Lecture outline

■ Modeling
■ The state variables
■ The decision variables
■ Exogenous information
■ The transition function
■ The objective function
Modeling

- Determining when to store energy from wind turbines or the grid:
Which driver to assign to each load
How to find our way home?
Modeling

- Each of these problems can be modeled as a sequential, stochastic optimization problem:
  - Managing a fleet of locomotives
  - Managing the storage of energy
  - Load matching for truckload trucking
  - Finding the shortest path through a network

- Below, we argue that each problem represents a good application of a fundamentally different solution strategy.

- But before we solve these problems, we have to learn how to model them.
Modeling

For deterministic problems, we speak the language of mathematical programming

» For static problems

$$\min cx$$
$$Ax = b$$
$$x \geq 0$$

» For time-staged problems

$$\min \sum_{t=0}^{T} c_t x_t$$
$$A_t x_t - B_{t-1} x_{t-1} = b_t$$
$$D_t x_t \leq u_t$$
$$x_t \geq 0$$

The entire world models deterministic optimization problems using the same language. It is arguably Dantzig’s most enduring achievement.

We lack a similar foundation when we introduce random variables.
Modeling

Almost all sequential decision problems can be modeled using five core components:

» State variables
  • What is in the state variable? What do we need to know at time t?

» Decision variables
  • What are our decisions?

» Exogenous information
  • What do we learn for the first time between t and t+1?

» Transition function
  • How does the problem evolve from t to t+1.

» Objective function
  • What are we minimizing and how?
Lecture outline

- A modeling problem
- The state variables
- The decision variables
- Exogenous information
- The transition function
- The objective function
The state variables

What is a state variable?

» Surprisingly, the academic community has generally avoided defining a state variable.

» Bellman’s classic text on dynamic programming (1957) describes the state variable with:
  • “… we have a physical system characterized at any stage by a small set of parameters, the state variables.”

» The most popular book on dynamic programming (Puterman, 2005, p.18) “defines” a state variable with the following sentence:
  • “At each decision epoch, the system occupies a state.”

» Needless to say, these are not “definitions” in any formal sense of the word.
The state variables

What is a state variable?

» Wikipedia:
  • “State commonly refers to either the present condition of a system or entity” or.
  • A state variable is one of the set of variables that are used to describe the mathematical ‘state’ of a dynamical system

» Kirk (2004), an introduction to control theory, offers the definition:
  • A state variable is a set of quantities $x_1(t), x_2(t), ...$ which if known at time $t = t_0$ are determined for $t \geq t_0$ by specifying the inputs for the system for $t \geq t_0$.
  • True, but too vague to be useful.

» What was the state variable for our cash balance problem?
The state variables

Illustrating state variables
» A deterministic graph

\[ S_t = (N_t) = 6 \]
The state variables

- Illustrating state variables
  - A stochastic graph

\[ S_t = ? \]
The state variables

- Illustrating state variables
  - A stochastic graph

\[ S_t = \left( N_t, (c_{t, N_t, j})_j \right) = (6, (12.7, 8.9, 13.5)) \]
The state variables

Illustrating state variables

» A stochastic graph with left turn penalties

\[
S_t = \left( R_t, \left( I_t, N_{t-1} \right), (c_t, N_t, j) \right) = \left( 6, (12, 7, 8.9, 13.5), 3 \right)
\]
The state variables

- Illustrating state variables
  - A stochastic graph with generalized learning

\[ S_t = ? \]

© 2013 W.B. Powell
The state variables

Illustrating state variables

» A stochastic graph with generalized learning

\[ S_t = \left( \begin{array}{c} N_t, \\ R_t \\ I_t \\ K_t \end{array}, \left( c_{t,N_t,j} \right)_j \right) \]
Let’s illustrate our new notation with a new optimization problem... selling a stock.

Let $p_t$ be the price of the stock at time $t$. One way to model the evolution of our stock price is using

$$p_t = p_{t-1} + \hat{p}_t$$

In this representation, we view $\hat{p}_t$ as the exogenous information. We would write $\hat{p}_t$ as a random variable (that is, it is random before time $t$), and $\hat{p}_t(\omega)$ as a sample realization when we are following sample path $\omega$. 
The state variables

■ Selling a stock

When should we sell our stock? It makes sense to sell when the price gets too much higher than a long term trend. Estimate the trend using

$$\bar{p}_t = (1-\alpha)p_{t-1} + \alpha p_t$$

For this problem, our state variable is

$$S_t = (p_t, \bar{p}_t, R_t) \quad R_t = 1 \text{ if we are still holding the stock, 0 otherwise}$$

Let

$$X(S_t) = \begin{cases} 
1 & \text{If we sell at time } t \text{ (requires } R_t = 1) \\
0 & \text{Otherwise} 
\end{cases}$$

One possible rule for selling the stock might be to use

$$X^\beta(S_t) = 1 \text{ if } p_t \geq \bar{p}_t + \beta$$

So our policy $\pi$ represents both this type of rule, plus the parameter $\beta$. 
The state variables

- We can use different ways for tracking prices:
  » Instead of using the smoothing
    \[ \bar{p}_t = (1 - \alpha)\bar{p}_{t-1} + \alpha p_t \]
  » ... we used a moving average over the last three periods:
    \[ \bar{p}_t = \frac{1}{3}(p_t + p_{t-1} + p_{t-2}) \]
  » In this case, what is the state of our system?
The state variables

■ Moving average pricing problem
  » State variable
    \[ S_t = (p_t, p_{t-1}, p_{t-2}, R_t) \]
  » Transition function
    \[ p_{t+1} = p_t + \hat{p}_{t+1} \]
    \[ R_{t+1} = R_t - x_t \]
    \[ S_{t+1} = (p_{t+1}, p_t, p_{t-1}, R_{t+1}) \]
  » Compare this state and transition with the first pricing problem using exponentially weighted prices.
The state variables

What is a state variable?

» From Powell (2011):

Definition 5.4.1 A state variable is the minimally dimensioned function of history that is necessary and sufficient to compute the decision function, the transition function, and the contribution function.

» In plain English:
  • All the information you need, and nothing more.

» Some communities call this a sufficient statistic; we require that the state variable be necessary and sufficient.

» Where do we need information:
  • To compute the decision function
  • To compute the cost/contribution function
  • To compute the transition function

» We include the transition function because we may need information in the future (e.g. at the end of our horizon, such as the maturation of an option).
The state variables

The information we need typically consists of:

» Static information that does not vary over time
  • The tax rate on long term capital gains.
  • The dosage needed to reduce your blood sugar.
  • The amount of energy lost when being stored in a battery
  • The location of Florida

» Dynamic information that varies over time
  • The demand for a product
  • The price of a stock
  • The number of people with a disease

» By convention, we only include dynamic information in the state variable. Static information is listed separately to keep the model as compact as possible.
The state variables

The state variable (dynamic information) comes in three flavors:

- **Endogenously controllable**
  - The cash in a mutual fund
  - The water in a reservoir (which can include exogenous rainfall)

- **Exogenous information**
  - The weather
  - The price of a stock (assuming I cannot influence the market)
  - The number of patients needing blood

- **Exogenous information that can be endogenously influenced**
  - The price of a stock if I am a big player
  - The rate of new HIV infections (if I am working on policies to reduce transmission)
The state variables

■ To illustrate

» Consider our pricing problem where our decision uses the three most recent prices. The state variable is:

\[ S_t = (p_t, p_{t-1}, p_{t-2}, R_t) \]

» So what if we chose to write it as:

\[ S_t = (p_t, p_{t-1}, R_t) \]

• Or

\[ S_t = (p_t, p_{t-1}, p_{t-2}) \]

» What if we write it as

\[ S_t = (p_t, p_{t-1}, p_{t-2}, p_{t-3}, R_t) \]

• What is wrong with this? It has all the information we need.
The state variables

■ A decision tree problem
  » Do we hold or sell an asset?
  » Each node in the decision tree represents a state, which is also equivalent to the history of the process at that time.
The state variables

- An energy storage problem
  - In the battery arbitrage problem, we can store energy in a battery from the grid when prices are low, or sell back to the grid when prices are high.
  - Let
    - \( p_t \) = Purchase price for electricity from the grid
    - \( x_t \) = How much we decide to buy (\( >0 \)) or sell (\( <0 \)) between \( t \) and \( t+1 \).
    - \( R_t \) = Amount of energy in the battery at time \( t \)
    - \( R_{\text{max}} \) = Capacity of our battery
  - We might make decisions using
    \[
    x_t = \arg \max_{-R_t \leq x \leq R_{\text{max}} - R_t} \left( -p_t x + V_{t+1} (R_t + x_t) \right)
    \]
  - Our decision depends on \( p_t \) and \( R_t \).
More on state variables

- The state variable has to include everything we need to compute the decision function, the transition function and the objective function.

- But it is pointless to carry unnecessary information. That is what we mean by “minimally dimensioned.”

- Many in the academic community would view, for our pricing problem, $p_t$ would be viewed as the “state” of this system (along with $R_t$), while $p_{t-1}$ and $p_{t-2}$ would be considered the “history.”
The state variables

- Markovian vs. “history dependent” systems
  » Many in the academic community would view, for our pricing problem, $p_t$ would be viewed as the “state” of this system (along with $R_t$), while $p_{t-1}$ and $p_{t-2}$ would be considered the “history.”
  » Our requirement that a state variable be necessary and sufficient means that all dynamic systems are Markovian!
The state variables

Dimensions of a state variable:

» I often find it useful to use three perspectives of a state variable:
  • Physical state
    – This is a snapshot of the state of the physical system at a point in time
  • Information state
    – This is any information needed that is not in the physical state that we need to model the system
  • Belief state
    – This captures what we believe (in the form of probability distributions) about unobservable parameters.

» The issue of Markovian vs. history dependent systems arises when people equate state with physical state. Different communities handle this issue differently.
The state variables

■ The physical state
  » What is the price of the stock (right now)?
  » At what node are you located on a graph?
  » How much inventory do you have in stock?
  » These are all variables indexed by $t$.

■ The information state
  » We use this to include all the information other than what is in the physical state:
    • Your model that forecasts future weather, demands, stock prices.
    • Any history (variables index by $t-1$, $t-2$, …) needed to make decisions, compute costs or model the future.

■ The belief state
  » This consists of probability distributions about unknown parameters
    • How many books will I sell at price $p$?
    • What is the average global temperature?
    • How will a patient respond to a particular medication?
  » Each of these are unknown parameters. We might assume that we model our belief using normal distributions. Our belief state would include the normality assumption along with means and variances.
The state variables

Forecasting

Imagine that we are forecasting demand using the model:

\[ f_{tt'} = \theta_{t_0} + \theta_{t_1} (t' - t) \]

Our forecast of demand for time \( t+1 \) is

\[ f_{t,t+1} = \theta_{t_0} + \theta_{t_1} (t + 1 - t) = \theta_{t_0} + \theta_{t_1} \]

Assume we observe \( \hat{D}_{t+1} \) at time \( t+1 \), and we believe that

\[ \hat{D}_{t+1} = \hat{D}_t + \varepsilon_{t+1} \quad \text{where} \quad \varepsilon_{t+1} \sim N(0, \sigma^2) \]

We can update our parameters using

\[ \theta_{t+1,0} = (1 - \alpha) \theta_{t_0} + \alpha \hat{D}_{t+1} \]

\[ \theta_{t+1,1} = (1 - \alpha) \theta_{t_1} + \alpha \left( \hat{D}_{t+1} - \hat{D}_t \right) \]

The state of our system would be

\[ S_t = (\hat{D}_t, \theta_{t_0}, \theta_{t_1}) \]
The state variables

- Closing notes

  » Our definition of a state variable is made in the context of making a decision. In a nutshell, it is the information needed to make a decision (now or in the future), and we insist that it also include the information we need to compute our cost/contribution function.

  » The probability/stochastic process community has not tried to define a state variable, because there is no need when you do not have a specific decision to make.
Lecture outline

- A modeling problem
- The state variables
- The decision variables
- Exogenous information
- The transition function
- The objective function
The decision variables

- How to land a plane:

  » We have to manipulate the controls of an aircraft to land it safely.
The decision variables

Where to send a plane:

» We have to find the best path to get a military transport aircraft from Los Angeles to the Philippines
The decision variables

- Where to send a plane:

» We have to manage an entire fleet of aircraft
The decision variables

There are three common notational systems for decisions:

Computer science

\[ a_t = \text{Discrete action} \]

Control theory

\[ u_t = \text{Low-dimensional continuous vector} \]

Operations research

\[ x_t = \text{Usually a discrete or continuous but high-dimensional vector of decisions}. \]
The decision variables

■ Notes
  » It is extremely important that making a decision at time \( t \) can only use information in the state \( S_t \) at time \( t \).

■ A definition of a “decision”
  » An endogenously controllable information class.

■ How do we make decisions?
  » We use policies, which are rules for making decisions. We will use notation such as:
    \[ A^\pi(s) = \text{The policy for determining an action } a \]
    \[ U^\pi(s) = \text{The policy for determining a control } u \]
    \[ X^\pi(s) = \text{The policy for determining a decision } x \]
  Here, \( \pi \) is a label that determines the type of function.
The decision variables

■ Defining decisions

» In some settings (e.g. finance), it is relatively easy to identify decisions:
  • What stock to buy?
  • How much to allocate into different assets?
  • What price to sell a stock?

» In more complex settings (business, policy), it is easier to identify goals (improve cost/service, improve health coverage), but it is not always easy to identify decisions.
  • How should Xerox improve its profits?
  • How should we increase health coverage?

We know the goals, but we do not know how…
Lecture outline

- A modeling problem
- The state variables
- The decision variables
- Exogenous information
- The transition function
- The objective function
Exogenous information

- Guessing what is happening in the real world
  » Is the stock market going (further) down?
  » What do you think?
Exogenous information

- Guessing what is happening in the real world
  » What about the price of the barrel of crude oil?
Exogenous information

Often, we need to estimate what will happen in the future to make a decision now. Let:

\[ D_t = \text{Our forecast of the customer demand in future time period } t. \]
\[ p_t = \text{Our forecast of the market price at time } t. \]

The future looks like:
\[ \{(D_1, p_1), (D_2, p_2), \ldots, (D_t, p_t)\} \]

We say that \((D_t, p_t)\) is the information arriving at time \(t\). It is sometimes useful to have a single variable to represent the new information arriving to the system at time \(t\). There is not standard notation for modeling information. Some people let:

\[ \omega_t = \text{The information arriving in time } t. \]

\(\omega_t\) represents a realization of the information that arrives in time period \(t\).
Exogenous information

- We need a system for indexing time. In particular, it is important to know the mapping between discrete and continuous time.

It is useful to think of information as arriving continuously over time. Functions (states, decisions) are measured at a point in time. At time $t$, anything $t' \leq t$ is known, anything $t' > t$ is unknown.
Exogenous information

All the information arriving would then be:

$$\omega = (\omega_1, \omega_2, ..., \omega_t, ...)$$

$\omega_t$ is not a random variable (although some people treat it as one). We sometimes need a random variable which is a function providing the information that might arrive during time period $t$. If we do not have a specific variable such as a price, quantity or demand, we can use the generic notation:

$$W_t = \text{The information arriving in time } t.$$ 

We use $W_t$ because it "looks like" $\omega_t$. We can also write:

$$\omega_t = W_t(\omega)$$
Exogenous information

For our example, we would write:

\[ W_t = (D_t, p_t) \]

In the future, we do not know what might happen. Assume that the only type of new information arriving is customer demands. That is, \( W_t = (D_t) \).

Assume that there are only 10 possible sets of demands that might happen in the future:

<table>
<thead>
<tr>
<th>( \omega )</th>
<th>D_1</th>
<th>D_2</th>
<th>D_3</th>
<th>D_4</th>
<th>D_5</th>
<th>D_6</th>
<th>D_7</th>
<th>D_8</th>
<th>D_9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>10</td>
<td>17</td>
<td>6</td>
<td>4</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>7</td>
<td>17</td>
<td>15</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>18</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>13</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>11</td>
<td>16</td>
<td>16</td>
<td>1</td>
<td>2</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>17</td>
<td>8</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>18</td>
<td>5</td>
<td>20</td>
<td>13</td>
<td>16</td>
<td>18</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>14</td>
<td>4</td>
<td>11</td>
<td>19</td>
<td>3</td>
<td>20</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>10</td>
<td>5</td>
<td>19</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>15</td>
<td>14</td>
<td>4</td>
<td>6</td>
<td>17</td>
<td>16</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

It is absolutely standard notation to index these outcomes by \( \omega \) (don't ask why). The set of outcomes (the sample space) is referred to as \( \Omega \). So an element \( \omega \in \Omega \) refers to a particular set of potential outcomes.
We would say that $D_t$ is a random variable because we do not know the demand $D_t$ right now. We might, for example, assume that $D_t$ follows some probability distribution so that we can describe the range of possible outcomes.

Sometimes we need to refer to a particular realization. For this, we let:

$D_t(\omega) = \text{A sample realization of the demand at time } t.$

$D_t(\omega)$ is not a random variable. Let's say $\omega=6$. Then,

$$D_t(6) = \begin{array}{cccccc}
\omega & D_1 & D_2 & D_3 & D_4 \\
6 & 3 & 18 & 5 & 20 & 13 & 16 & 18 & 11 & 10
\end{array}$$
Exogenous information

- At this point, we have:
  - Illustrated some actual stochastic optimization problems
  - Reviewed the core dimensions of a problem
  - Illustrated some policies
  - Hinted at how you can find the best policy.

  - But we have brushed by some concepts that deserve a little more time.
Lecture outline

- A modeling problem
- The state variables
- The decision variables
- Exogenous information
- The transition function
- The objective function
The transition function

- The transition function captures the evolution over time:

\[ S_{t+1} = S^M \left( S_t, x_t, W_{t+1} \right) \]

- The transition function goes by many names:
  - System model
  - Plant model
  - Plant equation
  - Law of motion
  - Transition law
  - Transfer function
  - “Model”
The transition function

The transition function captures the evolution over time:

\[ S_{t+1} = S^M \left( S_t, x_t, W_{t+1} \right) \]

» At time \( t \):

- \( S_t \) is known (deterministic)
- \( x_t \) is a deterministic function of \( S_t \)
- \( W_{t+1} \) is random
The transition function

- Some communities (e.g. control theory, heavily used in finance) will write this equation as:

$$S_{t+1} = S^M (S_t, x_t, W_t)$$

where at time $t$:

- $S_t$ is known (deterministic)
- $x_t$ is a deterministic function of $S_t$
- $W_t$ is random

- The reason is that this makes the most sense when modeling problems in *continuous* time.

In continuous time, $W_t$ is the information arriving between $t$ and $t + dt$. This is most naturally indexed $W_t$. 
The transition function

Illustrations

» Our deterministic shortest path problem

\[ S_t = \text{node} = i \]
Decision \( x_{ij} = 1 \)
\[ S_{t+1} = j \]

» The stochastic shortest path problem

\[ S_t = \left( R_t, (\hat{c}_{ij})_j \right) \]
Decision \( x_{ij} = 1 \)
\[ R_{t+1} = j \]
\[ S_{t+1} = \left( R_{t+1}, (\hat{c}_{jk})_k \right) \]
The transition function

Illustrations

» The first pricing problem
  • State variable
    \[ S_t = (p_t, \overline{p}_t, R_t) \]
  • Transition function
    \[ p_{t+1} = p_t + \hat{p}_{t+1} \]
    \[ \overline{p}_{t+1} = (1 - \alpha) \overline{p}_t + \alpha p_{t+1} \]
    \[ R_{t+1} = R_t - x_t \]
Lecture outline

- A modeling problem
- The state variables
- The decision variables
- Exogenous information
- The transition function
- The objective function
The objective function

\[
\min_{\pi} \mathbb{E}^\pi \left\{ \sum_{t} \gamma^t C(S_t, X^{\pi}(S_t)) \right\}
\]

- **Finding the best policy**
- **Expectation over all random outcomes**
- **State variable**
- **Cost function**
- **Decision function (policy)**
The objective function

- There are different objectives that we can:
  - Expectations
    \[
    \min_x \mathbb{E} F(x, W)
    \]
  - Risk measures
    \[
    \min_x \mathbb{E} F(x, W) + \theta \mathbb{E} \left[ F(x, W) - f_\alpha \right]_+^2
    \]
    \[
    \min_x \rho (F(x, W)) \quad \rho \in \text{Convex/coherent risk measures}
    \]
  - Worst case (“robust optimization”)
    \[
    \min_x \max_w F(x, w)
    \]
The objective function

Consider the problem of selling an asset

» State variable
\[ S_t = (R_t, p_t, \bar{p}_t) \]

» Transition function
\[
\begin{align*}
    p_{t+1} &= p_t + \hat{p}_{t+1} \\
    \bar{p}_{t+1} &= (1 - \alpha)\bar{p}_t + \alpha p_{t+1} \\
    R_{t+1} &= R_t - x_t
\end{align*}
\]

» Suggested policy
\[
X^\pi(S_t) = \begin{cases} 
1 & \text{if } p_t > \bar{p}_t + \beta \\
0 & \text{Otherwise}
\end{cases}
\]

• The policy is parameterized by \( \beta \)
The objective function

- Estimating how well we did

  » We start by defining a class of policies (as we have done). For our class, we have to find $\beta$.

  » So we could write our problem as

\[
\max_{\pi} E \sum_{t=0}^{T} \gamma^t p_t X^\pi(S_t) \quad \text{where } \pi=\beta.
\]

  » We generally cannot compute this expectation. Let’s use our policy to follow a single sample path:

\[
F^\pi(\omega) = \sum_{t=0}^{T} \gamma^t p_t(\omega) X^\pi(S_t(\omega)) \quad = \text{Discounted price at time of sale.}
\]
The objective function

- Simulated price process
The objective function

- Search for the best value of beta

<table>
<thead>
<tr>
<th>Alpha</th>
<th>Beta</th>
<th>Price</th>
<th>Smoothed</th>
<th>R</th>
<th>X</th>
<th>C</th>
<th>sum= 9.492115</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2013 W.B. Powell
The objective function

- Evaluating a policy
  - We simulate a policy $N$ times and take an average:
    \[
    \overline{F}^\pi = \frac{1}{N} \sum_{n=1}^{N} F^\pi(\omega^n)
    \]
    - If we simulate policies $\pi_1$ and $\pi_2$, we would like to conclude that $\pi_1$ is better than $\pi_2$ if
    \[
    \overline{F}^{\pi_1} > \overline{F}^{\pi_2}
    \]
  - There are some technical issues we deal with later:
    - How large should $N$ be?
    - How do we deal with the fact that this is at best a statistically noisy measurement?
      - Need to compute confidence intervals
    - How do we search over different policies?
      - Stochastic search (or “optimal learning”)