m} \left\{ (y^*)^T y - \psi^*(-y^*) \right\}.
\]

Consequently, the conjugate of \( \vartheta^* \) is
\[
\vartheta^{**}(y) = \sup_{y^* \in \mathbb{R}^m} \left\{ (y^*)^T y - \psi^*(0, y^*) \right\}.
\]

This leads to the following dual of (7.33):
\[
\max_{y^* \in \mathbb{R}^m} \left\{ (y^*)^T y - \psi^*(0, y^*) \right\}.
\]

In the above formulation of problem (7.33) and its (conjugate) dual (7.35) we have that \( \vartheta(y) \) and \( \vartheta^{**}(y) \) are optimal values of (7.33) and (7.35), respectively. Suppose that \( \vartheta(\cdot) \) is convex. Then we have by the Fenchel–Moreau Theorem that either \( \vartheta^{**}(\cdot) \) is identically \( -\infty \), or
\[
\vartheta^{**}(y) = (\text{lsc} \ \vartheta)(y), \quad \forall y \in \mathbb{R}^m.
\]

It follows that \( \vartheta^{**}(y) \leq \vartheta(y) \) for any \( y \in \mathbb{R}^m \). It is said that there is no duality gap between (7.33) and its dual (7.35) if \( \vartheta^{**}(y) = \vartheta(y) \).

Suppose now that the function \( \psi(x, y) \) is convex (as a function of \( (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \)). Then it is straightforward to verify that the optimal value function \( \vartheta(y) \) is also convex. It is said that the problem (7.33) is \( \text{subconsistent} \), for a given value of \( y \), if \( \text{lsc} \ \vartheta(y) < +\infty \). If problem (7.33) is feasible, i.e., \( \text{dom} \ \psi(\cdot, y) \) is nonempty, then \( \vartheta(y) < +\infty \), and hence (7.33) is subconsistent.

**Theorem 7.8.** Suppose that the function \( \psi(\cdot, \cdot) \) is convex. Then the following holds: (i) The optimal value function \( \vartheta(\cdot) \) is convex. (ii) If problem (7.33) is subconsistent, then \( \vartheta^{**}(y) = \vartheta(y) \) if and only if the optimal value function \( \vartheta(\cdot) \) is lower semicontinuous at \( y \). (iii) If \( \vartheta^{**}(y) \) is finite, then the set of optimal solutions of the dual problem (7.35) coincides with \( \partial \vartheta^{**}(y) \). (iv) The set of optimal solutions of the dual problem (7.35) is nonempty and bounded if and only if \( \vartheta(y) \) is finite and \( \vartheta(\cdot) \) is continuous at \( y \).
A few words about the above statements are now in order. Assertion (ii) follows by the Fenchel–Moreau Theorem. Assertion (iii) follows from formula (7.24). If \( \vartheta(\cdot) \) is continuous at \( y \), then it is lower semicontinuous at \( y \), and hence \( \vartheta^{**}(y) = \vartheta(y) \). Moreover, in that case \( \partial \vartheta^{**}(y) = \partial \vartheta(y) \) and is nonempty and bounded provided that \( \vartheta(y) \) is finite. It follows then that the set of optimal solutions of the dual problem (7.35) is nonempty and bounded. Conversely, if the set of optimal solutions of (7.35) is nonempty and bounded, then, by (iii), \( \partial \vartheta^{**}(y) \) is nonempty and bounded, and hence by convex analysis \( \vartheta(\cdot) \) is continuous at \( y \). Note also that if \( \partial \vartheta(y) \) is nonempty, then \( \vartheta^{**}(y) = \vartheta(y) \) and \( \partial \vartheta^{**}(y) = \partial \vartheta(y) \).

The above analysis can be also used in order to describe differentiability properties of the optimal value function \( \vartheta(\cdot) \) in terms of its subdifferentials.

**Theorem 7.9.** Suppose that the function \( \psi(\cdot, \cdot) \) is convex and let \( y \in \mathbb{R}^m \) be a given point. Then the following holds: (i) The optimal value function \( \vartheta(\cdot) \) is subdifferentiable at \( y \) if and only if \( \vartheta(\cdot) \) is lower semicontinuous at \( y \) and the dual problem (7.35) possesses an optimal solution. (ii) The subdifferential \( \partial \vartheta(y) \) is nonempty and bounded if and only if \( \vartheta(y) \) is finite and the set of optimal solutions of the dual problem (7.35) is nonempty and bounded. (iii) In both above cases \( \partial \vartheta(y) \) coincides with the set of optimal solutions of the dual problem (7.35).

Since \( \vartheta(\cdot) \) is convex, we also have that \( \partial \vartheta(y) \) is nonempty and bounded iff \( \vartheta(y) \) is finite and \( y \in \text{int}(\text{dom } \vartheta) \). The condition \( y \in \text{int}(\text{dom } \vartheta) \) means the following: there exists a neighborhood \( N \) of \( y \) such that for any \( y' \in N \) the domain of \( \psi(\cdot, y') \) is nonempty.

As an example let us consider the following problem

\[
\begin{align*}
\text{Min}_{x \in X} & \quad f(x) \\
\text{subject to} & \quad g_i(x) + y_i \leq 0, \ i = 1, \ldots, m,
\end{align*}
\]

where \( X \) is a subset of \( \mathbb{R}^n \), \( f(x) \) and \( g_i(x) \) are real valued functions, and \( y = (y_1, \ldots, y_m) \) is a vector of parameters. We can formulate this problem in the form (7.33) by defining

\[
\psi(x, y) := f(x) + F(G(x) + y),
\]

where \( f(x) := f(x) + I_X(x) \) (recall that \( I_X \) denotes the indicator function of the set \( X \)) and \( F(\cdot) \) is the indicator function of the negative orthant, i.e., \( F(z) := 0 \) if \( z_i \leq 0, \ i = 1, \ldots, m \), and \( F(z) := +\infty \) otherwise, and \( G(x) := (g_1(x), \ldots, g_m(x)) \).

Suppose that the problem (7.37) is convex, that is, the set \( X \) and the functions \( f(x) \) and \( g_i(x), \ i = 1, \ldots, m \), are convex. Then it is straightforward to verify that the function \( \psi(x, y) \) is also convex. Let us calculate the conjugate of the function \( \psi(x, y) \).

\[
\psi^*(x^*, y^*) = \sup_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m} \left\{ ((x^*)^T x + (y^*)^T y - f(x) - F(G(x) + y) \right\} = \sup_{x \in \mathbb{R}^n} \left\{ ((x^*)^T x - f(x) - (y^*)^T G(x) + \sup_{y \in \mathbb{R}^m} [(y^*)^T(G(x) + y) - F(G(x) + y)] \right\}.
\]

By change of variables \( z = G(x) + y \) we obtain that

\[
\sup_{y \in \mathbb{R}^m} [(y^*)^T(G(x) + y) - F(G(x) + y)] = \sup_{z \in \mathbb{R}^m} [(y^*)^T z - F(z)] = I_{\mathbb{R}^n^*}(y^*).
\]
Therefore we obtain
\[ \psi^*(x^*, y^*) = \sup_{x \in X} \left\{ (x^*)^T x - L(x, y^*) \right\} + I_{\mathbb{R}^m}(y^*) , \]
where \( L(x, y^*) := f(x) + \sum_{i=1}^m y_i^* g_i(x) \), is the Lagrangian of the problem. Consequently the dual of the problem (7.37) can be written in the form
\[ \max_{\lambda \geq 0} \left\{ \lambda^T y + \inf_{x \in X} L(x, \lambda) \right\} . \tag{7.38} \]

Note that we changed the notation from \( y^* \) to \( \lambda \) in order to emphasize that the above problem (7.38) is the standard Lagrangian dual of (7.37) with \( \lambda \) being vector of Lagrange multipliers. The results of Propositions 7.8 and 7.9 can be applied to problem (7.37) and its dual (7.38) in a straightforward way.

As another example consider a function \( L : \mathbb{R}^n \times Y \to \mathbb{R} \), where \( Y \) is a vector space (not necessarily finite dimensional), and the corresponding pair of dual problems (7.28) and (7.29). Define
\[ \varphi(y, z) := \sup_{x \in \mathbb{R}^n} \left\{ z^T x - L(x, y) \right\} , \quad (y, z) \in Y \times \mathbb{R}^n . \tag{7.39} \]
Note that the problem
\[ \max_{y \in Y} \{ -\varphi(y, 0) \} \tag{7.40} \]
coinsides with the problem (7.29). Note also that for every \( y \in Y \) the function \( \varphi(y, \cdot) \) is the conjugate of \( L(\cdot, y) \). Suppose that for every \( y \in Y \) the function \( L(\cdot, y) \) is convex and lsc. Then by the Fenchel-Moreau Theorem we have that the conjugate of the conjugate of \( L(\cdot, y) \) coincides with \( L(\cdot, y) \). Consequently the dual of (7.40), of the form (7.35), coincides with the problem (7.28). This leads to the following result.

**Theorem 7.10.** Let \( Y \) be an abstract vector space and \( L : \mathbb{R}^n \times Y \to \mathbb{R} \). Suppose that: (i) for every \( x \in \mathbb{R}^n \) the function \( L(x, \cdot) \) is concave, (ii) for every \( y \in Y \) the function \( L(\cdot, y) \) is convex and lower semicontinuous, (iii) problem (7.28) has a nonempty and bounded set of optimal solutions. Then the optimal values of problems (7.28) and (7.29) are equal to each other.

**Proof.** Consider function \( \varphi(y, z) \), defined in (7.39), and the corresponding optimal value function
\[ \vartheta(z) := \inf_{y \in Y} \varphi(y, z) . \tag{7.41} \]
Since \( \varphi(y, z) \) is given by maximum of convex in \( (y, z) \) functions, it is convex, and hence \( \vartheta(z) \) is also convex. We have that \(-\vartheta(0)\) is equal to the optimal value of the problem (7.29) and \(-\vartheta^*(0)\) is equal to the optimal value of (7.28). We also have that
\[ \vartheta^*(z^*) = \sup_{y \in Y} L(z^*, y) , \]
and (see (7.24))
\[ \vartheta^{**}(0) = -\arg\min_{z^* \in \mathbb{R}^n} \vartheta^*(z^*) = -\arg\min_{z^* \in \mathbb{R}^n} \{ \sup_{y \in Y} L(z^*, y) \} . \]
That is, \( -\partial \vartheta^*(0) \) coincides with the set of optimal solutions of the problem (7.28). It follows by assumption (iii) that \( \partial \vartheta^*(0) \) is nonempty and bounded. Since \( \vartheta^* : \mathbb{R}^n \to \mathbb{R} \) is a convex function, this in turn implies that \( \vartheta^*(\cdot) \) is continuous in a neighborhood of 0 \( \in \mathbb{R}^n \). It follows that \( \vartheta(\cdot) \) is also continuous in a neighborhood of 0 \( \in \mathbb{R}^n \), and hence \( \vartheta^*(0) = \vartheta(0) \). This completes the proof. \( \square \)

**Remark 24.** Note that it follows from the lower semicontinuity of \( L(\cdot, y) \) that the max-function \( f(x) = \sup_{y \in Y} L(x, y) \) is also lsc. Indeed, the epigraph of \( f(\cdot) \) is given by the intersection of the epigraphs of \( L(\cdot, y) \), \( y \in Y \), and hence is closed. Therefore, if in addition, the set \( X \subset \mathbb{R}^n \) is compact and problem (7.28) has a finite optimal value, then the set of optimal solutions of (7.28) is nonempty and compact, and hence bounded.

**Hoffman’s Lemma**

The following result about Lipschitz continuity of linear systems is known as Hoffman’s lemma. For a vector \( a = (a_1, \ldots, a_m)^T \in \mathbb{R}^m \), we use notation \((a)_+ \) componentwise, i.e., \((a)_+ := ([a_1]_+, \ldots, [a_m]_+)^T \), where \([a_i]_+ := \max\{0, a_i\} \).

**Theorem 7.11 (Hoffman).** Consider the multifunction \( M(b) := \{ x \in \mathbb{R}^n : Ax \leq b \} \), where \( A \) is a given \( m \times n \) matrix. Then there exists a positive constant \( \kappa \), depending on \( A \), such that for any \( x \in \mathbb{R}^n \) and any \( b \in \text{dom} M \),

\[
\text{dist}(x, M(b)) \leq \kappa \| (Ax - b)_+ \|. \tag{7.42}
\]

**Proof.** Suppose that \( b \in \text{dom} M \), i.e., the system \( Ax \leq b \) has a feasible solution. Note that for any \( a \in \mathbb{R}^n \) we have that \( \|a\| = \sup_{\|z\|_1 \leq 1} z^T a \), where \( \| \cdot \|_1 \) is the dual of the norm \( \| \cdot \|. \) Then we have

\[
\text{dist}(x, M(b)) = \inf_{x' \in M(b)} \|x - x'\| = \inf_{\|z\|_1 \leq 1} \sup_{Ax' \leq b} z^T(x - x') = \sup_{\|z\|_1 \leq 1} \inf_{Ax \leq b} z^T(x - x'),
\]

where the interchange of the ‘min’ and ‘max’ operators can be justified, for example, by applying theorem 7.10 (see Remark 24, page 346). By making change of variables \( y = x - x' \) and using linear programming duality we obtain

\[
\inf_{Ax \leq b} z^T(x - x') = \inf_{A^T \lambda \geq Ax - b} z^T y = \sup_{\lambda \geq 0, A^T \lambda = z} \lambda^T(Ax - b).
\]

It follows that

\[
\text{dist}(x, M(b)) = \sup_{\lambda \geq 0, \|A^T \lambda\|_1 \leq 1} \lambda^T (Ax - b). \tag{7.43}
\]

Since any two norms on \( \mathbb{R}^n \) are equivalent, we can assume without loss of generality that \( \| \cdot \| \) is the \( \ell_1 \) norm, and hence its dual is the \( \ell_\infty \) norm. For such choice of a polyhedral norm, we have that the set \( S := \{ \lambda : \lambda \geq 0, \|A^T \lambda\|_1 \leq 1 \} \) is polyhedral. We obtain that the right hand side of (7.43) is given by a maximization of a linear function over the
polyhedral set $S$ and has a finite optimal value (since the left hand side of (7.43) is finite), and hence has an optimal solution $\lambda$. It follows that
\[
\text{dist}(x, S(b)) = \lambda^T(Ax - b) \leq \lambda^T(Ax - b)_{+} \leq \|\lambda\|\|(Ax - b)_{+}\|.
\]
It remains to note that the polyhedral set $S$ depends only on $A$, and can be represented as the direct sum $S = S_0 + C$ of a bounded polyhedral set $S_0$ and a polyhedral cone $C$, and that optimal solution $\lambda$ can be taken to be an extreme point of the polyhedral set $S_0$. Consequently, (7.42) follows with $\kappa := \max_{\lambda \in S_0} \|\lambda\|\ast$. □

The term $\|(Ax - b)_{+}\|$, in the right hand side of (7.42), measures the infeasibility of the point $x$.

Consider now the following linear programming problem
\[
\begin{align*}
\text{Min} & \quad c^Tx \\
\text{subject to} & \quad Ax \leq b, \\
& \quad c^Tx - t \leq 0.
\end{align*}
\]
(7.44)

A slight variation of the proof of Hoffman's lemma leads to the following result.

**Theorem 7.12.** Let $S(b)$ be the set of optimal solutions of problem (7.44). Then there exists a positive constant $\gamma$, depending only on $A$, such that for any $b, b' \in \text{dom} S$ and any $x \in S(b)$,
\[
\text{dist}(x, S(b')) \leq \gamma \|b - b'\|. \tag{7.45}
\]

**Proof.** Problem (7.44) can be written in the following equivalent form:
\[
\begin{align*}
\text{Min} & \quad t \\
\text{subject to} & \quad Ax \leq b, \\
& \quad c^Tx - t \leq 0.
\end{align*}
\]
(7.46)

Denote by $M(b)$ the set of feasible points of problem (7.46), i.e.,
\[
M(b) := \{(x, t) : Ax \leq b, c^Tx - t \leq 0\}.
\]

Let $b, b' \in \text{dom} S$ and consider a point $(x, t) \in M(b)$. Proceeding as in the proof of Theorem 7.11 we can write
\[
\text{dist}((x, t), M(b')) = \sup_{\|z, a\|\leq 1, \lambda \geq 0, A^T \lambda + a = z} z^T(x - x') + a(t - t').
\]

By changing variables $y = x - x'$ and $s = t - t'$ and using linear programming duality, we have
\[
\begin{align*}
\inf_{Ax' \leq b', c^Tx' \leq t'} z^T(x - x') + a(t - t') &= \sup_{\lambda \geq 0, A^T \lambda + a = z} \lambda^T(Ax - b') + a(c^Tx - t) \\
\end{align*}
\]
for $a \geq 0$, and for $a < 0$ the above minimum is $-\infty$. By using $\ell_1$ norm $\|\cdot\|$, and hence $\ell_\infty$ norm $\|\cdot\|\ast$, we obtain that
\[
\text{dist}((x, t), M(b')) = \bar{\lambda}^T(Ax - b') + \bar{a}(c^Tx - t),
\]
where \((\bar{\lambda}, \bar{a})\) is an optimal solution of the problem

\[
\max_{\lambda \geq 0, a \geq 0} \lambda^T (Ax - b') + a(c^T x - t) \quad \text{subject to} \quad \|A^T \lambda + ac\|^* \leq 1, \ a \leq 1. \tag{7.47}
\]

By normalizing \(c\) we can assume without loss of generality that \(\|c\|^* \leq 1\). Then by replacing the constraint \(\|A^T \lambda + ac\|^* \leq 1\) with the constraint \(\|A^T \lambda\|^* \leq 2\) we increase the feasible set of problem (7.47), and hence increase its optimal value. Let \((\bar{\lambda}, \bar{a})\) be an optimal solution of the obtained problem. Note that \(\bar{\lambda}\) can be taken to be an extreme point of the polyhedral set \(S := \{\lambda : \|A^T \lambda\|^* \leq 2\}\). The polyhedral set \(S\) depends only on \(A\) and has a finite number of extreme points. Therefore \(\|\bar{\lambda}\|^*\) can be bounded by a constant \(\gamma\) which depends only on \(A\). Since \((x, t) \in M(b)\), and hence \(Ax - b \leq 0\) and \(c^T x - t \leq 0\), we have

\[
\bar{\lambda}^T (Ax - b') = \bar{\lambda}^T (Ax - b) + \bar{\lambda}^T (b - b') \leq \|\bar{\lambda}\|^* \|b - b'\|
\]

and \(\bar{a}(c^T x - t) \leq 0\), and hence

\[
\text{dist}((x, t), M(b')) \leq \|\bar{\lambda}\|^* \|b - b'\| \leq \gamma \|b - b'\|. \quad \tag{7.48}
\]

The above inequality implies (7.45). \(\square\)

### 7.1.4 Optimality Conditions

Consider optimization problem

\[
\min_{x \in X} f(x), \tag{7.49}
\]

where \(X \subset \mathbb{R}^n\) and \(f : \mathbb{R}^n \to \mathbb{R}\) is an extended real valued function.

#### First Order Optimality Conditions

**Convex case.** Suppose that the function \(f : \mathbb{R}^n \to \mathbb{R}\) is convex. It follows immediately from the definition of the subdifferential that if \(f(\bar{x})\) is finite for some point \(\bar{x} \in \mathbb{R}^n\), then \(f(x) \geq f(\bar{x})\) for all \(x \in \mathbb{R}^n\) iff

\[
0 \in \partial f(\bar{x}). \tag{7.50}
\]

That is, condition (7.50) is necessary and sufficient for the point \(\bar{x}\) to be a (global) minimizer of \(f(x)\) over \(x \in \mathbb{R}^n\).

Suppose, further, that the set \(X \subset \mathbb{R}^n\) is convex and closed and the function \(f : \mathbb{R}^n \to \mathbb{R}\) is proper and convex, and consider a point \(\bar{x} \in X \cap \text{dom} f\). It follows that the function \(f(\bar{x}) := f(x) + \|X(x)\) is convex, and of course the point \(\bar{x}\) is an optimal solution of the problem (7.49) iff \(\bar{x}\) is a (global) minimizer of \(f(x)\). Suppose that

\[
\text{ri}(X) \cap \text{ri}(\text{dom} f) \neq \emptyset. \tag{7.51}
\]

Then by the Moreau-Rockafellar Theorem we have that \(\partial f(\bar{x}) = \partial f(\bar{x}) + \partial I_X(\bar{x})\). Recalling that \(\partial I_X(\bar{x}) = N_X(\bar{x})\), we obtain that \(\bar{x}\) is an optimal solution of problem (7.49) iff

\[
0 \in \partial f(\bar{x}) + N_X(\bar{x}), \tag{7.52}
\]
provided that the regularity condition \( (7.51) \) holds. Note that \( (7.51) \) holds, in particular, if \( \bar{x} \in \text{int}(\text{dom} f) \).

**Nonconvex case.** Assume that the function \( f : \mathbb{R}^n \to \mathbb{R} \) is real valued continuously differentiable and the set \( X \) is closed (not necessarily convex).

**Definition 7.13.** The contingent (Bouligand) cone to \( X \) at \( x \in X \), denoted \( T_X(x) \), is formed by vectors \( h \in \mathbb{R}^n \) such that there exist sequences \( h_k \to h \) and \( t_k \downarrow 0 \) such that \( x + t_k h_k \in X \).

Note that \( T_X(x) \) is nonempty only if \( x \in X \). If the set \( X \) is convex, then the contingent cone \( T_X(x) \) coincides with the corresponding tangent cone. We have the following simple necessary condition for a point \( \bar{x} \in X \) to be a locally optimal solution of problem \( (7.49) \).

**Proposition 7.14.** Let \( \bar{x} \in X \) be a locally optimal solution of problem \( (7.49) \). Then

\[
h^T \nabla f(\bar{x}) \geq 0, \quad \forall h \in T_X(\bar{x}).
\]

**Proof.** Consider \( h \in T_X(\bar{x}) \) and let \( h_k \to h \) and \( t_k \downarrow 0 \) be sequences such that \( x_k := \bar{x} + t_k h_k \in X \). Since \( \bar{x} \in X \) is a local minimizer of \( f(x) \) over \( x \in X \), we have that \( f(x_k) - f(\bar{x}) \geq 0 \). We also have that

\[
f(x_k) - f(\bar{x}) = t_k h^T \nabla f(\bar{x}) + o(t_k),
\]

and hence \( (7.53) \) follows. \( \square \)

Condition \( (7.53) \) means that \( \nabla f(\bar{x}) \in -[T_X(\bar{x})]^* \). If the set \( X \) is convex, then the polar \( [T_X(\bar{x})]^* \) of the tangent cone \( T_X(\bar{x}) \) coincides with the normal cone \( N_X(\bar{x}) \). Therefore, if \( f(\cdot) \) is convex and differentiable and \( X \) is convex, then optimality conditions \( (7.52) \) and \( (7.53) \) are equivalent.

Suppose now that the set \( X \) is given in the following form

\[
X := \{ x \in \mathbb{R}^n : G(x) \in K \},
\]

where \( G(\cdot) = (g_1(\cdot), \ldots, g_m(\cdot)) : \mathbb{R}^n \to \mathbb{R}^m \) is a continuously differentiable mapping and \( K \subset \mathbb{R}^m \) is a closed convex cone. In particular, if \( K := \{ q \} \times \mathbb{R}^{m-q} \), where \( q \in \mathbb{R}^q \) is the null vector and \( \mathbb{R}^{m-q} = \{ y \in \mathbb{R}^m : y \leq 0 \} \), then formulation \( (7.54) \) becomes

\[
X = \{ x \in \mathbb{R}^n : g_i(x) = 0, \ i = 1, \ldots, q, \ g_i(x) \leq 0, \ i = q + 1, \ldots, m \}.
\]

Under some regularity conditions (called *constraint qualifications*) we have the following formula for the contingent cone \( T_X(\bar{x}) \) at a feasible point \( \bar{x} \in X \): \n
\[
T_X(\bar{x}) = \{ h \in \mathbb{R}^n : [\nabla G(\bar{x})] h \in T_K(G(\bar{x})) \},
\]

where \( \nabla G(\bar{x}) = [\nabla g_1(\bar{x}), \ldots, \nabla g_m(\bar{x})]^T \) is the corresponding \( m \times n \) Jacobian matrix. The following condition is called *Robinson constraint qualification*

\[
[\nabla G(\bar{x})]\mathbb{R}^n + T_K(G(\bar{x})) = \mathbb{R}^m.
\]
If the cone $K$ has a nonempty interior, Robinson constraint qualification is equivalent to the following condition
\[
\exists h : \ G(\bar{x}) + [\nabla G(\bar{x})]h \in \text{int}(K).
\] (7.58)

In case $X$ is given in the form (7.55), Robinson constraint qualification is equivalent to the Mangasarian-Fromovitz constraint qualification:
\[
\nabla g_i(\bar{x}), \ i = 1, ..., q, \text{ are linearly independent,} \\
\exists h : \ h^T \nabla g_i(\bar{x}) = 0, \ i = 1, ..., q, \\
h^T \nabla g_i(\bar{x}) < 0, \ i \in I(\bar{x}),
\] (7.59)

where $I(\bar{x}) := \{i \in \{q + 1, ..., m\} : g_i(\bar{x}) = 0\}$ denotes the set of active at $\bar{x}$ inequality constraints.

Consider the Lagrangian
\[
L(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)
\]
associated with problem (7.49) and the constraint mapping $G(x)$. Under a constraint qualification ensuring validity of formula (7.56), first order necessary optimality condition (7.53) can be written in the following dual form: there exists a vector $\lambda \in \mathbb{R}^m$ of Lagrange multipliers such that
\[
\nabla_x L(\bar{x}, \lambda) = 0, \ G(\bar{x}) \in K, \ \lambda \in K^*, \ \lambda^T G(\bar{x}) = 0.
\] (7.60)

Denote by $\Lambda(\bar{x})$ the set of Lagrange multipliers vectors $\lambda$ satisfying (7.60).

**Theorem 7.15.** Let $\bar{x}$ be a locally optimal solution of problem (7.49). Then the set $\Lambda(\bar{x})$ Lagrange multipliers is nonempty and bounded iff Robinson constraint qualification holds.

In particular, if $\Lambda(\bar{x})$ is a singleton (i.e., there exists unique Lagrange multiplier vector), then Robinson constraint qualification holds. If the set $X$ is defined by a finite number of constraints in the form (7.55), then optimality conditions (7.60) are often referred to as the Karush-Kuhn-Tucker (KKT) necessary optimality conditions.

**Second Order Optimality Conditions**

We assume in this section that the function $f(x)$ is real valued twice continuously differentiable and denote by $\nabla^2 f(x)$ the Hessian matrix of second order partial derivatives of $f$ at $x$. Let $\bar{x}$ be a locally optimal solution of problem (7.49). Consider the set (cone)
\[
C(\bar{x}) := \{h \in T_X(\bar{x}) : h^T \nabla f(\bar{x}) = 0\}.
\] (7.61)

The cone $C(\bar{x})$ represents those feasible directions along which the first order approximation of $f(x)$ at $\bar{x}$ is zero, and is called the critical cone. The set
\[
T_{\bar{x}}^2(x, h) := \{z \in \mathbb{R}^n : \text{dist} (x + th + \frac{1}{2}t^2z, X) = o(t^2), \ t \geq 0\}
\] (7.62)
is called the (inner) second order tangent set to $X$ at the point $x \in X$ in the direction $h$. That is, the set $T^2_X(x, h)$ is formed by vectors $z$ such that $x + th + \frac{1}{2}t^2z + r(t) \in X$ for some $r(t) = o(t^2)$, $t \geq 0$. Note that this implies that $x + th + o(t) \in X$, and hence $T^2_X(x, h)$ can be nonempty only if $h \in T_X(x)$.

**Proposition 7.16.** Let $\bar{x}$ be a locally optimal solution of problem (7.49). Then

$$h^T \nabla^2 f(\bar{x}) h - s (-\nabla f(\bar{x}), T^2_X(\bar{x}, h)) \geq 0, \quad \forall h \in C(\bar{x}).$$

(7.63)

**Proof.** For some $h \in C(\bar{x})$ and $z \in T^2_X(\bar{x}, h)$ consider the (parabolic) curve $x(t) := \bar{x} + th + \frac{1}{2}t^2z$. By the definition of the second order tangent set, we have that there exists $r(t) = o(t^2)$ such that $x(t) + r(t) \in X$, $t \geq 0$. It follows by local optimality of $\bar{x}$ that $f(x(t) + r(t)) - f(\bar{x}) \geq 0$ for all $t \geq 0$ small enough. Since $r(t) = o(t^2)$, by the second order Taylor expansion we have

$$f(x(t) + r(t)) - f(\bar{x}) = th^T \nabla f(\bar{x}) + \frac{1}{2}t^2 [z^T \nabla f(\bar{x}) + h^T \nabla^2 f(\bar{x}) h] + o(t^2).$$

Since $h \in C(\bar{x})$ the first term in the right hand side of the above equation vanishes. It follows that

$$z^T \nabla f(\bar{x}) + h^T \nabla^2 f(\bar{x}) h \geq 0, \quad \forall h \in C(\bar{x}), \quad \forall z \in T^2_X(\bar{x}, h).$$

(7.64)

Condition (7.64) can be written in the form:

$$\inf_{z \in T^2_X(\bar{x}, h)} \{ z^T \nabla f(\bar{x}) + h^T \nabla^2 f(\bar{x}) h \} \geq 0, \quad \forall h \in C(\bar{x}).$$

(7.65)

Since

$$\inf_{z \in T^2_X(\bar{x}, h)} z^T \nabla f(\bar{x}) = - \sup_{z \in T^2_X(\bar{x}, h)} z^T (-\nabla f(\bar{x})) = -s (-\nabla f(\bar{x}), T^2_X(\bar{x}, h)), $$

the second order necessary conditions (7.64) can be written in the form (7.63).

If the set $X$ is polyhedral, then for $\bar{x} \in X$ and $h \in T_X(\bar{x})$ the second order tangent set $T^2_X(\bar{x}, h)$ is equal to the sum of $T_X(\bar{x})$ and the linear space generated by vector $h$. Since for $h \in C(\bar{x})$ we have that $h^T \nabla f(\bar{x}) = 0$ and because of the first order optimality conditions (7.53), it follows that if the set $X$ is polyhedral, then the term $s (-\nabla f(\bar{x}), T^2_X(\bar{x}, h))$ in (7.63) vanishes. In general, this term is nonpositive and corresponds to a curvature of the set $X$ at $\bar{x}$.

If the set $X$ is given in the form (7.54) with the mapping $G(x)$ being twice continuously differentiable, then the second order optimality conditions (7.63) can be written in the following dual form.

**Theorem 7.17.** Let $\bar{x}$ be a locally optimal solution of problem (7.49). Suppose that Robinson constraint qualification (7.57) is fulfilled. Then the following second order necessary conditions hold

$$\sup_{\lambda \in \Lambda(\bar{x})} \{ h^T \nabla^2_{xx} L(\bar{x}, \lambda) h - s (\lambda, T(h)) \} \geq 0, \quad \forall h \in C(\bar{x}),$$

(7.66)

Recall that $s(v, A) = \sup_{z \in A} z^T v$ denotes the support function of set $A$. 

---
where $T(h) := T_K^2(G(\bar{x}), [\nabla G(\bar{x})]h)$.

Note that if the cone $K$ is polyhedral, then the curvature term $s(\lambda, T(h))$ in (7.66) vanishes. In general, $s(\lambda, T(h)) \leq 0$ and the second order necessary conditions (7.66) are stronger than the "standard" second order conditions:

$$
\sup_{\lambda \in \Lambda(\bar{x})} h^T \nabla_x^2 L(\bar{x}, \lambda) h \geq 0, \quad \forall h \in C(\bar{x}).
$$

(7.67)

**Second order sufficient conditions.** Consider condition

$$
h^T \nabla^2 f(\bar{x}) h - s\left(-\nabla f(\bar{x}), T_X^2(\bar{x}, h)\right) > 0, \quad \forall h \in C(\bar{x}), h \neq 0.
$$

(7.68)

This condition is obtained from the second order necessary condition (7.63) by replacing "$\geq 0$" sign with the strict inequality sign "$> 0$". Necessity of second order conditions (7.63) was derived by verifying optimality of $\bar{x}$ along parabolic curves. There is no reason *a priori* that verification of (local) optimality along parabolic curves is sufficient to ensure local optimality of $\bar{x}$. Therefore in order to verify sufficiency of condition (7.68) we need an additional condition.

**Definition 7.18.** It is said that the set $X$ is second order regular at $\bar{x} \in X$ if for any sequence $x_k \in X$ of the form $x_k = \bar{x} + t_k h + \frac{1}{2} t_k^2 r_k$, where $t_k \downarrow 0$ and $t_k r_k \to 0$, it follows that

$$
\lim_{k \to \infty} \text{dist}\left(r_k, T_X^2(\bar{x}, h)\right) = 0.
$$

(7.69)

Note that in the above definition the term $\frac{1}{2} t_k^2 r_k = o(t_k)$, and hence such a sequence $x_k \in X$ can exist only if $h \in T_X(\bar{x})$. It turns out that second order regularity can be verified in many interesting cases. In particular, any polyhedral set is second order regular, the cone of positive semidefinite symmetric matrices is second order regular, etc. We refer to [22, section 3.3] for a discussion of this concept.

Recall that it is said that the *quadratic growth condition* holds at $\bar{x} \in X$ if there exist constant $c > 0$ and a neighborhood $N$ of $\bar{x}$ such that

$$
f(x) \geq f(\bar{x}) + c\|x - \bar{x}\|^2, \quad \forall x \in X \cap N.
$$

(7.70)

Of course, the quadratic growth condition implies that $\bar{x}$ is a locally optimal solution of problem (7.49).

**Proposition 7.19.** Let $\bar{x} \in X$ be a feasible point of problem (7.49) satisfying first order necessary conditions (7.53). Suppose that $X$ is second order regular at $\bar{x}$. Then the second order conditions (7.68) are necessary and sufficient for the quadratic growth at $\bar{x}$ to hold.

**Proof.** Suppose that conditions (7.68) hold. In order to verify the quadratic growth condition we argue by a contradiction, so suppose that it does not hold. Then there exists a sequence $x_k \in X \setminus \{\bar{x}\}$ converging to $\bar{x}$ and a sequence $c_k \downarrow 0$ such that

$$
f(x_k) - f(\bar{x}) \leq c_k \|x_k - \bar{x}\|^2.
$$

(7.71)
Denote \( t_k := \|x_k - \bar{x}\| \) and \( h_k := t_k^{-1}(x_k - \bar{x}) \). By passing to a subsequence if necessary we can assume that \( h_k \) converges to a vector \( h \). Clearly \( h \neq 0 \) and by the definition of \( \mathcal{T}_X(\bar{x}) \) it follows that \( h \in \mathcal{T}_X(\bar{x}) \). Moreover, by \eqref{7.71} we have

\[
eq \frac{\kappa}{2} \geq f(x_k) - f(\bar{x}) = t_k h^T \nabla f(\bar{x}) + o(t_k),
\]

and hence \( h^T \nabla f(\bar{x}) \leq 0 \). Because of the first order necessary conditions it follows that \( h^T \nabla f(\bar{x}) = 0 \), and hence \( h \in C(\bar{x}) \).

Denote \( r_k := 2t_k^{-1}(h_k - h) \). We have that \( x_k = \bar{x} + t_k h + \frac{1}{2} t_k^2 r_k \in X \) and \( t_k r_k \to 0 \). Consequently it follows by the second order regularity that there exists a sequence \( z_k \in \mathcal{T}_X^2(\bar{x}, h) \) such that \( r_k - z_k \to 0 \). Since \( h^T \nabla f(\bar{x}) = 0 \), by the second order Taylor expansion we have

\[
f(x_k) = f(\bar{x} + t_k h + \frac{1}{2} t_k^2 r_k) = f(\bar{x}) + \frac{1}{2} t_k^2 \left[ z_k^T \nabla f(\bar{x}) + h^T \nabla^2 f(\bar{x}) r_k \right] + o(t_k^2).
\]

Moreover, since \( z_k \in \mathcal{T}_X^2(\bar{x}, h) \) we have that

\[
z_k^T \nabla f(\bar{x}) + h^T \nabla^2 f(\bar{x}) r_k \geq c,
\]

where \( c \) is equal to the left hand side of \eqref{7.68}, which by the assumption is positive. It follows that

\[
f(x_k) \geq f(\bar{x}) + \frac{1}{2} c \|x_k - \bar{x}\|^2 + o(\|x_k - \bar{x}\|^2),
\]

a contradiction with \eqref{7.71}.

Conversely, suppose that the quadratic growth condition holds at \( \bar{x} \). It follows that the function \( \phi(x) := \frac{1}{2} c \|x - \bar{x}\|^2 \) also attains its local minimum over \( X \) at \( \bar{x} \). Note that \( \nabla \phi(\bar{x}) = \nabla f(\bar{x}) \) and \( h^T \nabla^2 \phi(\bar{x}) = h^T \nabla^2 f(\bar{x}) - c \|h\|^2 \). Therefore, by the second order necessary conditions \eqref{7.63}, applied to the function \( \phi \), it follows that the left hand side of \eqref{7.68} is greater than or equal to \( c \|h\|^2 \). This completes the proof. \( \square \)

If the set \( X \) is given in the form \eqref{7.54}, then similar to Theorem 7.17 it is possible to formulate second order sufficient conditions \eqref{7.68} in the following dual form.

**Theorem 7.20.** Let \( \bar{x} \in X \) be a feasible point of problem \eqref{7.49} satisfying first order necessary conditions \eqref{7.60}. Suppose that Robinson constraint qualification \eqref{7.57} is fulfilled and the set (cone) \( K \) is second order regular at \( G(\bar{x}) \). Then the following conditions are necessary and sufficient for the quadratic growth at \( \bar{x} \) to hold:

\[
\sup_{\lambda \in \Lambda(\bar{x})} \left\{ h^T \nabla^2 L(\bar{x}, \lambda) h - s(\lambda, T(h)) \right\} > 0, \quad \forall h \in C(\bar{x}), \ h \neq 0,
\]

where \( T(h) := \mathcal{T}_X^2(G(\bar{x}), |\nabla G(\bar{x})| h) \).

Note again that if the cone \( K \) is polyhedral, then \( K \) is second order regular and the curvature term \( s(\lambda, T(h)) \) in \eqref{7.72} vanishes.
7.1.5 Perturbation Analysis

Differentiability Properties of Max-Functions

We often have to deal with optimal value functions, say max-functions of the form

\[ \phi(x) := \sup_{\theta \in \Theta} g(x, \theta), \]  \hspace{1cm} (7.73)

where \( g : \mathbb{R}^n \times \Theta \to \mathbb{R} \). In applications the set \( \Theta \) usually is a subset of a finite dimensional vector space. At this point, however, this is not important and we can assume that \( \Theta \) is an abstract topological space. Denote

\[ \bar{\Theta}(x) := \text{arg max}_{\theta \in \Theta} g(x, \theta). \]

The following result about directional differentiability of the max-function is often called Danskin Theorem.

**Theorem 7.21 (Danskin).** Let \( \Theta \) be a nonempty, compact topological space and \( g : \mathbb{R}^n \times \Theta \to \mathbb{R} \) be such that \( g(\cdot, \theta) \) is differentiable for every \( \theta \in \Theta \) and \( \nabla_x g(x, \theta) \) is continuous on \( \mathbb{R}^n \times \Theta \). Then the corresponding max-function \( \phi(x) \) is locally Lipschitz continuous, directionally differentiable and

\[ \phi'(x, h) = \sup_{\theta \in \bar{\Theta}(x)} h^T \nabla_x g(x, \theta). \]  \hspace{1cm} (7.74)

In particular, if for some \( x \in \mathbb{R}^n \) the set \( \bar{\Theta}(x) = \{ \bar{\theta} \} \) is a singleton, then the max-function is differentiable at \( x \) and

\[ \nabla \phi(x) = \nabla_x g(x, \bar{\theta}). \]  \hspace{1cm} (7.75)

In the convex case we have the following result giving a description of subdifferentials of max-functions.

**Theorem 7.22 (Levin-Valadier).** Let \( \Theta \) be a nonempty compact topological space and \( g : \mathbb{R}^n \times \Theta \to \mathbb{R} \) be a real valued function. Suppose that: (i) for every \( \theta \in \Theta \) the function \( g(\cdot, \theta) \) is convex on \( \mathbb{R}^n \), (ii) for every \( x \in \mathbb{R}^n \) the function \( g(x, \cdot) \) is upper semicontinuous on \( \Theta \). Then the max-function \( \phi(x) \) is convex real valued and

\[ \partial \phi(x) = \text{cl} \left\{ \text{conv} \left( \bigcup_{\theta \in \bar{\Theta}(x)} \partial g(\cdot, \theta) \right) \right\}. \]  \hspace{1cm} (7.76)

Let us make the following observations regarding the above theorem. Since \( \Theta \) is compact and by the assumption (ii), we have that the set \( \bar{\Theta}(x) \) is nonempty and compact. Since the function \( \phi(\cdot) \) is convex real valued, it is subdifferentiable at every \( x \in \mathbb{R}^n \) and its subdifferential \( \partial \phi(x) \) is a convex, closed bounded subset of \( \mathbb{R}^n \). It follows then from (7.76) that the set \( A := \bigcup_{\theta \in \bar{\Theta}(x)} \partial g(\cdot, \theta) \) is bounded. Suppose further that:

(iii) For every \( x \in \mathbb{R}^n \) the function \( g(x, \cdot) \) is continuous on \( \Theta \).
Then the set $A$ is closed, and hence is compact. Indeed, consider a sequence $z_k \in A$. Then, by the definition of the set $A$, $z_k \in \partial g_{\theta_k}(x)$ for some sequence $\theta_k \in \Theta(x)$. Since $\Theta(x)$ is compact and $A$ is bounded, by passing to a subsequence if necessary, we can assume that $\theta_k$ converges to a point $\bar{\theta} \in \Theta(x)$ and $z_k$ converges to a point $\bar{z} \in \mathbb{R}^n$. By the definition of subgradients $z_k$ we have that for any $x' \in \mathbb{R}^n$ the following inequality holds

$$g_{\theta_k}(x') - g_{\theta_k}(x) \geq z_k^T(x' - x).$$

By passing to the limit in the above inequality as $k \to \infty$, we obtain that $\bar{z} \in \partial g_{\bar{\theta}}(x)$. It follows that $\bar{z} \in A$, and hence $A$ is closed. Now since convex hull of a compact subset of $\mathbb{R}^n$ is also compact, and hence is closed, we obtain that if the assumption (ii) in the above theorem is strengthened to the assumption (iii), then the set inside the parentheses in (7.76) is closed, and hence formula (7.76) takes the form

$$\partial \phi(x) = \text{conv} \left( \bigcup_{\theta \in \Theta(x)} \partial g_{\theta}(x) \right).$$

(7.77)

**Second Order Perturbation Analysis**

Consider the following parameterization of problem (7.49):

$$\min_{x \in X} f(x) + t \eta_t(x),$$

(7.78)

depending on parameter $t \in \mathbb{R}_+$. We assume that the set $X \subset \mathbb{R}^n$ is nonempty and compact and consider a convex compact set $U \subset \mathbb{R}^n$ such that $X \subset \text{int}(U)$. It follows, of course, that the set $U$ has a nonempty interior. Consider the space $W^{1,\infty}(U)$ of Lipschitz continuous functions $\psi : U \to \mathbb{R}$ equipped with the norm

$$\|\psi\|_{1,U} := \sup_{x \in U} |\psi(x)| + \sup_{x \in U'} \|\nabla \psi(x)\|,$$

(7.79)

with $U' \subset \text{int}(U)$ being the set of points where $\psi(\cdot)$ is differentiable. Recall that by Rademacher Theorem, a function $\psi(\cdot) \in W^{1,\infty}(U)$ is differentiable at almost every point of $U$. We assume that the functions $f(\cdot)$ and $\eta_t(\cdot)$, $t \in \mathbb{R}_+$, are Lipschitz continuous on $U$, i.e., $f, \eta_t \in W^{1,\infty}(U)$. We also assume that $\eta_t$ converges (in the norm topology) to a function $\delta \in W^{1,\infty}(U)$, that is, $\|\eta_t - \delta\|_{1,U} \to 0$ as $t \downarrow 0$.

Denote by $v(t)$ the optimal value and by $\bar{x}(t)$ an optimal solution of (7.78), i.e.,

$$v(t) := \inf_{x \in X} \{ f(x) + t \eta_t(x) \} \quad \text{and} \quad \bar{x}(t) \in \arg\min_{x \in X} \{ f(x) + t \eta_t(x) \}.$$

We will be interested in second order differentiability properties of $v(t)$ and first order differentiability properties of $\bar{x}(t)$ at $t = 0$. We assume that $f(x)$ has unique minimizer $\bar{x}$ over $x \in X$, i.e., the set of optimal solutions of the unperturbed problem (7.49) is the singleton $\{\bar{x}\}$. Moreover, we assume that $\delta(\cdot)$ is differentiable at $\bar{x}$ and $f(x)$ is twice continuously differentiable at $\bar{x}$. Since $X$ is compact and the objective function of problem (7.78) is continuous, it has an optimal solution for any $t$.

The following result is taken from [22, section 4.10.3].

**Theorem 7.23.** Let $\bar{x}$ be unique optimal solution of problem (7.49). Suppose that: (i) the set $X$ is compact and second order regular at $\bar{x}$, (ii) $\eta_t$ converges (in the norm topology)
to $\delta \in W^{1,\infty}(U)$ as $t \downarrow 0$, (iii) $\delta(x)$ is differentiable at $\bar{x}$ and $f(x)$ is twice continuously differentiable at $\bar{x}$, (iv) the quadratic growth condition (7.70) holds. Then

$$v(t) = v(0) + tv_{\eta}(\bar{x}) + \frac{1}{2}t^2\mathfrak{Y}_f(\delta) + o(t^2), \quad t \geq 0,$$

(7.80)

where $\mathfrak{Y}_f(\delta)$ is the optimal value of the auxiliary problem

$$\min_{h \in C(\bar{x})} \left\{ 2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 f(\bar{x})h - s\left(-\nabla f(\bar{x}), T_X^2(\bar{x}, h)\right) \right\}.$$ 

(7.81)

Moreover, if (7.81) has unique optimal solution $h$, then

$$\bar{x}(t) = \bar{x} + th + o(t), \quad t \geq 0.$$ 

(7.82)

**Proof.** Since the minimizer $\bar{x}$ is unique and the set $X$ is compact, it is not difficult to show that, under the specified assumptions, $\bar{x}(t)$ tends to $\bar{x}$ as $t \downarrow 0$. Moreover, we have that $\|\bar{x}(t) - \bar{x}\| = O(t)$, $t > 0$. Indeed, by the quadratic growth condition, for $t > 0$ small enough and some $c > 0$ it follows that

$$v(t) = f(\bar{x}(t)) + tv_{\eta}(\bar{x}(t)) \geq f(\bar{x}) + c\|\bar{x}(t) - \bar{x}\|^2 + tv_{\eta}(\bar{x}(t)).$$

Since $\bar{x} \in X$ we also have that $v(t) \leq f(\bar{x}) + tv_{\eta}(\bar{x})$. Consequently,

$$t |\eta_{\bar{x}}(\bar{x}(t)) - \eta_{\bar{x}}(\bar{x})| \geq c\|\bar{x}(t) - \bar{x}\|^2.$$

Moreover, $|\eta_{\bar{x}}(\bar{x}(t)) - \eta_{\bar{x}}(\bar{x})| = O(\|\bar{x}(t) - \bar{x}\|)$, and hence $\|\bar{x}(t) - \bar{x}\| = O(t)$.

Let $h \in C(\bar{x})$ and $w \in T_X^2(\bar{x}, h)$. By the definition of the second order tangent set it follows that there is a path $x(t) \in X$ of the form $x(t) = \bar{x} + th + \frac{1}{2}t^2w + o(t^2)$. Since $x(t) \in X$ we have that $v(t) \leq f(x(t)) + tv_{\eta}(x(t))$. Moreover, by using the second order Taylor expansion of $f(x)$ at $x = \bar{x}$ we have

$$f(x(t)) = f(\bar{x}) + th^T \nabla f(\bar{x}) + \frac{1}{2}t^2w^T \nabla^2 f(\bar{x})h + \frac{1}{2}th^T \nabla^2 f(\bar{x})h + o(t^2),$$

and since $h \in C(\bar{x})$ we have that $h^T \nabla f(\bar{x}) = 0$. Also since $\|\eta_{\bar{x}} - \delta\|_{1,\infty} \to 0$, we have by the Mean Value Theorem that

$$\eta_{\bar{x}}(x(t)) - \delta(x(t)) = \eta_{\bar{x}}(\bar{x}) - \delta(\bar{x}) + o(t),$$

and since $\delta(x)$ is differentiable at $\bar{x}$ that

$$\delta(x(t)) = \delta(\bar{x}) + th^T \nabla \delta(\bar{x}) + o(t).$$

Putting this all together and noting that $f(\bar{x}) = v(0)$ we obtain that

$$f(x(t)) + tv_{\eta}(x(t)) = v(0) + tv_{\eta}(\bar{x}) + t^2h^T \nabla \delta(\bar{x}) + \frac{1}{2}t^2h^T \nabla^2 f(\bar{x})h + \frac{1}{2}t^2w^T \nabla^2 f(\bar{x})h + o(t^2).$$

Consequently

$$\limsup_{t \downarrow 0} \frac{v(t) - v(0) - tv_{\eta}(\bar{x})}{\frac{1}{2}t^2} \leq 2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 f(\bar{x})h + w^T \nabla f(\bar{x}).$$

(7.83)
7.1. Optimization and Convex Analysis

Since the above inequality (7.83) holds for any \( w \in T^2_X(\bar{x}, h) \), by taking minimum (with respect to \( w \)) in the right hand side of (7.83) we obtain for any \( h \in C(\bar{x}) \):

\[
\limsup_{t \downarrow 0} \frac{v(t) - v(0) - t \eta_h(\bar{x})}{t^2} \leq 2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 f(\bar{x})h - s(- \nabla f(\bar{x}), T^2_X(\bar{x}, h)).
\]

In order to show the converse estimate we argue as follows. Consider a sequence \( t_k \downarrow 0 \) and \( x_k := \bar{x}(t_k) \). Since \( \| \bar{x}(t) - \bar{x} \| = O(t) \), we have that \( (x_k - \bar{x})/t_k \) is bounded, and hence by passing to a subsequence if necessary we can assume that \( (x_k - \bar{x})/t_k \) converges to a vector \( h \). Since \( x_k \in X \), it follows that \( h \in T_X(\bar{x}) \). Moreover,

\[
v(t_k) = f(x_k) + t_k \eta_h(x_k) = f(\bar{x}) + t_k h^T \nabla f(\bar{x}) + t_k \delta(\bar{x}) + o(t_k),
\]

and by Danskin Theorem \( \eta'(0) = \delta(\bar{x}) \). It follows that \( h^T \nabla f(\bar{x}) = 0 \), and hence \( h \in C(\bar{x}) \). Consider \( r_k := 2(x_k - \bar{x} - t_k h)/t_k^2 \), i.e., \( r_k \) are such that \( x_k = \bar{x} + t_k h + \frac{1}{2} t_k^2 r_k \). Note that \( t_k r_k \to 0 \) and \( x_k \in X \) and hence, by the second order regularity of \( X \), there exists \( w_k \in T^2_X(\bar{x}, h) \) such that \( \| r_k - w_k \| \to 0 \). Finally,

\[
v(t_k) = f(x_k) + t_k \eta_h(x_k) \\
= f(\bar{x}) + t_k \eta_h(\bar{x}) + t_k^2 h^T \nabla \delta(\bar{x}) + \frac{1}{2} t_k^2 h^T \nabla^2 f(\bar{x})h + \frac{1}{2} t_k^2 w_k^T \nabla f(\bar{x}) + o(t_k^2) \\
\geq v(0) + t_k \eta_h(\bar{x}) + t_k^2 h^T \nabla \delta(\bar{x}) + \frac{1}{2} t_k^2 h^T \nabla^2 f(\bar{x})h \\
\quad + \frac{1}{2} t_k^2 \inf_{w \in T_X^2(\bar{x}, h)} w^T \nabla f(\bar{x}) + o(t_k^2).
\]

It follows that

\[
\liminf_{t \downarrow 0} \frac{v(t) - v(0) - t \eta_h(\bar{x})}{t^2} \geq 2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 f(\bar{x})h - s(- \nabla f(\bar{x}), T^2_X(\bar{x}, h)).
\]

This completes the proof of (7.80).

Also by the above analysis we have that any accumulation point of \( (\bar{x}(t) - \bar{x})/t \), as \( t \downarrow 0 \), is an optimal solution of problem (7.81). Since \( (\bar{x}(t) - \bar{x})/t \) is bounded, the assertion (7.82) follows by compactness arguments.

As in the case of second order optimality conditions, we have here that if the set \( X \) is polyhedral, then the curvature term \( s(- \nabla f(\bar{x}), T^2_X(\bar{x}, h)) \) in (7.81) vanishes.

Suppose now that the set \( X \) is given in the form (7.54) with the mapping \( G(\lambda) \) being twice continuously differentiable. Suppose further that Robinson constraint qualification (7.57), for the unperturbed problem, holds. Then the optimal value of problem (7.81) can be written in the following dual form

\[
\mathfrak{M}_f(\delta) = \inf_{h \in C(\bar{x})} \sup_{\lambda \in \Delta(X)} \left\{ 2h^T \nabla \delta(\bar{x}) + h^T \nabla^2 L(\bar{x}, \lambda)h - s(\lambda, T(h)) \right\}, \quad (7.84)
\]

where \( T(h) := T^2_K(G(\bar{x}), [\nabla G(x)]h) \). Note again that if the set \( K \) is polyhedral, then the curvature term \( s(\lambda, T(h)) \) in (7.84) vanishes.

Minimax Problems

In this section we consider the following minimax problem

\[
\min_{x \in X} \left\{ \phi(x) := \sup_{y \in Y} f(x, y) \right\}, \quad (7.85)
\]
and its dual

$$\max_{y \in Y} \left\{ \varsigma(y) := \inf_{x \in X} f(x, y) \right\}. \quad (7.86)$$

We assume that the sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ are convex and compact, and the function $f : X \times Y \to \mathbb{R}$ is continuous\(^4\), i.e., $f \in C(X, Y)$. Moreover, assume that $f(x, y)$ is convex in $x \in X$ and concave in $y \in Y$. Under these conditions there is no duality gap between problems (7.85) and (7.86), i.e., the optimal values of these problems are equal to each other. Moreover, the max-function $\phi(x)$ is continuous on $X$ and problem (7.85) has a nonempty set of optimal solutions, denoted $X^*$, the min-function $\iota(y)$ is continuous on $Y$ and problem (7.86) has a nonempty set of optimal solutions, denoted $Y^*$, and $X^* \times Y^*$ forms the set of saddle points of the minimax problems (7.85) and (7.86).

Consider the following perturbation of the minimax problem (7.85):

$$\min \sup_{x \in X, y \in Y} \left\{ f(x, y) + t \eta_t(x, y) \right\}, \quad (7.87)$$

where $\eta_t \in C(X, Y)$, $t \geq 0$. Denote by $v(t)$ the optimal value of the above parameterized problem (7.87). Clearly $v(0)$ is the optimal value of the unperturbed problem (7.85). We assume that $\eta_t$ converges uniformly (i.e., in the sup-norm) as $t \downarrow 0$ to a function $\gamma \in C(X, Y)$, that is

$$\lim_{t \downarrow 0} \sup_{x \in X, y \in Y} \left| \eta_t(x, y) - \gamma(x, y) \right| = 0.$$

**Theorem 7.24.** Suppose that: (i) the sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ are convex and compact, (ii) for all $t \geq 0$ the function $\zeta_t := f + t \eta_t$ is continuous on $X \times Y$, convex in $x \in X$ and concave in $y \in Y$, (iii) $\eta_t$ converges uniformly as $t \downarrow 0$ to a function $\gamma \in C(X, Y)$. Then

$$\lim_{t \downarrow 0} \frac{v(t) - v(0)}{t} = \inf_{x \in X^*} \sup_{y \in Y^*} \gamma(x, y). \quad (7.88)$$

**Proof.** Consider a sequence $t_k \downarrow 0$. Denote $\eta_k := \eta_{t_k}$ and $\zeta_k := \zeta_{t_k} = f + t_k \eta_k$. By the assumption (ii) we have that functions $\zeta_k(x, y)$ are continuous and convex-concave on $X \times Y$. Also by the definition

$$v(t_k) = \inf_{x \in X} \sup_{y \in Y} \zeta_k(x, y).$$

For a point $x^* \in X^*$ we can write

$$v(0) = \sup_{y \in Y} f(x^*, y) \quad \text{and} \quad v(t_k) \leq \sup_{y \in Y} \zeta_k(x^*, y).$$

Since the set $Y$ is compact and function $\zeta_k(x^*, \cdot)$ is continuous, we have that the set $\arg \max_{y \in Y} \zeta_k(x^*, y)$ is nonempty. Let $y_k \in \arg \max_{y \in Y} \zeta_k(x^*, y)$. We have that

$$\arg \max_{y \in Y} f(x^*, y) = Y^*$$

\(^4\)Recall that $C(X, Y)$ denotes the space of continuous functions $\psi : X \times Y \to \mathbb{R}$ equipped with the sup-norm $\|\psi\| = \sup_{(x, y) \in X \times Y} |\psi(x, y)|$. 
The assertion of the theorem follows from (7.89) and (7.90).

Since \( y_k \) tends uniformly to \( y \), it follows that \( \eta_k(x^*, y_k) \to \gamma(x^*, y^*) \). Consequently

\[
v(t_k) - v(0) \leq \zeta_k(x^*, y_k) - f(x^*, y_k) = t_k \eta_k(x^*, y_k) = t_k \gamma(x^*, y^*) + o(t_k).
\]

We obtain that for any \( x^* \in X^* \) there exists \( y^* \in Y^* \) such that

\[
\limsup_{k \to \infty} \frac{v(t_k) - v(0)}{t_k} \leq \gamma(x^*, y^*).
\]

It follows that

\[
\limsup_{k \to \infty} \frac{v(t_k) - v(0)}{t_k} \leq \inf_{x \in X^*} \sup_{y \in Y^*} \gamma(x, y). \tag{7.89}
\]

In order to prove the converse inequality we proceed as follows. Consider a sequence \( x_k \in \arg\min_{x \in X} \theta_k(x) \), where \( \theta_k(x) := \sup_{y \in Y} \zeta_k(x, y) \). We have that \( \theta_k : X \to \mathbb{R} \) are continuous functions converging uniformly in \( x \in X \) to the max-function \( \phi(x) = \sup_{y \in Y} f(x, y) \). Consequently \( x_k \) converges in distance to the set \( \arg \min_{x \in X} \phi(x) \), which is equal to \( X^* \). By passing to a subsequence if necessary we can assume that \( x_k \) converges to a point \( x^* \in X^* \). For any \( y \in Y^* \) we have \( v(0) \leq f(x_k, y) \). Since \( \zeta_k(x, y) \) is convex-concave, it has a nonempty set of saddle points \( X_k^* \times Y_k^* \). We have that \( x_k \in X_k^* \), and hence \( v(t_k) \geq \zeta_k(x_k, y) \) for any \( y \in Y^* \). It follows that for any \( y \in Y^* \) the following holds

\[
v(t_k) - v(0) \geq \zeta_k(x_k, y) - f(x_k, y) = t_k \gamma_k(x^*, y) + o(t_k),
\]

and hence

\[
\liminf_{k \to \infty} \frac{v(t_k) - v(0)}{t_k} \geq \gamma(x^*, y).
\]

Since \( y \) was an arbitrary element of \( Y^* \), we obtain that

\[
\liminf_{k \to \infty} \frac{v(t_k) - v(0)}{t_k} \geq \sup_{y \in Y^*} \gamma(x^*, y),
\]

and hence

\[
\liminf_{k \to \infty} \frac{v(t_k) - v(0)}{t_k} \geq \inf_{x \in X^*} \sup_{y \in Y^*} \gamma(x, y). \tag{7.90}
\]

The assertion of the theorem follows from (7.89) and (7.90). \( \square \)
7.1.6 Epiconvergence

Consider a sequence \( f_k : \mathbb{R}^n \to \overline{\mathbb{R}} \), \( k = 1, \ldots \), of extended real valued functions. It is said that the functions \( f_k \) epiconverge to a function \( f : \mathbb{R}^n \to \overline{\mathbb{R}} \), written \( f_k \overset{e}{\to} f \), if the epigraphs of the functions \( f_k \) converge, in a certain set-valued sense, to the epigraph of \( f \). It is also possible to define the epiconvergence in the following equivalent way.

**Definition 7.25.** It is said that \( f_k \) epiconverge to \( f \) if for any point \( x \in \mathbb{R}^n \) the following two conditions hold:

(i) for any sequence \( x_k \) converging to \( x \) one has

\[
\liminf_{k \to \infty} f_k(x_k) \geq f(x)
\]  
(7.91)

(ii) there exists a sequence \( x_k \) converging to \( x \) such that

\[
\limsup_{k \to \infty} f_k(x_k) \leq f(x).
\]  
(7.92)

Epiconvergence \( f_k \overset{e}{\to} f \) implies that the function \( f \) is lower semicontinuous.

For \( \varepsilon \geq 0 \) we say that a point \( \bar{x} \in \mathbb{R}^n \) is an \( \varepsilon \)-minimizer\(^5\) of \( f \) if \( f(\bar{x}) \leq \inf_{x \in \mathbb{R}^n} f(x) + \varepsilon \) (we write here ‘inf \( f(x) \)’ for ‘\( \inf_{x \in \mathbb{R}^n} f(x) \)’). Clearly, for \( \varepsilon = 0 \) the set of \( \varepsilon \)-minimizers of \( f \) coincides with the set \( \arg \min f \) (of minimizers of \( f \)).

**Proposition 7.26.** Suppose that \( f_k \overset{e}{\to} f \). Then

\[
\limsup_{k \to \infty} [\inf_{x \in \mathbb{R}^n} f_k(x)] \leq \inf_{x \in \mathbb{R}^n} f(x).
\]  
(7.93)

Suppose, further, that: (i) for some \( \varepsilon_k \downarrow 0 \) there exists an \( \varepsilon_k \)-minimizer \( x_k \) of \( f_k(\cdot) \) such that the sequence \( x_k \) converges to a point \( \bar{x} \). Then \( \bar{x} \in \arg \min f \) and

\[
\lim_{k \to \infty} [\inf_{x \in \mathbb{R}^n} f_k(x)] = \inf_{x \in \mathbb{R}^n} f(x).
\]  
(7.94)

**Proof.** Consider a point \( \bar{x} \in \mathbb{R}^n \) and let \( x_k \) be a sequence converging to \( \bar{x} \) such that the inequality (7.92) holds. Clearly \( f_k(x_k) \geq \inf f_k(x) \) for all \( k \). Together with (7.92) this implies that

\[
f(\bar{x}) \geq \limsup_{k \to \infty} f_k(x_k) \geq \limsup_{k \to \infty} [\inf_{x \in \mathbb{R}^n} f_k(x)].
\]

Since the above holds for any \( \bar{x} \), the inequality (7.93) follows.

Now let \( x_k \) be a sequence of \( \varepsilon_k \)-minimizers of \( f_k \) converging to a point \( \bar{x} \). We have then that \( f_k(x_k) \leq \inf f_k(x) + \varepsilon_k \), and hence by (7.93) we obtain

\[
\liminf_{k \to \infty} [\inf_{x \in \mathbb{R}^n} f_k(x)] = \liminf_{k \to \infty} [\inf_{x \in \mathbb{R}^n} f_k(x) + \varepsilon_k] \geq \liminf_{k \to \infty} f_k(x_k) \geq f(\bar{x}) \geq \inf_{x \in \mathbb{R}^n} f(x).
\]

\(^5\)Note that here some (all) points \( x_k \) can be equal to \( x \).

\(^6\)For the sake of convenience we allow in this section for a minimizer, or \( \varepsilon \)-minimizer, \( \bar{x} \) to be such that \( f(\bar{x}) \) is not finite, i.e., can be equal to \( +\infty \) or \( -\infty \).
Together with (7.93) this implies (7.94) and \( f(\bar{x}) = \inf f(x) \). This completes the proof. 

Assumption (i) in the above proposition can be ensured by various boundedness conditions. Proof of the following theorem can be found in [181, Theorem 7.17].

**Theorem 7.27.** Let \( f_k : \mathbb{R}^n \to \mathbb{R} \) be a sequence of convex functions and \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex lower semicontinuous function such that \( \text{dom}(f) \) has a nonempty interior. Then the following are equivalent: (i) \( f_k \overset{e}{\to} f \), (ii) there exists a dense subset \( D \) of \( \mathbb{R}^n \) such that \( f_k(x) \to f(x) \) for all \( x \in D \), (iii) \( f_k(\cdot) \) converges uniformly to \( f(\cdot) \) on every compact set \( C \) that does not contain a boundary point of \( \text{dom}(f) \).

### 7.2 Probability

#### 7.2.1 Probability Spaces and Random Variables

Let \( \Omega \) be an abstract set. It is said that a set \( \mathcal{F} \) of subsets of \( \Omega \) is a *sigma algebra* (also called sigma field) if: (i) it is closed under standard set theoretic operations (i.e., if \( A, B \in \mathcal{F} \), then \( A \cap B \in \mathcal{F} \), \( A \cup B \in \mathcal{F} \), and \( A \setminus B \in \mathcal{F} \)), (ii) the set \( \Omega \) belongs to \( \mathcal{F} \), and (iii) if \( i \), \( A_i \in \mathcal{F} \), \( i \in \mathbb{N} \), then \( \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F} \). The set \( \Omega \) equipped with a sigma algebra \( \mathcal{F} \) is called a *sample space* and denoted \((\Omega, \mathcal{F})\). A set \( A \subseteq \Omega \) is said to be \( \mathcal{F} \)-measurable if \( A \in \mathcal{F} \). It is said that the sigma algebra \( \mathcal{F} \) is generated by its subset \( \mathcal{G} \) if any \( \mathcal{F} \)-measurable set can be obtained from sets belonging to \( \mathcal{G} \) by set theoretic operations and by taking the union of a countable family of sets from \( \mathcal{G} \). That is, \( \mathcal{F} \) is generated by \( \mathcal{G} \) if \( \mathcal{F} \) is the smallest sigma algebra containing \( \mathcal{G} \). If we have two sigma algebras \( \mathcal{F}_1 \) and \( \mathcal{F}_2 \) defined on the same set \( \Omega \), then it is said that \( \mathcal{F}_1 \) is a subalgebra of \( \mathcal{F}_2 \) if \( \mathcal{F}_1 \subseteq \mathcal{F}_2 \). The smallest possible sigma algebra on \( \Omega \) consists of two elements \( \emptyset \) and \( \Omega \), such a sigma algebra is called *trivial*. An \( \mathcal{F} \)-measurable set \( A \) is said to be elementary if any \( \mathcal{F} \)-measurable subset of \( A \) is either the empty set or the set \( A \). If the sigma algebra \( \mathcal{F} \) is finite, then it is generated by a family \( A_i \subseteq \Omega \), \( i = 1, \ldots, n \), of disjoint elementary sets and has \( 2^n \) elements. The sigma algebra generated by the set of open (or closed) subsets of a finite dimensional space \( \mathbb{R}^n \) is called its Borel sigma algebra. An element of this sigma algebra is called a Borel set. For a considered set \( \Xi \subseteq \mathbb{R}^n \) we denote by \( \mathcal{B}(\Xi) \) the sigma algebra of all Borel subsets of \( \Xi \).

A function \( P : \mathcal{F} \to \mathbb{R}_+ \) is called a (sigma-additive) *measure* on \((\Omega, \mathcal{F})\) if for every collection \( A_i \in \mathcal{F} \), \( i \in \mathbb{N} \), such that \( A_i \cap A_j = \emptyset \) for all \( i \neq j \), we have

\[
P\left( \bigcup_{i \in \mathbb{N}} A_i \right) = \sum_{i \in \mathbb{N}} P(A_i). \tag{7.95}
\]

In this definition it is assumed that for every \( A \in \mathcal{F} \), and in particular for \( A = \Omega \), \( P(A) \) is finite. Sometimes such measures are called *finite*. An important example of a measure which is not finite is the Lebesgue measure on \( \mathbb{R}^n \). Unless stated otherwise we assume that a considered measure is finite. A measure \( P \) is said to be a *probability measure* if \( P(\Omega) = 1 \). A sample space \((\Omega, \mathcal{F})\) equipped with a probability measure \( P \) is called a *probability space* and denoted \((\Omega, \mathcal{F}, P)\). Recall that \( \mathcal{F} \) is said to be \( P \)-complete if \( A \subseteq B \), \( B \in \mathcal{F} \) and \( P(B) = 0 \), implies that \( A \in \mathcal{F} \), and hence \( P(A) = 0 \). Since it is always possible to enlarge

---

3By \( \mathbb{N} \) we denote the set of positive integers.
the sigma algebra and extend the measure in such a way as to get complete space, we can assume without loss of generality that considered probability measures are complete. It is said that an event \( A \in \mathcal{F} \) happens \( P \)-almost surely (a.s.) or almost everywhere (a.e.) if \( P(A) = 1 \), or equivalently \( P(\Omega \setminus A) = 0 \). We also sometimes say that such an event happens with probability one (w.p.1).

Let \( P \) and \( Q \) be two measures on a measurable space \((\Omega, \mathcal{F})\). It is said that \( Q \) is absolutely continuous with respect to \( P \) if \( A \in \mathcal{F} \) and \( P(A) = 0 \) implies that \( Q(A) = 0 \). If the measure \( Q \) is finite, this is equivalent to condition: for every \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that if \( P(A) < \delta \), then \( Q(A) < \varepsilon \).

**Theorem 7.28 (Radon-Nikodym).** If \( P \) and \( Q \) are measures on \((\Omega, \mathcal{F})\), then \( Q \) is absolutely continuous with respect to \( P \) if and only if there exists a function \( f : \Omega \rightarrow \mathbb{R}^+ \) such that \( Q(A) = \int_A f \, dP \) for every \( A \in \mathcal{F} \).

The function \( f \) in the representation \( Q(A) = \int_A f \, dP \) is called density of measure \( Q \) with respect to measure \( P \). If the measure \( Q \) is a probability measure, then \( f \) is called the probability density function (pdf). The Radon-Nikodym Theorem says that measure \( Q \) has a density with respect to \( P \) if and only if \( Q \) is absolutely continuous with respect to \( P \). We write this as \( f = dQ/dP \) or \( dQ = f \, dP \).

A mapping \( V : \Omega \rightarrow \mathbb{R}^m \) is said to be measurable if for any Borel set \( A \in \mathcal{B} \), its inverse image \( V^{-1}(A) := \{ \omega \in \Omega : V(\omega) \in A \} \) is \( \mathcal{F} \)-measurable.\(^8\) A measurable mapping \( V(\omega) \) from probability space \((\Omega, \mathcal{F}, P)\) into \( \mathbb{R}^m \) is called a random vector. Note that the mapping \( V \) generates the probability measure\(^9\) (also called the probability distribution) \( P(A) := P(V^{-1}(A)) \) on \((\mathbb{R}^m, \mathcal{B})\). The smallest closed set \( \Xi \subset \mathbb{R}^m \) such that \( P(\Xi) = 1 \) is called the support of measure \( P \). We can view the space \((\Xi, \mathcal{B})\) equipped with probability measure \( P \) as a probability space \((\Xi, \mathcal{B}, P)\). This probability space provides all relevant probabilistic information about the considered random vector. In that case we write \( \Pr(A) \) for the probability of the event \( A \in \mathcal{B} \). We often denote by \( \xi \) data vector of a considered problem. Sometimes we view \( \xi \) as a random vector \( \xi : \Omega \rightarrow \mathbb{R}^m \) supported on a set \( \Xi \subset \mathbb{R}^m \), and sometimes as an element \( \xi \in \Xi \), i.e., as a particular realization of the random data vector. Usually, the meaning of such notation will be clear from the context and will not cause any confusion. If in doubt, in order to emphasize that we view \( \xi \) as a random vector, we sometimes write \( \xi = \xi(\omega) \).

A measurable mapping (function) \( Z : \Omega \rightarrow \mathbb{R} \) is called a random variable. Its probability distribution is completely defined by the cumulative distribution function (cdf) \( H_Z(z) := \Pr(Z \leq z) \). Note that since the Borel sigma algebra of \( \mathbb{R} \) is generated by the family of half line intervals \((-\infty, a]\), in order to verify measurability of \( Z(\omega) \) it suffices to verify measurability of sets \( \{ \omega \in \Omega : Z(\omega) \leq z \} \) for all \( z \in \mathbb{R} \). We denote random vectors (variables) by capital letters, like \( V, Z \) etc., or \( \xi(\omega) \), and often suppress their explicit dependence on \( \omega \in \Omega \). The coordinate functions \( V_1(\omega), ..., V_m(\omega) \) of the \( m \)-dimensional random vector \( V(\omega) \) are called its components. While considering a random vector \( V \) we often talk about its probability distribution as the joint distribution of its components

---

\(^8\)In fact it suffices to verify \( \mathcal{F} \)-measurability of \( V^{-1}(A) \) for any family of sets generating the Borel sigma algebra of \( \mathbb{R}^m \).

\(^9\)With some abuse of notation we also denote here by \( P \) the probability distribution induced by the probability measure \( P \) on \((\Omega, \mathcal{F})\).
7.2. Probability

(random variables) $V_1, ..., V_n$.

Since we often deal with random variables which are given as optimal values of optimization problems we need to consider random variables $Z(\omega)$ which can also take values $+\infty$ or $-\infty$, i.e., functions $Z : \Omega \to \mathbb{R}$, where $\mathbb{R}$ denotes the set of extended real numbers. Such functions $Z : \Omega \to \mathbb{R}$ are referred to as extended real valued functions. Operations between real numbers and symbols $\pm \infty$ are clear except for such operations as adding $+\infty$ and $-\infty$ which should be avoided. Measurability of an extended real valued function $Z(\omega)$ is defined in the standard way, i.e., $Z(\omega)$ is measurable if the set $\{\omega \in \Omega : Z(\omega) \leq z\}$ is $\mathcal{F}$-measurable for any $z \in \mathbb{R}$. A measurable extended real valued function is called an (extended) random variable. Note that here $\lim_{n \to +\infty} F_Z(z)$ is equal to the probability of the event $\{\omega \in \Omega : Z(\omega) < +\infty\}$ and can be less than one if the event $\{\omega \in \Omega : Z(\omega) = +\infty\}$ has a positive probability.

The expected value or expectation of an (extended) random variable $Z : \Omega \to \mathbb{R}$ is defined by the integral

$$E_P[Z] := \int_{\Omega} Z(\omega) dP(\omega). \quad \tag{7.96}$$

When there is no ambiguity as to what probability measure is considered, we omit the subscript $P$ and simply write $E[Z]$. For a nonnegative valued measurable function $Z(\omega)$ such that the event $\Upsilon := \{\omega \in \Omega : Z(\omega) = +\infty\}$ has zero probability the above integral is defined in the usual way and can take value $+\infty$. If probability of the event $\Upsilon$ is positive, then, by definition, $E[Z] = +\infty$. For a general (not necessarily nonnegative valued) random variable we would like to define $E[Z] := E[Z_+] - E[(-Z)_+]$. In order to do that we have to ensure that we do not add $+\infty$ and $-\infty$. We say that the expected value $E[Z]$ of an (extended real valued) random variable $Z(\omega)$ is well defined if it does not happen that both $E[Z_+]$ and $E[(-Z)_+]$ are $+\infty$, in which case $E[Z] = E[Z_+] - E[(-Z)_+]$. That is, in order to verify that the expected value of $Z(\omega)$ is well defined one has to check that $Z(\omega)$ is measurable and either $E[Z_+] < +\infty$ or $E[(-Z)_+] < +\infty$. Note that if $Z(\omega)$ and $Z'(\omega)$ are two (extended) random variables such that their expectations are well defined and $Z(\omega) = Z'(\omega)$ for all $\omega \in \Omega$ except possibly on a set of measure zero, then $E[Z] = E[Z']$. It is said that $Z(\omega)$ is $P$-integrable if the expected value $E[Z]$ is well defined and finite. The expected value of a random vector is defined componentwise.

If the random variable $Z(\omega)$ can take only a countable (finite) number of different values, say $z_1, z_2, ..., z_n$, then it is said that $Z(\omega)$ has a discrete distribution (discrete distribution with a finite support). In such cases all relevant probabilistic information is contained in the probabilities $p_i := \Pr\{Z = z_i\}$. In that case $E[Z] = \sum_i p_i z_i$.

Let $f_n(\omega)$ be a sequence of real valued measurable functions on a probability space $(\Omega, \mathcal{F}, P)$. By $f_n \uparrow f$ a.e. we mean that for almost every $\omega \in \Omega$ the sequence $f_n(\omega)$ is monotonically nondecreasing and hence converges to a limit denoted $f(\omega)$, where $f(\omega)$ can be equal to $+\infty$. We have the following classical results about convergence of integrals.

**Theorem 7.29 (Monotone Convergence Theorem).** Suppose that $f_n \uparrow f$ a.e. and there exists a $P$-integrable function $g(\omega)$ such that $f_n(\cdot) \geq g(\cdot)$. Then $\int_{\Omega} f dP$ is well defined and $\int_{\Omega} f_n dP \uparrow \int_{\Omega} f dP$.

$^{10}$Recall that $Z_+ := \max\{0, Z\}$.
Theorem 7.30 (Fatou's lemma). Suppose that there exists a $P$-integrable function $g(\omega)$ such that $f_n(\omega) \geq g(\omega)$. Then

$$\liminf_{n \to \infty} \int_{\Omega} f_n \, dP \geq \int_{\Omega} \liminf_{n \to \infty} f_n \, dP.$$  

(7.97)

Theorem 7.31 (Lebesgue Dominated Convergence Theorem). Suppose that there exists a $P$-integrable function $g(\omega)$ such that $|f_n| \leq g$ a.e., and that $f_n(\omega)$ converges to $f(\omega)$ for almost every $\omega \in \Omega$. Then $\int_{\Omega} f \, dP$ is well defined and $\int_{\Omega} f_n \, dP \to \int_{\Omega} f \, dP$.

We also have the following useful result. Unless stated otherwise we always assume that considered measures are finite and nonnegative, i.e., $\mu(A)$ is a finite nonnegative number for every $A \in \mathcal{F}$.

Theorem 7.32 (Richter-Rogosinski). Let $(\Omega, \mathcal{F})$ be a measurable space, $f_1, \ldots, f_m$ be measurable on $(\Omega, \mathcal{F})$ real valued functions, and $\mu$ be a measure on $(\Omega, \mathcal{F})$ such that $f_1, \ldots, f_m$ are $\mu$-integrable. Suppose that every finite subset of $\Omega$ is $\mathcal{F}$-measurable. Then there exists a measure $\eta$ on $(\Omega, \mathcal{F})$ with a finite support of at most $m$ points such that $\int_{\Omega} f_i \, d\mu = \int_{\Omega} f_i \, d\eta$ for all $i = 1, \ldots, m$.

Proof. The proof proceeds by induction on $m$. It can be easily shown that the assertion holds for $m = 1$. Consider the set $S \subset \mathbb{R}^m$ generated by vectors of the form $(\int f_1 d\mu', \ldots, \int f_m d\mu')$ with $\mu'$ being a measure on $\Omega$ with a finite support. We have to show that vector $a := (\int f_1 d\mu, \ldots, \int f_m d\mu)$ belongs to $S$. Note that the set $S$ is a convex cone. Suppose that $a \notin S$. Then, by the separation theorem, there exists $c \in \mathbb{R}^m \setminus \{0\}$ such that $c^T a \leq c^T x$, for all $x \in S$. Since $S$ is a cone, it follows that $c^T a \leq 0$. This implies that for $f := \sum_{i=1}^m c_i f_i$ we have that $\int f d\mu \leq 0$ and $\int f d\mu \leq \int f d\mu'$ for any measure $\mu'$ with a finite support. In particular, by taking measures of the form $\mu' = \alpha \Delta(\omega)$, with $\alpha > 0$ and $\omega \in \Omega$, we obtain from the second inequality that $\int f d\mu \leq \alpha f(\omega)$. This implies that $f(\omega) \geq 0$ for all $\omega \in \Omega$, since otherwise if $f(\omega) < 0$ we can make a $\alpha$ arbitrary small by taking $a$ large enough. Together with the first inequality this implies that $\int f d\mu = 0$.

Consider the set $\Omega' := \{\omega \in \Omega : f(\omega) = 0\}$. Note that the function $f$ is measurable and hence $\Omega' \in \mathcal{F}$. Since $\int f d\mu = 0$ and $f(\cdot)$ is nonnegative valued, it follows that $\Omega'$ is a support of $\mu$, i.e., $\mu(\Omega') = \mu(\Omega)$. If $\mu(\Omega) = 0$, then the assertion clearly holds. Therefore suppose that $\mu(\Omega) > 0$. Then $\mu(\Omega') > 0$, and hence $\Omega'$ is nonempty. Moreover, the functions $f_i$, $i = 1, \ldots, m$, are linearly dependent on $\Omega'$. Consequently, by the induction assumption there exists a measure $\mu'$ with a finite support on $\Omega'$ such that $\int_{\Omega'} f_i d\mu' = 0$ for all $i = 1, \ldots, m$, where $\mu'$ is the restriction of the measure $\mu$ to the set $\Omega'$. Moreover, since $\mu$ is supported on $\Omega'$ we have that $\int_{\Omega'} f_i d\mu' = \int_{\Omega} f_i d\mu$, and hence the proof is complete. 

Let us remark that if the measure $\mu$ is a probability measure, i.e., $\mu(\Omega) = 1$, then by adding the constraint $\int_{\Omega} d\eta = 1$, we obtain in the above theorem that there exists a probability measure $\eta$ on $(\Omega, \mathcal{F})$ with a finite support of at most $m + 1$ points such that

---

11We denote by $\Delta(\omega)$ measure of mass one at the point $\omega$, and refer to such measures as Dirac measures.
7.2. Conditional Probability and Conditional Expectation

Also let us recall two famous inequalities. Chebyshev inequality\(^{12}\) says that if \(Z : \Omega \to \mathbb{R}_+\) is a nonnegative valued random variable, then
\[
\Pr(Z \geq \alpha) \leq \alpha^{-1} \mathbb{E}[Z], \quad \forall \alpha > 0. \tag{7.98}
\]

Proof of (7.98) is rather simple, we have
\[
\Pr(Z \geq \alpha) = \mathbb{E}[\mathbf{1}_{[\alpha, +\infty)}(Z)] \leq \mathbb{E}[\alpha^{-1}Z] = \alpha^{-1} \mathbb{E}[Z].
\]

Chebyshev inequality says that if \(V : \Omega \to \mathbb{R}^m\) is a random vector, \(\nu := \mathbb{E}[V]\) and \(f : \mathbb{R}^m \to \mathbb{R}\) is a convex function, then
\[
\mathbb{E}[f(V)] \geq f(\nu), \tag{7.99}
\]

provided the above expectations are finite. Indeed, for a subgradient \(g \in \partial f(\nu)\) we have that
\[
f(V) \geq f(\nu) + g^T(V - \nu). \tag{7.100}
\]

By taking expectation of the both sides of (7.100) we obtain (7.99).

Finally, let us mention the following simple inequality. Let \(Y_1, Y_2 : \Omega \to \mathbb{R}\) be random variables and \(a_1, a_2\) be numbers. Then the intersection of the events \(\{\omega : Y_1(\omega) < a_1\}\) and \(\{\omega : Y_2(\omega) < a_2\}\) is included in the event \(\{\omega : Y_1(\omega) + Y_2(\omega) < a_1 + a_2\}\), or equivalently the event \(\{\omega : Y_1(\omega) + Y_2(\omega) \geq a_1 + a_2\}\) is included in the union of the events \(\{\omega : Y_1(\omega) \geq a_1\}\) and \(\{\omega : Y_2(\omega) \geq a_2\}\). It follows that
\[
\Pr\{Y_1 + Y_2 \geq a_1 + a_2\} \leq \Pr\{Y_1 \geq a_1\} + \Pr\{Y_2 \geq a_2\}. \tag{7.101}
\]

7.2.2 Conditional Probability and Conditional Expectation

For two events \(A\) and \(B\) the conditional probability of \(A\) given \(B\) is
\[
P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{7.102}
\]

provided that \(P(B) \neq 0\). Now let \(X\) and \(Y\) be discrete random variables with joint mass function \(p(x, y) := P(X = x, Y = y)\). Of course, since \(X\) and \(Y\) are discrete, \(p(x, y)\) is nonzero only for a finite or countable number of values of \(x\) and \(y\). The marginal mass functions of \(X\) and \(Y\) are \(p_x(x) := P(X = x) = \sum_y p(x, y)\) and \(p_y(y) := P(Y = y) = \sum_x p(x, y)\), respectively. It is natural to define conditional mass function of \(X\) given that \(Y = y\) as
\[
p_{x|Y}(x|y) := P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_y(y)} \tag{7.103}
\]

\(^{12}\) Sometimes (7.98) is called Markov inequality, while Chebyshev inequality is referred to the inequality (7.98) applied to the function \((Z - \mathbb{E}[Z])^2\).
for all values of \( y \) such that \( p_y (y) > 0 \). We have that \( X \) is independent of \( Y \) iff \( p(x, y) = p_x (x)p_y (y) \) holds for all \( x \) and \( y \), which is equivalent to that \( p_{X|Y} (x|y) = p_x (x) \) for all \( y \) such that \( p_y (y) > 0 \).

If \( X \) and \( Y \) have continuous distribution with a joint pdf \( f(x, y) \), then the conditional pdf of \( X \), given that \( Y = y \), is defined in a way similar to (7.103) for all values of \( y \) such that \( f_y (y) > 0 \) as

\[
f_{x|y}(x|y) := \frac{f(x, y)}{f_y (y)}. \tag{7.104}\]

Here \( f_y (y) := \int_{-\infty}^{+\infty} f(x, y)dx \) is the marginal pdf of \( Y \). In the continuous case the conditional expectation of \( X \), given that \( Y = y \), is defined for all values of \( y \) such that \( f_y (y) > 0 \) as

\[
\mathbb{E}[X|Y = y] := \int_{-\infty}^{+\infty} xf_{x|y}(x|y)dx. \tag{7.105}\]

In the discrete case it is defined in a similar way.

Note that \( \mathbb{E}[X|Y = y] \) is a function of \( y \), say \( h(y) := \mathbb{E}[X|Y = y] \). Let us denote by \( \mathbb{E}[X|Y] \) that function of random variable \( Y \), i.e., \( \mathbb{E}[X|Y] := h(Y) \). We have then the following important formula

\[
\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]. \tag{7.106}\]

In the continuous case, for example, we have

\[
\mathbb{E}[X] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y)dxdy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf_{x|y}(x|y)dxf_y (y)dy,
\]

and hence

\[
\mathbb{E}[X] = \int_{-\infty}^{+\infty} \mathbb{E}[X|Y = y]f_y (y)dy. \tag{7.107}\]

The above definitions can be extended to the case where \( X \) and \( Y \) are two random vectors in a straightforward way.

It is also useful to define conditional expectation in the following abstract form. Let \( X \) be a nonnegative valued integrable random variable on a probability space \( (\Omega, \mathcal{F}, P) \), and let \( \mathcal{G} \) be a subalgebra of \( \mathcal{F} \). Define a measure on \( \mathcal{G} \) by \( \nu(G) := \int_G XdP \) for any \( G \in \mathcal{G} \). This measure is finite because \( X \) is integrable, and is absolutely continuous with respect to \( P \). Hence by the Radon-Nikodym Theorem there is a \( \mathcal{G} \)-measurable function \( h(\omega) \) such that \( \nu(G) = \int_G h dP \). This function \( h(\omega) \), viewed as a random variable has the following properties: (i) \( h(\omega) \) is \( \mathcal{G} \)-measurable and integrable, (ii) it satisfies the equation \( \int_G h dP = \int_G XdP \) for any \( G \in \mathcal{G} \). By definition we say that a random variable, denoted \( \mathbb{E}[X|\mathcal{G}] \), is said to be the conditional expected value of \( X \) given \( \mathcal{G} \), if it satisfies the following two properties:

(i) \( \mathbb{E}[X|\mathcal{G}] \) is \( \mathcal{G} \)-measurable and integrable,

(ii) \( \mathbb{E}[X|\mathcal{G}] \) satisfies the functional equation

\[
\int_G \mathbb{E}[X|\mathcal{G}]dP = \int_G XdP, \; \forall G \in \mathcal{G}. \tag{7.108}\]
7.2. Probability

The above construction showed existence of such random variable for nonnegative $X$. If $X$ is not necessarily nonnegative, apply the same construction to the positive and negative part of $X$.

There will be many random variable satisfying the above properties (i) and (ii). Any one of them is called a version of the conditional expected value. We sometimes write it as $E[X|G](\omega)$ or $E[X|G]_\omega$ to emphasize that this a random variable. Any two versions of $E[X|G]$ are equal to each other with probability one. Note that, in particular, for $G = \Omega$ it follows from (ii) that

$$E[X] = \int_\Omega E[X|G]dP = E[E[X|G]].$$

(7.109)

Note also that if the sigma algebra $G$ is trivial, i.e., $G = \{\emptyset, \Omega\}$, then $E[X|G]$ is constant equal to $E[X]$.

Conditional probability $P(A|G)$ of event $A \in \mathcal{F}$ can be defined as $P(A|G) = E[1_A|G]$. In that case the corresponding properties (i) and (ii) take the form:

(i') $P(A|G)$ is $G$-measurable and integrable,

(ii') $P(A|G)$ satisfies the functional equation

$$\int_G P(A|G)dP = P(A \cap G), \forall G \in \mathcal{G}.$$ 

(7.110)

7.2.3 Measurable Multifunctions and Random Functions

Let $G$ be a mapping from $\Omega$ into the set of subsets of $\mathbb{R}^n$, i.e., $G$ assigns to each $\omega \in \Omega$ a subset (possibly empty) $G(\omega)$ of $\mathbb{R}^n$. We refer to $G$ as a multifunction and write $G : \Omega \rightrightarrows \mathbb{R}^n$. It is said that $G$ is closed valued if $G(\omega)$ is a closed subset of $\mathbb{R}^n$ for every $\omega \in \Omega$. A closed valued multifunction $G$ is said to be measurable, if for every closed set $A \subset \mathbb{R}^n$ one has that the inverse image $G^{-1}(A) := \{\omega \in \Omega : G(\omega) \cap A \neq \emptyset\}$ is $\mathcal{F}$-measurable. Note that measurability of $G$ implies that the domain

$$\text{dom } G := \{\omega \in \Omega : G(\omega) \neq \emptyset\} = G^{-1}(\mathbb{R}^n)$$

of $G$ is an $\mathcal{F}$-measurable subset of $\Omega$.

**Proposition 7.33.** A closed valued multifunction $G : \Omega \rightrightarrows \mathbb{R}^n$ is measurable iff the (extended real valued) function $d(\omega) := \text{dist}(x, G(\omega))$ is measurable for any $x \in \mathbb{R}^n$.

**Proof.** Recall that by the definition $\text{dist}(x, G(\omega)) = +\infty$ if $G(\omega) = \emptyset$. Note also that $\text{dist}(x, G(\omega)) = \|x - y\|$ for some $y \in G(\omega)$, because of closedness of set $G(\omega)$. Therefore, for any $t \geq 0$ and $x \in \mathbb{R}^n$ we have that

$$\{\omega \in \Omega : \text{dist}(x, G(\omega)) \leq t\} = G^{-1}(x + tB),$$

where $B := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. It remains to note that it suffices to verify the measurability of $G^{-1}(A)$ for closed sets of the form $A = x + tB$, $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$. \qed
Remark 25. Suppose now that $\Omega$ is a Borel subset of $\mathbb{R}^m$ equipped with its Borel sigma algebra. Suppose, further, that the multifunction $G : \Omega \rightrightarrows \mathbb{R}^n$ is closed. That is, if $\omega_k \rightarrow \omega$, $x_k \in G(\omega_k)$ and $x_k \rightarrow x$, then $x \in G(\omega)$. Of course, any closed multifunction is closed valued. It follows that for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$ the level set $\{ \omega \in \Omega : \text{dist}(x, G(\omega)) \leq t \}$ is closed, and hence the function $d(\omega) := \text{dist}(x, G(\omega))$ is measurable. Consequently we obtain that any closed multifunction $G : \Omega \rightrightarrows \mathbb{R}^n$ is measurable.

It is said that a mapping $G : \text{dom} \; G \rightarrow \mathbb{R}^n$ is a selection of $G$ if $G(\omega) \in G(\omega)$ for all $\omega \in \text{dom} \; G$. If, in addition, the mapping $G$ is measurable, it is said that $G$ is a measurable selection of $G$.

Theorem 7.34 (Measurable selection theorem). A closed valued multifunction $G : \Omega \rightrightarrows \mathbb{R}^n$ is measurable iff its domain is an $\mathcal{F}$-measurable subset of $\Omega$ and there exists a countable family $\{ G_i \}_{i \in \mathbb{N}}$, of measurable selections of $G$, such that for every $\omega \in \Omega$, the set $\{ G_i(\omega) : i \in \mathbb{N} \}$ is dense in $G(\omega)$.

In particular, we have by the above theorem that if $G : \Omega \rightrightarrows \mathbb{R}^n$ is a closed valued measurable multifunction, then there exists at least one measurable selection of $G$. In [181, Theorem 14.5] the result of the above theorem is called Castaing representation.

Consider a function $F : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$. We say that $F$ is a random function if for every fixed $x \in \mathbb{R}^n$, the function $F(x, \cdot)$ is $\mathcal{F}$-measurable. For a random function $F(x, \omega)$ we can define the corresponding expected value function

$$f(x) := \mathbb{E}[F(x, \omega)] = \int_{\Omega} F(x, \omega) dP(\omega).$$

We say that $f(x)$ is well defined if the expectation $\mathbb{E}[F(x, \omega)]$ is well defined for every $x \in \mathbb{R}^n$. Also for every $\omega \in \Omega$ we can view $F(\cdot, \omega)$ as an extended real valued function.

Definition 7.35. It is said that the function $F(x, \omega)$ is random lower semicontinuous if the associated epigraphical multifunction $\omega \mapsto \text{epi} \; F(\cdot, \omega)$ is closed valued and measurable.

In some publications random lower semicontinuous functions are called normal integrands. It follows from the above definitions that if $F(x, \omega)$ is random lsc, then the multifunction $\omega \mapsto \text{dom} \; F(\cdot, \omega)$ is measurable, and $F(x, \cdot)$ is measurable for every fixed $x \in \mathbb{R}^n$. Close valuedness of the epigraphical multifunction means that for every $\omega \in \Omega$, the epigraph $\text{epi} \; F(\cdot, \omega)$ is a closed subset of $\mathbb{R}^{n+1}$, i.e., $F(\cdot, \omega)$ is lower semicontinuous. Note, however, that the lower semicontinuity in $x$ and measurability in $\omega$ does not imply measurability of the corresponding epigraphical multifunction and random lower semicontinuity of $F(x, \omega)$. A large class of random lower semicontinuous is given by the so-called Carathéodory functions, i.e., real valued functions $F : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$ such that $F(x, \cdot)$ is $\mathcal{F}$-measurable for every $x \in \mathbb{R}^n$ and $F(\cdot, \cdot)$ continuous for a.e. $\omega \in \Omega$.

Theorem 7.36. Suppose that the sigma algebra $\mathcal{F}$ is $\mathcal{P}$-complete. Then an extended real valued function $F : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$ is random lsc if the following two properties hold: (i) for every $\omega \in \Omega$, the function $F(\cdot, \omega)$ is lsc, (ii) the function $F(\cdot, \cdot)$ is measurable with respect to the sigma algebra of $\mathbb{R}^n \times \Omega$ given by the product of the sigma algebras $\mathcal{B}$ and $\mathcal{F}$.
With a random function \( F(x, \omega) \) we associate its **optimal value function** \( \vartheta(\omega) := \inf_{x \in \mathbb{R}^n} F(x, \omega) \) and the optimal solution multifunction \( X^*(\omega) := \arg \min_{x \in \mathbb{R}^n} F(x, \omega) \).

**Theorem 7.37.** Let \( F : \mathbb{R}^n \times \Omega \to \mathbb{R} \) be a random lsc function. Then the optimal value function \( \vartheta(\omega) \) and the optimal solution multifunction \( X^*(\omega) \) are both measurable.

Note that it follows from lower semicontinuity of \( F(\cdot, \omega) \) that the optimal solution multifunction \( X^*(\omega) \) is closed valued. Note also that if \( F(x, \omega) \) is random lsc and \( \bar{G} : \Omega \rightrightarrows \mathbb{R}^n \) is a closed valued measurable multifunction, then the function

\[
\bar{F}(x, \omega) := \begin{cases} 
F(x, \omega), & \text{if } x \in \bar{G}(\omega), \\
+\infty, & \text{if } x \notin \bar{G}(\omega),
\end{cases}
\]

is also random lsc. Consequently the corresponding optimal value \( \omega \mapsto \inf_{x \in \bar{G}(\omega)} F(x, \omega) \) and the optimal solution multifunction \( \omega \mapsto \arg \min_{x \in \bar{G}(\omega)} F(x, \omega) \) are both measurable, and hence by the measurable selection theorem, there exists a measurable selection \( \bar{x}(\omega) \in \arg \min_{x \in \bar{G}(\omega)} F(x, \omega) \).

**Theorem 7.38.** Let \( F : \mathbb{R}^{n+m} \times \Omega \to \mathbb{R} \) be a random lsc function and

\[
\vartheta(x, \omega) := \inf_{y \in \mathbb{R}^m} F(x, y, \omega)
\]

be the associated optimal value function. Suppose that there exists a bounded set \( S \subset \mathbb{R}^m \) such that \( \text{dom} F(x, \cdot, \omega) \subset S \) for all \( (x, \omega) \in \mathbb{R}^n \times \Omega \). Then the optimal value function \( \vartheta(x, \omega) \) is random lsc.

Let us observe that the above framework of random lsc functions is aimed at minimization problems. Of course, the problem of maximization of \( \mathbb{E}[F(x, \omega)] \) is equivalent to minimization of \( \mathbb{E}[-F(x, \omega)] \). Therefore, for maximization problems one would need the comparable concept of random upper semicontinuous functions.

Consider a multifunction \( \bar{G} : \Omega \rightrightarrows \mathbb{R}^n \). Denote

\[
\|\bar{G}(\omega)\| := \sup \{ \|G(\omega)\| : G(\omega) \in \bar{G}(\omega) \},
\]

and by \( \text{conv} \bar{G}(\omega) \) the convex hull of set \( \bar{G}(\omega) \). If the set \( \Omega = \{\omega_1, \ldots, \omega_K\} \) is finite and equipped with respective probabilities \( p_k \), \( k = 1, \ldots, K \), then it is natural to define the integral

\[
\int_{\Omega} \bar{G}(\omega) dP(\omega) := \sum_{k=1}^{K} p_k \bar{G}(\omega_k),
\]

where the sum of two sets \( A, B \subset \mathbb{R}^n \) and multiplication by a scalar \( \gamma \in \mathbb{R} \) are defined in the natural way, \( A + B := \{a+b : a \in A, b \in B\} \) and \( \gamma A := \{\gamma a : a \in A\} \). For a general measure \( P \) on a sample space \((\Omega, \mathcal{F})\) the corresponding integral is defined as follows.

**Definition 7.39.** The integral \( \int_{\Omega} \bar{G}(\omega) dP(\omega) \) is defined as the set of all points of the form \( \int_{\Omega} G(\omega) dP(\omega) \), where \( G(\omega) \) is a \( P \)-integrable selection of \( \bar{G}(\omega) \), i.e., \( G(\omega) \in \bar{G}(\omega) \) for a.e. \( \omega \in \Omega \), \( G(\omega) \) is measurable and \( \int_{\Omega} ||G(\omega)|| dP(\omega) \) is finite.
Theorem 7.40 (Aumann). Suppose that the measure \( P \) is nonatomic and let \( G : \Omega \rightarrow \mathbb{R}^n \) be a multifunction. Then the set \( \int_\Omega G \, dP \) is convex. Suppose, further, that \( G(\omega) \) is closed valued and measurable, and there exists a \( P \)-integrable function \( g(\omega) \) such that \( \| G(\omega) \| \leq g(\omega) \) for a.e. \( \omega \in \Omega \). Then
\[
\int_\Omega G(\omega) \, dP(\omega) = \int_\Omega (\text{conv } G(\omega)) \, dP(\omega). \tag{7.113}
\]

The above theorem is a consequence of a theorem due to A.A. Lyapunov (1940).

Theorem 7.41 (Lyapunov). Let \( \mu_1, ..., \mu_n \) be a finite collection of nonatomic measures on a measurable space \((\Omega, \mathcal{F})\). Then the set \( \{(\mu_1(S), ..., \mu_n(S)) : S \in \mathcal{F}\} \) is a closed and convex subset of \( \mathbb{R}^n \).

7.2.4 Expectation Functions

Consider a random function \( F : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R} \) and the corresponding expected value (or simply expectation) function \( f(x) = \mathbb{E}[F(x, \omega)] \). Recall that by assuming that \( F(x, \omega) \) is a random function we assume that \( F(x, \cdot) \) is measurable for every \( x \in \mathbb{R}^n \). We have that the function \( f(x) \) is well defined on a set \( X \subset \mathbb{R}^n \) if for every \( x \in X \) either \( \mathbb{E}[F(x, \omega) \mathbb{1}_X] < +\infty \) or \( \mathbb{E}[|F(x, \omega)| \mathbb{1}_X] < +\infty \). The expectation function inherits various properties of the functions \( F(\cdot, \omega) \), \( \omega \in \Omega \). As it is shown in the next proposition the lower semicontinuity of the expected value function follows from the lower semicontinuity of \( F(\cdot, \omega) \).

Theorem 7.42. Suppose that for \( P \)-almost every \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is lsc at a point \( x_0 \) and there exists \( P \)-integrable function \( Z(\omega) \) such that \( F(x, \omega) \geq Z(\omega) \) for \( P \)-almost all \( \omega \in \Omega \) and all \( x \) in a neighborhood of \( x_0 \). Then for all \( x \) in a neighborhood of \( x_0 \) the expected value function \( f(x) := \mathbb{E}[F(x, \omega)] \) is well defined and lsc at \( x_0 \).

**Proof.** It follows from the assumption that \( F(x, \omega) \) is bounded from below by a \( P \)-integrable function, that \( f(\cdot) \) is well defined in a neighborhood of \( x_0 \). Moreover, by Fatou’s lemma we have
\[
\liminf_{x \to x_0} \int_\Omega F(x, \omega) \, dP(\omega) \geq \int_\Omega \liminf_{x \to x_0} F(x, \omega) \, dP(\omega). \tag{7.114}
\]
Together with lsc of \( F(\cdot, \omega) \) this implies lower semicontinuity of \( f \) at \( x_0 \).

With a little stronger assumptions we can show that the expectation function is continuous.

---

\(^{13}\)It is said that measure \( P \), and the space \((\Omega, \mathcal{F}, P)\), is nonatomic if any set \( A \in \mathcal{F} \), such that \( P(A) > 0 \), contains a subset \( B \in \mathcal{F} \) such that \( P(A) > P(B) > 0 \).
Theorem 7.43. Suppose that for P-almost every \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is continuous at \( x_0 \) and there exists \( P \)-integrable function \( Z(\omega) \) such that \( |F(x, \omega)| \leq Z(\omega) \) for \( P \)-almost every \( \omega \in \Omega \) and all \( x \) in a neighborhood of \( x_0 \). Then for all \( x \) in a neighborhood of \( x_0 \) the expected value function \( f(x) \) is well defined and continuous at \( x_0 \).

Proof. It follows from the assumption that \( |F(x, \omega)| \) is dominated by a \( P \)-integrable function, that \( f(x) \) is well defined and finite valued for all \( x \) in a neighborhood of \( x_0 \). Moreover, by the Lebesgue Dominated Convergence Theorem we can take the limit inside the integral, which together with the continuity assumption implies

\[
\lim_{x \to x_0} \int_{\Omega} F(x, \omega)dP(\omega) = \int_{\Omega} \lim_{x \to x_0} F(x, \omega)dP(\omega) = \int_{\Omega} F(x_0, \omega)dP(\omega). \quad (7.115)
\]

This shows the continuity of \( f(x) \) at \( x_0 \). \( \square \)

Consider, for example, the characteristic function \( F(x, \omega) := 1_{[-\infty, x]}(\xi(\omega)) \), with \( x \in \mathbb{R} \) and \( \xi = \xi(\omega) \) being a real valued random variable. We have then that \( f(x) = \Pr(\xi \leq x) \), i.e., that \( f(\cdot) \) is the cumulative distribution function of \( \xi \). It follows that in this example the expected value function is continuous at a point \( x_0 \) iff the probability of the event \( \{\xi = x_0\} \) is zero. Note that \( x = \xi(\omega) \) is the only point at which the function \( F(\cdot, \omega) \) is discontinuous.

We say that random function \( F(x, \omega) \) is convex, if the function \( F(\cdot, \omega) \) is convex for a.e. \( \omega \in \Omega \). Convexity of \( F(\cdot, \omega) \) implies convexity of the expectation function \( f(x) \). Indeed, if \( F(x, \omega) \) is convex and the measure \( P \) is discrete, then \( f(x) \) is a weighted sum, with positive coefficients, of convex functions and hence is convex. For general measures convexity of the expectation function follows by passing to the limit. Recall that if \( f(x) \) is convex, then it is continuous on the interior of its domain. In particular, if \( f(x) \) is real valued for all \( x \in \mathbb{R}^n \), then it is continuous on \( \mathbb{R}^n \).

We discuss now differentiability properties of the expected value function \( f(x) \). We sometimes write \( F_\omega(\cdot) \) for the function \( F(\cdot, \omega) \) and denote by \( F'_\omega(x_0, h) \) the directional derivative of \( F_\omega(\cdot) \) at the point \( x_0 \) in the direction \( h \). Definitions and basic properties of directional derivatives are given in section 7.1.1. Consider the following conditions:

(A1) The expectation \( f(x_0) \) is well defined and finite valued at a given point \( x_0 \in \mathbb{R}^n \).

(A2) There exists a positive valued random variable \( C(\omega) \) such that \( \mathbb{E}[C(\omega)] < +\infty \), and for all \( x_1, x_2 \) in a neighborhood of \( x_0 \) and almost every \( \omega \in \Omega \) the following inequality holds

\[
|F(x_1, \omega) - F(x_2, \omega)| \leq C(\omega)\|x_1 - x_2\|. \quad (7.116)
\]

(A3) For almost every \( \omega \) the function \( F_\omega(\cdot) \) is directionally differentiable at \( x_0 \).

(A4) For almost every \( \omega \) the function \( F_\omega(\cdot) \) is differentiable at \( x_0 \).

Theorem 7.44. We have the following: (a) If conditions (A1) and (A2) hold, then the expected value function \( f(x) \) is Lipschitz continuous in a neighborhood of \( x_0 \). (b) If conditions (A1)–(A3) hold, then the expected value function \( f(x) \) is directionally differentiable
at $x_0$, and
\[
f'(x_0, h) = \mathbb{E} \[ F'_\omega(x_0, h) \] \quad \text{for all } h.
\] (7.117)

(c) If conditions (A1), (A2) and (A4) hold, then $f(x)$ is differentiable at $x_0$ and
\[
\nabla f(x_0) = \mathbb{E} \[ \nabla_x F(x_0, \omega) \].
\] (7.118)

Proof. It follows from (7.116) that for any $x_1, x_2$ in a neighborhood of $x_0$,
\[
|f(x_1) - f(x_2)| \leq \int_{\Omega} |F(x_1, \omega) - F(x_2, \omega)| dP(\omega) \leq c\|x_1 - x_2\|
\]
where $c := \mathbb{E}[C(\omega)]$. Together with assumption (A1) this implies that $f(x)$ is well defined, finite valued and Lipschitz continuous in a neighborhood of $x_0$.

Suppose now that assumptions (A1)–(A3) hold. For $t \neq 0$ consider the ratio
\[
R_t(\omega) := t^{-1} \[ F(x_0 + th, \omega) - F(x_0, \omega) \].
\]
By assumption (A2) we have that $|R_t(\omega)| \leq C(\omega)\|h\|$, and by assumption (A3) that
\[
\lim_{t \downarrow 0} R_t(\omega) = F'_\omega(x_0, h) \quad \text{w.p.1.}
\]
Therefore, it follows by the Lebesgue Dominated Convergence Theorem that
\[
\lim_{t \downarrow 0} \int_{\Omega} R_t(\omega) dP(\omega) = \int_{\Omega} \lim_{t \downarrow 0} R_t(\omega) dP(\omega).
\]
Together with assumption (A3) this implies formula (7.117). This proves the assertion (b).

Finally, if $F'_\omega(x_0, h)$ is linear in $h$ for almost every $\omega$, i.e., the function $F'_\omega(\cdot)$ is differentiable at $x_0$ w.p.1, then (7.117) implies that $f'(x_0, h)$ is linear in $h$, and hence (7.118) follows. Note that since $f(x)$ is locally Lipschitz continuous, we only need to verify linearity of $F'(x_0, \cdot)$ in order to establish (Fréchet) differentiability of $f(x)$ at $x_0$ (see theorem 7.2). This completes proof of (c). □

The above analysis shows that two basic conditions for interchangeability of the expectation and differentiation operators, i.e., for the validity of formula (7.118), are the above conditions (A2) and (A4). The following lemma shows that if, in addition to the assumptions (A1)–(A3), the directional derivative $F'_\omega(x_0, h)$ is convex in $h$ w.p.1, then $f(x)$ is differentiable at $x_0$ if and only if $F'(\cdot, \omega)$ is differentiable at $x_0$ w.p.1.

Lemma 7.45. Let $\psi : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be a random function such that for almost every $\omega \in \Omega$ the function $\psi(\cdot, \omega)$ is convex and positively homogeneous, and the expected value function $\phi(h) := \mathbb{E}[\psi(h, \omega)]$ is well defined and finite valued. Then the expected value function $\phi(\cdot)$ is linear if and only if the function $\psi(\cdot, \omega)$ is linear w.p.1.

Proof. We have here that the expected value function $\phi(\cdot)$ is convex and positively homogeneous. Moreover, it immediately follows from the linearity properties of the expectation operator that if the function $\psi(\cdot, \omega)$ is linear w.p.1, then $\phi(\cdot)$ is also linear.
Conversely, let $e_1, \ldots, e_n$ be a basis of the space $\mathbb{R}^n$. Since $\phi(\cdot)$ is convex and positively homogeneous, it follows that $\phi(e_i) + \phi(-e_i) \geq \phi(0) = 0$, $i = 1, \ldots, n$. Furthermore, since $\phi(\cdot)$ is finite valued, it is the support function of a convex compact set. This convex set is a singleton iff
\[ \phi(e_i) + \phi(-e_i) = 0, \quad i = 1, \ldots, n. \tag{7.119} \]
Therefore, $\phi(\cdot)$ is linear iff condition (7.119) holds. Consider the sets
\[ A_i := \{ \omega \in \Omega : \psi(e_i, \omega) + \psi(-e_i, \omega) > 0 \}. \]
Thus the set of $\omega \in \Omega$ such that $\psi(\cdot, \omega)$ is not linear coincides with the set $\bigcup_{i=1}^n A_i$. If $P(\bigcup_{i=1}^n A_i) > 0$, then at least one of the sets $A_i$ has a positive measure. Let, for example, $P(A_i)$ be positive. Then $\phi(e_i) + \phi(-e_i) > 0$, and hence $\phi(\cdot)$ is not linear. This completes the proof. \[ \square \]

Regularity conditions which are required for formula (7.117) to hold are simplified further if the random function $F(x, \omega)$ is convex. In that case, by using the Monotone Convergence Theorem instead of the Lebesgue Dominated Convergence Theorem, it is possible to prove the following result.

**Theorem 7.46.** Suppose that the random function $F(x, \omega)$ is convex and the expected value function $f(x)$ is well defined and finite valued in a neighborhood of a point $x_0$. Then $f(x)$ is convex, directionally differentiable at $x_0$ and formula (7.117) holds. Moreover, $f(x)$ is differentiable at $x_0$ if and only if $F(x)$ is differentiable at $x_0$ w.p.1, in which case formula (7.118) holds.

**Proof.** The convexity of $f(x)$ follows from convexity of $F(x, \omega)$. Since $f(x)$ is convex and finite valued near $x_0$ it follows that $f(x)$ is directionally differentiable at $x_0$ with finite directional derivative $f'(x_0, h)$ for every $h \in \mathbb{R}^n$. Consider a direction $h \in \mathbb{R}^n$. Since $f(x)$ is finite valued near $x_0$, we have that $f(x_0)$ and, for some $t_0 > 0$, $f(x_0 + t_0 h)$ are finite. It follows from the convexity of $F(x, \cdot)$ that the ratio
\[ R_t(\omega) := t^{-1} [F(x_0 + th, \omega) - F(x_0, \omega)] \]
is monotonically decreasing to $F'_+(x_0, h)$ as $t \downarrow 0$. Also we have that
\[ \mathbb{E} |R_t(\omega)| \leq t_0^{-1} \left( \mathbb{E} |F(x_0 + t_0 h, \omega)| + \mathbb{E} |F(x_0, \omega)| \right) < +\infty. \]
Then it follows by the Monotone Convergence Theorem that
\[ \lim_{t \downarrow 0} \mathbb{E} [R_t(\omega)] = \mathbb{E} \left[ \lim_{t \downarrow 0} R_t(\omega) \right] = \mathbb{E} [F'_+(x_0, h)]. \tag{7.120} \]
Since $\mathbb{E} [R_t(\omega)] = t^{-1} [f(x_0 + th) - f(x_0)]$, we have that the left hand side of (7.120) is equal to $f'(x_0, h)$, and hence formula (7.117) follows.

The last assertion follows then from Lemma 7.45. \[ \square \]

**Remark 26.** It is possible to give a version of the above result for a particular direction $h \in \mathbb{R}^n$. That is, suppose that: (i) the expected value function $f(x)$ is well defined in a
neighborhood of a point \( x_0 \), (ii) \( f(x_0) \) is finite, (iii) for almost every \( \omega \in \Omega \) the function \( F_\omega(\cdot) := F(\cdot, \omega) \) is convex, (iv) \( \mathbb{E}[F(x_0 + t_0 h, \omega)] < +\infty \) for some \( t_0 > 0 \). Then \( f'(x_0, h) < +\infty \) and formula (7.117) holds. Note also that if the above assumptions (i)–(iii) are satisfied and \( \mathbb{E}[F(x_0 + th, \omega)] = +\infty \) for any \( t > 0 \), then clearly \( f'(x_0, h) = +\infty \).

Often the expectation operator smoothes the integrand \( F(x, \omega) \). Consider, for example, \( F(x, \omega) := |x - \xi(\omega)| \) with \( x \in \mathbb{R} \) and \( \xi(\omega) \) being a real valued random variable. Suppose that \( f(x) = \mathbb{E}[F(x, \omega)] \) is finite valued. We have here that \( F(\cdot, \omega) \) is convex and \( F(\cdot, \omega) \) is differentiable everywhere except \( x = \xi(\omega) \). The corresponding derivative is given by \( \partial F(x, \omega)/\partial x = 1 \) if \( x > \xi(\omega) \) and \( \partial F(x, \omega)/\partial x = -1 \) if \( x < \xi(\omega) \). Therefore, \( f(x) \) is differentiable at \( x_0 \) if the event \( \{\xi(\omega) = x_0\} \) has zero probability, in which case

\[
\frac{df(x_0)}{dx} = \mathbb{E}[\partial F(x_0, \omega)/\partial x] = \Pr(\xi < x_0) - \Pr(\xi > x_0). \tag{7.121}
\]

If the event \( \{\xi(\omega) = x_0\} \) has positive probability, then the directional derivatives \( f'(x_0, h) \) exist but are not linear in \( h \), that is,

\[
f'(x_0, -1) + f'(x_0, 1) = 2\Pr(\xi = x_0) > 0. \tag{7.122}
\]

We can also investigate differentiability properties of the expectation function by studying the subdifferentiability of the integrand. Suppose for the moment that the set \( \Omega \) is finite, say \( \Omega := \{\omega_1, \ldots, \omega_K\} \) with \( P(\omega = \omega_k) = p_k > 0 \), and that the functions \( F(\cdot, \omega_k), \omega \in \Omega \), are proper. Then \( f(x) = \sum_{k=1}^K p_k F(x, \omega_k) \) and \( \partial f = \bigcap_{k=1}^K \partial F_k \), where \( F_k(\cdot) := F(\cdot, \omega_k) \). The Moreau–Rockafellar Theorem (Theorem 7.4) allows us to express the subdifferential of \( f(x) \) as the sum of subdifferentials of \( p_k F(x, \omega_k) \). That is, suppose that: (i) the set \( \Omega = \{\omega_1, \ldots, \omega_K\} \) is finite, (ii) for every \( \omega_k \in \Omega \) the function \( F_k(\cdot) := F(\cdot, \omega_k) \) is proper and convex, (iii) the sets \( \partial F_k \), \( k = 1, \ldots, K \), have a common point. Then for any \( x_0 \in \partial f \),

\[
\partial f(x_0) = \sum_{k=1}^K p_k \partial F(x_0, \omega_k). \tag{7.123}
\]

Note that the above regularity assumption (iii) holds, in particular, if the interior of \( \partial f \) is nonempty.

The subdifferentials at the right hand side of (7.123) are taken with respect to \( x \). Note that \( \partial F(x_0, \omega_k) \), and hence \( \partial f(x_0) \), in (7.123) can be unbounded or empty. Suppose that all probabilities \( p_k \) are positive. It follows then from (7.123) that \( \partial f(x_0) \) is a singleton iff all subdifferentials \( \partial F(x_0, \omega_k) \), \( k = 1, \ldots, K \), are singletons. That is, \( f(\cdot) \) is differentiable at a point \( x_0 \in \partial f \) iff all \( F(\cdot, \omega_k) \) are differentiable at \( x_0 \).

**Remark 27.** In the case of a finite set \( \Omega \) we didn’t have to worry about the measurability of the multifunction \( \omega \mapsto \partial F(x, \omega) \). Consider now a general case where the measurable space does not need to be finite. Suppose that the function \( F(x, \omega) \) is random lower semi-continuous and for a.e. \( \omega \in \Omega \) the function \( F(\cdot, \omega) \) is convex and proper. Then for any \( x \in \mathbb{R}^n \), the multifunction \( \omega \mapsto \partial F(x, \omega) \) is measurable. Indeed, consider the conjugate

\[
F^*(z, \omega) := \sup_{x \in \mathbb{R}^n} \left\{ z^T x - F(x, \omega) \right\}
\]
of the function $F(\cdot, \omega)$. It is possible to show that the function $F^*(z, \omega)$ is also random lower semicontinuous. Moreover, by the Fenchel–Moreau Theorem, $F^{**} = F$ and by convex analysis (see (7.24))

$$\partial F(x, \omega) = \arg\max_{z \in \mathbb{R}^n} \{ z^T x - F^*(z, \omega) \}.$$ 

Then it follows by Theorem 7.37 that the multifunction $\omega \mapsto \partial F(x, \omega)$ is measurable.

In general we have the following extension of formula (7.123).

**Theorem 7.47.** Suppose that: (i) the function $F(x, \omega)$ is random lower semicontinuous, (ii) for a.e. $\omega \in \Omega$ the function $F(\cdot, \omega)$ is convex, (iii) the expectation function $f$ is proper, (iv) the domain of $f$ has a nonempty interior. Then for any $x_0 \in \text{dom } f$,

$$\partial f(x_0) = \int_{\Omega} \partial F(x_0, \omega) \, dP(\omega) + \mathcal{N}_{\text{dom } f}(x_0). \tag{7.124}$$

**Proof.** Consider a point $z \in \int_{\Omega} \partial F(x_0, \omega) \, dP(\omega)$. By the definition of that integral we have then that there exists a $P$-integrable selection $G(\omega) \in \partial F(x_0, \omega)$ such that $z = \int_{\Omega} G(\omega) \, dP(\omega)$. Consequently, for a.e. $\omega \in \Omega$ the following holds

$$F(x, \omega) - F(x_0, \omega) \geq G(\omega)^T (x - x_0), \quad \forall x \in \mathbb{R}^n.$$ 

By taking the integral of the both sides of the above inequality we obtain that $z$ is a subgradient of $f$ at $x_0$. This shows that

$$\int_{\Omega} \partial F(x_0, \omega) \, dP(\omega) \subset \partial f(x_0). \tag{7.125}$$

In particular, it follows from (7.125) that if $\partial f(x_0)$ is empty, then the set at the right hand side of (7.124) is also empty. If $\partial f(x_0)$ is nonempty, i.e., $f$ is subdifferentiable at $x_0$, then $\mathcal{N}_{\text{dom } f}(x_0)$ forms the recession cone of $\partial f(x_0)$. In any case it follows from (7.125) that

$$\int_{\Omega} \partial F(x_0, \omega) \, dP(\omega) + \mathcal{N}_{\text{dom } f}(x_0) \subset \partial f(x_0). \tag{7.126}$$

Note that inclusion (7.126) holds irrespective of assumption (iv).

Proving the converse of inclusion (7.126) is a more delicate problem. Let us outline main steps of such a proof based on the interchangeability property of the directional derivative and integral operators. We can assume that both sets at the left and right hand sides of (7.125) are nonempty. Since the subdifferentials $\partial F(x_0, \omega)$ are convex, it is quite easy to show that the set $\int_{\Omega} \partial F(x_0, \omega) \, dP(\omega)$ is convex. With some additional effort it is possible to show that this set is closed. Let us denote by $s_1(\cdot)$ and $s_2(\cdot)$ the support functions of the sets at the left and right hand sides of (7.126), respectively. By virtue of inclusion (7.125), $\mathcal{N}_{\text{dom } f}(x_0)$ forms the recession cone of the set at the left hand side of (7.126) as well. Since the tangent cone $T_{\text{dom } f}(x_0)$ is the polar of $\mathcal{N}_{\text{dom } f}(x_0)$, it follows that $s_1(h) = s_2(h) = +\infty$ for any $h \notin T_{\text{dom } f}(x_0)$. Suppose now that (7.124) does not
hold, i.e., inclusion (7.126) is strict. Then \( s_1(h) < s_2(h) \) for some \( h \in \text{dom } f(x_0) \). Moreover, by assumption (iv), the tangent cone \( \text{dom } f(x_0) \) has a nonempty interior and there exists \( h \) in the interior of \( \text{dom } f(x_0) \) such that \( s_1(h) < s_2(h) \). For such \( h \) the directional derivative \( f'(x_0, h) \) is finite for all \( h \) in a neighborhood of \( \bar{h} \), \( f'(x_0, \bar{h}) = s_2(\bar{h}) \) and (see Remark 26, page 373)

\[
f'(x_0, \bar{h}) = \int_\Omega F'_x(x_0, \bar{h}) \, dP(\omega).
\]

Also, \( F'_x(x_0, h) \) is finite for a.e. \( \omega \) and for all \( h \) in a neighborhood of \( \bar{h} \), and hence \( F'_x(x_0, h) = \bar{h}^T G(\omega) \) for some \( G(\omega) \in \partial F(x_0, \omega) \). Moreover, since the multifunction \( \omega \mapsto \partial F(x_0, \omega) \) is measurable, we can choose a measurable \( G(\omega) \) here. Consequently,

\[
\int_\Omega F'_x(x_0, \bar{h}) \, dP(\omega) = \bar{h}^T \int_\Omega G(\omega) \, dP(\omega).
\]

Since \( \int_\Omega G(\omega) \, dP(\omega) \) is a point of the set at the left hand side of (7.125), we obtain that \( s_1(\bar{h}) \geq f'(x_0, \bar{h}) = s_2(\bar{h}) \), a contradiction. \( \square \)

In particular, if \( x_0 \) is an interior point of the domain of \( f \), then under the assumptions of the above theorem we have that

\[
\partial f(x_0) = \int_\Omega \partial F(x_0, \omega) \, dP(\omega). \tag{7.127}
\]

Also it follows from formula (7.124) that \( f(\cdot) \) is differentiable at \( x_0 \) iff \( x_0 \) is an interior point of the domain of \( f \) and \( \partial F(x_0, \omega) \) is a singleton for a.e. \( \omega \in \Omega \), i.e., \( F(\cdot, \omega) \) is differentiable at \( x_0 \) w.p.1.

### 7.2.5 Uniform Laws of Large Numbers

Consider a sequence \( \xi^i = \xi^i(\omega) \), \( i \in \mathbb{N} \), of \( d \)-dimensional random vectors defined on a probability space \((\Omega, \mathcal{F}, P)\). As it was discussed in section 7.2.1 we can view \( \xi^i \) as random vectors supported on a (closed) set \( \Xi \subset \mathbb{R}^d \), equipped with its Borel sigma algebra \( \mathcal{B} \). We say that \( \xi^i, i \in \mathbb{N} \), are identically distributed if each \( \xi^i \) has the same probability distribution on \( (\Xi, \mathcal{B}) \). If, moreover, \( \xi^i, i \in \mathbb{N} \), are independent, we say that they are independent identically distributed (iid). Consider a measurable function \( F : \Xi \to \mathbb{R} \) and the sequence \( F(\xi^i), i \in \mathbb{N}, \) of random variables. If \( \xi^i \) are identically distributed, then \( F(\xi^i), i \in \mathbb{N} \), are also identically distributed and hence their expectations \( \mathbb{E}[F(\xi^i)] \) are constant, i.e., \( \mathbb{E}[F(\xi^i)] = \mathbb{E}[F(\xi^1)] \) for all \( i \in \mathbb{N} \). The Law of Large Numbers (LLN) says that if \( \xi^i \) are identically distributed and the expectation \( \mathbb{E}[F(\xi^1)] \) is well defined, then, under some regularity conditions\(^{14}\),

\[
N^{-1} \sum_{i=1}^N F(\xi^i) \to \mathbb{E}[F(\xi^1)] \text{ w.p.1 as } N \to \infty. \tag{7.128}
\]

\(^{14}\) Sometimes (7.128) is referred to as the strong LLN to distinguish it from the weak LLN where the convergence is ensured ‘in probability’ instead of ‘with probability one’. Unless stated otherwise we deal with the strong LLN.
7.2. Probability

In particular, the classical LLN states that the convergence (7.128) holds if the sequence $\xi^i$ is iid.

Consider now a random function $F : X \times \Xi \to \mathbb{R}$, where $X$ is a nonempty subset of $\mathbb{R}^n$ and $\xi = \xi(\omega)$ is a random vector supported on the set $\Xi$. Suppose that the corresponding expected value function $f(x) := \mathbb{E}[F(x, \xi)]$ is well defined and finite valued for every $x \in X$. Let $\xi^i = \xi^i(\omega)$, $i \in \mathbb{N}$, be an iid sequence of random vectors having the same distribution as the random vector $\xi$, and

$$\hat{f}_N(x) := N^{-1} \sum_{i=1}^N F(x, \xi^i)$$  \hspace{1cm} (7.129)

be the so-called sample average functions. Note that the sample average function $\hat{f}_N(x)$ depends on the random sequence $\xi^1, \ldots, \xi^N$, and hence is a random function. Since we assumed that all $\xi^i = \xi^i(\omega)$ are defined on the same probability space, we can view $\hat{f}_N(x) = f_N(x, \omega)$ as a sequence of functions of $x \in X$ and $\omega \in \Omega$.

We have that for every fixed $x \in X$ the LLN holds, i.e.,

$$\hat{f}_N(x) \to f(x) \text{ w.p.1 as } N \to \infty. \hspace{1cm} (7.130)$$

This means that for a.e. $\omega \in \Omega$, the sequence $f_N(x, \omega)$ converges to $f(x)$. That is, for any $\varepsilon > 0$ and a.e. $\omega \in \Omega$ there exists $N^* = N^*(\varepsilon, \omega, x)$ such that $|f_N(x) - f(x)| < \varepsilon$ for any $N \geq N^*$. It should be emphasized that $N^*$ depends on $\varepsilon$ and $\omega$, and also on $x \in X$. We may refer to (7.130) as a pointwise LLN. In some applications we will need a stronger form of LLN where the number $N^*$ can be chosen independent of $x \in X$. That is, we say that $f_N(x)$ converges to $f(x)$ w.p.1 uniformly on $X$ if

$$\sup_{x \in X} \left| f_N(x) - f(x) \right| \to 0 \text{ w.p.1 as } N \to \infty, \hspace{1cm} (7.131)$$

and refer to this as the uniform LLN. Note that maximum of a countable number of measurable functions is measurable. Since maximum (supremum) in (7.131) can be taken over a countable and dense subset of $X$, this supremum is a measurable function on $(\Omega, \mathcal{F})$.

We have the following basic result. It is said that $F(x, \xi)$, $x \in X$, is dominated by an integrable function if there exists a nonnegative valued measurable function $g(\xi)$ such that $\mathbb{E}[g(\xi)] < +\infty$ and for every $x \in X$ the inequality $|F(x, \xi)| \leq g(\xi)$ holds w.p.1.

**Theorem 7.48.** Let $X$ be a nonempty compact subset of $\mathbb{R}^n$ and suppose that: (i) for any $x \in X$ the function $F(\cdot, \xi)$ is continuous at $x$ for almost every $\xi \in \Xi$, (ii) $F(x, \xi)$, $x \in X$, is dominated by an integrable function, (iii) the sample is iid. Then the expected value function $f(x)$ is finite valued and continuous on $X$, and $f_N(x)$ converges to $f(x)$ w.p.1 uniformly on $X$.

**Proof.** It follows from the assumption (ii) that $|f(x)| \leq \mathbb{E}[g(\xi)]$, and consequently $|f(x)| < +\infty$ for all $x \in X$. Consider a point $x \in X$ and let $x_k$ be a sequence of points in $X$ converging to $x$. By the Lebesgue Dominated Convergence Theorem, assumption (ii) implies that

$$\lim_{k \to \infty} \mathbb{E}[F(x_k, \xi)] = \mathbb{E} \left[ \lim_{k \to \infty} F(x_k, \xi) \right].$$
Since, by (i), $F(x_k, \xi) \to F(x, \xi)$ w.p.1, it follows that $f(x_k) \to f(x)$, and hence $f(x)$ is continuous.

Choose now a point $\bar{x} \in X$, a sequence $\gamma_k$ of positive numbers converging to zero, and define $V_k := \{ x \in X : \|x - \bar{x}\| \leq \gamma_k \}$ and

$$\delta_k(\xi) := \sup_{x \in V_k} \|F(x, \xi) - F(\bar{x}, \xi)\|.$$  \hfill (7.132)

By the assumption (i) we have that for a.e. $\xi \in \Xi$, $\delta_k(\xi)$ tends to zero as $k \to \infty$. Moreover, by the assumption (ii) we have that $\delta_k(\xi)$, $k \in \mathbb{N}$, are dominated by an integrable function, and hence by the Lebesgue Dominated Convergence Theorem we have that

$$\lim_{k \to \infty} \mathbb{E}[\delta_k(\xi)] = \mathbb{E} \left[ \lim_{k \to \infty} \delta_k(\xi) \right] = 0. \hfill (7.133)$$

We also have that

$$\left| \hat{f}_N(x) - \hat{f}_N(\bar{x}) \right| \leq \frac{1}{N} \sum_{i=1}^{N} \left| F(x, \xi^i) - F(\bar{x}, \xi^i) \right|,$$

and hence

$$\sup_{x \in V_k} \left| \hat{f}_N(x) - \hat{f}_N(\bar{x}) \right| \leq \frac{1}{N} \sum_{i=1}^{N} \delta_k(\xi^i). \hfill (7.134)$$

Since the sequence $\xi^i$ is iid, it follows by the LLN that the right hand side of (7.134) converges w.p.1 to $\mathbb{E}[\delta_k(\xi)]$ as $N \to \infty$. Together with (7.133) this implies that for any given $\varepsilon > 0$ there exists a neighborhood $W$ of $\bar{x}$ such that w.p.1 for sufficiently large $N$,

$$\sup_{x \in W \cap X} \left| \hat{f}_N(x) - \hat{f}_N(\bar{x}) \right| < \varepsilon. \hfill (7.135)$$

Since $X$ is compact, there exists a finite number of points $x_1, \ldots, x_m \in X$ and corresponding neighborhoods $W_1, \ldots, W_m$ covering $X$ such that w.p.1 for $N$ large enough the following holds

$$\sup_{x \in W_j \cap X} \left| \hat{f}_N(x) - \hat{f}_N(x_j) \right| < \varepsilon, \quad j = 1, \ldots, m. \hfill (7.135)$$

Furthermore, since $f(x)$ is continuous on $X$, these neighborhoods can be chosen in such a way that

$$\sup_{x \in W_j \cap X} \left| f(x) - f(x_j) \right| < \varepsilon, \quad j = 1, \ldots, m. \hfill (7.136)$$

Again by the LLN we have that $\hat{f}_N(x_j)$ converges pointwise to $f(x)$ w.p.1. Therefore,

$$\left| \hat{f}_N(x_j) - f(x_j) \right| < \varepsilon, \quad j = 1, \ldots, m, \hfill (7.137)$$

w.p.1 for $N$ large enough. It follows from (7.135)–(7.137) that w.p.1 for $N$ large enough

$$\sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| < 3\varepsilon. \hfill (7.138)$$
Since \( \varepsilon > 0 \) was arbitrary, we obtain that (7.131) follows and hence the proof is complete.

\[
\square
\]

**Remark 28.** It could be noted that the assumption (i) in the above theorem means that \( F(\cdot, \xi) \) is continuous at any given point \( x \in X \) w.p.1. This does not mean, however, that \( F(\cdot, \xi) \) is continuous on \( X \) w.p.1. Take, for example, \( F(x, \xi) := 1_{\mathbb{R}_+}(x - \xi), \ x, \xi \in \mathbb{R}, \) i.e., \( F(x, \xi) = 1 \) if \( x \geq \xi \) and \( F(x, \xi) = 0 \) otherwise. We have here that \( F(\cdot, \xi) \) is always discontinuous at \( x = \xi \), and that the expectation \( \mathbb{E}[F(x, \xi)] \) is equal to the probability \( \Pr(\xi \leq x) \), i.e., \( f(x) = \mathbb{E}[F(x, \xi)] \) is the cumulative distribution function (cdf) of \( \xi \). The assumption (i) means here that for any given \( x \), probability of the event \( \{ x = \xi \} \) is zero, i.e., that the cdf of \( \xi \) is continuous at \( x \). In this example, the sample average function \( \hat{f}_N(\cdot) \) is just the empirical cdf of the considered random sample. The fact that the empirical cdf converges to its true counterpart uniformly on \( \mathbb{R} \) w.p.1 is known as the Glivenko-Cantelli Theorem. In fact the Glivenko-Cantelli Theorem states that the uniform convergence holds even if the corresponding cdf is discontinuous.

The analysis simplifies further if for a.e. \( \xi \in \Xi \) the function \( F(\cdot, \xi) \) is convex, i.e., the random function \( F(x, \xi) \) is convex. We can view \( \hat{f}_N(x) = f_N(x, \omega) \) as a sequence of random functions defined on a common probability space \((\Omega, \mathcal{F}, P)\). Recall definition 7.25 of epiconvergence of extended real valued functions. We say that functions \( \hat{f}_N \) epiconverge to \( f \) w.p.1, written \( \hat{f}_N \rightharpoonup f \) w.p.1, if for a.e. \( \omega \in \Omega \) the functions \( \hat{f}_N(\cdot, \omega) \) epiconverge to \( f(\cdot) \). In the following theorem we assume that function \( F(x, \xi) : \mathbb{R}^n \times \Xi \to \mathbb{R} \) is an extended real valued function, i.e., can take values \( \pm \infty \).

**Theorem 7.49.** Suppose that for almost every \( \xi \in \Xi \) the function \( F(\cdot, \xi) \) is an extended real valued convex function, the expected value function \( f(\cdot) \) is lower semicontinuous and its domain, \( \text{dom} f, \) has a nonempty interior, and the pointwise LLN holds. Then \( \hat{f}_N \rightharpoonup f \) w.p.1.

**Proof.** It follows from the assumed convexity of \( F(\cdot, \xi) \) that the function \( f(\cdot) \) is convex and that w.p.1 the functions \( \hat{f}_N(\cdot) \) are convex. Let us choose a countable and dense subset \( D \) of \( \mathbb{R}^n \). By the pointwise LLN we have that for any \( x \in D, \hat{f}_N(x) \) converges to \( f(x) \) w.p.1 as \( N \to \infty \). This means that there exists a set \( \Upsilon_x \subset \Omega \) of \( P \)-measure zero such that for any \( \omega \in \Omega \setminus \Upsilon_x, \hat{f}_N(x, \omega) \) tends to \( f(x) \) as \( N \to \infty \). Consider the set \( \Upsilon := \cup_{x \in D} \Upsilon_x \). Since the set \( D \) is countable and \( P(\Upsilon_x) = 0 \) for every \( x \in D \), we have that \( P(\Upsilon) = 0 \). We also have that for any \( \omega \in \Omega \setminus \Upsilon, \hat{f}_N(x, \omega) \) converges to \( f(x) \), as \( N \to \infty \), pointwise on \( D \). It follows then by Theorem 7.27 that \( \hat{f}_N(\cdot, \omega) \rightharpoonup f(\cdot) \) for any \( \omega \in \Omega \setminus \Upsilon \). That is, \( \hat{f}_N(\cdot) \rightharpoonup f(\cdot) \) w.p.1. \( \square \)

We also have the following result. It can be proved in a way similar to the proof of the above theorem by using Theorem 7.27.

**Theorem 7.50.** Suppose that the random function \( F(x, \xi) \) is convex and let \( X \) be a compact subset of \( \mathbb{R}^n \). Suppose that the expectation function \( f(x) \) is finite valued on a neighborhood
of $X$ and that the pointwise LLN holds for every $x$ in a neighborhood of $X$. Then $\hat{f}_N(x)$ converges to $f(x)$ w.p.1 uniformly on $X$.

It is worthwhile to note that in some cases the pointwise LLN can be verified by ad hoc methods, and hence the above epi-convergence and uniform LLN for convex random functions can be applied, without the assumption of independence.

For iid random samples we have the following version of epi-convergence LLN.

**Theorem 7.51.** Suppose that: (a) the function $F(x, \xi)$ is random lower semicontinuous, (b) for every $\bar{x} \in \mathbb{R}^n$ there exists a neighborhood $V$ of $\bar{x}$ and $P$-integrable function $h : \Xi \to \mathbb{R}$ such that $F(x, \xi) \geq h(\xi)$ for all $x \in V$ and a.e. $\xi \in \Xi$, (c) the sample is iid. Then $\hat{f}_N \to f$ w.p.1.

**Proof.** Because of the assumption (b) and since $F(x, \cdot)$ is measurable, we have that $f(x)$ is well defined and $f(x) > -\infty$. We need to verify that conditions (i) and (ii) of Definition 7.25 hold w.p.1 at an arbitrary point $\bar{x} \in \mathbb{R}^n$. By the (strong) LLN we have that $\hat{f}_N(\bar{x}) \to f(\bar{x})$ w.p.1. This implies that condition (ii) of Definition 7.25 holds w.p.1 (simply take all points of the requires sequence to be equal to $\bar{x}$).

In order to verify condition (i) we proceed as follows. For some sequence $\gamma_k \downarrow 0$ of positive numbers consider $V_k := \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq \gamma_k\}$, and let

$$
\iota_k(\xi) := \inf_{x \in V_k} F(x, \xi), \quad k \in \mathbb{N}.
$$

Note that $\iota_k(\xi)$ is measurable and for all $k$ large enough $\iota_k(\xi) \geq h(\xi)$ for a.e. $\xi$. By Fatou’s lemma and lower semicontinuity of $F(\cdot, \xi)$

$$
\liminf_{k \to \infty} E[\iota_k(\xi)] \geq E[\liminf_{k \to \infty} \iota_k(\xi)] \geq E[F(\bar{x}, \xi)] = f(\bar{x}). \quad (7.139)
$$

Consequently for any $\varepsilon > 0$ there is $\bar{k} \in \mathbb{N}$ such that

$$
E[\iota_k(\xi)] \geq f(\bar{x}) - \varepsilon, \quad \forall k \geq \bar{k}. \quad (7.140)
$$

We also have that for every $k \in \mathbb{N}$,

$$
\inf_{x \in V_k} \hat{f}_N(x) \geq \frac{1}{N} \sum_{i=1}^{N} \inf_{x \in V_k} F(x, \xi^i), \quad (7.141)
$$

and by the strong LLN that

$$
\frac{1}{N} \sum_{i=1}^{N} \iota_k(\xi^i) \to E[\iota_k(\xi)] \quad \text{w.p.1.} \quad (7.142)
$$

It follows that for any $\varepsilon > 0$ the following inequality holds w.p.1 for all $k$ large enough

$$
\inf_{x \in V_k} \hat{f}_N(x) \geq f(\bar{x}) - \varepsilon. \quad (7.143)
$$
Consequently for any sequence \( x_N \) converging to \( x \) we have \( \operatorname{w.p.1} \) that
\[
\lim_{N \to \infty} \inf f_N(x_N) \geq f(x) - \varepsilon. \tag{7.144}
\]
Since \( \varepsilon > 0 \) is arbitrary, this proves the required condition (i).

**Uniform LLN for Derivatives**

Let us discuss now uniform LLN for derivatives of the sample average function. By Theorem 7.44 we have that, under the corresponding assumptions \((A1),(A2)\) and \((A4)\), the expectation function is differentiable at the point \( x_0 \) and the derivatives can be taken inside the expectation, i.e., formula (7.118) holds. Now if we assume that the expectation function is well defined and finite valued, \( \nabla_x F(\cdot, \xi) \) is continuous on \( X \) for \( \operatorname{a.e.} \xi \in \Xi \), and \( \| \nabla_x F(x, \xi) \|, \ x \in X, \) is dominated by an integrable function, then the assumptions \((A1),(A2)\) and \((A4)\) hold and by Theorem 7.48 we obtain that \( f(x) \) is continuously differentiable on \( X \) and \( \nabla f_N(x) \) converges to \( \nabla f(x) \) \( \operatorname{w.p.1} \) uniformly on \( X \). However, in many interesting applications the function \( F(\cdot, \xi) \) is not everywhere differentiable for any \( \xi \in \Xi \), and yet the expectation function is smooth. Such simple example of \( F(x, \xi) := |x - \xi| \) was discussed after Remark 26, page 373.

**Theorem 7.52.** Let \( U \subset \mathbb{R}^n \) be an open set, \( X \) be a nonempty compact subset of \( U \) and \( F : U \times \Xi \to \mathbb{R} \) be a random function. Suppose that: (i) \( \{ F(x, \xi) \} \) is dominated by an integrable function, (ii) there exists an integrable function \( C(\xi) \) such that
\[
|F(x', \xi) - F(x, \xi)| \leq C(\xi) \| x' - x \| \quad \text{a.e. } \xi \in \Xi, \ \forall x, x' \in U. \tag{7.145}
\]
(iii) for every \( x \in X \) the function \( F(\cdot, \xi) \) is continuously differentiable at \( x \) \( \operatorname{w.p.1} \). Then the following holds: (a) the expectation function \( f(x) \) is finite valued and continuously differentiable on \( X \), (b) for all \( x \in X \) the corresponding derivatives can be taken inside the integral, i.e.,
\[
\nabla f(x) = \mathbb{E} [\nabla_x F(x, \xi)], \tag{7.146}
\]
(c) Clarke generalized gradient \( \partial^c f_N(x) \) converges to \( \nabla f(x) \) \( \operatorname{w.p.1} \) uniformly on \( X \), i.e.,
\[
\lim_{N \to \infty} \sup_{x \in X} \mathbb{D} \left( \partial^c f_N(x), \{ \nabla f(x) \} \right) = 0 \quad \text{w.p.1.} \tag{7.147}
\]

**Proof.** Assumptions (i) and (ii) imply that the expectation function \( f(x) \) is finite valued for all \( x \in U \). Note that assumption (ii) is basically the same as assumption (A2) and, of course, assumption (iii) implies assumption (A4) of Theorem 7.44. Consequently, it follows by Theorem 7.44 that \( f(\cdot) \) is differentiable at every point \( x \in X \) and the interchangeability formula (7.146) holds. Moreover, it follows from (7.145) that \( \| \nabla_x F(x, \xi) \| \leq C(\xi) \) for \( \operatorname{a.e.} \xi \) and all \( x \in U \) where \( \nabla_x F(x, \xi) \) is defined. Hence by assumption (iii) and the Lebesgue Dominated Convergence Theorem, we have that for any sequence \( x_k \) in \( U \) converging to a point \( x \in X \) it follows that
\[
\lim_{k \to \infty} \nabla f(x_k) = \mathbb{E} \left[ \lim_{k \to \infty} \nabla_x F(x_k, \xi) \right] = \mathbb{E} [\nabla_x F(x, \xi) - \nabla f(x)].
\]
We obtain that \( f(\cdot) \) is continuously differentiable on \( X \).

The assertion (c) can be proved by following the same steps as in the proof of Theorem 7.48. That is, consider a point \( \bar{x} \in X \), a sequence \( V_k \) of shrinking neighborhoods of \( \bar{x} \) and

\[
\delta_k(\xi) := \sup_{x \in V_k^*} \| \nabla_x F(x, \xi) - \nabla_x F(\bar{x}, \xi) \|.
\]

Here \( V_k^* \) denotes the set of points of \( V_k \) where \( F(\cdot, \xi) \) is differentiable. By assumption (iii) we have that \( \delta_k(\xi) \to 0 \) for a.e. \( \xi \). Also

\[
\delta_k(\xi) \leq \| \nabla_x F(\bar{x}, \xi) \| + \sup_{x \in V_k^*} \| \nabla_x F(x, \xi) \| \leq 2C(\xi),
\]

and hence \( \delta_k(\xi), k \in \mathbb{N} \), are dominated by the integrable function \( 2C(\xi) \). Consequently

\[
\lim_{k \to \infty} E[\delta_k(\xi)] = E\left[ \lim_{k \to \infty} \delta_k(\xi) \right] = 0,
\]

and the remainder of the proof can be completed in the same way as the proof of Theorem 7.48 using compactness arguments.

### 7.2.6 Law of Large Numbers for Random Sets and Subdifferentials

Consider a measurable multifunction \( A : \Omega \to \mathbb{R}^n \). Assume that \( A \) is compact valued, i.e., \( A(\omega) \) is a nonempty compact subset of \( \mathbb{R}^n \) for every \( \omega \in \Omega \). Let us denote by \( \mathcal{C}_n \) the space of nonempty compact subsets of \( \mathbb{R}^n \). Equipped with the Hausdorff distance between two sets \( A, B \in \mathcal{C}_n \), the space \( \mathcal{C}_n \) becomes a metric space. We equip \( \mathcal{C}_n \) with the sigma algebra \( B \) of its Borel subsets (generated by the family of closed subsets of \( \mathcal{C}_n \)). This makes \( (\mathcal{C}_n, B) \) a sample (measurable) space. Of course, we can view the multifunction \( A \) as a mapping from \( \Omega \) into \( \mathcal{C}_n \). We have that the multifunction \( A : \Omega \to \mathbb{R}^n \) is measurable iff the corresponding mapping \( A : \Omega \to \mathcal{C}_n \) is measurable.

We say \( A_i : \Omega \to \mathcal{C}_n, i \in \mathbb{N} \), is an iid sequence of realizations of \( A \) if each \( A_i(\omega) \) has the same probability distribution on \( (\mathcal{C}_n, B) \) as \( A(\omega) \), and \( A_i, i \in \mathbb{N} \), are independent. We have the following (strong) LLN for an iid sequence of random sets.

**Theorem 7.53 (Artstein-Vitale).** Let \( A_i, i \in \mathbb{N} \), be an iid sequence of realizations of a measurable mapping \( A : \Omega \to \mathcal{C}_n \) such that \( E[\| A(\omega) \|] < \infty \). Then

\[
N^{-1}(A_1 + \ldots + A_N) \to E[\text{conv}(A)] \quad \text{w.p.1 as } N \to \infty,
\]

where the convergence is understood in the sense of the Hausdorff metric.

In order to understand the above result let us make the following observations. There is a one-to-one correspondence between convex sets \( A \in \mathcal{C}_n \) and finite valued convex positively homogeneous functions on \( \mathbb{R}^n \), defined by \( A \mapsto s_A \), where \( s_A(h) := \sup_{z \in \text{conv}(A)} z^T h \) is the support function of \( A \). Note that for any two convex sets \( A, B \in \mathcal{C}_n \) we have that
7.2. Probability

\[ s_{A_n}(\cdot) = s_A(\cdot) + s_{B_n}(\cdot), \quad A \subset B \text{ iff } s_A(\cdot) \leq s_{B_n}(\cdot). \]
Consequently, for convex sets \( A_1, A_2 \in \mathcal{C}_n \) and \( B_r := \{ x : \| x \| \leq r \}, \) \( r \geq 0, \) we have

\[ D(A_1, A_2) = \inf \{ r \geq 0 : A_1 \subset A_2 + B_r \} \quad (7.149) \]

and

\[ \inf \{ r \geq 0 : A_1 \subset A_2 + B_r \} = \inf \{ r \geq 0 : s_{A_1}(\cdot) \leq s_{A_2}(\cdot) + s_{B_r}(\cdot) \}. \quad (7.150) \]

Moreover, \( s_{B_r}(h) = \sup_{\| z \| \leq r} z^T h = r \| h \|^*, \) where \( \| \cdot \|^* \) is the dual of the norm \( \| \cdot \|. \) We obtain that

\[ \mathbb{E}(A_1, A_2) = \sup_{\| h \|^* \leq 1} \left| s_{A_1}(h) - s_{A_2}(h) \right|. \quad (7.151) \]

It follows that if the multifunction \( A(\omega) \) is compact and convex valued, then the convergence assertion (7.148) is equivalent to:

\[ \sup_{\| h \|^* \leq 1} \left| N^{-1} \sum_{i=1}^{N} s_{A_i}(h) - \mathbb{E}[s_A(h)] \right| \to 0 \text{ w.p.1 as } N \to \infty. \quad (7.152) \]

Therefore, for compact and convex valued multifunction \( A(\omega), \) Theorem 7.53 is a direct consequence of Theorem 7.50. For general compact valued multifunctions, the averaging operation (in the left hand side of (7.148)) makes a “convexification” of the limiting set.

Consider now a random lsc convex function \( F : \mathbb{R}^n \times \Xi \to \mathbb{R} \) and the corresponding sample average function \( \hat{f}_N(x) \) based on an iid sequence \( \xi^i = \xi^i(\omega), \) \( i \in \mathbb{N} \) (see (7.129)). Recall that for any \( x \in \mathbb{R}^n, \) the multifunction \( \xi \mapsto \partial F(x, \xi) \) is measurable (see Remark 27, page 374). In a sense the following result can be viewed as a particular case of Theorem 7.53 for compact convex valued multifunctions.

**Theorem 7.54.** Let \( F : \mathbb{R}^n \times \Xi \to \mathbb{R} \) be a random lsc convex function and \( \hat{f}_N(x) \) be the corresponding sample average function based on an iid sequence \( \xi^i. \) Suppose that the expectation function \( f(x) \) is well defined and finite valued in a neighborhood of a point \( \bar{x} \in \mathbb{R}^n. \) Then

\[ \mathbb{E}(\partial \hat{f}_N(\bar{x}), \partial f(\bar{x})) \to 0 \text{ w.p.1 as } N \to \infty. \quad (7.153) \]

**Proof.** By Theorem 7.46 we have that \( f(x) \) is directionally differentiable at \( \bar{x} \) and

\[ f'(\bar{x}, h) = \mathbb{E}\left[ F_\xi'_{\bar{x}}(\bar{x}, h) \right]. \quad (7.154) \]

Note that since \( f(\cdot) \) is finite valued near \( \bar{x}, \) the directional derivative \( f'(\bar{x}, \cdot) \) is finite valued as well. We also have that

\[ \hat{f}_N(\bar{x}, h) = N^{-1} \sum_{i=1}^{N} F_{\xi^i}'_{\bar{x}}(\bar{x}, h). \quad (7.155) \]

Therefore, by the LLN it follows that \( \hat{f}_N(\bar{x}, \cdot) \) converges to \( f'(\bar{x}, \cdot) \) pointwise w.p.1 as \( N \to \infty. \) Consequently, by Theorem 7.50 we obtain that \( \hat{f}_N(\bar{x}, \cdot) \) converges to \( f'(\bar{x}, \cdot) \)
w.p.1 uniformly on the set \( \{ h : \| h \|^* \leq 1 \} \). Since \( \hat{f}_N(x, \cdot) \) is the support function of the set \( \partial f_N(x) \), it follows by (7.151) that \( \hat{f}_N(x) \) converges (in the Hausdorff metric) w.p.1 to \( \mathbb{E}[\partial F(x, \xi)] \). It remains to note that by Theorem 7.47 we have \( \mathbb{E}[\partial F(x, \xi)] = \partial f(x) \).

The problem in trying to extend the pointwise convergence (7.153) to a uniform type of convergence is that the multifunction \( x \mapsto \partial f(x) \) is not continuous even if \( f(x) \) is convex real valued.\(^{15}\)

Let us consider now the \( \varepsilon \)-subdifferential, \( \varepsilon \geq 0 \), of a convex real valued function \( f : \mathbb{R}^n \to \mathbb{R} \), defined as

\[
\partial_\varepsilon f(x) := \{ z \in \mathbb{R}^n : f(x) - f(\bar{x}) \geq z^\top (x - \bar{x}) - \varepsilon, \ \forall x \in \mathbb{R}^n \}.
\]

Clearly for \( \varepsilon = 0 \), the \( \varepsilon \)-subdifferential coincides with the usual subdifferential (at the respective point). It is possible to show that for \( \varepsilon > 0 \) the multifunction \( x \mapsto \partial_\varepsilon f(x) \) is continuous (in the Hausdorff metric) on \( \mathbb{R}^n \).

**Theorem 7.55.** Let \( g_k : \mathbb{R}^n \to \mathbb{R} \), \( k \in \mathbb{N} \), be a sequence of convex real valued (deterministic) functions. Suppose that for every \( x \in \mathbb{R}^n \) the sequence \( g_k(x) \), \( k \in \mathbb{N} \), converges to a finite limit \( g(x) \), i.e., functions \( g_k(\cdot) \) converge pointwise to the function \( g(\cdot) \). Then the function \( g(x) \) is convex, and for any \( \varepsilon > 0 \) the \( \varepsilon \)-subdifferentials \( \partial_\varepsilon g_k(\cdot) \) converge uniformly to \( \partial_\varepsilon g(\cdot) \) on any nonempty compact set \( X \subset \mathbb{R}^n \), i.e.,

\[
\lim_{k \to \infty} \sup_{x \in X} \mathbb{H}(\partial_\varepsilon g_k(x), \partial_\varepsilon g(x)) = 0.
\]

**Proof.** Convexity of \( g(\cdot) \) means that

\[
g(tx_1 + (1 - t)x_2) \leq tg(x_1) + (1 - t)g(x_2), \ \forall x_1, x_2 \in \mathbb{R}^n, \forall t \in [0, 1].
\]

This follows from convexity of functions \( g_k(\cdot) \) by passing to the limit.

By continuity and compactness arguments we have that in order to prove (7.157) it suffices to show that if \( x_k \) is a sequence of points converging to a point \( \bar{x} \), then the Hausdorff distance \( \mathbb{H}(\partial_\varepsilon g_k(x_k), \partial_\varepsilon g(\bar{x})) \) tends to zero as \( k \to \infty \). Consider the \( \varepsilon \)-directional derivative of \( g \) at \( x \):

\[
g'_\varepsilon(x, h) := \inf_{t > 0} \frac{g(x + th) - g(x) + \varepsilon}{t}.
\]

It is known that \( g'_\varepsilon(x, \cdot) \) is the support function of the set \( \partial_\varepsilon g(x) \). Therefore, since convergence of a sequence of nonempty convex compact sets in the Hausdorff metric is equivalent to the pointwise convergence of the corresponding support functions, it suffices to show that for any given \( h \in \mathbb{R}^n \):

\[
\lim_{k \to \infty} g'_{\varepsilon k}(x_k, h) = g'_{\varepsilon}(\bar{x}, h).
\]

Let us fix \( t > 0 \). Then

\[
\lim_{k \to \infty} \sup_{k \to \infty} g'_{\varepsilon k}(x_k, h) \leq \lim_{k \to \infty} \sup_{k \to \infty} \frac{g_k(x_k + th) - g_k(x_k) + \varepsilon}{t} = \frac{g(\bar{x} + th) - g(\bar{x}) + \varepsilon}{t}.
\]

\(^{15}\)This multifunction is upper semicontinuous in the sense that if the function \( f(\cdot) \) is convex and continuous at \( \bar{x} \), then \( \lim_{x \to \bar{x}} \mathbb{H}(\partial f(x), \partial f(\bar{x})) = 0 \).
Since \( t > 0 \) was arbitrary this implies that
\[
\limsup_{k \to \infty} g_{k\varepsilon}(x_k, h) \leq g_{\varepsilon}(\bar{x}, h).
\]

Now let us suppose for a moment that the minimum of \( t^{-1}[g(\bar{x} + th) - g(\bar{x}) + \varepsilon] \), over \( t > 0 \), is attained on a bounded set \( T_\varepsilon \subset \mathbb{R}_+ \). It follows then by convexity that for \( k \) large enough, \( t^{-1}[g_k(x_k + th) - g_k(x_k) + \varepsilon] \) attains its minimum over \( t > 0 \), say at a point \( t_k \), and \( \text{dist}(t_k, T_\varepsilon) \to 0 \). Consequently
\[
\liminf_{k \to \infty} g_{k\varepsilon}(x_k, h) = \liminf_{k \to \infty} \frac{g_k(x_k + t_k h) - g_k(x_k) + \varepsilon}{t_k} \geq g_{\varepsilon}(\bar{x}, h).
\]

In the general case the proof can be completed by adding the term \( \alpha\|x - \bar{x}\|^2 \), \( \alpha > 0 \), to the functions \( g_k(x) \) and \( g(x) \) and passing to the limit \( \alpha \downarrow 0 \).

The above result is deterministic. It can be easily translated into the stochastic framework as follows.

**Theorem 7.56.** Suppose that the random function \( F(x, \xi) \) is convex, and for every \( x \in \mathbb{R}^n \) the expectation \( f(x) \) is well defined and finite and the sample average \( \hat{f}_N(x) \) converges to \( f(x) \) w.p.1. Then for any \( \varepsilon > 0 \) the \( \varepsilon \)-subdifferentials \( \partial_\varepsilon \hat{f}_N(x) \) converge uniformly to \( \partial_\varepsilon f(x) \) w.p.1 on any nonempty compact set \( X \subset \mathbb{R}^n \), i.e.,
\[
\sup_{x \in X} H\left( \partial_\varepsilon \hat{f}_N(x), \partial_\varepsilon f(x) \right) \to 0 \quad \text{w.p.1 as } N \to \infty.
\]

**Proof.** In a way similar to the proof of Theorem 7.50 it can be shown that for a.e. \( \omega \in \Omega, f_N(x) \) converges pointwise to \( f(x) \) on a countable and dense subset of \( \mathbb{R}^n \). By the convexity arguments it follows that w.p.1, \( \hat{f}_N(x) \) converges pointwise to \( f(x) \) on \( \mathbb{R}^n \) (see Theorem 7.27), and hence the proof can be completed by applying Theorem 7.55.

Note that the assumption that the expectation function \( f(\cdot) \) is finite valued on \( \mathbb{R}^n \) implies that \( F(\cdot, \xi) \) is finite valued for a.e. \( \xi \), and since \( F(\cdot, \xi) \) is convex it follows that \( F(\cdot, \xi) \) is continuous. Consequently, it follows that \( F(x, \xi) \) is a Carathéodory function, and hence is random lower semicontinuous. Note also that the equality \( \partial_\varepsilon \hat{f}_N(x) = N^{-1} \sum_{i=1}^N \partial_\varepsilon F(x, \xi^i) \) holds for \( \varepsilon = 0 \) (by the Moreau-Rockafellar Theorem), but does not hold for \( \varepsilon > 0 \) and \( N > 1 \).

### 7.2.7 Delta Method

In this section we discuss the so-called Delta method approach to asymptotic analysis of stochastic problems. Let \( Z_k, k \in \mathbb{N} \), be a sequence of random variables converging in distribution to a random variable \( Z \), denoted \( Z_k \xrightarrow{D} Z \).

**Remark 29.** It can be noted that convergence in distribution does not imply convergence of the expected values \( \mathbb{E}[Z_k] \) to \( \mathbb{E}[Z] \), as \( k \to \infty \), even if all these expected values are finite.
This implication holds under the additional condition that \( Z_k \) are uniformly integrable, that is
\[
\lim_{c \to \infty} \sup_{k \in \mathbb{N}} \mathbb{E}[Z_k(c)] = 0, \tag{7.160}
\]
where \( Z_k(c) := |Z_k| \) if \( |Z_k| \geq c \), and \( Z_k(c) := 0 \) otherwise. A simple sufficient condition ensuring uniform integrability, and hence the implication that \( Z_k \overset{\mathcal{D}}{\to} Z \) implies \( \mathbb{E}[Z_k] \to \mathbb{E}[Z] \), is the following: there exists \( \varepsilon > 0 \) such that \( \sup_{k \in \mathbb{N}} \mathbb{E}[|Z_k|^{1+\varepsilon}] < \infty \). Indeed, for \( c > 0 \) we have
\[
\mathbb{E}[Z_k(c)] \leq c^{-\varepsilon} \mathbb{E}[Z_k(c)^{1+\varepsilon}] \leq c^{-\varepsilon} \mathbb{E}[|Z_k|^{1+\varepsilon}],
\]
from which the assertion follows.

**Remark 30 (stochastic order notation).** The notation \( O_p(\cdot) \) and \( o_p(\cdot) \) stands for a probabilistic analogue of the usual order notation \( O(\cdot) \) and \( o(\cdot) \), respectively. That is, let \( X_k \) and \( Z_k \) be sequences of random variables. It is written that \( Z_k = O_p(X_k) \) if for any \( \varepsilon > 0 \) there exists \( c > 0 \) such that \( \Pr(|Z_k/X_k| > c) \leq \varepsilon \) for all \( k \in \mathbb{N} \). It is written that \( Z_k = o_p(X_k) \) if for any \( \varepsilon > 0 \) it holds that \( \lim_{k \to \infty} \Pr(|Z_k/X_k| > \varepsilon) = 0 \). Usually this is used with the sequence \( X_k \) being deterministic. In particular, the notation \( Z_k = O_p(1) \) asserts that the sequence \( Z_k \) is bounded in probability, and \( Z_k = o_p(1) \) means that the sequence \( Z_k \) converges in probability to zero.

**First Order Delta Method**

In order to investigate asymptotic properties of sample estimators it will be convenient to use the Delta method, which we are going to discuss now. Let \( Y_N \in \mathbb{R}^d \) be a sequence of random vectors, converging in probability to a vector \( \mu \in \mathbb{R}^d \). Suppose that there exists a sequence \( \tau_N \) of positive numbers, tending to infinity, such that \( \tau_N(Y_N - \mu) \) converges in distribution to a random vector \( Y \), i.e., \( \tau_N(Y_N - \mu) \overset{\mathcal{D}}{\to} Y \). Let \( G : \mathbb{R}^d \to \mathbb{R}^m \) be a vector valued function, differentiable at \( \mu \). That is
\[
G(y) - G(\mu) = J(y - \mu) + r(y), \tag{7.161}
\]
where \( J := \nabla G(\mu) \) is the \( m \times d \) Jacobian matrix of \( G \) at \( \mu \), and the remainder \( r(y) \) is of order \( o(||y - \mu||) \), i.e., \( r(y)/||y - \mu|| \to 0 \) as \( y \to \mu \). It follows from (7.161) that
\[
\tau_N [G(Y_N) - G(\mu)] = J \tau_N(Y_N - \mu) + \tau_N r(Y_N). \tag{7.162}
\]
Since \( \tau_N(Y_N - \mu) \) converges in distribution, it is bounded in probability, and hence \( ||Y_N - \mu|| \) is of stochastic order \( O_p(\tau_N^{-1}) \). It follows that
\[
r(Y_N) = o(||Y_N - \mu||) = o_p(\tau_N^{-1}),
\]
and hence \( \tau_N r(Y_N) \) converges in probability to zero. Consequently we obtain by (7.162) that
\[
\tau_N [G(Y_N) - G(\mu)] \overset{\mathcal{D}}{\to} JY. \tag{7.163}
\]
This formula is routinely employed in multivariate analysis and is known as the (finite dimensional) Delta Theorem. In particular, suppose that \( N^{1/2}(Y_N - \mu) \) converges in distribution to a (multivariate) normal distribution with zero mean vector and covariance matrix \( \Sigma \), written \( N^{1/2}(Y_N - \mu) \overset{D}{\rightarrow} N(0, \Sigma) \). Often, this can be ensured by an application of the Central Limit Theorem. Then it follows by (7.163) that
\[
N^{1/2}[G(Y_N) - G(\mu)] \overset{D}{\rightarrow} N(0, J \Sigma J^T).
\]
(7.164)

We need to extend this method in several directions. The random functions \( \hat{f}_N(\cdot) \) can be viewed as random elements in an appropriate functional space. This motivates us to extend formula (7.163) to a Banach space setting. Let \( B_1 \) and \( B_2 \) be two Banach spaces, and \( G : B_1 \rightarrow B_2 \) be a mapping. Suppose that \( G \) is directionally differentiable at a considered point \( \mu \in B_1 \), i.e., the limit
\[
G'_\mu(d) := \lim_{t \downarrow 0} \frac{G(\mu + td) - G(\mu)}{t}
\]
exists for all \( d \in B_1 \). If, in addition, the directional derivative \( G'_\mu : B_1 \rightarrow B_2 \) is linear and continuous, then it is said that \( G \) is Gâteaux differentiable at \( \mu \). Note that, in any case, the directional derivative \( G'_\mu(\cdot) \) is positively homogeneous, that is \( G'_\mu(\alpha d) = \alpha G'_\mu(d) \) for any \( \alpha \geq 0 \) and \( d \in B_1 \).

It follows from (7.165) that
\[
G(\mu + d) - G(\mu) = G'_\mu(d) + r(d)
\]
with the remainder \( r(d) \) being “small” along any fixed direction \( d \), i.e., \( r(td)/t \rightarrow 0 \) as \( t \downarrow 0 \). This property is not sufficient, however, to neglect the remainder term in the corresponding asymptotic expansion and we need a stronger notion of directional differentiability. It is said that \( G \) is directionally differentiable at \( \mu \) in the sense of Hadamard if the directional derivative \( G'_\mu(d) \) exists for all \( d \in B_1 \) and, moreover,
\[
G'_\mu(d) = \lim_{t \downarrow 0} \frac{G(\mu + td) - G(\mu)}{t}
\]
(7.166)

**Proposition 7.57.** Let \( B_1 \) and \( B_2 \) be Banach spaces, \( G : B_1 \rightarrow B_2 \) and \( \mu \in B_1 \). Then the following holds. (i) If \( G(\cdot) \) is Hadamard directionally differentiable at \( \mu \), then the directional derivative \( G'_\mu(\cdot) \) is continuous. (ii) If \( G(\cdot) \) is Lipschitz continuous in a neighborhood of \( \mu \) and directionally differentiable at \( \mu \), then \( G(\cdot) \) is Hadamard directionally differentiable at \( \mu \).

The above properties can be proved in a way similar to the proof of Theorem 7.2. We also have the following chain rule.

**Proposition 7.58 (Chain Rule).** Let \( B_1, B_2 \) and \( B_3 \) be Banach spaces, and \( G : B_1 \rightarrow B_2 \) and \( F : B_2 \rightarrow B_3 \) be mappings. Suppose that \( G \) is directionally differentiable at a point \( \mu \in B_1 \) and \( F \) is Hadamard directionally differentiable at \( \eta := G(\mu) \). Then the composite mapping \( F \circ G : B_1 \rightarrow B_3 \) is directionally differentiable at \( \mu \) and
\[
(F \circ G)'(\mu, d) = F'(\eta, G'(\mu, d)), \quad \forall d \in B_1.
\]
(7.167)
Proof. Since $G$ is directionally differentiable at $\mu$, we have for $t \geq 0$ and $d \in B_1$ that

$$G(\mu + td) = G(\mu) + tG'(\mu, d) + o(t).$$

Since $F$ is Hadamard directionally differentiable at $\eta := G(\mu)$, it follows that

$$F(G(\mu + td)) = F(G(\mu) + tG'(\mu, d) + o(t)) = F(\eta) + tF'(\eta, G'(\mu, d)) + o(t).$$

This implies that $F \circ G$ is directionally differentiable at $\mu$ and formula (7.167) holds. \qed

Now let $B_1$ and $B_2$ be equipped with their Borel $\sigma$-algebras $B_1$ and $B_2$, respectively. An $\mathcal{F}$-measurable mapping from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into $B_1$ is called a random element of $B_1$. Consider a sequence $X_N$ of random elements of $B_1$. It is said that $X_N$ converges in distribution (weakly) to a random element $Y$ of $B_1$, and denoted $X_N \overset{D}{\rightarrow} Y$, if the expected values $\mathbb{E}[f(X_N)]$ converge to $\mathbb{E}[f(Y)]$, as $N \rightarrow \infty$, for any bounded and continuous function $f : B_1 \rightarrow \mathbb{R}$. Let us formulate now the first version of the Delta Theorem. Recall that a Banach space is said to be separable if it has a countable dense subset.

**Theorem 7.59 (Delta Theorem).** Let $B_1$ and $B_2$ be Banach spaces, equipped with their Borel $\sigma$-algebras, $Y_N$ a sequence of random elements of $B_2$, $G : B_1 \rightarrow B_2$ be a mapping, and $\tau_N$ be a sequence of positive numbers tending to infinity as $N \rightarrow \infty$. Suppose that the space $B_1$ is separable, the mapping $G$ is Hadamard directionally differentiable at a point $\mu \in B_1$, and the sequence $X_N := \tau_N(Y_N - \mu)$ converges in distribution to a random element $Y$ of $B_1$. Then

$$\tau_N [G(Y_N) - G(\mu)] \overset{D}{\rightarrow} G'(\mu),$$

and

$$\tau_N [G(Y_N) - G(\mu)] = G'(\mu)(X_N) + o_p(1).$$

Note that, because of the Hadamard directional differentiability of $G$, the mapping $G'(\mu) : B_1 \rightarrow B_2$ is continuous, and hence is measurable with respect to the Borel $\sigma$-algebras of $B_1$ and $B_2$. The above infinite dimensional version of the Delta Theorem can be proved easily by using the following Skorohod-Dudley almost sure representation theorem.

**Theorem 7.60 (Representation Theorem).** Suppose that a sequence of random elements $X_N$, of a separable Banach space $B$, converges in distribution to a random element $Y$. Then there exists a sequence $X'_N$, $Y'$, defined on a single probability space, such that $X'_N \overset{D}{\rightarrow} X_N$, for all $N$, $Y' \overset{D}{\rightarrow} Y$ and $X'_N \rightarrow Y'$ with probability one.

Here $Y' \overset{D}{\rightarrow} Y$ means that the probability measures induced by $Y'$ and $Y$ coincide.

Proof. [of Theorem 7.59] Consider the sequence $X_N := \tau_N(Y_N - \mu)$ of random elements of $B_1$. By the Representation Theorem, there exists a sequence $X'_N$, $Y'$, defined on a single probability space, such that $X'_N \overset{D}{\rightarrow} X_N$, $Y' \overset{D}{\rightarrow} Y$ and $X'_N \rightarrow Y'$ w.p.1. Consequently for
Let us now formulate the second version of the Delta Theorem where the mapping $G$ is restricted to a subset $K$ of the space $B_1$. We say that $G$ is Hadamard directionally differentiable at a point $\mu$ tangentially to the set $K$ if for any sequence $d_N$ of the form $d_N := (y_N - \mu)/t_N$, where $y_N \in K$ and $t_N \downarrow 0$, and such that $d_N \to d$, the following limit exists
\[
G'_\mu(d) = \lim_{N \to \infty} \frac{G(\mu + t_N d_N) - G(\mu)}{t_N}. \tag{7.172}
\]
Equivalently the above condition (7.172) can be written in the form
\[
G'_\mu(d) = \lim_{d' \to d} \frac{G(\mu + t d') - G(\mu)}{t}, \tag{7.173}
\]
where the notation $d' \to d$ means that $d' \to d$ and $\mu + t d' \in K$.

Since $y_N \in K$, and hence $\mu + t_N d_N \in K$, the mapping $G$ needs only to be defined on the set $K$. Recall that the contingent (Bouligand) cone to $K$ at $\mu$, denoted $T_K(\mu)$, is formed by vectors $d \in B$ such that there exist sequences $d_N \to d$ and $t_N \downarrow 0$ such that $\mu + t_N d_N \in K$. Note that $T_K(\mu)$ is nonempty only if $\mu$ belongs to the topological closure of the set $K$. If the set $K$ is convex, then the contingent cone $T_K(\mu)$ coincides with the corresponding tangent cone. By the above definitions we have that $G'_\mu(\cdot)$ is defined on the set $T_K(\mu)$. The following “tangential” version of the Delta Theorem can be easily proved in a way similar to the proof of Theorem 7.59.

**Theorem 7.61 (Delta Theorem).** Let $B_1$ and $B_2$ be Banach spaces, $K$ be a subset of $B_1$, $G : K \to B_2$ be a mapping, and $Y_N$ be a sequence of random elements of $B_2$. Suppose that: (i) the space $B_1$ is separable, (ii) the mapping $G$ is Hadamard directionally differentiable at a point $\mu$ tangentially to the set $K$, (iii) for some sequence $\tau_N$ of positive numbers tending to infinity, the sequence $X_N := \tau_N (Y_N - \mu)$ converges in distribution to a random element $Y$, (iv) $Y_N \in K$, with probability one, for all $N$ large enough. Then
\[
\tau_N [G(Y_N) - G(\mu)] \overset{D}{\to} G'_\mu(Y). \tag{7.174}
\]
Moreover, if the set $K$ is convex, then equation (7.169) holds.

Note that it follows from the assumptions (iii) and (iv) that the distribution of $Y$ is concentrated on the contingent cone $T_K(\mu)$, and hence the distribution of $G'_\mu(Y)$ is well defined.

**Second Order Delta Theorem**

Our third variant of the Delta Theorem deals with a second order expansion of the mapping $G$. That is, suppose that $G$ is directionally differentiable at $\mu$ and define

$$G''_\mu(d) := \lim_{t \to 0} \frac{G(\mu + td) - G(\mu) - tG'_\mu(d)}{\frac{1}{2}t^2}. \tag{7.175}$$

If the mapping $G$ is twice continuously differentiable, then this second order directional derivative $G''_\mu(d)$ coincides with the second order term in the Taylor expansion of $G(\mu + d)$. The above definition of $G''_\mu(d)$ makes sense for directionally differentiable mappings. However, in interesting applications, where it is possible to calculate $G''_\mu(d)$, the mapping $G$ is actually (Gâteaux) differentiable. We say that $G$ is second order Hadamard directionally differentiable at $\mu$ if the second order directional derivative $G''_\mu(d)$, defined in (7.175), exists for all $d \in B_1$. We say that $G$ is second order Hadamard directionally differentiable at $\mu$ tangentially to a set $K \subset B_1$ if for all $d \in T_K(\mu)$ the limit

$$G''_\mu(d) = \lim_{t \to 0} \frac{G(\mu + td) - G(\mu) - tG'_\mu(d)}{\frac{1}{2}t^2} \tag{7.176}$$

exists.

Note that if $G$ is first and second order Hadamard directionally differentiable at $\mu$ tangentially to $K$, then $G'_\mu(\cdot)$ and $G''_\mu(\cdot)$ are continuous on $T_K(\mu)$, and that $G''_\mu(\alpha d) = \alpha^2 G''_\mu(d)$ for any $\alpha \geq 0$ and $d \in T_K(\mu)$.

**Theorem 7.62 (second-order Delta Theorem).** Let $B_1$ and $B_2$ be Banach spaces, $K$ be a convex subset of $B_1$, $Y_N$ be a sequence of random elements of $B_2$, $G : K \to B_2$ be a mapping, and $\tau_N$ be a sequence of positive numbers tending to infinity as $N \to \infty$. Suppose that: (i) the space $B_1$ is separable, (ii) $G$ is first and second order Hadamard directionally differentiable at $\mu$ tangentially to the set $K$, (iii) the sequence $X_N := \tau_N(Y_N - \mu)$ converges in distribution to a random element $Y$ of $B_1$, (iv) $Y_N$ is w.p.1 for $N$ large enough. Then

$$\tau_N^2 \left[G(Y_N) - G(\mu) - G'_\mu(Y_N - \mu)\right] \overset{D}{\to} \frac{1}{2} G''_\mu(Y), \tag{7.177}$$

and

$$G(Y_N) = G(\mu) + G'_\mu(Y_N - \mu) + \frac{1}{2} G''_\mu(Y_N - \mu) + o_p(\tau_N^{-2}). \tag{7.178}$$

**Proof.** Let $X'_N$, $Y'$ and $Y'_N$ be elements as in the proof of Theorem 7.59. Recall that their existence is guaranteed by the Representation Theorem. Then by the definition of $G''_\mu$ we
have
\[ \tau_n^2 [G(Y_N') - G(\mu) - \tau_n^{-1} G'(X_N')] \rightarrow \frac{1}{2} G''(Y') \quad \text{w.p.1.} \]

Note that \( G''(\cdot) \) is defined on \( T_K(\mu) \) and, since \( K \) is convex, \( X_N' = \tau_N(Y_N' - \mu) \in T_K(\mu) \). Therefore the expression in the left hand side of the above limit is well defined. Since convergence w.p.1 implies convergence in distribution, formula (7.177) follows. Since \( G''(\cdot) \) is continuous on \( T_K(\mu) \), and, by convexity of \( K \), \( Y_N' - \mu \in T_K(\mu) \) w.p.1, we have that \( \tau_n^2 G''(Y_N' - \mu) \rightarrow G''(Y') \) w.p.1. Since convergence w.p.1 implies convergence in probability, formula (7.178) then follows. \( \square \)

### 7.2.8 Exponential Bounds of the Large Deviations Theory

Consider an iid sequence \( Y_1, \ldots, Y_N \) of replications of a real valued random variable \( Y \), and let \( Z_N := \sum_{i=1}^N Y_i \) be the corresponding sample average. Then for any real numbers \( a \) and \( t > 0 \) we have that \( \Pr(Z_N \geq a) \leq e^{-ta} \mathbb{E}[e^{tZ_N}] = e^{-ta}[M(t/N)]^N \), where \( M(t) := \mathbb{E}[e^{tY}] \) is the moment generating function of \( Y \). Suppose that \( Y \) has finite mean \( \mu := \mathbb{E}[Y] \) and \( a \geq \mu \). By taking the logarithm of both sides of the above inequality, changing variables \( t' = t/N \) and minimizing over \( t' > 0 \), we obtain

\[ \frac{1}{N} \ln \Pr(Z_N \geq a) \leq -I(a), \quad (7.179) \]

where

\[ I(z) := \sup_{t \in \mathbb{R}} \{tz - \Lambda(t)\} \quad (7.180) \]

is the conjugate of the logarithmic moment generating function \( \Lambda(t) := \ln M(t) \). In the LD theory, \( I(z) \) is called the (large deviations) rate function, and the inequality (7.179) corresponds to the upper bound of Cramèr’s LD Theorem.

Note that the moment generating function \( M(\cdot) \) is convex, positive valued, \( M(0) = 1 \) and its domain \( \text{dom} M \) is a subinterval of \( \mathbb{R} \) containing zero. It follows by Theorem 7.44 that \( M(\cdot) \) is infinitely differentiable at every interior point of its domain. Moreover, if \( a := \inf(\text{dom} M) \) is finite, then \( M(\cdot) \) is right side continuous at \( a \), and similarly for the \( b := \sup(\text{dom} M) \). It follows that \( M(\cdot) \), and hence \( \Lambda(\cdot) \), are proper lower semicontinuous functions. The logarithmic moment generating function \( \Lambda(\cdot) \) is also convex. Indeed, \( \text{dom} \Lambda = \text{dom} M \) and at an interior point \( t \) of \( \text{dom} \Lambda \),

\[ \Lambda''(t) = \frac{\mathbb{E}[Y^2 e^{tY}] \mathbb{E}[e^{tY}] - \mathbb{E}[Y e^{tY}]^2}{M(t)^2}. \quad (7.181) \]

Moreover, the matrix
\[
\begin{pmatrix}
Y^2 e^{tY} & Y e^{tY} \\
Y e^{tY} & e^{tY}
\end{pmatrix}
\]
is positive semidefinite, and hence its expectation is also a positive semidefinite matrix. Consequently the determinant of the later matrix is nonnegative, i.e.,

\[ \mathbb{E}[Y^2 e^{tY}] \mathbb{E}[e^{tY}] - \mathbb{E}[Y e^{tY}]^2 \geq 0. \]
We obtain that $\Lambda''(\cdot)$ is nonnegative at every point of the interior of $\text{dom}\Lambda$, and hence $\Lambda(\cdot)$ is convex.

Note that the constraint $t > 0$ is removed in the above definition of the rate function $I(\cdot)$. This is because of the following. Consider the function $\psi(t) := ta - \Lambda(t)$. The function $\Lambda(t)$ is convex, and hence $\psi(t)$ is concave. Suppose that the moment generating function $M(\cdot)$ is finite valued at some $t > 0$. Then $M(t)$ is finite for all $t \in [0, t]$ and right side differentiable at $t = 0$. Moreover, the right side derivative of $M(t)$ at $t = 0$ is $\mu$, and hence the right side derivative of $\psi(t)$ at $t = 0$ is positive if $a > \mu$. Consequently in that case $\psi(t) > \psi(0)$ for all $t > 0$ small enough, and hence $I(a) > 0$ and the supremum in (7.180) is not changed if the constraint $t > 0$ is removed. If $a = \mu$, then the supremum in (7.180) is attained at $t = 0$ and hence $I(a) = 0$. In that case the inequality (7.179) trivially holds. Now if $M(t) = +\infty$ for all $t > 0$, then $I(a) = 0$ for any $a \geq \mu$ and the inequality (7.179) trivially holds.

For $a \leq \mu$ the upper bound (7.179) takes the form

$$\frac{1}{N} \ln \Pr(Z_N \leq a) \leq -I(a),$$

(7.182)

which of course can be written as

$$\Pr(Z_N \leq a) \leq e^{-I(a)N}$$

(7.183)

The rate function $I(z)$ is convex and has the following properties. Suppose that the random variable $Y$ has finite mean $\mu := \mathbb{E}[Y]$. Then $\Lambda'(0) = \mu$ and hence the maximum in the right hand side of (7.180) is attained at $t^* = 0$. It follows that $I(\mu) = 0$ and

$$I'(\mu) = t^*\mu - \Lambda'(t^*) = -\Lambda(0) = 0,$$

and hence $I(z)$ attains its minimum at $z = \mu$. Suppose, further, that the moment generating function $M(t)$ is finite valued for all $t$ in a neighborhood of $t = 0$. Then $\Lambda(t)$ is infinitely differentiable at $t = 0$, and $\Lambda'(0) = \mu$ and $\Lambda''(0) = \sigma^2$, where $\sigma^2 := \text{Var}[Y]$. It follows by the above discussion that in that case $I(a) > 0$ for any $a \neq \mu$. We also have then that $I'(\mu) = 0$ and $I''(\mu) = \sigma^{-2}$, and hence by Taylor’s expansion,

$$I(a) = \frac{(a - \mu)^2}{2\sigma^2} + o(|a - \mu|^2).$$

(7.184)

If $Y$ has normal distribution $N(\mu, \sigma^2)$, then its logarithmic moment generating function is $\Lambda(t) = \mu t + \sigma^2 t^2/2$. In that case

$$I(a) = \frac{(a - \mu)^2}{2\sigma^2}.$$ 

(7.185)

The constant $I(a)$ in (7.179) gives, in a sense, the best possible exponential rate at which the probability $\Pr(Z_N \geq a)$ converges to zero. This follows from the lower bound

$$\liminf_{N \to \infty} \frac{1}{N} \ln \Pr(Z_N \geq a) \geq -I(a)$$

(7.186)

of Cramér’s LD Theorem, which holds for $a \geq \mu$. 
Another, closely related, exponential type inequalities can be derived for bounded random variables.

**Proposition 7.63.** Let $Y$ be a random variable such that $a \leq Y \leq b$, for some $a, b \in \mathbb{R}$, and $\mathbb{E}[Y] = 0$. Then

\[
\mathbb{E}[e^{tY}] \leq e^{t^2(b-a)^2 / 8}, \quad \forall t \geq 0.
\]  

(7.187)

**Proof.** If $Y$ is identically zero, then (7.187) obviously holds. Therefore we can assume that $Y$ is not identically zero. Since $\mathbb{E}[Y] = 0$, it follows that $a < 0$ and $b > 0$.

Any $Y \in [a, b]$ can be represented as convex combination $Y = \tau a + (1 - \tau)b$, where $\tau = (b - Y)/(b - a)$. Since $e^y$ is a convex function, it follows that

\[
e^Y \leq \frac{b - Y}{b - a} e^a + \frac{Y - a}{b - a} e^b.
\]  

(7.188)

Taking expectation from both sides of (7.188) and using $\mathbb{E}[Y] = 0$, we obtain

\[
\mathbb{E}[e^Y] \leq \frac{b}{b - a} e^a - \frac{a}{b - a} e^b.
\]  

(7.189)

The right hand side of (7.188) can be written as $e^{g(u)}$, where $u := b - a$, $g(x) := -\alpha x + \ln(1 - \alpha + \alpha e^x)$ and $\alpha := -a/(b - a)$. Note that $\alpha > 0$ and $1 - \alpha > 0$.

Let us observe that $g(0) = g'(0) = 0$ and

\[
g''(x) = \frac{\alpha(1 - \alpha)}{(1 - \alpha)^2 e^{-x} + \alpha^2 e^x + 2\alpha(1 - \alpha)}.
\]  

(7.190)

Moreover, $(1 - \alpha)^2 e^{-x} + \alpha^2 e^x \geq 2\alpha(1 - \alpha)$, and hence $g''(x) \leq 1/4$ for any $x$. By Taylor expansion of $g(\cdot)$ at zero, we have $g(u) = u^2 g''(\tilde{u})/2$ for some $\tilde{u} \in (0, u)$. It follows that $g(u) \leq u^2/8 = (b - a)^2/8$, and hence

\[
\mathbb{E}[e^Y] \leq e^{(b-a)^2/8}.
\]  

(7.191)

Finally, (7.187) follows from (7.191) by rescaling $Y$ to $tY$ for $t \geq 0$. \qed

In particular, if $|Y| \leq b$ and $\mathbb{E}[Y] = 0$, then $| - Y| \leq b$ and $\mathbb{E}[ - Y] = 0$ as well, and hence by (7.187) we have

\[
\mathbb{E}[e^{tY}] \leq e^{t^2b^2/2}, \quad \forall t \in \mathbb{R}.
\]  

(7.192)

Let $Y$ be a (real valued) random variable supported on a bounded interval $[a, b] \subset \mathbb{R}$, and $\mu := \mathbb{E}[Y]$. Then it follows from (7.187) that the rate function of $Y - \mu$ satisfies

\[
I(\varepsilon) \geq \sup_{\varepsilon \in \mathbb{R}} \{ t\varepsilon - t^2(b - a)^2/8 \} = 2\varepsilon^2/(b - a)^2.
\]  

Together with (7.183) this implies the following. Let $Y_1, \ldots, Y_N$ be an iid sequence of realizations of $Y$ and $Z_N$ be the corresponding average. Then for $\tau > 0$ it holds that

\[
\Pr (Z_N \geq \mu + \tau) \leq e^{-2\tau^2 N/(b - a)^2}.
\]  

(7.193)
Chapter 7. Background Material

The bound (7.193) is often referred to as 
Hoeffding inequality.

In particular, let $W \sim B(n,p)$ be a random variable having Binomial distribution, i.e., $\Pr(W = k) = \binom{n}{k} p^k (1 - p)^{n-k}, k = 0, \ldots, n$. Recall that $W$ can be represented as $W = Y_1 + \ldots + Y_n$, where $Y_1, \ldots, Y_n$ is an iid sequence of Bernoulli random variables with $\Pr(Y_i = 1) = p$ and $\Pr(Y_i = 0) = 1 - p$. It follows from Hoeffding’s inequality that for a nonnegative integer $k \leq np,$

$$\Pr(W \leq k) \leq \exp \left\{ -\frac{2(np - k)^2}{n} \right\}. \quad (7.194)$$

For small $p$ it is possible to improve the above estimate as follows. For $Y \sim \text{Bernoulli}(p)$ we have

$$E[e^{tY}] = pe^t + 1 - p = 1 - p(e^t - 1).$$

By using the inequality $e^{-x} \geq 1 - x$ with $x := p(1 - e^t)$, we obtain

$$E[e^{tY}] \leq \exp[p(e^t - 1)],$$

and hence for $z > 0$,

$$I(z) := \sup_{t \in \mathbb{R}} \{tz - \ln E[e^{tY}] \} \geq \sup_{t \in \mathbb{R}} \{tz - p(e^t - 1) \} = z \ln \frac{z}{p} - z + p.$$

Moreover, since $\ln(1 + x) \geq x - x^2/2$ for $x \geq 0$, we obtain

$$I(z) \geq \frac{(z - p)^2}{2p} \text{ for } z \geq p.$$

By (7.179) it follows that

$$\Pr\left(n^{-1}W \geq z\right) \leq \exp \left\{ -n(z - p)^2/(2p) \right\}, \text{ for } z \geq p. \quad (7.195)$$

Alternatively, this can be written as

$$\Pr(W \leq k) \leq \exp \left\{ -\frac{(np - k)^2}{2pn} \right\}, \quad (7.196)$$

for a nonnegative integer $k \leq np$. The above inequality (7.196) is often called Chernoff inequality. For small $p$ it can be significantly better than Hoeffding inequality (7.194).

The above, one dimensional, LD results can be extended to multivariate and even infinite dimensional settings, and also to non iid random sequences. In particular, suppose that $Y$ is a $d$-dimensional random vector and let $\mu := E[Y]$ be its mean vector. We can associate with $Y$ its moment generating function $M(t)$, of $t \in \mathbb{R}^d$, and the rate function $I(z)$ defined in the same way as in (7.180) with the supremum taken over $t \in \mathbb{R}^d$ and $tz$ denoting the standard scalar product of vectors $t, z \in \mathbb{R}^d$. Consider a (Borel) measurable set $A \subset \mathbb{R}^d$. Then, under certain regularity conditions, the following Large Deviations Principle holds:

$$-\inf_{z \in \text{int}(A)} I(z) \leq \liminf_{N \to \infty} N^{-1} \ln \Pr(Z_N \in A) \leq \limsup_{N \to \infty} N^{-1} \ln \Pr(Z_N \in A) \leq -\inf_{z \in \text{cl}(A)} I(z), \quad (7.197)$$
where \( \text{int}(A) \) and \( \text{cl}(A) \) denote the interior and topological closure, respectively, of the set \( A \). In the above one dimensional setting, the LD principle (7.197) was derived for sets \( A := [a, +\infty) \).

We have that if \( \mu \in \text{int}(A) \) and the moment generating function \( M(t) \) is finite valued for all \( t \) in a neighborhood of \( 0 \in \mathbb{R}^d \), then \( \inf_{z \in \mathbb{R}^d \setminus \text{int}(A)} I(z) \) is positive. Moreover, if the sequence is iid, then
\[
\lim_{N \to \infty} \sup_{N} N^{-1} \ln [\Pr(Z_N \notin A)] < 0,
\]
i.e., the probability \( \Pr(Z_N \in A) = 1 - \Pr(Z_N \notin A) \) approaches one exponentially fast as \( N \) tends to infinity.

Finally, let us derive the following useful result.

**Proposition 7.64.** Let \( \xi^1, \xi^2, \ldots \) be a sequence of iid random variables (vectors), \( \sigma_t > 0 \), \( t = 1, \ldots \), be a sequence of deterministic numbers and \( \phi_t = \phi_t(\xi_t) \) be (measurable) functions of \( \xi_t = (\xi^1, \ldots, \xi^t) \) such that
\[
\mathbb{E} [\phi_t | \xi_{t-1}] = 0 \quad \text{and} \quad \mathbb{E} [\exp\{\phi_t^2/\sigma_t^2\} | \xi_{t-1}] \leq \exp\{1\} \text{ w.p.1.}
\]
(7.199)

Then for any \( \Theta > 0 \):
\[
\Pr \left\{ \sum_{t=1}^N \phi_t \geq \Theta \sqrt{\sum_{t=1}^N \sigma_t^2} \right\} \leq \exp\{-\Theta^2/3\}.
\]
(7.200)

**Proof.** Let us set \( \tilde{\phi}_t := \phi_t/\sigma_t \). By condition (7.199) we have that \( \mathbb{E} [\tilde{\phi}_t | \xi_{t-1}] = 0 \) and \( \mathbb{E} [\exp\{\tilde{\phi}_t^2\} | \xi_{t-1}] \leq \exp\{1\} \text{ w.p.1.} \) By Jensen inequality it follows that for any \( a \geq 1 \):
\[
\mathbb{E} [\exp\{a\tilde{\phi}_t^2\} | \xi_{t-1}] \leq \left( \mathbb{E} [\exp\{\tilde{\phi}_t^2\} | \xi_{t-1}] \right)^a \leq \exp\{a\}.
\]
We also have that \( \exp\{x\} \leq x + \exp\{9x^2/16\} \) for all \( x \) (this inequality can be verified by direct calculations), and hence for any \( \lambda \in [0, 4/3] \):
\[
\mathbb{E} [\exp\{\lambda \tilde{\phi}_t\} | \xi_{t-1}] \leq \mathbb{E} [\exp\{(9\lambda^2/16)\tilde{\phi}_t^2\} | \xi_{t-1}] \leq \exp\{9\lambda^2/16\}.
\]
(7.201)

Moreover, we have that \( \lambda x \leq \frac{3}{8} \lambda^2 + \frac{2}{5} x^2 \) for any \( \lambda \) and \( x \), and hence
\[
\mathbb{E} [\exp\{\lambda \tilde{\phi}_t\} | \xi_{t-1}] \leq \exp\{3\lambda^2/8\} \mathbb{E} [\exp\{2\tilde{\phi}_t^2/3\} | \xi_{t-1}] \leq \exp\{2/3 \leq 3\lambda^2/8\}.
\]

Combining the latter inequality with (7.201), we get
\[
\mathbb{E} [\exp\{\lambda \tilde{\phi}_t\} | \xi_{t-1}] \leq \exp\{3\lambda^2/4\}, \quad \forall \lambda \geq 0.
\]

Going back to \( \phi_t \), the above inequality reads
\[
\mathbb{E} [\exp\{\gamma \phi_t\} | \xi_{t-1}] \leq \exp\{3\gamma^2 \sigma_t^2/4\}, \quad \forall \gamma \geq 0.
\]
(7.202)
Now, since $\phi_t$ is a deterministic function of $\xi_i\xi'$ and using (7.202), we obtain for any $\gamma \geq 0$:

$$
E \left[ \exp \left\{ \gamma \sum_{t=1}^{T} \phi_t \right\} \right] = E \left[ \exp \left\{ \gamma \sum_{t=1}^{T} \phi_t \right\} E \left( \exp(\gamma \phi_t) | \xi_{t-1} \right) \right] \\
\leq \exp\{3\gamma^2 \sigma^2 / 4\} E \left[ \exp\{\gamma \sum_{t=1}^{T} \phi_t \} \right],
$$

and hence

$$
E \left[ \exp \left\{ \gamma \sum_{t=1}^{T} \phi_t \right\} \right] \leq \exp\{3\gamma^2 \sum_{t=1}^{T} \sigma^2 / 4\}. \tag{7.203}
$$

By Chebyshev’s inequality, we have for $\gamma > 0$ and $\Theta$:

$$
Pr \left\{ \sum_{t=1}^{T} \phi_t \geq \Theta \sqrt{\sum_{t=1}^{T} \sigma_t^2} \right\} = Pr \left\{ \exp \left[ \gamma \sum_{t=1}^{T} \phi_t \right] \geq \exp \left[ \gamma \Theta \sqrt{\sum_{t=1}^{T} \sigma_t^2} \right] \right\} \\
\leq \exp \left\{ -\gamma \Theta \sqrt{\sum_{t=1}^{T} \sigma_t^2} \right\} E \left\{ \exp \left[ \gamma \sum_{t=1}^{T} \phi_t \right] \right\}.
$$

Together with (7.203) this implies for $\Theta \geq 0$:

$$
Pr \left\{ \sum_{t=1}^{T} \phi_t \geq \Theta \sqrt{\sum_{t=1}^{T} \sigma_t^2} \right\} \leq \inf_{\gamma > 0} \exp \left\{ \frac{3}{4} \gamma^2 \sum_{t=1}^{T} \sigma_t^2 - \gamma \Theta \sqrt{\sum_{t=1}^{T} \sigma_t^2} \right\} = \exp \left\{ -\Theta^2 / 3 \right\}.
$$

This completes the proof. \qed

### 7.2.9 Uniform Exponential Bounds

Consider the setting of section 7.2.5 with a sequence $\xi_i, i \in \mathbb{N}$, of random realizations of an $d$-dimensional random vector $\xi = \xi(\omega)$, a function $F : X \times \Xi \to \mathbb{R}$ and the corresponding sample average function $f_N(x)$. We assume here that the sequence $\xi_i, i \in \mathbb{N}$, is iid, the set $X \subset \mathbb{R}^n$ is nonempty and compact and the expectation function $f(x) = E[F(x, \xi)]$ is well defined and finite valued for all $x \in X$. We discuss now uniform exponential rates of convergence of $f_N(x)$ to $f(x)$. Denote by

$$
M_x(t) := E \left[ e^{t(F(x, \xi) - f(x))} \right]
$$

the moment generating function of the random variable $F(x, \xi) - f(x)$. Let us make the following assumptions.

(C1) For every $x \in X$ the moment generating function $M_x(t)$ is finite valued for all $t$ in a neighborhood of zero.

(C2) There exists a (measurable) function $\kappa : \Xi \to \mathbb{R}_+$ such that

$$
|F(x', \xi) - F(x, \xi)| \leq \kappa(\xi) \|x' - x\| \tag{7.204}
$$

for all $\xi \in \Xi$ and all $x', x \in X$.

(C3) The moment generating function $M_{\kappa}(t) := E \left[ e^{t\kappa(\xi)} \right]$ of $\kappa(\xi)$ is finite valued for all $t$ in a neighborhood of zero.
Theorem 7.65. Suppose that conditions (C1)–(C3) hold and the set X is compact. Then for any $\varepsilon > 0$ there exist positive constants $C$ and $\beta = \beta(\varepsilon)$, independent of $N$, such that

$$ \Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq Ce^{-N\beta}. $$

(7.205)

**Proof.** By the upper bound (7.179) of Cramér’s Large Deviation Theorem we have that for any $x \in X$ and $\varepsilon > 0$ it holds that

$$ \Pr \{ \hat{f}_N(x) - f(x) \geq \varepsilon \} \leq \exp\{-NI(x)(\varepsilon)\}, $$

(7.206)

where

$$ I(x) := \sup_{t \in \mathbb{R}} \left\{ xt - \ln M_x(t) \right\} $$

(7.207)

is the LD rate function of random variable $F(x, \xi) - f(x)$. Similarly

$$ \Pr \{ \hat{f}_N(x) - f(x) \leq -\varepsilon \} \leq \exp\{-NI(x)(-\varepsilon)\}, $$

and hence

$$ \Pr \left\{ \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq \exp\{-NI(x)(\varepsilon)\} + \exp\{-NI(x)(-\varepsilon)\}. $$

(7.208)

By assumption (C1) we have that both $I(x)(\varepsilon)$ and $I(x)(-\varepsilon)$ are positive for every $x \in X$.

For a $\nu > 0$, let $\bar{x}_1, \ldots, \bar{x}_K \in X$ be such that for every $x \in X$ there exists $\bar{x}_i$, $i \in \{1, \ldots, K\}$, such that $\|x - \bar{x}_i\| \leq \nu$, i.e., $\{\bar{x}_1, \ldots, \bar{x}_K\}$ is a $\nu$-net in $X$. We can choose this net in such a way that

$$ K \leq \lceil gD/\nu \rceil^n, $$

(7.209)

where

$$ D := \sup_{x, x' \in X} \|x' - x\| $$

is the diameter of $X$ and $g$ is a constant depending on the chosen norm $\|\cdot\|$. By (7.204) we have that

$$ |f(x') - f(x)| \leq L \|x' - x\|, $$

(7.210)

where $L := \mathbb{E}[\kappa(\xi)]$ is finite by assumption (C3). Moreover,

$$ |\hat{f}_N(x') - \hat{f}_N(x)| \leq \kappa_N \|x' - x\|, $$

(7.211)

where $\kappa_N := N^{-1} \sum_{i=1}^{N} \kappa(\xi_i)$. Again, because of condition (C3), by Cramér’s LD Theorem we have that for any $L' > L$ there is a constant $\ell > 0$ such that

$$ \Pr \{ \kappa_N \geq L' \} \leq \exp\{-N\ell\}. $$

(7.212)

Consider

$$ Z_i := \hat{f}_N(\bar{x}_i) - f(\bar{x}_i), \quad i = 1, \ldots, K. $$

We have that the event $\{\max_{1 \leq i \leq K} |Z_i| \geq \varepsilon\}$ is equal to the union of the events $\{|Z_i| \geq \varepsilon\}$, $i = 1, \ldots, K$, and hence

$$ \Pr \{\max_{1 \leq i \leq K} |Z_i| \geq \varepsilon\} \leq \sum_{i=1}^{K} \Pr (|Z_i| \geq \varepsilon). $$
Together with (7.208) this implies that
\[
\Pr \left\{ \max_{1 \leq i \leq K} \left| \hat{f}_N(x_i) - f(x) \right| \geq \varepsilon \right\} \leq 2 \sum_{i=1}^{K} \exp \left\{ -N[I_{x_{i}}(\varepsilon) \land I_{x_{i}}(-\varepsilon)] \right\}. \tag{7.213}
\]

For an \( x \in X \) let \( i(x) \in \arg \min_{1 \leq i \leq K} \| x - x_i \| \). By construction of the \( \nu \)-net we have that \( \| x - x_{i(x)} \| \leq \nu \) for every \( x \in X \). Then
\[
\left| \hat{f}_N(x) - f(x) \right| \leq \left| \hat{f}_N(x) - \hat{f}_N(x_{i(x)}) \right| + \left| \hat{f}_N(x_{i(x)}) - f(x_{i(x)}) \right| + \left| f(x_{i(x)}) - f(x) \right| \leq \kappa_N \nu + \left| \hat{f}_N(x_{i(x)}) - f(x_{i(x)}) \right| + L \nu.
\]

Let us take now a \( \nu \)-net with such \( \nu \) that \( L \nu = \varepsilon/4 \), i.e., \( \nu := \varepsilon/(4L) \). Then
\[
\Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq \Pr \left\{ \kappa_N \nu + \max_{1 \leq i \leq K} \left| \hat{f}_N(x_i) - f(x_i) \right| \geq 3\varepsilon/4 \right\}.
\]

Moreover, we have that
\[
\Pr \{ \kappa_N \nu \geq \varepsilon/2 \} \leq \exp \{ -N\ell \},
\]
where \( \ell \) is a positive constant specified in (7.212) for \( L' := 2L \). Consequently
\[
\Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq \exp \{ -N\ell \} + \Pr \left\{ \max_{1 \leq i \leq K} \left| \hat{f}_N(x_i) - f(x_i) \right| \geq \varepsilon/4 \right\} \leq \exp \{ -N\ell \} + 2 \sum_{i=1}^{K} \exp \left\{ -N[I_{x_i}(\varepsilon/4) \land I_{x_i}(-\varepsilon/4)] \right\}. \tag{7.214}
\]

Since the above choice of the \( \nu \)-net does not depend on the sample (although it depends on \( \varepsilon \)), and both \( I_{x_i}(\varepsilon/4) \) and \( I_{x_i}(-\varepsilon/4) \) are positive, \( i = 1, \ldots, K \), we obtain that (7.214) implies (7.205), and hence completes the proof. \( \square \)

In the convex case the (Lipschitz continuity) condition (C2) holds, in a sense, automatically. That is, we have the following result.

**Theorem 7.66.** Let \( U \subset \mathbb{R}^n \) be a convex open set. Suppose that: (i) for a.e. \( \xi \in \Xi \) the function \( F(\xi) : U \to \mathbb{R} \) is convex, and (ii) for every \( x \in U \) the moment generating function \( M_t(x) \) is finite valued for all \( t \) in a neighborhood of zero. Then for every compact set \( X \subset U \) and \( \varepsilon > 0 \) there exist positive constants \( C \) and \( \beta = \beta(\varepsilon) \), independent of \( N \), such that
\[
\Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq Ce^{-N\beta^2}. \tag{7.215}
\]

**Proof.** We have here that the expectation function \( f(x) \) is convex and finite valued for all \( x \in U \). Let \( X \) be a (nonempty) compact subset of \( U \). For \( \gamma \geq 0 \) consider the set
\[
X_{\gamma} := \{ x \in \mathbb{R}^n : \text{dist}(x, X) \leq \gamma \}.
\]

Since the set \( U \) is open, we can choose \( \gamma > 0 \) such that \( X_{\gamma} \subset U \). The set \( X_{\gamma} \) is compact and by convexity of \( f(\cdot) \) we have that \( f(\cdot) \) is continuous and hence is bounded on \( X_{\gamma} \). That
is, there is constant $c > 0$ such that $|f(x)| \leq c$ for all $x \in X_\gamma$. Also by convexity of $f(\cdot)$ we have for any $\tau \in [0, 1], x \in X$ and $y \in \mathbb{R}^n$ such that $x - y/\tau \in U$:

$$f(x) = f\left(\frac{1}{1+\tau}(x+y) + \frac{\tau}{1+\tau}(x - y/\tau)\right) \leq \frac{1}{1+\tau}f(x+y) + \frac{\tau}{1+\tau}f(x - y/\tau).$$

It follows that if $\|y\| \leq \tau \gamma$, and hence $x - y/\tau \in X_\gamma$, then

$$f(x+y) \geq (1 + \tau)f(x) - \tau f(x - y/\tau) \geq f(x) - 2\tau c.$$  

(7.216)

Now we proceed similar to the proof of Theorem 7.65. Let $\varepsilon > 0$ and $\nu > 0$, and $ar{x}_1, \ldots, \bar{x}_K \in U$ be a $\nu$-net for $X_\gamma$. As in the proof of Theorem 7.65, this $\nu$-net will be independent of $\varepsilon$, but not on the random sample $\xi^1, \ldots, \xi^N$. Consider the event

$$A_N := \left\{ \max_{1 \leq i \leq K} |\hat{f}_N(\bar{x}_i) - f(\bar{x}_i)| \leq \varepsilon \right\}.$$

By (7.206) and (7.208) we have similar to (7.213) that $\Pr(A_N) \geq 1 - \alpha_N$, where

$$\alpha_N := 2 \sum_{i=1}^K \exp\left\{ - N[I_{\varepsilon}\nu] \cap I_{\varepsilon}[-\varepsilon] \right\}.$$

Consider a point $x \in X$ and let $\mathcal{I} \subset \{1, ..., K\}$ be such index set that $x$ is a convex combination of points $\bar{x}_i, i \in \mathcal{I}$, i.e., $x = \sum_{i \in \mathcal{I}} t_i \bar{x}_i$, for some positive numbers $t_i$ summing up to one. Moreover, let $\mathcal{I}$ be such that $\|x - \bar{x}_i\| \leq a \nu$, for all $i \in \mathcal{I}$, where $a > 0$ is a constant independent of $x$ and the net. By convexity of $\hat{f}_N(\cdot)$ we have that $\hat{f}_N(x) \leq \sum_{i \in \mathcal{I}} t_i \hat{f}_N(\bar{x}_i)$. It follows that the event $A_N$ is included in the event $\left\{ \hat{f}_N(x) \leq \sum_{i \in \mathcal{I}} t_i f(\bar{x}_i) + \varepsilon \right\}$. By (7.216) we also have that

$$f(x) \geq f(\bar{x}_i) - 2\tau c, \ \forall i \in \mathcal{I},$$

provided that $a \nu \leq \tau \gamma$. Setting $\tau := \varepsilon/(2\varepsilon)$, we obtain that the event $A_N$ is included in the event $B_x := \left\{ \hat{f}_N(x) \leq f(x) + 2\varepsilon \right\}$, provided that $\nu \leq O(1)\varepsilon$. It follows that the event $A_N$ is included in the event $\bigcap_{x \in X} B_x$, and hence

$$\Pr\left\{ \sup_{x \in X} \left( \hat{f}_N(x) - f(x) \right) \leq 2\varepsilon \right\} = \Pr\left\{ \bigcap_{x \in X} B_x \right\} \geq \Pr\{A_N\} \geq 1 - \alpha_N,$$  

(7.217)

provided that $\nu \leq O(1)\varepsilon$.

In order to derive the converse to (7.217) estimate let us observe that by convexity of $\hat{f}_N(\cdot)$ we have with probability at least $1 - \alpha_N$ that $\sup_{x \in X_\gamma} \hat{f}_N(x) \leq c + \varepsilon$. Also by using (7.216) we have with probability at least $1 - \alpha_N$ that $\inf_{x \in X_\gamma} \hat{f}_N(x) \geq -(c + \varepsilon)$, provided that $\nu \leq O(1)\varepsilon$. That is, with probability at least $1 - 2\alpha_N$ we have that

$$\sup_{x \in X_\gamma} |\hat{f}_N(x)| \leq c + \varepsilon,$$

\[\text{16Recall that } O(1) \text{ denotes a generic constant, here } O(1) = \gamma/(2ca).\]
provided that \( \nu \leq O(1)\varepsilon \). We can now proceed in the same way as above to show that

\[
\Pr \left\{ \sup_{x \in X} \left( f(x) - \hat{f}_N(x) \right) \leq 2\varepsilon \right\} \geq 1 - 3\alpha_N. \tag{7.218}
\]

Since, by the condition (ii), \( I_{\varepsilon}(\varepsilon) \) and \( I_{\varepsilon}(-\varepsilon) \) are positive, this completes the proof. \( \square \)

Now let us strengthen condition (C1) to the following condition:

(C4) There exists constant \( \sigma > 0 \) such that for any \( x \in X \), the following inequality holds:

\[
M_x(t) \leq \exp \left\{ \sigma^2 t^2 / 2 \right\}, \quad \forall t \in \mathbb{R}. \tag{7.219}
\]

It follows from condition (7.219) that \( \ln M_x(t) \leq \sigma^2 t^2 / 2 \), and hence

\[
I_x(z) \geq z^2 / 2\sigma^2, \quad \forall z \in \mathbb{R}. \tag{7.220}
\]

Consequently, inequality (7.214) implies

\[
\Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq \exp\left\{ -N\ell \right\} + 2K \exp\left\{ -N\varepsilon^2 / 32\sigma^2 \right\}, \tag{7.221}
\]

where \( \ell \) is a constant specified in (7.212) with \( L' := 2L, K = [\varrho D / \nu]^n, \nu = \varepsilon / (4L) \) and hence

\[
K = \left[\frac{4\varrho DL}{\varepsilon}\right]^n. \tag{7.222}
\]

If we assume further that the Lipschitz constant in (7.204) does not depend on \( \xi \), i.e., \( \kappa(\xi) \equiv L \), then the first term in the right hand side of (7.221) can be omitted. Therefore we obtain the following result.

**Theorem 7.67.** Suppose that conditions (C2)–(C4) hold and that the set \( X \) has finite diameter \( D \). Then

\[
\Pr \left\{ \sup_{x \in X} \left| \hat{f}_N(x) - f(x) \right| \geq \varepsilon \right\} \leq \exp\left\{ -N\ell \right\} + 2 \left[\frac{4\varrho DL}{\varepsilon}\right]^n \exp\left\{ -N\varepsilon^2 / 32\sigma^2 \right\}. \tag{7.223}
\]

Moreover, if \( \kappa(\xi) \equiv L \) in condition (C2), then condition (C3) holds automatically and the term \( \exp\{-N\ell\} \) in the right hand side of (7.223) can be omitted.

As it was shown in the proof of Theorem 7.66, in the convex case estimates of the form (7.223), with different constants, can be obtained without assuming the (Lipschitz continuity) condition (C2).

\( ^{17} \) Recall that if random variable \( F(x, \xi) - f(x) \) has normal distribution with variance \( \sigma^2 \), then its moment generating function is equal to the right hand side of (7.219), and hence the inequalities (7.219) and (7.220) hold as equalities.
7.2. Probability

Exponential Convergence of Generalized Gradients

The above results can be also applied to establishing rates of convergence of directional derivatives and generalized gradients (subdifferentials) of $f_N(x)$ at a given point $x \in X$. Consider the following condition.

(C5) For a.e. $\xi \in \Xi$, the function $F_{\xi}(\cdot) = F(\cdot, \xi)$ is directionally differentiable at a point $\bar{x} \in X$.

Consider the expected value function $f(x) = \mathbb{E}[F(x, \xi)] = \int_{\Xi} F(x, \xi) dP(\xi)$. Suppose that $f(\bar{x})$ is finite and condition (C2) holds with the respective Lipschitz constant $\kappa(\xi)$ being $P$-integrable, i.e., $\mathbb{E}[\kappa(\xi)] < +\infty$. Then it follows that $f(x)$ is finite valued and Lipschitz continuous on $X$ with Lipschitz constant $\mathbb{E}[\kappa(\xi)]$. Moreover, the following result for Clarke generalized gradient of $f(x)$ holds (cf., [38, Theorem 2.7.2]).

Theorem 7.68. Suppose that condition (C2) holds with $\mathbb{E}[\kappa(\xi)] < +\infty$, and let $\bar{x}$ be an interior point of the set $X$ such that $f(\bar{x})$ is finite. If, moreover, $F(\cdot, \xi)$ is Clarke-regular at $\bar{x}$ for a.e. $\xi \in \Xi$, then $f$ is Clarke-regular at $\bar{x}$ and

$$\partial^\circ f(\bar{x}) = \int_{\Xi} \partial^\circ F(\bar{x}, \xi) dP(\xi), \tag{7.224}$$

where Clarke generalized gradient $\partial^\circ F(\bar{x}, \xi)$ is taken with respect to $x$.

The above result can be extended to an infinite dimensional setting with the set $X$ being a subset of a separable Banach space $\mathcal{X}$. Formula (7.224) can be interpreted in the following way. For every $\gamma \in \partial^\circ f(\bar{x})$, there exists a measurable selection $\Gamma(\xi) \in \partial^\circ F(\bar{x}, \xi)$ such that for every $\nu \in \mathcal{X}^*$, the function $\langle \nu, \Gamma(\cdot) \rangle$ is integrable and

$$\langle \nu, \gamma \rangle = \int_{\Xi} \langle \nu, \Gamma(\xi) \rangle dP(\xi).$$

In this way, $\gamma$ can be considered as an integral of a measurable selection from $\partial^\circ F(\bar{x}, \cdot)$.

Theorem 7.69. Let $\bar{x}$ be an interior point of the set $X$. Suppose that $f(\bar{x})$ is finite and conditions (C2)–(C3) and (C5) hold. Then for any $\varepsilon > 0$ there exist positive constants $C$ and $\beta = \beta(\varepsilon)$, independent of $N$, such that\footnote{By $S^{n-1} := \{d \in \mathbb{R}^n : \|d\| = 1\}$ we denote the unit sphere taken with respect to a norm $\|\cdot\|$ on $\mathbb{R}^n$.}

$$\Pr\left\{ \sup_{d \in S^{n-1}} \left| \hat{f}_N(\bar{x}, d) - f(\bar{x}, d) \right| > \varepsilon \right\} \leq C e^{-N\beta}. \tag{7.225}$$

Moreover, suppose that for a.e. $\xi \in \Xi$ the function $F(\cdot, \xi)$ is Clarke-regular at $\bar{x}$. Then

$$\Pr\left\{ \mathbb{H} \left( \partial^\circ f_N(\bar{x}), \partial^\circ f(\bar{x}) \right) > \varepsilon \right\} \leq C e^{-N\beta}. \tag{7.226}$$

Furthermore, if in condition (C2) $\kappa(\xi) \equiv L$ is constant, then

$$\Pr\left\{ \mathbb{H} \left( \partial^\circ f_N(\bar{x}), \partial^\circ f(\bar{x}) \right) > \varepsilon \right\} \leq 2 \left[ \frac{4L}{\varepsilon} \right]^n \exp \left\{ - \frac{N\varepsilon^2}{12L^2} \right\}. \tag{7.227}$$
Proof. Since \( f(\bar{x}) \) is finite, conditions (C2)-(C3) and (C5) imply that \( f(\cdot) \) is finite valued and Lipschitz continuous in a neighborhood of \( \bar{x} \), \( f(\cdot) \) is directionally differentiable at \( \bar{x} \), its directional derivative \( f'(\bar{x}, \cdot) \) is Lipschitz continuous, and \( f'(\bar{x}, \cdot) = \mathbb{E}[\eta(\cdot, \xi)] \), where \( \eta(\cdot, \xi) := f'(\bar{x}, \cdot) \) (see Theorem 7.44). We also have here that \( \hat{f}_N(\bar{x}, \cdot) = \hat{\eta}_N(\cdot), \) where

\[
\hat{\eta}_N(d) := \frac{1}{N} \sum_{i=1}^{N} \eta(d, \xi_i), \quad d \in \mathbb{R}^n, \tag{7.228}
\]

and \( \mathbb{E}[\hat{\eta}_N(d)] = f'(\bar{x}, d), \) for all \( d \in \mathbb{R}^n \). Moreover, conditions (C2) and (C5) imply that \( \eta(\cdot, \xi) \) is Lipschitz continuous on \( \mathbb{R}^n \), with Lipschitz constant \( \kappa(\xi) \), and in particular that \( |\eta(d, \xi)| \leq \kappa(\xi) \|d\| \) for any \( d \in \mathbb{R}^n \) and \( \xi \in \Xi \). Hence together with condition (C3) this implies that, for every \( d \in \mathbb{R}^n \), the moment generating function of \( \eta(d, \xi) \) is finite valued in a neighborhood of zero.

Consequently, the estimate (7.225) follows directly from Theorem 7.65. If \( F_\xi(\cdot) \) is Clarke-regular for a.e. \( \xi \in \Xi \), then \( \hat{f}_N(\cdot) \) is also Clarke-regular and

\[
\partial^\circ \hat{f}_N(\bar{x}) = N^{-1} \sum_{i=1}^{N} \partial^\circ F_\xi(\bar{x}).
\]

By applying (7.225) together with equation (7.151) for sets \( A_1 := \partial^\circ \hat{f}_N(\bar{x}) \) and \( A_2 := \partial^\circ f(\bar{x}) \), we obtain (7.226).

Now if \( \kappa(\xi) \equiv L \) is constant, then \( \eta(\cdot, \xi) \) is Lipschitz continuous on \( \mathbb{R}^n \), with Lipschitz constant \( L \), and \( \|\eta(d, \xi)\| \leq L \) for every \( d \in S^{n-1} \) and \( \xi \in \Xi \). Consequently, for any \( d \in S^{n-1} \) and \( \xi \in \Xi \) we have that \( |\eta(d, \xi) - \mathbb{E}[\eta(d, \xi)]| \leq 2L \), and hence for every \( d \in S^{n-1} \) the moment generating function \( M_d(t) \) of \( \eta(d, \xi) - \mathbb{E}[\eta(d, \xi)] \) is bounded \( M_d(t) \leq \exp\{2t^2L^2\} \), for all \( t \in \mathbb{R} \) (see (7.192)). It follows by Theorem 7.67 that

\[
\Pr \left\{ \sup_{d \in S^{n-1}} |\hat{f}_N(\bar{x}, d) - f'(\bar{x}, d)| > \varepsilon \right\} \leq 2 \left[ \frac{4gL}{\varepsilon} \right]^n \exp \left\{ -\frac{N\varepsilon^2}{128L^2} \right\}, \tag{7.229}
\]

and hence (7.227) follows. \( \square \)

### 7.3 Elements of Functional Analysis

A linear space \( Z \) equipped with a norm \( \| \cdot \| \) is said to be a **Banach space** if it is complete, i.e., every Cauchy sequence in \( Z \) has a limit. Let \( Z \) be a Banach space. Unless stated otherwise, all topological statements (like convergence, continuity, lower continuity etc) will be made with respect to the norm topology of \( Z \).

The space of all linear continuous functionals \( \zeta : Z \to \mathbb{R} \) forms the dual space of \( Z \) and is denoted \( Z^* \). For \( \zeta \in Z^* \) and \( z \in Z \) we denote \( \langle \zeta, z \rangle := \zeta(z) \) and view it as a scalar product on \( Z^* \times Z \). The space \( Z^* \), equipped with the dual norm

\[
\|\zeta\|_* := \sup_{\|z\| \leq 1} \langle \zeta, z \rangle, \tag{7.230}
\]

is also a Banach space. Consider the dual \( Z^{**} \) of the space \( Z^* \). There is a natural embedding of \( Z \) into \( Z^{**} \) given by identifying \( z \in Z \) with linear functional \( \langle \cdot, z \rangle \) on \( Z^* \). In that
7.3. Elements of Functional Analysis

sense, \( Z \) can be considered as a subspace of \( Z^{**} \). It is said that Banach space \( Z \) is reflexive if \( Z \) coincides with \( Z^{**} \).

It follows from the definition of the dual norm that

\[
||\langle \zeta, z \rangle|| \leq ||\zeta|| \cdot ||z||, \quad z \in Z, \ \zeta \in Z^*.
\]  

(7.231)

Also to every \( z \in Z \) corresponds set

\[
\mathcal{S}_z := \text{arg max} \{ \langle \zeta, z \rangle : \zeta \in Z^*, ||\zeta|| \leq 1 \}.
\]  

(7.232)

The set \( \mathcal{S}_z \) is always nonempty and will be referred to as the set of contact points of \( z \in Z \). Every point of \( \mathcal{S}_z \) will be called a contact point of \( z \).

An important class of Banach spaces are \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) spaces, where \( (\Omega, \mathcal{F}) \) is a sample space, equipped with sigma algebra \( \mathcal{F} \) and probability measure \( \mathcal{P} \), and \( p \in [1, +\infty) \). The space \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) consists of all \( \mathcal{F} \)-measurable functions \( \phi : \Omega \to \mathbb{R} \) such that \( \int_{\Omega} |\phi(\omega)|^p \ d\mathcal{P}(\omega) < +\infty \). More precisely, an element of \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) is a class of such functions \( \phi(\omega) \) which may differ from each other on sets of \( \mathcal{P} \)-measure zero. Equipped with the norm

\[
||\phi||_p := \left( \int_{\Omega} |\phi(\omega)|^p \ d\mathcal{P}(\omega) \right)^{1/p},
\]  

(7.233)

\( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) becomes a Banach space.

We also use the space \( \mathcal{L}_\infty(\Omega, \mathcal{F}, \mathcal{P}) \) of functions (or rather classes of functions which may differ on sets of \( \mathcal{P} \)-measure zero) \( \phi : \Omega \to \mathbb{R} \) which are \( \mathcal{F} \)-measurable and essentially bounded. A function \( \phi \) is said to be essentially bounded if its sup-norm

\[
||\phi||_\infty := \text{ess sup}_{\omega \in \Omega} |\phi(\omega)|
\]  

(7.234)

is finite, where

\[
\text{ess sup}_{\omega \in \Omega} |\phi(\omega)| := \inf \left\{ \sup_{\omega \in \Omega} |\psi(\omega)| : \phi(\omega) = \psi(\omega) \ \text{a.e.} \ \omega \in \Omega \right\}.
\]

In particular, suppose that the set \( \Omega := \{\omega_1, \ldots, \omega_K\} \) is finite, and let \( \mathcal{F} \) be the sigma algebra of all subsets of \( \Omega \) and \( p_1, \ldots, p_K \) be (positive) probabilities of the corresponding elementary events. In that case every element \( z \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) can be viewed as a finite dimensional vector \( (z(\omega_1), \ldots, z(\omega_K)) \), and \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) can be identified with the space \( \mathbb{R}^K \) equipped with the corresponding norm

\[
||z||_p := \left( \sum_{k=1}^K p_k |z(\omega_k)|^p \right)^{1/p}.
\]  

(7.235)

We also use spaces \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}; \mathbb{R}^m) \), with \( p \in [1, +\infty] \). For \( p \in [1, +\infty] \) this space is formed by all \( \mathcal{F} \)-measurable functions (mappings) \( \psi : \Omega \to \mathbb{R}^m \) such that \( \int_{\Omega} ||\psi(\omega)||^p \ d\mathcal{P}(\omega) < +\infty \), with the corresponding norm \( ||\cdot|| \) on \( \mathbb{R}^m \) being, for example, the Euclidean norm. For \( p = \infty \), the corresponding space consists of all essentially bounded functions \( \psi : \Omega \to \mathbb{R}^m \).

For \( p \in (1, +\infty) \) the dual of \( \mathcal{L}_p(\Omega, \mathcal{F}, \mathcal{P}) \) is the space \( \mathcal{L}_q(\Omega, \mathcal{F}, \mathcal{P}) \), where \( q \in (1, +\infty) \) is such that \( 1/p + 1/q = 1 \), and these spaces are reflexive. This duality is derived
by Hölder inequality
\[ \int_{\Omega} |\zeta(\omega) z(\omega)| dP(\omega) \leq \left( \int_{\Omega} |\zeta(\omega)|^q dP(\omega) \right)^{1/q} \left( \int_{\Omega} |z(\omega)|^p dP(\omega) \right)^{1/p}. \quad (7.236) \]
For points \( z \in L_p(\Omega, \mathcal{F}, P) \) and \( \zeta \in L_q(\Omega, \mathcal{F}, P) \), their scalar product is defined as
\[ \langle \zeta, z \rangle := \int_{\Omega} \zeta(\omega) z(\omega) dP(\omega). \quad (7.237) \]

The dual of \( L_1(\Omega, \mathcal{F}, P) \) is the space \( L_\infty(\Omega, \mathcal{F}, P) \), and these spaces are not reflexive.

If \( z(\omega) \) is not zero for a.e. \( \omega \in \Omega \), then the equality in (7.236) holds iff \( \zeta(\omega) \) is proportional\(^{19} \) to \( \text{sign}(z(\omega)) z(\omega)^{1/(q-1)} \). It follows that for \( p \in (1, +\infty) \), with every nonzero \( z \in L_p(\Omega, \mathcal{F}, P) \) is associated unique contact point, denoted \( \tilde{\zeta}_z \), which can be written in the form
\[ \tilde{\zeta}_z(\omega) = \frac{\text{sign}(z(\omega)) z(\omega)^{1/(q-1)}}{\|z\|_q^{1/p}}. \quad (7.238) \]
In particular, for \( p = 2 \) and \( q = 2 \) the contact point is \( \tilde{\zeta}_z = \|z\|_2^{-1} z \). Of course, if \( z = 0 \), then \( \mathcal{G}_0 = \{ \zeta \in Z^* : \|\zeta\|_* \leq 1 \} \).

For \( p = 1 \) and \( z \in L_1(\Omega, \mathcal{F}, P) \) the corresponding set of contact points can be described as follows
\[ \mathcal{G}_z = \left\{ \zeta \in L_\infty(\Omega, \mathcal{F}, P) : \begin{array}{ll} \zeta(\omega) = 1, & \text{if } z(\omega) > 0, \\ \zeta(\omega) = -1, & \text{if } z(\omega) < 0, \\ \zeta(\omega) \in [-1,1], & \text{if } z(\omega) = 0. \end{array} \right\} \quad (7.239) \]
It follows that \( \mathcal{G}_z \) is a singleton iff \( z(\omega) \neq 0 \) for a.e. \( \omega \in \Omega \).

Together with the strong (norm) topology of \( Z \) we sometimes need to consider its weak topology, which is the weakest topology in which all linear functionals \( \langle \cdot, \cdot \rangle, \zeta \in Z^* \), are continuous. The dual space \( Z^* \) can be also equipped with its weak* topology, which is the weakest topology in which all linear functionals \( \langle \cdot, z \rangle, z \in Z, \) are continuous. If the space \( Z \) is reflexive, then \( Z^* \) is also reflexive and its weak* and weak topologies do coincide. Note also that a convex subset of \( Z \) is closed in the strong topology iff it is closed in the weak topology of \( Z \).

**Theorem 7.70 (Banach-Alaoglu).** Let \( Z \) be Banach space. The closed unit ball \( \{ \zeta \in Z^* : \|\zeta\|_* \leq 1 \} \) is compact in the weak* topology of \( Z^* \).

It follows that any bounded (in the dual norm \( \| \cdot \|_* \)) and weakly* closed subset of \( Z^* \) is weakly* compact.

### 7.3.1 Conjugate Duality and Differentiability

Let \( Z \) be a Banach space, \( Z^* \) be its dual space and \( f : Z \to \mathbb{R} \) be an extended real valued function. Similar to the final dimensional case we define the **conjugate function** of \( f \) as
\[ f^*(\zeta) := \sup_{z \in Z} \{ \langle \zeta, z \rangle - f(z) \}. \quad (7.240) \]

\(^{19} \) For \( a \in \mathbb{R}, \text{sign}(a) \) is equal to 1 if \( a > 0 \), to -1 if \( a < 0 \), and to 0 if \( a = 0 \).
7.3. Elements of Functional Analysis

The conjugate function $f^* : Z^* \rightarrow \mathbb{R}$ is always convex and lower semicontinuous. The biconjugate function $f^{**} : Z \rightarrow \mathbb{R}$, i.e., the conjugate of $f^*$, is

$$f^{**}(z) := \sup_{\zeta \in Z^*} \{ \langle \zeta, z \rangle - f^*(\zeta) \}.$$  \hfill (7.241)

The basic duality theorem still holds in the considered infinite dimensional framework.

**Theorem 7.71 (Fenchel-Moreau).** Let $Z$ be a Banach space and $f : Z \rightarrow \mathbb{R}$ be a proper extended real valued convex function. Then

$$f^{**} = \text{lsc } f.$$  \hfill (7.242)

It follows from (7.242) that if $f$ is proper and convex, then $f^{**} = f$ iff $f$ is lower semicontinuous. A basic difference between finite and infinite dimensional frameworks is that in the infinite dimensional case a proper convex function can be discontinuous at an interior point of its domain. As the following result shows, for a convex proper function continuity and lower semicontinuity properties on the interior of its domain are the same.

**Proposition 7.72.** Let $Z$ be a Banach space and $f : Z \rightarrow \mathbb{R}$ be a convex lower semicontinuous function having a finite value in at least one point. Then $f$ is proper and is continuous on $\text{int}(\text{dom } f)$.

In particular, it follows from the above proposition that if $f : Z \rightarrow \mathbb{R}$ is real valued convex and lower semicontinuous, then $f$ is continuous on $Z$.

The subdifferential of a function $f : Z \rightarrow \mathbb{R}$, at a point $z_0$ such that $f(z_0)$ is finite, is defined in a way similar to the final dimensional case. That is,

$$\partial f(z_0) := \{ \zeta \in Z^* : f(z) - f(z_0) \geq \langle \zeta, z - z_0 \rangle, \forall z \in Z \}.$$  \hfill (7.243)

It is said that $f$ is subdifferentiable at $z_0$ if $\partial f(z_0)$ is nonempty. Clearly, if $f$ is subdifferentiable at some point $z_0 \in Z$, then $f$ is proper and lower semicontinuous at $z_0$. Similar to the finite dimensional case we have the following.

**Proposition 7.73.** Let $Z$ be a Banach space and $f : Z \rightarrow \mathbb{R}$ be a convex function and $z \in Z$ be such that $f^{**}(z)$ is finite. Then

$$\partial f^{**}(z) = \arg \max_{\zeta \in Z^*} \{ \langle \zeta, z \rangle - f^*(\zeta) \}.$$  \hfill (7.244)

Moreover, if $f^{**}(z) = f(z)$, then $\partial f^{**}(z) = \partial f(z)$.

**Proposition 7.74.** Let $Z$ be a Banach space, $f : Z \rightarrow \mathbb{R}$ be a convex function. Suppose that $f$ is finite valued and continuous at a point $z_0 \in Z$. Then $f$ is subdifferentiable at $z_0$, $\partial f(z_0)$ is nonempty, convex, bounded and weakly* compact subset of $Z^*$, $f$ is Hadamard directionally differentiable at $z_0$ and

$$f'(z_0, h) = \sup_{\zeta \in \partial f(z_0)} \langle \zeta, h \rangle.$$  \hfill (7.245)
Note that by the definition, every element of the subdifferential $\partial f(z_0)$ (called subgradient) is a continuous linear functional on $Z$. A linear (not necessarily continuous) functional $\ell : Z \to \mathbb{R}$ is called an algebraic subgradient of $f$ at $z_0$ if

$$f(z_0 + h) - f(z_0) \geq \ell(h), \quad \forall h \in Z.$$  \hspace{1cm} (7.246)

Of course, if the algebraic subgradient $\ell$ is also continuous, then $\ell \in \partial f(z_0)$.

**Proposition 7.75.** Let $Z$ be a Banach space and $f : Z \to \mathbb{R}$ be a proper convex function. Then the set of algebraic subgradients at any point $z_0 \in \text{int}(\text{dom} f)$ is nonempty.

**Proof.** Consider the directional derivative function $\delta(h) := f'(z_0, h)$. The directional derivative is defined here in the same way as in section 7.1.1. Since $f$ is convex we have that

$$f'(z_0, h) = \inf_{t > 0} \frac{f(z_0 + th) - f(z_0)}{t},$$

and $\delta(.)$ is convex, positively homogeneous. Moreover, since $z_0 \in \text{int}(\text{dom} f)$ and hence $f(z)$ is finite valued for all $z$ in a neighborhood of $z_0$, it follows by (7.247) that $\delta(h)$ is finite valued for all $h \in Z$. That is, $\delta(.)$ is a real valued subadditive and positively homogeneous function. Consequently, by Hahn-Banach Theorem we have that there exists a linear functional $\ell : Z \to \mathbb{R}$ such that $\delta(h) \geq \ell(h)$ for all $h \in Z$. Since $f(z_0 + h) \geq f(z_0) + \delta(h)$ for any $h \in Z$, it follows that $\ell$ is an algebraic subgradient of $f$ at $z_0$. \hfill $\square$

There is also the following version of Moreau-Rockafellar Theorem in the infinite dimensional setting.

**Theorem 7.76 (Moreau-Rockafellar).** Let $f_1, f_2 : Z \to \mathbb{R}$ be convex proper lower semicontinuous functions, $f := f_1 + f_2$ and $\bar{z} \in \text{dom}(f_1) \cap \text{dom}(f_2)$. Then

$$\partial f(\bar{z}) = \partial f_1(\bar{z}) + \partial f_2(\bar{z}),$$

provided that the following regularity condition holds

$$0 \in \text{int} \{\text{dom}(f_1) - \text{dom}(f_2)\}.$$  \hspace{1cm} (7.249)

In particular, equation (7.248) holds if $f_1$ is continuous at $\bar{z}$.

**Remark 31.** It is possible to derive the following (first order) necessary optimality condition from the above theorem. Let $S$ be a convex closed subset of $Z$ and $f : Z \to \mathbb{R}$ be a convex proper lower semicontinuous function. We have that a point $z_0 \in S$ is a minimizer of $f(z)$ over $z \in S$ iff $z_0$ is a minimizer of $\psi(z) := f(z) + \mathbb{I}_S(z)$ over $z \in Z$. The last condition is equivalent to the condition that $0 \in \partial \psi(z_0)$. Since $S$ is convex and closed, the indicator function $\mathbb{I}_S(\cdot)$ is convex lower semicontinuous, and $\partial \mathbb{I}_S(z_0) = \mathcal{N}_S(z_0)$. Therefore, we have the following.

If $z_0 \in S \cap \text{dom}(f)$ is a minimizer of $f(z)$ over $z \in S$, then

$$0 \in \partial f(z_0) + \mathcal{N}_S(z_0),$$

(7.250)
provided that $0 \in \text{int} \{\text{dom}(f) - S\}$. In particular, (7.250) holds, if $f$ is continuous at $z_0$.

It is also possible to apply the conjugate duality theory to dual problems of the form (7.33) and (7.35) in an infinite dimensional setting. That is, let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, $\psi: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ and $\vartheta(y) := \inf_{x \in \mathcal{X}} \psi(x, y)$.

**Theorem 7.77.** Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces. Suppose that the function $\psi(x, y)$ is proper convex and lower semicontinuous, and that $\vartheta(y)$ is finite. Then $\vartheta(y)$ is continuous at $y$ if for every $y$ in a neighborhood of $y$, $\vartheta(y) < +\infty$, i.e., $y \in \text{int}(\text{dom} \vartheta)$.

If $\vartheta(y)$ is continuous at $y$, then there is no duality gap between the corresponding primal and dual problems and the set of optimal solutions of the dual problem coincides with $\partial \vartheta(y)$ and is nonempty and weakly$^*$ compact.

### 7.3.2 Lattice Structure

Let $C \subset Z$ be a closed convex pointed cone. It defines an order relation on the space $Z$. That is, $z_1 \succeq z_2$ if $z_1 - z_2 \in C$. It is not difficult to verify that this order relation defines a partial order on $Z$, i.e., the following conditions hold for any $z, z', z'' \in Z$: (i) $z \succeq z$, (ii) if $z \succeq z'$ and $z' \succeq z''$, then $z \succeq z''$ (transitivity), (iii) if $z \succeq z'$ and $z' \succeq z$, then $z = z'$. This partial order relation is also compatible with the algebraic operations, i.e., the following conditions hold: (iv) if $z \succeq z'$ and $t \geq 0$, then $tz \succeq tz'$, (v) if $z' \succeq z''$ and $z \in Z$, then $z' + z \succeq z'' + z$.

It is said that $u \in Z$ is the least upper bound (or supremum) of $z, z' \in Z$, written $u = z \lor z'$, if $u \succeq z$ and $u \succeq z'$ and, moreover, if $u' \succeq z$ and $u' \succeq z'$ for some $u' \in Z$, then $u' \succeq u$. By the above property (iii) we have that if the least upper bound $z \lor z'$ exists, then it is unique. It is said that the considered partial order induces a lattice structure on $Z$ if the least upper bound $z \lor z'$ exists for any $z, z' \in Z$. Denote $z_0 := z \lor 0, z_{\omega} := (-z) \lor 0$, and $|z| := z_0 \lor z_{\omega} = z \lor (-z)$. It is said that Banach space $Z$ with lattice structure is a *Banach lattice* if $z, z' \in Z$ and $|z| \succeq |z'|$ implies $\|z\| \geq \|z'\|$.

For $p \in [1, +\infty]$, consider Banach space $Z := L_p(\Omega, F, P)$ and cone $C := L_p^+(\Omega, F, P)$, where

$$L_p^+(\Omega, F, P) := \{z \in L_p(\Omega, F, P) : z(\omega) \geq 0 \text{ for a.e. } \omega \in \Omega\}. \quad (7.251)$$

This cone $C$ is closed, convex and pointed. The corresponding partial order means that $z \succeq z'$ iff $z(\omega) \geq z'(\omega)$ for a.e. $\omega \in \Omega$. It has a lattice structure with

$$(z \lor z')(\omega) = \max\{z(\omega), z'(\omega)\},$$

and $|z(\omega)| = |z(\omega)|$. Also the property: “if $z \ni z' \ni 0$, then $\|z\| \leq \|z'\|$”, clearly holds. It follows that space $L_p(\Omega, F, P)$ with cone $L_p^+(\Omega, F, P)$ forms a Banach lattice.

**Theorem 7.78 (Klee-Nachbin-Namioka).** Let $Z$ be a Banach lattice and $\ell : Z \to \mathbb{R}$ be a linear functional. Suppose that $\ell$ is positive, i.e., $\ell(z) \geq 0$ for any $z \geq 0$. Then $\ell$ is continuous.

---

30 Recall that cone $C$ is said to be pointed if $z \in C$ and $-z \in C$ implies that $z = 0$. 


Proof. We have that linear functional $\ell$ is continuous iff it is bounded on the unit ball of $\mathcal{Z}$, i.e., iff there exists positive constant $c$ such that $|\ell(z)| \leq c\|z\|$ for all $z \in \mathcal{Z}$. First, let us show that there exists $c > 0$ such that $\ell(z) \leq c\|z\|$ for all $z \succeq 0$. Recall that $z \succeq 0$ iff $z \in \mathcal{C}$. We argue by a contradiction. Suppose that this is incorrect. Then there exists a sequence $z_k \in \mathcal{C}$ such that $\|z_k\| = 1$ and $\ell(z_k) \geq 2^k$ for all $k \in \mathbb{N}$. Consider $\bar{z} := \sum_{k=1}^{\infty} 2^{-k} z_k$. Note that $\sum_{k=1}^{n} 2^{-k} z_k$ forms a Cauchy sequence in $\mathcal{Z}$ and hence is convergent, i.e., the point $\bar{z}$ is well defined. Note also that since $\mathcal{C}$ is closed, it follows that $\sum_{k=m}^{\infty} 2^{-k} z_k \in \mathcal{C}$, and hence it follows by positivity of $\ell$ that $\ell\left(\sum_{k=m}^{\infty} 2^{-k} z_k\right) \geq 0$ for any $m \in \mathbb{N}$. Therefore we have

$$
\ell(\bar{z}) = \ell\left(\sum_{k=1}^{n} 2^{-k} z_k\right) + \ell\left(\sum_{k=n+1}^{\infty} 2^{-k} z_k\right) \geq \ell\left(\sum_{k=1}^{n} 2^{-k} z_k\right) \geq n,
$$

for any $n \in \mathbb{N}$. This gives a contradiction.

Now for any $z \in \mathcal{Z}$ we have

$$|z| = z_+ \lor z_- \succeq z_+ = |z_+|.$$

It follows that for $v = |z|$ we have that $\|v\| \geq \|z_+\|$, and similarly $\|v\| \geq \|z_-\|$. Since $z = z_+ - z_-$ and $\ell(z_-) \geq 0$ by positivity of $\ell$, it follows that

$$\ell(z) = \ell(z_+) - \ell(z_-) \leq c\|z_+\| \leq c\|z\|,$$

and similarly

$$-\ell(z) = -\ell(z_+) + \ell(z_-) \leq c\|z_-\| \leq c\|z\|.$$ 

It follows that $|\ell(z)| \leq c\|z\|$, which completes the proof. $\square$

Suppose that Banach space $\mathcal{Z}$ has a lattice structure. It is said that a function $f : \mathcal{Z} \rightarrow \mathbb{R}$ is monotone if $z \succeq z'$ implies that $f(z) \geq f(z')$.

Theorem 7.79. Let $\mathcal{Z}$ be a Banach lattice and $f : \mathcal{Z} \rightarrow \mathbb{R}$ be proper convex and monotone. Then $f(\cdot)$ is continuous and subdifferentiable on the interior of its domain.

Proof. Let $z_0 \in \text{int}(\text{dom} f)$. By Proposition 7.75, function $f$ possesses an algebraic subgradient $\ell$ at $z_0$. It follows from monotonicity of $f$ that $\ell$ is positive. Indeed, if $\ell(h) < 0$ for some $h \in \mathcal{C}$, then it follows by (7.246) that

$$f(z_0 - h) \geq f(z_0) - \ell(h) > f(z_0),$$

which contradicts monotonicity of $f$. It follows by Theorem 7.78 that $\ell$ is continuous, and hence $\ell \in \partial f(z_0)$. This shows that $f$ is subdifferentiable at every point of $\text{int}(\text{dom} f)$. This, in turn, implies that $f$ is lower semicontinuous on $\text{int}(\text{dom} f)$, and hence by Proposition 7.72 is continuous on $\text{int}(\text{dom} f)$. $\square$

The above result can be applied to any space $\mathcal{Z} := \mathcal{L}^p_p(\Omega, \mathcal{F}, P)$, $p \in [1, +\infty]$, equipped with the lattice structure induced by the cone $\mathcal{C} := \mathcal{L}^p_+(\Omega, \mathcal{F}, P)$. 
Interchangeability Principle

Let $(\Omega, \mathcal{F}, P)$ be a probability space. It is said that a linear space $\mathcal{M}$ of $\mathcal{F}$-measurable functions (mappings) $\psi : \Omega \to \mathbb{R}^m$ is decomposable if for every $\psi \in \mathcal{M}$ and $A \in \mathcal{F}$, and every bounded and $\mathcal{F}$-measurable function $\gamma : \Omega \to \mathbb{R}^m$, the space $\mathcal{M}$ also contains the function $\eta(\cdot) := 1_{\Omega \setminus A}(\cdot)\psi(\cdot) + 1_A(\cdot)\gamma(\cdot)$. For example, spaces $\mathcal{M} := L_p(\Omega, \mathcal{F}, P; \mathbb{R}^m)$, with $p \in [1, +\infty]$, are decomposable. Proof of the following theorem can be found in [181, Theorem 14.60].

**Theorem 7.80.** Let $\mathcal{M}$ be a decomposable space and $f : \mathbb{R}^m \times \Omega \to \mathbb{R}$ be a random lower semicontinuous function. Then

\[
\mathbb{E} \left[ \inf_{x \in \mathbb{R}^m} f(x, \omega) \right] = \inf_{\chi \in \mathcal{M}} \mathbb{E}[F_{\chi}],
\]

where $F_{\chi}(\omega) := f(\chi(\omega), \omega)$, provided that the right hand side of (7.252) is less than $+\infty$. Moreover, if the common value of both sides in (7.252) is not $-\infty$, then

\[
\chi \in \mathrm{argmin}_{\chi \in \mathcal{M}} \mathbb{E}[F_{\chi}] \iff \chi(\omega) \in \mathrm{argmin}_{x \in \mathbb{R}^m} f(x, \omega), \text{ for a.e. } \omega \in \Omega, \text{ and } \chi \in \mathcal{M}. \quad (7.253)
\]

Clearly the above interchangeability principle can be applied to a maximization, rather than minimization, procedure simply by replacing function $f(x, \omega)$ with $-f(x, \omega)$. For an extension of this interchangeability principle to risk measures see Proposition 6.36.

---

**Exercises**

1. Show that function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lsc iff its epigraph $\text{epi} f$ is a closed subset of $\mathbb{R}^{n+1}$.

2. Show that a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is polyhedral iff its epigraph is a convex closed polyhedron, and $f(x)$ is finite for at least one $x$.

3. Give an example of a function $f : \mathbb{R}^2 \to \mathbb{R}$ which is Gâteaux but not Fréchet differentiable.

4. Show that if $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ is Hadamard directionally differentiable at $x_0 \in \mathbb{R}^n$, then $g'(x_0, \cdot)$ is continuous and $g$ is Fréchet directionally differentiable at $x_0$. Conversely, if $g$ is Fréchet directionally differentiable at $x_0$ and $g'(x_0, \cdot)$ is continuous, then $g$ is Hadamard directionally differentiable at $x_0$.

5. Show that if $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a convex function, finite valued at a point $x_0 \in \mathbb{R}^n$, then formula (7.17) holds and $f'(x_0, \cdot)$ is convex. If, moreover, $f(\cdot)$ is finite valued in a neighborhood of $x_0$, then $f'(x_0, h)$ is finite valued for all $h \in \mathbb{R}^n$.

6. Let $s(\cdot)$ be the support function of a nonempty set $C \subset \mathbb{R}^n$. Show that the conjugate of $s(\cdot)$ is the indicator function of the set $\text{cl}(\text{conv}(C))$.

7. Let $C \subset \mathbb{R}^n$ be a closed convex set and $x \in C$. Show that the normal cone $N_C(x)$ is equal to the subdifferential of the indicator function $1_C(\cdot)$ at $x$. 

7.8. Show that if multifunction $G : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is closed valued and upper semicontinuous, then it is closed. Conversely, if $G$ is closed and the set $\text{dom} \ G$ is compact, then $G$ is upper semicontinuous.

7.9. Consider function $F(x, \omega)$ used in Theorem 7.44. Show that if $F(\cdot, \omega)$ is differentiable for a.e. $\omega$, then condition (A2) of that theorem is equivalent to the condition: there exists a neighborhood $V$ of $x_0$ such that

$$
\mathbb{E} \left[ \sup_{x \in V} \| \nabla_x F(x, \omega) \| \right] < \infty.
$$

7.10. Show that if $f(x) := \mathbb{E}|x - \xi|$, then formula (7.121) holds. Conclude that $f(\cdot)$ is differentiable at $x_0 \in \mathbb{R}$ iff $\text{Pr}(\xi = x_0) = 0$.

7.11. Verify equalities (7.149) and (7.150), and hence conclude (7.151).

7.12. Show that the estimate (7.205) of Theorem 7.65 still holds if the bound (7.204) in condition (C2) is replaced by:

$$
|F(x', \xi) - F(x, \xi)| \leq \kappa(\xi) \|x' - x\|^\gamma
$$

(7.255)

for some constant $\gamma > 0$. Show how the estimate (7.223) of Theorem 7.67 should be corrected in that case.
Chapter 8

Bibliographical Remarks

Chapter 1

The Newsvendor (sometimes called Newsboy) problem, portfolio selection and supply chain models are classical with numerous papers written on each subject. It will be far beyond this monograph to give a complete review of all relevant literature. Our main purpose in discussing these models is to introduce such basic concepts as a recourse action, probabilistic (chance) constraints, here-and-now and wait-and-see solutions, the nonanticipativity principle, dynamic programming equations etc. We give below just a few basic references.

The Newsvendor Problem is a classical model used in Inventory Management. Its origin is going back to the paper by Edgeworth [62]. In the stochastic setting studying of the Newsvendor Problem started with the classical paper by Arrow, Harris and Marchak [5]. The optimality of the basestock policy for the multi-stage inventory model was first proved in Clark and Scarf [37]. The worst case distribution approach to the Newsvendor Problem was initiated by Scarf [193], where the case when only the mean and variance of the distribution of the demand are known was analyzed. For a thorough discussion and relevant references for single and multi-stage inventory models we can refer to Zipkin [230].

Modern portfolio theory was introduced by Markowitz [125, 126]. The concept of utility function has a long history. Its origins are going back as far as to work of Daniel Bernoulli (1738). Axiomatic approach to the expected utility theory was introduced by von Neumann and Morgenstern [221].

For an introduction to Supply Chain Network Design see, e.g., Nagurney [132]. The material of section 1.5 is based on Santosos et al [192].

For a thorough discussion of robust optimization we refer to the forthcoming book by Ben-Tal, El Ghaoui and Nemirovski [15].

Chapters 2 and 3

Stochastic programming with recourse originated in the works of Beale [14], Dantzig [41] and Tintner [215].
Properties of the optimal value \(Q(x, \xi)\) of the second stage linear programming problem and of its expectation \(\mathbb{E}[Q(x, \xi)]\) were first studied by Kall [99, 100], Walkup and Wets [223, 224], and Wets [226, 227]. Example 2.5 is discussed in Birge and Louveaux [19]. Polyhedral and convex two-stage problems, discussed in sections 2.2 and 2.3, are natural extensions of the linear two-stage problems. Many additional examples and analysis of particular models can be found in Birge and Louveaux [19], Kall and Wallace [102], Wallace and Ziemba [225]. For a thorough analysis of simple recourse models, see Kall and Mayer [101].

Duality analysis of stochastic problems, and in particular dualization of the nonanticipativity constraints, was developed by Eisner and Olsen [64], Wets [228] and Rockafellar and Wets [179, 180] (see also Rockafellar [176] and the references therein).

Expected value of perfect information is a classical concept in decision theory (see, Raiffa and Schlaifer [165], Raiffa [164]). In stochastic programming this and related concepts were analyzed by Madansky [122], Spivey [210], Avriel and Williams [11], Dempster [46], Huang, Vertinsky and Ziemba [95].

Numerical methods for solving two- and multi-stage stochastic programming problems are extensively discussed in Birge and Louveaux [19], Ruszczyński [186], Kall and Mayer [101], where the Reader can also find detailed references to original contributions.

There is also an extensive literature on constructing scenario trees for multistage models, encompassing various techniques using probability metrics, pseudo-random sequences, lower and upper bounding trees, and moment matching. The Reader is referred to Kall and Mayer [101], Heitsch and and Römisch [82], Hochreiter and Pflug [91], Casey and Sen [31], Pennanen [145], Dupačova, Grote-Kuska and Römisch [59], and the references therein.

An extensive stochastic programming bibliography can be found at the web site: http://mally.eco.rug.nl/spbib.html, maintained by Maarten van der Vlerk.

Chapter 4

Models involving probabilistic (chance) constraints were introduced by Charnes, Cooper and Symonds [32], Miller and Wagner [129], and Prékopa [154]. Problems with integrated chance constraints are considered in [81]. Models with stochastic dominance constraints are introduced and analyzed by Dentcheva and Ruszczyński in [52, 54, 55]. The notion of stochastic ordering or stochastic dominance of first order has been introduced in statistics in Mann and Whitney [124], and Lehmann [116], and further applied and developed in economics (see Quirk and Saposnik [163], Fishburn [67] and Hadar and Russell [80]).

An essential contribution to the theory and solutions of problems with chance constraints was the theory of \(\alpha\)-concave measures and functions. In Prékopa [155, 156] the concept of logarithmic concave measures was introduced and studied. This notion was generalized to \(\alpha\)-concave measures and functions in Borell [23, 24], Brascamp and Lieb [26], and Rinott [168], and further analyzed in Tamm [214], and Norkin [141]. Approximations of the probability function by Steklov-Sobolev transformation was suggested by Norkin in [139]. Differentiability properties of probability functions are studied in Uryasev [216, 217], Kibzun and Tretjakov [104], Kibzun and Uryasev [105], and Raik [166]. The first definition of \(\alpha\)-concave discrete multivariate distributions was introduced in Barndorff-Nielsen [13]. The generalized definition of \(\alpha\)-concave functions on a set,
which we have adopted here, was introduced in Dentcheva, Prekopa, and Ruszczyński [49]. It facilitates the development of optimality and duality theory of probabilistic optimization. Its consequences for probabilistic optimization are explored in Dentcheva, Prékopa, and Ruszczyński [50]. The notion of \( p \)-efficient points was first introduced in Prékopa [157]. Similar concept is used in Sen [195]. The concept was studied and applied in the context of discrete distributions and linear problems in the papers Dentcheva, Prékopa, and Ruszczyński [49, 50] and Prékopa and B. Vízvári and T. Badics [160] and in the context of general distributions in Dentcheva, Lai, and Ruszczyński [48]. Optimization problems with probabilistic set-covering constraint are investigated in Beraldí and Ruszczyński [16, 17], where efficient enumeration procedures of \( p \)-efficient points of 0–1 variable are employed.

There is a wealth of research on estimating probabilities of events. We refer to Boros and Prékopa [25], Bukszár [27], Bukszár and Prékopa [28], Dentcheva, Prékopa, Ruszczynski [50], Prékopa [158], and T. Szántai [212], where probability bounds are used in the context of chance constraints.

Statistical approximations of probabilistically constrained problems were analyzed in Sainetti [191], Kankova [103], Deák [43], and Gröwe [77]. Stability of models with probabilistic constraints is addresses in Dentcheva [47], Henrion [84, 83], and Henrion and Römisch [183, 85]. Nonlinear probabilistic problems are investigated in Dentcheva, Lai, and Ruszczyński [48], where optimality conditions are established.

Many applied models in engineering, where reliability is frequently a central issue (e.g., in telecommunication, transportation, hydrological network design and operation, engineering structure design, electronic manufacturing problems, etc.) include optimization under probabilistic constraints. We do not list these applied works here. In finance, the concept of Value-at-Risk enjoys great popularity (see, e.g., Dowd [57], Pflug [148], Pflug and Römisch [149]). The concept of stochastic dominance plays a fundamental role in economics and statistics. We refer to Mosler and Scarsini [131], Shaked and Shanthikumar [196], and Szekli [213] for more information and a general overview on stochastic orders.

**Chapter 5**

The concept of SAA estimators is closely related to the Maximum Likelihood (ML) method and M-estimators developed in Statistics literature. However, the motivation and scope of applications are quite different. In Statistics the involved constraints typically are of a simple nature and do not play such an essential role as in stochastic programming. Also in applications of Monte Carlo sampling techniques to stochastic programming the respective sample is generated in the computer and its size can be controlled, while in statistical applications the data are typically given and cannot be easily changed.

Starting with a pioneering work of Wald [222], consistency properties of the Maximum Likelihood and M-estimators were studied in numerous publications. Epi-convergence approach to studying consistency of statistical estimators was discussed in Dušapčová and Wets [60]. In the context of stochastic programming, consistency of SAA estimators was also investigated by tools of epi-convergence analysis in King and Wets [108] and Robinson [173].

Proposition 5.6 appeared in Norkin, Pflug and Ruszczyński [140] and Mak, Morton and Wood [123]. Theorems 5.7, 5.11 and 5.10 are taken from Shapiro [198] and [204], respectively. The approach to second order asymptotics, discussed in section 5.1.3, is based
on Dentcheva and Römisch [51] and Shapiro [199]. Starting with the classical asymptotic theory of the ML method, asymptotics of statistical estimators were investigated in numerous publications. Asymptotic normality of $M$-estimators was proved, under quite weak differentiability assumptions, in Huber [96]. An extension of the SAA method to stochastic generalized equations is a natural one. Stochastic variational inequalities were discussed by Gürkan, Özge and Robinson [79]. Proposition 5.14 and Theorem 5.15 are similar to the results obtained in [79, Theorems 1 and 2]. Asymptotics of SAA estimators of optimal solutions of stochastic programs are discussed in King and Rockafellar [107] and Shapiro [197].

The idea of using Monte Carlo sampling for solving stochastic optimization problems of the form (5.1) certainly is not new. There is a variety of sampling based optimization techniques which were suggested in the literature. It will be beyond the scope of this chapter to give a comprehensive survey of these methods, we mention a few approaches related to the material of this chapter. One approach uses the Infinitesimal Perturbation Analysis (IPA) techniques to estimate the gradients of $f(\cdot)$, which consequently are employed in the Stochastic Approximation (SA) method. For a discussion of the IPA and SA methods we refer to Ho and Cao [90], Glasserman [75], and Kushner and Clark [112], Nevelson and Hasminskii [137], respectively. For an application of this approach to optimization of queueing systems see Chong and Ramadge [36] and L'Ecuyer and Glynn [115], for example. Closely related to this approach is the Stochastic Quasi-Gradient method (see Ermoliev [65]).

Another class of methods uses sample average estimates of the values of the objective function, and may be its gradients (subgradients), in an “interior” fashion. Such methods are aimed at solving the true problem (5.1) by employing sampling estimates of $f(\cdot)$ and $\nabla f(\cdot)$ blended into a particular optimization algorithm. Typically, the sample is updated or a different sample is used each time function or gradient (subgradient) estimates are required at a current iteration point. In this respect we can mention, in particular, the statistical L-shaped method of Infanger [97] and the stochastic decomposition method of Higle and Sen [88].

In this chapter we mainly discussed an “exterior” approach, in which a sample is generated outside of an optimization procedure and consequently the constructed sample average approximation (SAA) problem is solved by an appropriate deterministic optimization algorithm. There are several advantages in such approach. The method separates sampling procedures and optimization techniques. This makes it easy to implement and, in a sense, universal. From the optimization point of view, given a sample $\xi^1, \ldots, \xi^N$, the obtained optimization problem can be considered as a stochastic program with the associated scenarios $\xi^1, \ldots, \xi^N$, each taken with equal probability $\frac{1}{N}$. Therefore any optimization algorithm which is developed for a considered class of stochastic programs can be applied to the constructed SAA problem in a straightforward way. Also the method is ideally suited for a parallel implementation. From the theoretical point of view there is available a quite well developed statistical inference of the SAA method. This, in turn, gives a possibility of error estimation, validation analysis and hence stopping rules. Finally, various variance reduction techniques can be conveniently combined with the SAA method.

It is difficult to point out an exact origin of the SAA method. The idea is simple indeed and it was used by various authors under different names. Variants of this approach are known as the stochastic counterpart method (Rubinstein and Shapiro [184],[185]) and
sample-path optimization (Plambeck, Fu, Robinson and Suri [151] and Robinson [173]), for example. Also similar ideas were used in statistics for computing maximum likelihood estimators by Monte Carlo techniques based on Gibbs sampling (see, e.g., Geyer and Thompson [72] and references therein). Numerical experiments with the SAA approach, applied to linear and discrete (integer) stochastic programming problems, can be also found in more recent publications [3],[120],[123],[220].

The complexity analysis of the SAA method, discussed in section 5.3, is motivated by the following observations. Suppose for the moment that components $\xi_i$, $i = 1, \ldots, d$, of the random data vector $\xi \in \mathbb{R}^d$ are independently distributed. Suppose, further, that we use $r$ points for discretization of the (marginal) probability distribution of each component $\xi_i$. Then the resulting number of scenarios is $K = r^d$, i.e., it grows exponentially with increase of the number of random parameters. Already with, say, $r = 4$ and $d = 20$ we will have an astronomically large number of scenarios $4^{20} \approx 10^{12}$. In such situations it seems hopeless just to calculate with a high accuracy the value $f(x) = \mathbb{E}[F(x, \xi)]$ of the objective function at a given point $x \in X$, much less to solve the corresponding optimization problem $\min_{x \in X} f(x)$. And, indeed, it is shown in Dyer and Stougie [61] that, under the assumption that the stochastic parameters are independently distributed, two-stage linear stochastic programming problems are $\mathcal{NP}$-hard. This indicates that, in general, two-stage stochastic programming problems cannot be solved with a high accuracy, as say with accuracy of order $10^{-3}$ or $10^{-4}$, as it is common in deterministic optimization. On the other hand, quite often in applications it does not make much sense to try to solve the corresponding stochastic problem with a high accuracy since the involved inaccuracies resulting from inexact modeling, distribution approximations etc, could be far bigger. In some situations the randomization approach based on Monte Carlo sampling techniques allows to solve stochastic programs with a reasonable accuracy and a reasonable computational effort.

The material of section 5.3.1 is based on Kleywegt, Shapiro and Homem-De-Mello [109]. The extension of that analysis to general feasible sets, given in section 5.3.2, is discussed in Shapiro [200, 202, 205] and Shapiro and Nemirovski [206]. The material of section 5.3.3 is based on Shapiro and Homem-de-Mello [208], where proof of Theorem 5.24 can be found.

In practical applications, in order to speed up the convergence, it is often advantageous to use Quasi-Monte Carlo techniques. Theoretical bounds for the error of numerical integration by Quasi-Monte Carlo methods are proportional to $(\log N)^d/N$, i.e., are of order $O((\log N)^d/N)$, with the respective proportionality constant $A_d$ depending on $d$. For small $d$ it is “almost” the same as of order $O(N^{-1})$, which of course is better than $O_p(N^{-1/2})$. However, the theoretical constant $A_d$ grows superexponentially with increase of $d$. Therefore, for larger values of $d$ one often needs a very large sample size $N$ for Quasi-Monte Carlo methods to become advantageous. It is beyond the scope of this chapter to give a thorough discussion of Quasi-Monte Carlo methods. A brief discussion of Quasi-Monte Carlo techniques is given in section 5.4. For a further readings on that topic we may refer to Niederreiter [138]. For applications of Quasi-Monte Carlo techniques to stochastic programming see, e.g., Koivu [110], Homem-de-Mello [94], Pennanen and Koivu [146].

---

21Of course, in some very specific situations it is possible to calculate $\mathbb{E}[F(x, \xi)]$ in a closed form. Also if $F(x, \xi)$ is decomposable into the sum $\sum_{i=1}^d F_i(x, \xi_i)$, then $\mathbb{E}[F(x, \xi)] = \sum_{i=1}^d \mathbb{E}[F_i(x, \xi_i)]$ and hence the problem is reduced to calculations of one dimensional integrals. This happens in the case of the so-called simple recourse.
For a discussion of variance reduction techniques in Monte Carlo sampling we refer to Fishman [68] and a survey paper by Avramidis and Wilson [10], for example. In the context of stochastic programming, variance reduction techniques were discussed in Rubinstein and Shapiro [185], Dantzig and Infanger [42], Higle [86] and Bailey, Jensen and Morton [12], for example.

The statistical bounds of section 5.6.1 were suggested in Norkin, Pflug and Ruszczyński [140], and developed in Mak, Morton and Wood [123]. The “common random numbers” estimator \( \hat{\gamma}^{N,M}_{N,M}(\bar{x}) \) of the optimality gap was introduced in [123]. The KKT statistical test, discussed in section 5.6.2, was developed in Shapiro and Homem-de-Mello [207], so that the material of that section is based on [207]. See also Higle and Sen [87].

The estimate of the sample size derived in Theorem 5.32 is due to Campi and Garatti [30]. This result builds on a previous work of Calafiore and Campi [29], and from the considered point of view gives a tightest possible estimate of the required sample size. Construction of upper and lower statistical bounds for chance constrained problems, discussed in section 5.7, is based on Nemirovski and Shapiro [134]. For some numerical experiments with these bounds see Luedtke and Ahmed [121].

The extension of the SAA method to multistage stochastic programming, discussed in section 5.8 and referred to as conditional sampling, is a natural one. A discussion of consistency of conditional sampling estimators is given, e.g., in Shapiro [201]. Discussion of the portfolio selection (Example 5.34) is based on Blomvall and Shapiro [21]. Complexity of the SAA approach to multistage programming is discussed in Shapiro and Nemirovski [206] and Shapiro [203].

Section 5.9 is based on Nemirovski et al [133]. The origins of the Stochastic Approximation algorithms are going back to the pioneering paper by Robbins and Monro [169]. For a thorough discussion of the asymptotic theory of the SA method we can refer to Kushner and Clark [112] and Nevelson and Hasminskii [137]. The robust SA approach was developed in Polyak [152] and Polyak and Juditsky [153]. The main ingredients of Polyak’s scheme (long steps and averaging) were, in a different form, proposed in Nemirovski and Yudin [135].

Chapter 6

Foundations of the expected utility theory were developed in von Neumann and Morgenstern [221]. The dual utility theory is developed in Quiggin [161, 162] and Yaari [229].

The mean–variance model was introduced and analyzed in Markowitz [125, 126, 127]. Deviations and semideviations in mean–risk analysis were analyzed in Kijima and Ohnishi [106], Konno [111], Ogryczak and Ruszczyński [142, 143], Ruszczyński and Vanderbei [190]. Weighted deviations from quantiles, relations to stochastic dominance and Lorenz curves are discussed in Ogryczak and Ruszczyński [144]. For Conditional (Average) Value at Risk see Acerbi and Tasche [1], Rockafellar and Uryasev [177], Pflug [148].

A general class of convex approximations of chance constraints is developed in Nemirovski and Shapiro [134].

The theory of coherent measures of risk was initiated in Artzner, Delbaen, Eber and Heath [8] and further developed, \textit{inter alia}, by Delbaen [44], Föllmer and Schied [69], Leitner [117], Rockafellar, Uryasev and Zabarankin [178]. Our presentation is based on Ruszczyński and Shapiro [187, 189]. Kusuoka representation of law invariant coherent
risk measures (Theorem 6.23) was derived in [113] for $L_\infty(\Omega, \mathcal{F}, P)$ spaces. For an extension to $L_p(\Omega, \mathcal{F}, P)$ spaces see, e.g., Pflug and Römisch [149]. Theory of consistency with stochastic orders was initiated in [142] and developed in [143, 144]. An alternative approach to asymptotic analysis of law invariant coherent risk measures (see section 6.5.3), was developed in Pflug and Wozabal [147] based on Kusuoka representation. Application to portfolio optimization is discussed in Miller and Ruszczyński [130].

The theory of conditional risk mappings was developed in Riedel [167], Ruszczyński and Shapiro [187, 188]. For the general theory of dynamic measures of risk, see Artzner, Delbaen, Eber, Heath and Ku [9], Cheridito, Delbaen and Kupper [33, 34], Frittelli and Rosazza Gianin [71, 70], Eichhorn and Römisch [63], Pflug and Römisch [149]. Our inventory example is based on Ahmed, Cakmak and Shapiro [2].

Chapter 7
There are many monographs where concepts of directional differentiability are discussed in detail, see, e.g., [22]. A thorough discussion of Clarke generalized gradient and regularity in the sense of Clarke can be found in Clarke [38]. Classical references on (finite dimensional) convex analysis are books by Rockafellar [174] and Hiriart-Urruty and Lemaréchal [89]. For a proof of Fenchel-Moreau Theorem (in an infinite dimensional setting) see, e.g., [175]. For a development of conjugate duality (in an infinite dimensional setting) we refer to Rockafellar [175]. Theorem 7.11 (Hoffman’s Lemma) appeared in [93].

Theorem 7.21 appeared in Danskin [40]. Theorem 7.22 is going back to Levin [118] and Valadier [218] (see also Ioffe and Tihomirov [98, p. 213]). For a general discussion of second order optimality conditions and perturbation analysis of optimization problems we can refer to Bonnans and Shapiro [22] and references therein. Theorem 7.24 is an adaptation of a result going back to Gol’shtein [76]. For a thorough discussion of epiconvergence we refer to Rockafellar and Wets [181]. Theorem 7.27 is taken from [181, Theorem 7.17].

There are many books on probability theory. Of course, it will be beyond the scope of this monograph to give a thorough development of that theory. In that respect we can mention the excellent book by Billingsley [18]. Theorem 7.32 appeared in Rogosinski [182]. A thorough discussion of measurable multifunctions and random lower semicontinuous functions can be found in Rockafellar and Wets [181, Chapter 14], to which the interested reader is referred for a further reading. For a proof of Aumann and Lyapunov Theorems (Theorems 7.40 and 7.41) see, e.g., [98, section 8.2].

Theorem 7.47 originated in Strassen [211], where the interchangeability of the sub-differential and integral operators was shown in the case when the expectation function is continuous. The present formulation of Theorem 7.47 is taken from [98, Theorem 4, p. 351].

Uniform Laws of Large Numbers take their origin in the Glivenko-Cantelli Theorem. For a further discussion of the uniform LLN we refer to van der Vaart and Wellner [219]. Epi-convergence Law of Large Numbers, derived in Theorem 7.51, is due to Artstein and Wets [7]. The uniform convergence w.p.1 of Clarke generalized gradients, specified in part (c) of Theorem 7.52, was obtained in [197]. Law of Large Numbers for random sets (Theorem 7.53) appeared in Artstein and Vitale [6]. The uniform convergence of $\varepsilon$-subdifferentials (Theorem 7.55) was derived in [209].

Finite dimensional Delta Method is well known and routinely used in theoretical
Statistics. The infinite dimensional version (Theorem 7.59) is going back to Gr"ubel [78], Gill [74] and King [107]. The tangential version (Theorem 7.61) appeared in [198].

There is a large literature on Large Deviations Theory (see, e.g., a book by Dembo and Zeitouni [45]). Hoeffding inequality appeared in [92] and Chernoff inequality in [35]. Theorem 7.68, about interchangeability of Clarke generalized gradient and integral operators, can be derived by using the interchangeability formula (7.117) for directional derivatives, Strassen’s Theorem 7.47 and the fact that in the Clarke regular case the directional derivative is the support function of the corresponding Clarke generalized gradient (see [38] for details).

A classical reference for functional analysis is Dunford and Schwartz [58]. The concept of algebraic subgradient and Theorem 7.78 are taken from Levin [119] (unfortunately this excellent book was not translated from Russian). Theorem 7.79 is from Ruszczyński and Shapiro [189]. The interchangeability principle (Theorem 7.80) is taken from [181, Theorem 14.60]. Similar results can be found in [98, Proposition 2, p.340] and [119, Theorem 0.9].
Bibliography


Bibliography


Bibliography


428 Bibliography


## Index

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>:=</td>
<td>equal by definition, 335</td>
</tr>
<tr>
<td>$A^T$</td>
<td>transpose of matrix (vector) $A$, 335</td>
</tr>
<tr>
<td>$C(X)$</td>
<td>space of continuous functions, 166</td>
</tr>
<tr>
<td>$C^*$</td>
<td>polar of cone $C$, 339</td>
</tr>
<tr>
<td>$C^1(V, \mathbb{R}^n)$</td>
<td>space of continuously differentiable mappings, 178</td>
</tr>
<tr>
<td>$IF$</td>
<td>influence function, 306</td>
</tr>
<tr>
<td>$L^+$</td>
<td>orthogonal of (linear) space $L$, 41</td>
</tr>
<tr>
<td>$O(1)$</td>
<td>generic constant, 189</td>
</tr>
<tr>
<td>$O_p(\cdot)$</td>
<td>term, 386</td>
</tr>
<tr>
<td>$S^\varepsilon$</td>
<td>the set of $\varepsilon$-optimal solutions of the true problem, 182</td>
</tr>
<tr>
<td>$V_d(A)$</td>
<td>Lebesgue measure of set $A \subset \mathbb{R}^d$, 196</td>
</tr>
<tr>
<td>$W^{1,\infty}(U)$</td>
<td>space of Lipschitz continuous functions, 167, 355</td>
</tr>
<tr>
<td>$[a]_+$</td>
<td>$\max{a, 0}$, 2</td>
</tr>
<tr>
<td>$1_A(\cdot)$</td>
<td>indicator function of set $A$, 336</td>
</tr>
<tr>
<td>$L_p(\Omega, \mathcal{F}, P)$</td>
<td>space, 403</td>
</tr>
<tr>
<td>$\Phi(z)$</td>
<td>cdf of standard normal distribution, 16</td>
</tr>
<tr>
<td>$\Pi_X$</td>
<td>metric projection onto set $X$, 232</td>
</tr>
<tr>
<td>$\mathcal{D}$</td>
<td>convergence in distribution, 164</td>
</tr>
<tr>
<td>$T^2_\varepsilon(x, h)$</td>
<td>second order tangent set, 351</td>
</tr>
<tr>
<td>$AV(\Omega)$</td>
<td>Average Value-at-Risk, 260</td>
</tr>
<tr>
<td>$\mathcal{P}$</td>
<td>set of probability measures, 308</td>
</tr>
<tr>
<td>$\mathbb{D}(A, B)$</td>
<td>deviation of set $A$ from set $B$, 336</td>
</tr>
<tr>
<td>$\mathbb{D}(</td>
<td>Z</td>
</tr>
<tr>
<td>$\mathbb{E}$</td>
<td>expectation, 363</td>
</tr>
<tr>
<td>$\mathbb{H}(A, B)$</td>
<td>Hausdorff distance between sets $A$ and $B$, 336</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>set of positive integers, 361</td>
</tr>
<tr>
<td>$\mathbb{R}^n$</td>
<td>$n$-dimensional space, 335</td>
</tr>
<tr>
<td>$\mathcal{A}$</td>
<td>domain of the conjugate of risk measure $\rho$, 265</td>
</tr>
<tr>
<td>$\mathcal{C}_n$</td>
<td>the space of nonempty compact subsets of $\mathbb{R}^n$, 382</td>
</tr>
<tr>
<td>$\mathcal{P}$</td>
<td>set of probability density functions, 266</td>
</tr>
<tr>
<td>$\mathcal{S}$</td>
<td>set of contact points, 403</td>
</tr>
<tr>
<td>$b(k; \alpha, N)$</td>
<td>cdf of binomial distribution, 216</td>
</tr>
<tr>
<td>$d$</td>
<td>distance generating function, 238</td>
</tr>
<tr>
<td>$g^+(x)$</td>
<td>right hand side derivative, 299</td>
</tr>
<tr>
<td>$cl(A)$</td>
<td>topological closure of set $A$, 336</td>
</tr>
<tr>
<td>$\mathcal{C}_n$</td>
<td>the space of nonempty compact subsets of $\mathbb{R}^n$, 382</td>
</tr>
<tr>
<td>$e_{\text{epi}}(\cdot)$</td>
<td>epiconvergence, 379</td>
</tr>
<tr>
<td>$\bar{S}_N$</td>
<td>the set of optimal solutions of the SAA problem, 158</td>
</tr>
<tr>
<td>$\hat{S}_N^\varepsilon$</td>
<td>the set of $\varepsilon$-optimal solutions of the SAA problem, 182</td>
</tr>
<tr>
<td>$\hat{f}_N(x)$</td>
<td>sample average function, 157</td>
</tr>
<tr>
<td>$1_A(\cdot)$</td>
<td>characteristic function of set $A$, 336</td>
</tr>
<tr>
<td>$\text{int}(C)$</td>
<td>interior of set $C$, 339</td>
</tr>
<tr>
<td>$[a]$</td>
<td>integer part of $a \in \mathbb{R}$, 221</td>
</tr>
</tbody>
</table>
Index

lsc $f$ lower semicontinuous hull of function $f$, 335
$\mathcal{R}_C$ radial cone to set $C$, 339
$\mathcal{T}_C$ tangent cone to set $C$, 339
$\nabla^2 f(x)$ Hessian matrix of second order partial derivatives, 180
$\partial$ subdifferential, 340
$\partial^\circ$ Clarke generalized gradient, 338
$\partial_\varepsilon$ epsilon subdifferential, 384
$\text{pos} W$ positive hull of matrix $W$, 29
$\Pr(A)$ probability of event $A$, 362
$\text{ri}$ relative interior, 339
$\sigma^+_p$ upper semideviation, 257
$\sigma^-_p$ lower semideviation, 257
$\mathbb{V}/\mathbb{R}_0$, Value-at-Risk, 258
$\mathbb{V}ar[X]$ variance of $X$, 15
$\vartheta^*$ optimal value of the true problem, 158
$\xi(t) = (\xi_1, \ldots, \xi_t)$ history of the process, 63
$a \lor b = \max\{a, b\}$, 187
$f^*$ conjugate of function $f$, 340
$f^\circ(x, d)$ generalized directional derivative, 338
$g'(x, h)$ directional derivative, 336
$\omega_p(\cdot)$ term, 386
$p$-efficient point, 117
iid independently identically distributed, 158
approximation
conservative, 260
Average Value-at-Risk, 259, 260, 262, 274
dual representation, 274
Banach lattice, 407
Borel set, 361
bounded in probability, 386
Bregman divergence, 239
Capacity Expansion, 31, 42, 59
chain rule, 387
chance constrained problem
ambiguous, 287
disjunctive semi-infinite formulation, 118
chance constraints, 6, 11, 16, 212

Clarke generalized gradient, 338
CLT Central Limit Theorem, 144
common random number generation method, 181
complexity
of multistage programs, 228
of two-stage programs, 182, 189
conditional expectation, 366
conditional probability, 365
conditional risk mapping, 312, 317
conditional sampling, 222
identical, 222
independent, 222
Conditional Value-at-Risk, 259, 260, 262
cone
contingent, 349, 389
critical, 179, 350
normal, 339
pointed, 407
polar, 29
recession, 29
tangent, 339
confidence interval, 164
conjugate duality, 343, 407
constraint
nonanticipativity, 54, 293
constraint qualification
linear independence, 170, 180
Mangasarian-Fromovitz, 350
Robinson, 349
Slater, 163
contingent cone, 349, 389
convergence
in distribution, 164, 385
in probability, 386
weak, 388
with probability one, 376
convex hull, 339
cumulative distribution function, 2
of random vector, 11
decision rule, 21
Delta Theorem, 388, 389
finite dimensional, 387
second-order, 390
deviation of a set, 336
Index

436

diameter
  of a set, 187
differential uniform dominance condition, 147
directional derivative, 336
  \( \varepsilon \)-directional derivative, 384
  generalized, 338
  Hadamard, 387
  second order, 390
tangentially to a set, 390
distribution
  asymptotically normal, 164
  Binomial, 394
  conditional, 365
  Dirichlet, 98
  discrete, 363
  discrete with a finite support, 363
  empirical, 158
  Gamma, 103
  log-concave, 97
  log-normal, 108
  multivariate normal, 16, 96
  multivariate Student, 152
  normal, 164
  Pareto, 153
  uniform, 96
  Wishart, 103
domain
  of multifunction, 367
  of a function, 335
dual feasibility condition, 129
duality gap, 342, 343
dynamic programming equations, 7, 64, 315
empirical cdf, 3
empirical distribution, 158
entropy function, 238
epiconvergence, 360
  with probability one, 379
epigraph of a function, 335
epsilon subdifferential, 384
estimator
  common random number, 207
  consistent, 159
  linear control, 202

expected value, 363
  well defined, 363
  expected value of perfect information, 61

Fatou’s lemma, 364
filtration, 71, 75, 311, 320
floating body of a probability measure, 105
Fréchet differentiability, 336
function
  \( \alpha \)-concave, 94
  \( \alpha \)-concave on a set, 106
  biconjugate, 264, 405
  Carathéodory, 158, 171, 368
  characteristic, 336
  Clarke-regular, 104, 338
  composite, 267
  conjugate, 264, 340, 404
  continuously differentiable, 338
  cost-to-go, 65, 67, 315
  cumulative distribution (cdf), 2, 362
distance generating, 238
disutility, 256, 273
esentially bounded, 403
extended real valued, 363
indicator, 29, 336
influence, 306
integrable, 363
likelihood ratio, 202
log-concave, 95
logarithmically concave, 95
lower semicontinuous, 335
moment generating, 391
monotone, 408
optimal value, 369
polyhedral, 28, 42, 335, 409
proper, 335
quasi-concave, 96
radical-inverse, 199
random, 368
random lower semicontinuous, 368
random polyhedral, 43
sample average, 377
strongly convex, 341
subdifferentiable, 340, 405
# Index

| Utility, 256, 273 | Lipschitz continuous, 337 |
| Well defined, 370 | Lower bound |
| Gâteaux differentiability, 336, 387 | Statistical, 204 |
| Generalized equation | Lyapunov condition, 145 |
| Sample average approximation, 176 | Mapping |
| Generic constant $O(1)$, 189 | Convex, 50 |
| Gradient, 338 | Measurable, 362 |
| Hadamard differentiability, 387 | Markov chain, 71 |
| Hausdorff distance, 336 | Markovian process, 63 |
| Here-and-now solution, 10 | Martingale, 326 |
| Hessian matrix, 350 | Mean absolute deviation, 257 |
| Higher order distribution functions, 90 | Measurable selection, 368 |
| Hoffman’s lemma, 346 | Measure |
| Identically distributed, 376 | $\alpha$-concave, 97 |
| Importance sampling, 203 | Absolutely continuous, 362 |
| Independent identically distributed, 376 | Complete, 361 |
| Inequality | Dirac, 364 |
| Chebyshev, 365 | Finite, 361 |
| Chernoff, 394 | Lebesgue, 361 |
| Hölder, 404 | N nonatomic, 370 |
| Hardy–Littlewood–Polya, 282 | Sigma-additive, 361 |
| Hoeffding, 394 | Metric projection, 232 |
| Jensen, 365 | Mirror Descent SA, 242 |
| Markov, 365 | Model state equations, 68 |
| Minkowski for matrices, 101 | Model state variables, 68 |
| Inf-compactness condition, 160 | Moment generating function, 391 |
| Interchangeability principle, 409 | Multifunction, 367 |
| For risk measures, 295 | Closed, 176, 368 |
| For two-stage programming, 49 | Closed valued, 176, 367 |
| Interior of a set, 339 | Convex, 50 |
| Inventory model, 1, 297 | Convex valued, 50, 370 |
| Jacobian matrix, 338 | Measurable, 367 |
| Lagrange multiplier, 350 | Optimal solution, 369 |
| Large deviations rate function, 391 | Upper semicontinuous, 384 |
| Law of Large Numbers, 2, 376 | Newsvendor problem, 1, 332 |
| For random sets, 382 | Node |
| Pointwise, 377 | Ancestor, 69 |
| Strong, 376 | Children, 69 |
| Uniform, 377 | Root, 69 |
| Weak, 376 | N nonatomicity, 7, 53, 63 |
| Least upper bound, 407 | N nonatomicity constraints, 72, 314 |
| N nonatomic probability space, 370 | Norm |
dual, 238, 402
normal cone, 339

optimality conditions
first order, 209, 348
Karush-Kuhn-Tucker (KKT), 175, 209, 350
second order, 180, 350

partial order, 407
point
contact, 403
saddle, 342

polar cone, 339

portfolio selection, 13, 300
positive hull, 29
positively homogeneous, 179
Preface, iii
probabilistic constraints, 6, 11, 87, 163
individual, 90
joint, 90
probabilistic liquidity constraint, 94
probability density function, 362
probability distribution, 362
probability measure, 361
probability vector, 311
problem
chance constrained, 87, 212
first stage, 10
of moments, 308
piecewise linear, 194
second stage, 10
semi-infinite programming, 310
subconsistent, 343
two stage, 10
prox-function, 239
prox-mapping, 239
quadratic growth condition, 192, 352
quantile, 16
left-side, 3, 258
right-side, 3, 258
radial cone, 339
random function
convex, 371
random variable, 362
random vector, 362
recession cone, 339
recourse
complete, 33
fixed, 33, 45
relatively complete, 10, 33
simple, 33
recourse action, 2
relative interior, 339
risk measure, 263
absolute semideviation, 303, 331
coherent, 264
composite, 314, 321
consistency with stochastic orders, 284
law based, 281
law invariant, 281
mean-deviation, 278
mean-upper-semideviation, 279
mean-upper-semideviation from a target, 280
mean-variance, 277
multiperiod, 323
proper, 263
version independent, 281
robust optimization, 12
saddle point, 342
sample
independently identically distributed (iid), 158
random, 157
sample average approximation (SAA), 157
multistage, 222
sample covariance matrix, 210
sampling
Latin Hypercube, 200
Monte Carlo, 181
Index

scenario tree, 69
scenarios, 3, 30
second order regularity, 352
second order tangent set, 351
semi-infinite probabilistic problem, 145
semideviation
  lower, 257
  upper, 257
separable space, 388
sequence
  Halton, 199
  log-concave, 107
  low-discrepancy, 198
  van der Corput, 199
set
  elementary, 361
  of contact points, 403
sigma algebra, 361
  Borel, 361
  trivial, 361
significance level, 6
simplex, 238
Slater condition, 163
solution
  ε-optimal, 182
  sharp, 192, 193
space
  Banach, 402
  decomposable, 409
  dual, 402
  Hilbert, 277
  measurable, 361
  probability, 361
  reflexive, 403
  sample, 361
stagewise independence, 8, 63
star discrepancy, 196
stationary point of α-concave function, 105
Stochastic Approximation, 232
stochastic dominance
  k-th order, 91
  first order, 90, 284
  higher order, 91
  second order, 285
stochastic dominance constraint, 91
stochastic generalized equations, 175
stochastic order, 90, 284
  increasing convex, 285
  usual, 284
stochastic ordering constraint, 91
stochastic programming
  nested risk averse multistage, 313, 320
stochastic programming problem
  minimax, 171
  multiperiod, 66
  multistage, 64
  multistage linear, 68
  two-stage convex, 50
  two-stage linear, 27
  two-stage polyhedral, 42
strict complementarity condition, 181, 211
strongly regular solution of a generalized equation, 177
subdifferential, 340, 405
subgradient, 340, 406
  algebraic, 406
  stochastic, 231
supply chain model, 22
support
  of a set, 339
  of measure, 362
support function, 28, 339, 340
support of a measure, 36
tangent cone, 339
theorem
  Artstein-Vitale, 382
  Aumann, 370
  Banach-Alaoglu, 404
  Birkhoff, 112
  Central Limit, 144
  Cramér’s Large Deviations, 391
  Danskin, 354
  Fenchel-Moreau, 265, 341, 405
  functional CLT, 166
  Glivenko-Cantelli, 379
  Helly, 339
  Hlawka, 198
  Klee-Nachbin-Namioka, 407
  Koksma, 197
Index

Kusuoka, 282
Lebesgue dominated convergence, 364
Levin-Valadier, 354
Lyapunov, 370
measurable selection, 368
monotone convergence, 363
Moreau–Rockafellar, 340, 406
Rademacher, 338, 355
Radon-Nikodym, 362
Richter-Rogosinski, 364
Skorohod-Dudley almost sure representation, 388
time consistency, 323
topology
  strong (norm), 404
  weak, 404
  weak*, 404
uncertainty set, 12, 308
uniformly integrable, 386
upper bound
  conservative, 206
  statistical, 205
utility model, 272
value function, 7
Value-at-Risk, 16, 258, 275
  constraint, 16
variation
  of a function, 197
variational inequality, 175
  stochastic, 175
Von Mises statistical functional, 306
wait-and-see solution, 10, 60
weighted mean deviation, 258