Lecture outline

- Midterm review

What we have covered

- The budgeting problem
  - » Problem assumptions
    - Single, fixed budget
    - Set of tasks/projects with known requirements
    - Function estimating the contribution from investing x_t in project *t*.
  - » Concepts:
    - Formulating the problem as a single optimization problem
      - – Helps if contributions are concave.
    - Solving the problem as a dynamic program.
      - – Requires discretizing the state variable.
    - The dynamic programming recursion
      - – What it is, how it works.
      - – What happens to your optimal solution when contribution functions are concave or convex?

What we have covered

- Economic order quantity
  - » Problem assumptions:
    - Deterministic demands
    - Instantaneous production
    - Fixed charge for positive production
  - » Concepts
    - Infinite horizon optimization
      - – Optimize cost per cycle divided by the length of the cycle
    - Understand the tradeoff between fixed cost (K) per unit time, and holding cost per unit of time
      - – At optimality, they are equal
    - Sensitivity of optimal costs to errors in model assumptions
      - – Formula for increase in cost given suboptimal order quantities.

What we have covered

- The newsvendor problem
  - » Problem assumptions
    - Two-stage process
    - Random demands
  - » Concepts
    - The contribution function
      - – Minimize cost of underage or overage
      - – Maximize expected profits
      - – *Know how to formulate the objective function for different problems*
    - The critical ratio for newsvendor with known distribution
      - – *Be able to derive this expression!*
      - – Requires knowing the demand distribution
    - The newsvendor problem and risk
      - – Optimal order quantities have more variability
    - Stochastic gradient algorithm for unknown distribution
      - – Does not require demand distribution
      - – Scaling the stepsize

What we have covered

- Censored demands
  - » Problem assumptions
    - We do not see real demand, only sales.
  - » Concepts
    - Measuring observed demand produces a downward bias.
    - We can overestimate demand to reduce the bias (but at a price).
    - Using stochastic gradients to avoid the observation problem.

What we have covered

- Modeling stochastic dynamic problems
  - » State variable S_t
    - All the information we need to compute the decision function, contribution function and transition function
    - Physical state, information state, belief state (belief state not covered yet)
  - » Decisions x_t
    - How we change the system
  - » Exogenous information W_t
    - Anything random
  - » Transition function $S_{t+1} = S^M\left(S_t, x_t, W_{t+1}\right)$
    - Given state, action and information, what is the new state
  - » Objective function
    - One period contribution $C(S_t, x_t)$
    - Objective over time: $\min_{\pi \in \Pi} E\left\{\sum_t \gamma^t C\left(S_t, X^\pi(S_t)\right)\right\}$

What we have covered

- Policies
  - » 1) Policy function approximations (PFAs)

    Let $\bar{X}(S_t)$ be a function that directly tells you an action given that you are in a state S_t.

  - » 2) Cost function approximations (CFAs)
    - Pure myopic policy: Take the action that maximizes contribution (or minimizes cost) for the current time period:

      $$X^M(S_t) = \arg\max_{x_t} C(S_t, x_t)$$

    - Use a modified cost designed to produce better long run behaviors

      $$X^M(S_t) = \arg\max_{x_t} \bar{C}^\pi(S_t, x_t)$$

  - » .

What we have covered

- Policies (cont'd)
  - » 3) Policies based on value function approximations (VFAs)

Let $\bar{V}_t(S_t)$ be an approximation of the value of being in state S_t

$$X^M(S_t) = \arg\max_{x_t}\left(C(S_t, x_t) + \gamma E\bar{V}_{t+1}(S_{t+1})\right)$$

  - » 4) Lookahead policies
    - Plan over the next T periods, but implement only the action it tells you to do now

$$X_t^{LA}(S_t) = \arg\max_{x_{tt}, x_{t,t+1}, \ldots, x_{t,t+T}} \sum_{t'=t}^{T} C(S_{t,t'}, x_{t,t'})$$

What we have covered

- Ways of approximating functions (policies or value functions)
  - » 1) Lookup tables
    - When in a (discrete) state, returns an action
    - When in a (discrete) state, returns the value of being in that state.
    - There is one value (parameter) to determine for each state.
  - » 2) Parametric models
    - A closed form, analytic function determined by one or more parameters.

$$X^M(S_t \mid \beta) = \begin{cases} 1 \text{ (sell)} & \text{If } S_t = p_t \geq \beta \\ 0 \text{ (hold)} & \text{Otherwise} \end{cases}$$

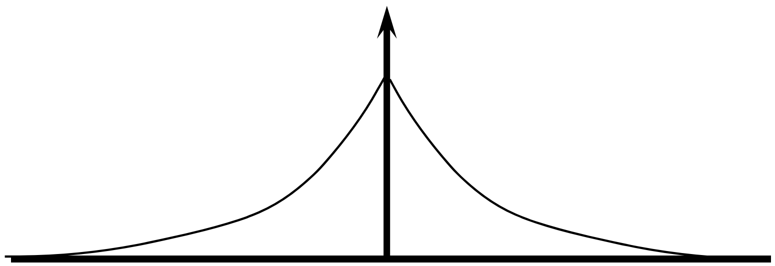
$$\bar{V}_t(S_t) = \theta_0 + \theta_1 S_t + \theta_2 (S_t)^2$$

What we have covered

- Ways of approximating functions (policies or value functions)
  - » 3) Nonparametric models (we did not cover this)

$$\bar{V}^n(s) = \frac{\sum_{i=1}^{n} \hat{v}^i k_h(s, S^i)}{\sum_{i=1}^{n} k_h(s, S^i)}$$

$$k_h(s, S^i) = \exp\left\{ -\frac{(s - S^i)^2}{h} \right\}$$



Nonparametric methods approximate a function by using a weighted sum of observations, where the weight declines with the distance between the observed state and the state we are trying to estimate.

What we have covered

- Evaluating a policy
  - » Generally, we cannot compute the expectation in

$$\min_{\pi \in \Pi} E\left\{ \sum_t \gamma^t C\left(S_t, X^\pi(S_t)\right) \right\}$$

  - » We can evaluate a policy using Monte Carlo methods:

$$\bar{F}^\pi = \frac{1}{N} \sum_{n=1}^{N} F^\pi(\omega^n)$$

  - » Then compute the variance of the estimate:

$$s^{2,\pi} = \frac{1}{N} \left(\frac{1}{N-1} \sum_{n=1}^{N} \left(F^\pi(\omega^n) - \bar{F}^\pi \right)^2 \right)$$

What we have covered

- Now construct confidence interval for the difference:
  - » $\bar{\delta} = \bar{F}^{\pi_1} - \bar{F}^{\pi_2} =$ Point estimate of difference
  - » Assume that the estimates of the value of each policy were performed independently. The variance of the difference is then
    - $s_\delta^2 = s^{2,\pi_1} + s^{2,\pi_2}$
  - » Now construct a confidence interval around the difference:
    - $\left(\bar{\delta} - z_{\alpha/2} s_\delta, \bar{\delta} + z_{\alpha/2} s_\delta\right)$

What we have covered

- Better way:
  - » Evaluate each policy using the same set of random variables (the same sample path)
  - » Compute a sample realization of the difference:
    - $$\delta(\omega) = F^{\pi_1}(\omega) - F^{\pi_2}(\omega)$$
      $$\bar{\delta} = \frac{1}{N}\sum_{n=1}^{N}\delta(\omega^n)$$
      $$s_\delta^2 = \frac{1}{N}\left(\frac{1}{N-1}\sum_{n=1}^{N}\left(\delta(\omega^n) - \bar{\delta}\right)^2\right)$$
    - Now compute confidence interval in the usual way.

What we have covered

- Optimizing continuous parameters
  - » The problem of finding the best policy can be written as a classic stochastic search problem:

$$\min_x E\left\{F(x,W)\right\}$$

  - » … where x is a vector of continuous parameters
  - » W represents all the random variables involved in evaluating a policy.

What we have covered

- We can find *x* using a classic stochastic gradient algorithm
  - » Let

$$F(x) = E\left\{F(x,W)\right\}$$

  - » Now assume that we can find the derivative with respect to each parameter in the policy (not always true). We would write this as

$$g(x,\omega) = \nabla F(x, W(\omega))$$

  - » The stochastic gradient algorithm is then

$$x^n = x^{n-1} - \alpha_{n-1} g(x^{n-1}, \omega^n)$$

  - » We then use x^n for iteration n+1 (for sample path ω^n)

What we have covered

- Notes:
  - If we are maximizing, we use

$$x^n = x^{n-1} + \alpha_{n-1} g(x^{n-1}, \omega^n)$$

  - This algorithm is provably convergent if we use a stepsize such as

$$\alpha_n = \frac{\alpha_0}{a+n-1} \qquad n = 1, 2, \ldots$$

  - Need to choose α_0 to solve the difference in units between the derivative and the parameters.

What we have covered

- Notes:
  - » Computing a gradient generally requires some insight into the structure of the problem.
  - » An alternative is to use a finite difference.
  - » Assume that *x* is a scalar. We can find a gradient using

$$g(x,\omega)=\frac{F(x+\delta,W(\omega))-F(x,W(\omega))}{\delta}$$

    - Very important: note that we are running the simulation twice using the same sample path.

What we have covered

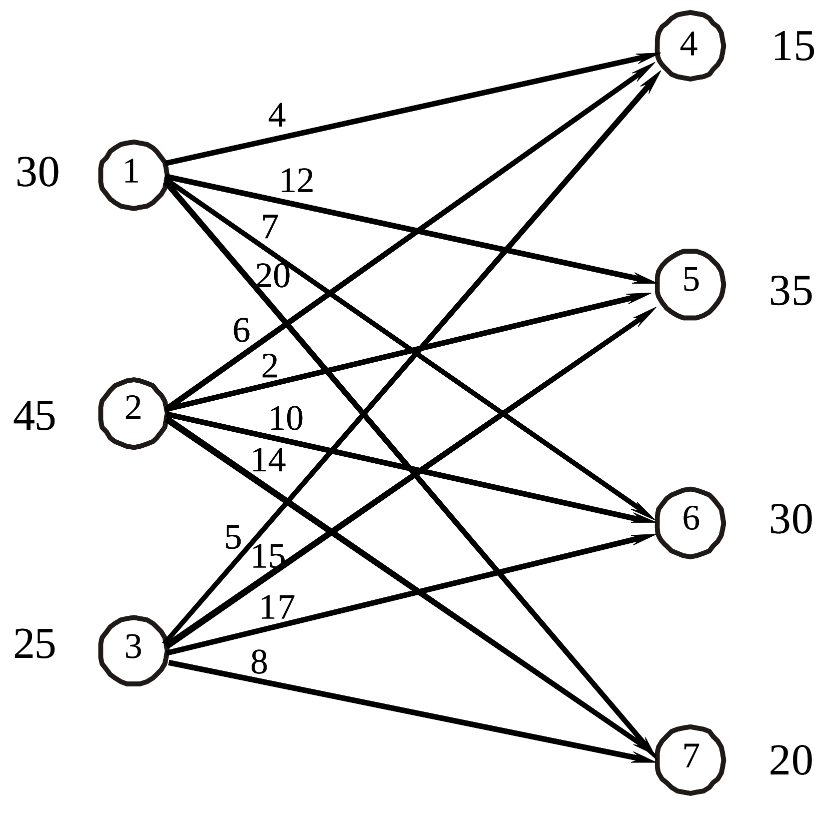
- Modeling a water reservoir
  - » Be able to define state variable, decisions, exogenous information, transition function, objective function.
  - » Be able to suggest a policy function approximation
  - » Show how to compute an optimal policy using Bellman's equation.
- Modeling mutual fund cash balance
  - » Be able to define state variable, decisions, exogenous information, transition function, objective function.
  - » Be able to suggest a policy function approximation
  - » Be able to explain the complexity of solving this problem using Bellman's equation. What is meant by the "curse of dimensionality"?

What we have covered

- Modeling natural gas forward contracts
  - » Key feature is storage – changes the entire nature of the problem.
  - » Understand how storage changes the state variable, decision variables.
- Modeling
  - » At time t, anything indexed by t is known, anything indexed by t+1 (or later) is not known at time t.
  - » Be able to define state variable, decisions, exogenous information, transition function, objective function.
  - » Be able to describe a myopic policy and a lookahead policy.

What we have covered

- Substitutable resources
  - » Know how to write out a basic "transportation problem" matching supplies to demands as a linear program.



What we have covered

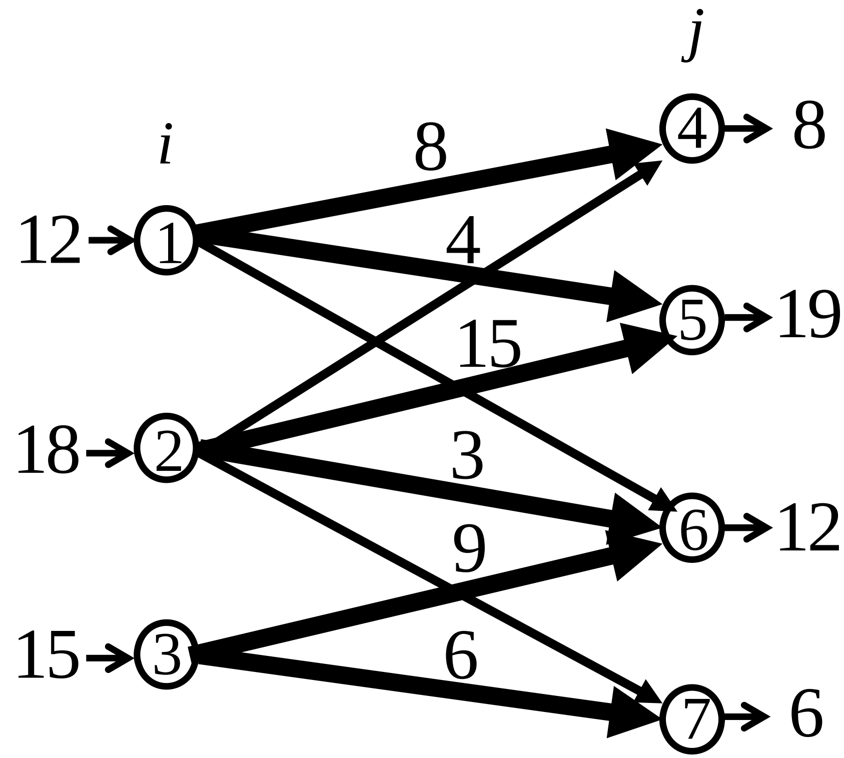
- Linear programming review
  - Understand that the *basis* is a spanning tree.
  - Be able to write out the basis matrix A^B and take its inverse.
  - Know how to calculate duals of each node and the reduced cost of a nonbasic link.
  - Know how to do a pivot.

What we have covered

- The basis matrix A^B:

$$
\begin{array}{c} \\ \end{array}
\qquad
\begin{array}{cccccc} 1-4 & 1-5 & 2-5 & 2-6 & 3-6 & 3-7 \end{array}
$$

$$
A^B = \begin{array}{c} 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{bmatrix}
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
-1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}
$$

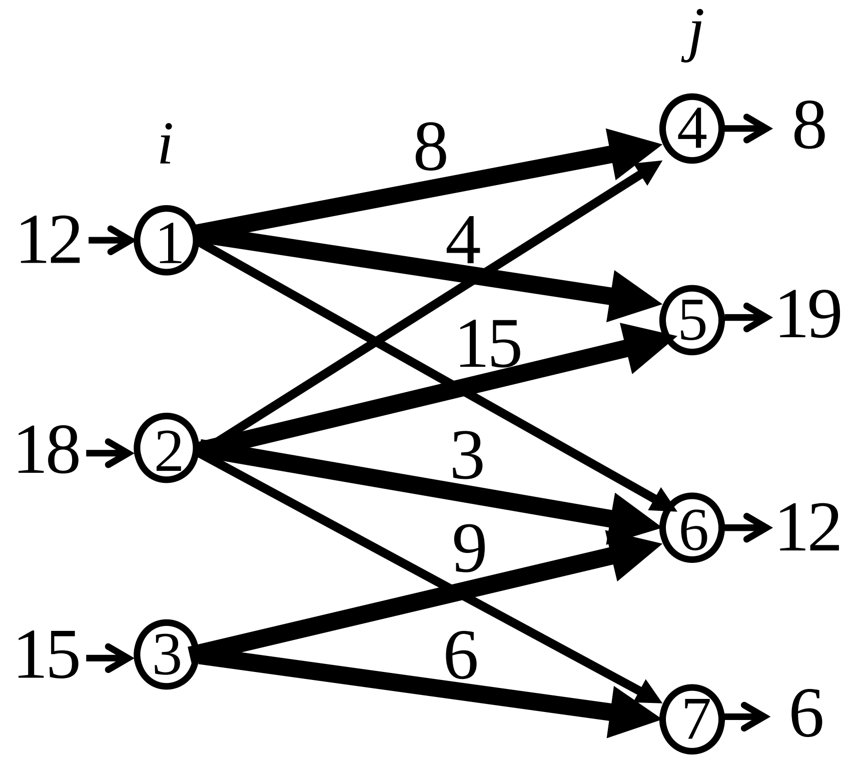


What we have covered

- The inverse

Links ↓ ← Paths →

$$
\left[A^B\right]^{-1} = \begin{array}{c} \\ 1-4 \\ 1-5 \\ 2-5 \\ 2-6 \\ 3-6 \\ 3-7 \end{array}
\begin{array}{c}
\begin{array}{cccccc} 2-1 & 3-1 & 4-1 & 5-1 & 6-1 & 7-1 \end{array} \\
\begin{bmatrix}
0 & 0 & -1 & 0 & 0 & 0 \\
-1 & -1 & 0 & -1 & -1 & -1 \\
1 & 1 & 0 & 0 & 1 & 1 \\
0 & -1 & 0 & 0 & -1 & -1 \\
0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}
\end{array}
$$

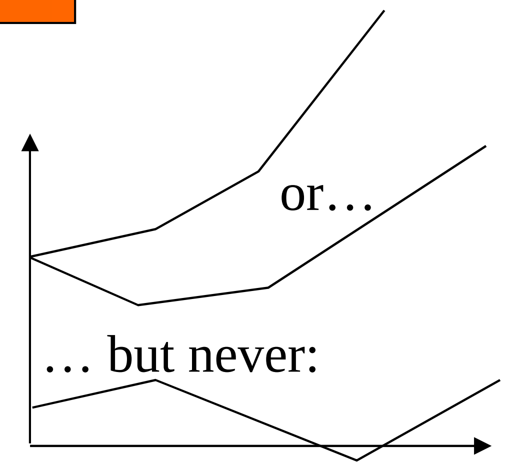


» Each column traces a path from a node to the root node.

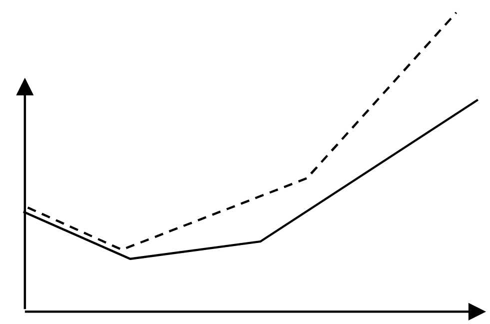
What we have covered

- Substitutable resources

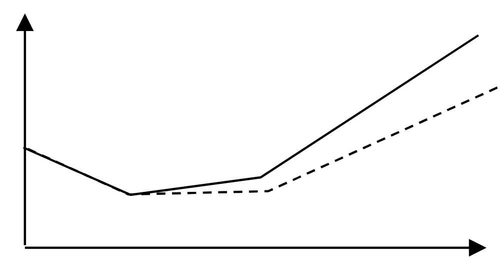
  » Understand the impact of adding more resources of a particular type.



  » How does this curve change if you increase the quantity of another type of resource?



  » How does this curve change if you increase the demand for a type of resource?



What we have covered

- Demand management
  - » *Extremely* important topic in industry. All major airlines and hotel chains use some sort of adaptive strategy to adjust prices to encourage early booking.
  - » Widely referred to as *revenue management*.
  - » Managing the *customer* is analogous to managing resources, but the knobs are different:
    - Primary control knob is pricing – lower prices encourage early booking. Use higher prices closer to the date that people need the resource.
    - Secondary knob is booking limits – control how much of the resource is available at different times/prices.

What we have covered

- Midterm rules
  - » Closed book
  - » Two sheets of paper (8 1/2 x 11, both sides, so four sides total!)