AM 121: Intro to Optimization Models and Methods

Lecture 19: Stochastic Optimization

Yiling Chen

Lesson Plan

• Two-stage stochastic optimization
  – Stage one, and stage two (recourse)
• The Farmer’s problem
• The optimal stochastic solution
• EVPI and VSS
• Analytic solution method
• Sample average approximation method

Reading: “A tutorial on stochastic programming,” Schapiro and Philpott, March 2007 (sections 1 and 2)
Stochastic Optimization

**MDP**: \( M = (S, A, P, R) \)
m states, \( n \) actions

Decision variables in the LP are \( \pi(s,a) \). Can only solve if \( m \times n \) is small.

But \( n \) may be large (e.g., we consider the “contractor’s problem” where the decision is the subset of projects to take.)

\( m \) may be very large (even infinite, e.g. realization of the time required to complete a project.)

---

Two-stage stochastic optimization:

Decision \( x \in X \) at time step 1

**External uncertainty**: some random event \( \xi \) occurs (irrespective of decision \( x \))

Decision \( y \) taken at time step 2.
Examples

- **Newsdelivery problem.** Buy $x$ papers at time 1, cost $c$. At time 2, uncertain demand $\xi$ realized and sell $\min(x, \xi)$ for price $p>c$; return $\max(x-\xi,0)$ for price $r<c$.

- **Farmer’s problem.** Plant $x$ crops at time 1. At time 2, uncertain weather $\xi$ realized and can buy and sell crops to ensure enough feed for animals, make profit.

- **Contractor problem.** Accept some projects $x$ at time 1. Each project brings a reward. At time 2, realize amount of resources $\xi$ needed for each project and can recruit additional workers at time 2 as necessary.

Two-stage stochastic optimization

```
stage 1         stage 2

x -> y(x, \xi_1) \quad \text{Scenario 1}
\quad \xi_2 \quad \quad \quad \xi_2 \quad \quad \quad \xi_3
y(x, \xi_2) \quad \text{Scenario 2}
\quad \xi_3 \quad \quad \quad \xi_3
y(x, \xi_3) \quad \text{Scenario 3}
```

initial decision  recourse decision
Example: Farmer’s problem  
\((\text{Birge \& Louveaux' 97})\)

- A farmer has 500 acres to plant wheat, grain and sugar beets. Wants to maximize profit.
- First suppose that there is no uncertainty, and the yield is exactly known to the farmer.

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield (T/acre)</td>
<td>2.5</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Cost ($/acre)</td>
<td>150</td>
<td>230</td>
<td>260</td>
</tr>
<tr>
<td>Sell price ($/T)</td>
<td>170</td>
<td>150</td>
<td>36 (under 6000 T) 10 (above 6000 T)</td>
</tr>
<tr>
<td>Purchase price ($/T)</td>
<td>238</td>
<td>210</td>
<td>-</td>
</tr>
<tr>
<td>Min requirement (T)</td>
<td>200</td>
<td>240</td>
<td>-</td>
</tr>
</tbody>
</table>

Formulation without Uncertainty

- **Stage 1:**  \(x_1, x_2, x_3 = \) acres to plant for each crop
- **Stage 2:**  
  - \(y_1^b, y_1^s = \) amount to buy, sell of crop 1
  - \(y_2^b, y_2^s = \) amount to buy, sell of crop 2
  - \(y_3^{s1}, y_3^{s2} = \) amount to sell of crop 3 at high, low price

\[
\begin{align*}
\text{max} \quad & -150x_1 - 230x_2 - 260x_3 \\
& - 238y_1^b + 170y_1^s - 210y_2^b + 150y_2^s + 36y_3^{s1} + 10y_3^{s2} \\
\text{s.t.} \quad & x_1 + x_2 + x_3 \leq 500 \quad \text{(land)} \\
& 2.5x_1 + y_1^b - y_1^s \geq 200 \quad \text{(quotas)} \\
& 3x_2 + y_2^b - y_2^s \geq 240 \\
& y_3^{s1} + y_3^{s2} \leq 20x_3 \\
& y_3^{s1} \leq 6000 \quad \text{(beets)} \\
& x_1, x_2, \ldots, y_1^b, \ldots, y_3^{s2} \geq 0
\end{align*}
\]
Optimal solution (No uncertainty)

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>acres</td>
<td>120</td>
<td>80</td>
<td>300</td>
</tr>
<tr>
<td>yield (T)</td>
<td>300</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>sales</td>
<td>100</td>
<td>-</td>
<td>6000</td>
</tr>
<tr>
<td>purchase</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Basic rationale in terms of profit per acre:
- wheat: $170(2.5) – $150 = $275
- corn: $150(3) – $230 = $220
- beets (high price): $36(20) – $260 = $460
- beets (low price): $10(20) – $260 = -$60

Farmer’s problem with uncertainty

- **Weather** may be “good” or “normal” or “poor.” This affects the yield on each crop
- Given this **uncertainty**, what is optimal **decision in stage one** about crops to plant?
- What will the farmer than do in the stage two recourse step?
Warm-up: Omniscience

Let’s first suppose the farmer can predict the weather...

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal decision if <strong>weather</strong> good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acres</td>
<td>183.3</td>
<td>66.7</td>
<td>250</td>
</tr>
<tr>
<td>yield (T)</td>
<td>550</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>sales</td>
<td>350</td>
<td>-</td>
<td>6000</td>
</tr>
<tr>
<td>purchase</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal decision if <strong>weather</strong> normal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acres</td>
<td>120</td>
<td>80</td>
<td>300</td>
</tr>
<tr>
<td>yield (T)</td>
<td>300</td>
<td>240</td>
<td>6000</td>
</tr>
<tr>
<td>sales</td>
<td>100</td>
<td>-</td>
<td>6000</td>
</tr>
<tr>
<td>purchase</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal decision if <strong>weather</strong> poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>acres</td>
<td>100</td>
<td>25</td>
<td>375</td>
</tr>
<tr>
<td>yield (T)</td>
<td>200</td>
<td>60</td>
<td>6000</td>
</tr>
<tr>
<td>sales</td>
<td>-</td>
<td>-</td>
<td>6000</td>
</tr>
<tr>
<td>purchase</td>
<td>-</td>
<td>180</td>
<td>-</td>
</tr>
</tbody>
</table>
Comparing the Omniscient solutions

<table>
<thead>
<tr>
<th></th>
<th>good</th>
<th>normal</th>
<th>poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>wheat (acres)</td>
<td>183.3</td>
<td>120</td>
<td>100</td>
</tr>
<tr>
<td>corn (acres)</td>
<td>66.7</td>
<td>80</td>
<td>25</td>
</tr>
<tr>
<td>beets (acres)</td>
<td>250</td>
<td>300</td>
<td>375</td>
</tr>
</tbody>
</table>

A lot of variation in the right thing to do: It would be best to plant between 183 and 100 acres of wheat, depending on the weather. The farmer is worried. What to do?!

Two-stage Stochastic Optimization

- Decision variables for stage two depend on the realization of the weather
  - $\xi_1$ :: good weather
  - $\xi_2$ :: normal weather
  - $\xi_3$ :: poor weather
- Suppose each uncertain event is equally likely
\[
\begin{align*}
\text{max} & \quad -150x_1 - 230x_2 - 260x_3 \\
& + \frac{1}{3} (-238y_{11}^b + 170y_{11}^s - 210y_{21}^b + 150y_{21}^s + 36y_{31}^{s1} + 10y_{31}^{s2}) \\
& + \frac{1}{3} (-238y_{12}^b + 170y_{12}^s - 210y_{22}^b + 150y_{22}^s + 36y_{32}^{s1} + 10y_{32}^{s2}) \\
& + \frac{1}{3} (-238y_{13}^b + 170y_{13}^s - 210y_{23}^b + 150y_{23}^s + 36y_{33}^{s1} + 10y_{33}^{s2}) \\
\text{s.t.} & \quad x_1 + x_2 + x_3 \leq 500 \\
& \begin{cases}
3x_1 + y_{11}^b - y_{11}^s \geq 200 \\
3.6x_2 + y_{21}^b - y_{21}^s \geq 240 \\
y_{31}^{s1} + y_{31}^{s2} \leq 24x_3 \\
y_{31} \leq 6000
\end{cases} \quad \text{stage 1} \\
& \begin{cases}
2.5x_1 + y_{12}^b - y_{12}^s \geq 200 \\
3x_2 + y_{22}^b - y_{22}^s \geq 240 \\
y_{32}^{s1} + y_{32}^{s2} \leq 20x_3 \\
y_{32} \leq 6000
\end{cases} \quad \text{stage 2: good weather} \\
& \begin{cases}
2x_1 + y_{13}^b - y_{13}^s \geq 200 \\
2.4x_2 + y_{23}^b - y_{23}^s \geq 240 \\
y_{33}^{s1} + y_{33}^{s2} \leq 16x_3 \\
y_{33} \leq 6000
\end{cases} \quad \text{stage 2: normal weather} \\
& \begin{cases}
2x_1 + y_{13}^b - y_{13}^s \geq 200 \\
2.4x_2 + y_{23}^b - y_{23}^s \geq 240 \\
y_{33}^{s1} + y_{33}^{s2} \leq 16x_3 \\
y_{33} \leq 6000
\end{cases} \quad \text{stage 2: bad weather} \\
x_1, x_2, \ldots, y_{11}^b, \ldots, y_{33}^{s2} \geq 0
\end{align*}
\]
## Optimal Solution with Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>Wheat</th>
<th>Corn</th>
<th>Beets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acres</strong></td>
<td>170</td>
<td>80</td>
<td>250</td>
</tr>
<tr>
<td><strong>Yield (T)</strong></td>
<td>510</td>
<td>288</td>
<td>6000</td>
</tr>
<tr>
<td><strong>Sales (T)</strong></td>
<td>310</td>
<td>48</td>
<td>6000</td>
</tr>
<tr>
<td><strong>Purchase (T)</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Yield (T)</strong></td>
<td>425</td>
<td>240</td>
<td>5000</td>
</tr>
<tr>
<td><strong>Sales (T)</strong></td>
<td>225</td>
<td>-</td>
<td>5000</td>
</tr>
<tr>
<td><strong>Purchase (T)</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Yield (T)</strong></td>
<td>340</td>
<td>192</td>
<td>4000</td>
</tr>
<tr>
<td><strong>Sales (T)</strong></td>
<td>140</td>
<td>-</td>
<td>4000</td>
</tr>
<tr>
<td><strong>Purchase (T)</strong></td>
<td>-</td>
<td>48</td>
<td>-</td>
</tr>
</tbody>
</table>

Profit: $108,390

### without uncertainty:

<table>
<thead>
<tr>
<th></th>
<th>good</th>
<th>normal</th>
<th>poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>wheat (acres)</td>
<td>183.3</td>
<td>120</td>
<td>100</td>
</tr>
<tr>
<td>corn (acres)</td>
<td>66.7</td>
<td>80</td>
<td>25</td>
</tr>
<tr>
<td>beets (acres)</td>
<td>250</td>
<td>300</td>
<td>375</td>
</tr>
</tbody>
</table>

### stochastic solution

- Wheat: 170 acres, expected yield 183.3, sales 120, purchase 100
- Corn: 80 acres, expected yield 66.7, sales 80, purchase 25
- Beets: 250 acres, expected yield 250, sales 300, purchase 375

### omniscient solution

- Wheat: 170 acres, expected yield 183.3, sales 120, purchase 100
- Corn: 80 acres, expected yield 66.7, sales 80, purchase 25
- Beets: 250 acres, expected yield 250, sales 300, purchase 375

### in expectation solution

- Wheat: 170 acres, expected yield 183.3, sales 120, purchase 100
- Corn: 80 acres, expected yield 66.7, sales 80, purchase 25
- Beets: 250 acres, expected yield 250, sales 300, purchase 375

### Expected value of Perfect Information (EVPI)

- $114,667 - $108,390 = $6,277

### Value of Stochastic Solution (VSS)

- $108,390 - $107,240 = $1,150
EVPI and VSS

• Let $V^*(\xi)$ denote value of optimal solution for scenario $\xi$
• Let $Q(x,\xi)$ denote value of optimal recourse decision given $x$ and $\xi$
• EVPI: expected value of perfect information
  \[ E_\xi[V^*(\xi)] - \max_x (c^T x + E_\xi[Q(x,\xi)]) \]
• VSS: expected value of stochastic solution
  \[ \max_x (c^T x + E_\xi[Q(x,\xi)]) - (c^T \bar{x} + E_\xi[Q(\bar{x},\xi)]) \]
  where $\bar{x}$ is the solution assuming $\xi = E_\xi[\xi]$

Stochastic Programming

• First stage decision $x \in \mathbb{R}^n$. Scenario $\xi=(q,T,W,h)$ defines data for second-stage problem.
• $y \in \mathbb{R}^m$ second-stage decision

\[
\begin{align*}
\max_x & \quad c^T x + E_\xi[Q(x,\xi)] \\
\text{s.t.} & \quad A x \leq b \\
& \quad x \geq 0
\end{align*}
\]

where
\[
Q(x,\xi) = \max_y q^T y
\]
\[
\begin{align*}
\text{s.t.} & \quad T x + W y \leq h \\
& \quad y \geq 0
\end{align*}
\]

First-stage problem:
maximize the sum profit of the first stage and the expected profit of the second stage.

Second-stage problem:
given $x$ and realized scenario $\xi=(q,T,W,h)$, maximize profit in second stage.
Example: Second-stage problem for Farmer’s problem

• Only $T$ matrix is uncertain in the farmer’s problem
• Let $t_i$ denote the yield of crop $i$; scenario $\xi = (t_1, t_2, t_3)$
• For a given scenario, the recourse problem is:

$\begin{align*}
Q(x, \xi) &= \max -238y_1^b + 170y_1^s - 210y_2^b + 150y_2^s + 36y_3^{s_1} + 10y_3^{s_2} \\
\text{s.t.} & \quad t_1x_1 + y_1^b - y_1^s \geq 200 \\
& \quad t_2x_2 + y_2^b - y_2^s \geq 240 \\
& \quad y_3^{s_1} + y_3^{s_2} \leq t_3x_3 \\
& \quad y_3^{s_1} \leq 6000
\end{align*}$

Computational approaches

• (1) small number of scenarios $\xi$, can enumerate to form a single LP and solve (e.g., farmer’s problem.)

• (2) infinite set of possible scenarios $\xi$, but can solve for $E_{\xi}[Q(x, \xi)]$ \textbf{analytically}. Do this, and solve first-stage decision analytically (e.g., newsdelivery problem.)

• (3) no analytic solution available, and cannot enumerate all $\xi$. Adopt the \textbf{sample average approximation (SAA)} (e.g., contractor’s problem.)
Computational approaches

• (1) **small number of scenarios** $\xi$, can enumerate to form a single LP and solve (e.g., farmer’s problem.)

• (2) infinite set of possible scenarios $\xi$, but can solve for $E_{\xi}[Q(x,\xi)]$ **analytically**. Do this, and solve first-stage decision analytically (e.g., newsdelivery problem.)

• (3) no analytic solution available, and cannot enumerate all $\xi$. Adopt the **sample average approximation (SAA)** (e.g., contractor’s problem.)

Analytic approach: Example

• Student group giving away a Harvard-Yale T-shirt
• Must decide how many to order $x$ at cost $c$. Demand $\xi$ uncertain; per-unit cost $b$ for “backorder cost” if $\xi>x$, per-unit holding cost $h$ if $x>\xi$.
• $c=1.0$, $b=1.5$, $h=0.1$, $\xi \sim U(0,100)$

• Solve: $\min cx + E_{\xi}[Q(x,\xi)] = \min cx + W(x)$
  
  $Q(x,\xi) = b \max(\xi-x,0) + h \max(x-\xi,0)$

• Can solve for $x^*$ by first-order optimality: $c+W'(x)=0$
Example: \( cx + Q(x, \xi) \) for \( \xi = 50 \)

Plot \( cx + Q(x, \xi) = 1x + 1.5 \max(50-x,0) + 0.1 \max(x-50,0) \)

- Let \( f \) denote prob density function on \( \xi \); \( F \) denote CDF.
- \( W(x) = E_\xi[b \max(\xi-x,0)] + E_\xi[h \max(x-\xi,0)] \)
- \( d/dx E_\xi[\max(\xi-x,0)] = d/dx \int_0^\infty (\xi-x)f(\xi)d\xi = -\int_0^\infty f(\xi)d\xi = -\Prob(\xi \geq x) \)
- \( d/dx E_\xi[\max(x-\xi,0)] = d/dx \int_0^\xi (x-\xi)f(\xi)d\xi = \int_0^\xi f(\xi)d\xi = \Prob(\xi \leq x) \)
- \( W'(x) = -b(1-F(x)) + hF(x) \)
- Set \( c - W'(x) = 0 \): need \( F(x) = (b-c/b+h) \)
- Solve, get \( x^* \approx F^{-1}(0.3125) \)
- If \( \xi \sim U(0,100) \); optimal \( x^* \approx 31.25 \).
• \( W(x) = W(0) + \int_0^x W'(z) \, dz \)
  \[ = b \, E_\xi [\xi] + \int_0^x [-b + (b+h)F(z)] \, dz \]
  \[ = b \, E_\xi [\xi] - bx + (b+h) \int_0^x F(z) \, dz \]

• \( cx + W(x) = bE_\xi [\xi] + (c-b)x + (b+h) \int_0^x F(z) \, dz \)
  \[ = (1.5)(50) + (-0.5)x + 1.6 \int_0^x z/100 \, dz \]
  \[ = 75 - 0.5x + 0.008x^2 \]
Solving stochastic programs

• (1) small number of scenarios $\xi$, can enumerate to form a single LP and solve (e.g., farmer’s problem.)

• (2) infinite set of possible scenarios $\xi$, but can solve for $E_{\xi}[Q(x,\xi)]$ analytically. Do this, and solve first-stage decision analytically (e.g., newsdelivery problem.)

• (3) no analytic solution available, and cannot enumerate all $\xi$. Adopt the sample average approximation (SAA) (e.g., contractor’s problem.)

Sample Average Approximation (Kleywegt et al. 2001)

• In practice, a closed form solution is rarely available.

• Can instead approximate $E_{\xi}[Q(x, \xi)]$ by sampling K scenarios, with

$$E_{\xi}[Q(x, \xi)] \approx \sum_{k=1}^{K} p_k Q(x, \xi_k)$$

where $p_k = 1/K$ is the prob of scenario k.

• Proceed as earlier, solve:

$$\min_{x,y} \quad c^T x + \sum_k p_k q_k^T y_k$$

s.t. $T_k x + W_k y_k \leq h_k$ for all $k$
Example: T-shirt problem

- Consider two demand scenarios
  - $\xi_1 = 20$, $\xi_2 = 80$
  - Objective: $cx + E_{\xi}[Q(x, \xi)] = cx + \frac{1}{2} Q(x, \xi_1) + \frac{1}{2} Q(x, \xi_2)$
- Solve as an LP

- Can also consider three (or more!) scenarios: $\xi_1 = 20$, $\xi_2 = 50$, $\xi_3 = 80$ with prob 2/5, 1/5 and 2/5.

Example: Approximating $cx + E_{\xi}[Q(x, \xi)]$

2 point approx

3 point approx
The Sample Average Approximation Method

(Kleywegt et al. 2001)

• Sample K scenarios $\xi_1, \ldots, \xi_K$ and solve
  
  \[
  \min_x \left[ c^T x + \frac{1}{K} \sum_k Q(x, \xi_k) \right]
  \]

  • Let $\hat{x}$ denote this solution

SAA: Theoretical properties

(Kleywegt et al. 2001)

• Let $\hat{\mu}_K$ denote expected value of solution $\hat{x}_K$ to SAA, based on K samples
• Let $S_k^d$ denote all solutions “close” to $\hat{x}_K$ (within distance d)
• Let $V^*$ denote expected value of optimal solution $x^*$
• Let $S^d$ denote all solutions “close” to optimal $x^*$ (within distance d)

Theorem:

• $\lim_{K \to \infty} \hat{\mu}_K = V^*$.  
• $S_k^d \subseteq S^d$ with probability 1 as $K \to \infty$, for any distance d.
Applications

• Contractor: Set of projects, each pay an amount, but take uncertain time. May need to take-on workers as necessary in recourse. Which set of projects to take on?

• Airline crew scheduling: Assign crew to routes, but routes take uncertain time. May need to pay over-time.

Empirical Results (Contractor)

Contractor’s problem with 1000 experiments; and 20 projects.
Summary

• Two-stage stochastic optimization

• Solve first stage in anticipation of distribution over scenarios. Maximize expected value.

• Second stage decision is the “recourse” decision. Made with knowledge of first stage decision and realized scenario.

• EVPI and VSS as measures of value of stochastic optimization

• Analytical method; and SAA method.