IE 495 – Lecture 15

Bounds in Stochastic Programming

Prof. Jeff Linderoth

March 5, 2003

Outline

- Review
  - ◇ Jensen's Lower Bound
  - ◇ What the @$\hat{!}$&*$\hat{!}$*! went wrong last time
- Upper Bounds
  - ◇ Edmundson-Madansky
- Using bounds in the LShaped Algorithm

Review

- *Why* do we care about bounds in stochastic programming?
- What's "wrong" with numerical integration for evaluating $\mathcal{Q}(x)$?
- What is Jensen's inequality?
- How is Jensen's Inequality used in stochastic programming?

Jensen's Inequality in Stochastic Programming

$$\mathbb{E}_\omega[Q(\hat{x}, \omega)] \geq Q(\hat{x}, \mathbb{E}_\omega[\omega])$$

- We get a *tight* lower bound on $\mathcal{Q}(\hat{x})$ by evauating $Q(\hat{x}, \bar{\omega})$.
- In our proof, we only used the fact that $Q(\hat{x}, \omega)$ was *convex* on Ω.
- So, in general, if ϕ is a convex function ω of a random variable over its support Ω, then

$$\mathbb{E}_\omega \phi(\omega) \geq \phi(\mathbb{E}_\omega(\omega))$$

A Recourse Formulation

minimize

$$\mathcal{Q}(x_1, x_2) = x_1 + x_2 + 5 \int_{\omega_1=1}^{4} \int_{\omega_2=1/3}^{2/3} y_1(\omega_1, \omega_2) + y_2(\omega_1, \omega_2) d\omega_1 d\omega_2$$

subject to

$$\begin{aligned} \omega_1 x_1 + x_2 + y_1(\omega_1, \omega_2) &\geq 7 \\ \omega_2 x_1 + x_2 + y_2(\omega_1, \omega_2) &\geq 4 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \\ y_1(\omega_1, \omega_2) &\geq 0 \\ y_2(\omega_1, \omega_2) &\geq 0 \end{aligned}$$

What Good Is This Stuff

- The real trick is that you recursively partitioning the region Ω and the bounds become tighter and tighter.

Let $\mathcal{S} = \{\Omega^l, l = 1, 2, \ldots v\}$ be some partition of Ω. Do you believe me that

$$\mathbb{E}_\omega[Q(\hat{x}, \omega)] \geq \sum_{l=1}^{v} P(\omega \in \Omega^l) Q(\hat{x}, \mathbb{E}_\omega(\omega | \omega \in \Omega^l))$$

Where We Left Off

- Let's (lower) bound $\mathcal{Q}(2, 2)$.
- \$1 Prize! First person who can tell me what was wrong about what we did last time...
- We'll set it right...

Edmundson-Madansky Inequality

- The Edmundson-Madansky Inequality is a way in which to get an upper bound of $Q(\hat{x}, \omega)$.
- Consider now the $1 - D$ case, where $\omega \in \Omega$

Your picture here

Edmundson-Madansky Inequality

- Since $Q(\hat{x},\omega)$ is convex in ω, the line segment between $(a, Q(\hat{x},a))$ and $(b, Q(\hat{x},b))$ is $\geq Q(\hat{x},\omega)\ \forall \omega \in \Omega$.
- Let $U(\hat{x},\omega)$ be the formula for this line segment.

$$U(\hat{x},\omega) = \frac{Q(\hat{x},b) - Q(\hat{x},a)}{b-a}(\omega - a) + Q(\hat{x},a).$$

With some algebra,

$$U(\hat{x},\omega) = \frac{Q(\hat{x},b) - Q(\hat{x},a)}{b-a}\omega + \frac{b}{b-a}Q(\hat{x},a) - \frac{a}{b-a}Q(\hat{x},b)$$

Edmundson-Madansky Inequality

$$
\begin{aligned}
\mathbb{E}_\omega U(\hat{x},\omega) &= \frac{Q(\hat{x},b)-Q(\hat{x},a)}{b-a}\mathbb{E}_\omega[\omega] + \frac{b}{b-a}Q(\hat{x},a) - \frac{a}{b-a}Q(\hat{x},b) \\
&= Q(\hat{x},a)\frac{b-\mathbb{E}_\omega[\omega]}{b-a} + Q(\hat{x},b)\frac{\mathbb{E}_\omega[\omega]-a}{b-a}.
\end{aligned}
$$

$$\text{Define } p = \frac{\bar{\omega}-a}{b-a}$$

- We get an upper bound (for an arbitrary distribution over Ω) by replacing with a two-point discrete distribution, where
  - $P(\omega = a) = 1 - p$
  - $P(\omega = b) = p$

It's a Lid!

- If Ω has a (finite) set of extreme points $\text{ext}(\Omega)$, then there is some probability measure p_i (convex multipliers) so that

$$\sum_{e \in \text{ext}(\Omega)} p_{e_i} Q(\hat{x}, e) \geq Q(\hat{x}, \omega).$$

Extending E-M to $\Omega \subseteq \Re^d$

- Extending to multiple dimensions, if Ω is a "hyper-rectangle",
  - $\Omega = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d]$
- and if the random variables in the d dimensions are independent...

$$U(\hat{x}, \omega) = \sum_{e \in \text{ext}(\Omega)} \prod_{i=1}^{d} \left(\frac{|\bar{\omega}_i - e_i|}{b_i - a_i} \right) Q(\hat{x}, e) \geq \mathbb{E}_\omega Q(w, \omega) = \mathcal{Q}(x)$$

Let's try it!

(ω_1, ω_2)	Value
(1, 1/3)	
(1, 1)	
(4, 1/3)	
(4, 1)	

$$0.25(v_1 + v_2 + v_3 + v_4) = ???$$

Partitioning

- Note that we *still* may need to evaluate $\mathcal{Q}(x)$ at 2^d points—Just to get an initial upper bound.
- The bound gets progressively tighter by doing the same "conditioning" trick as we did for our lower bound.
- We just "throw in" the extra points/evaluations as necessary in the algorithm.
- We'll do an example...

Using Bounds in LShaped Method

- All this rigamarole was just to get bounds on *ONE* point $\mathcal{Q}(\hat{x})$.
- We really want to optimize over $\mathcal{Q}(\hat{x})$
- Key idea of LShaped method is to underestimate $\mathcal{Q}(x)$ by one (or more) auxiliary variables θ.
- To use bounds, use the θ to underestimate a lower bounding function $\mathcal{Q}_L(x)$. (We called this $\mathbb{E}_\omega L(x,\omega)$ on Monday).

Bounds in LShaped Method

- Use the LShaped method to optimize the problem using $\mathcal{Q}_L(x)$.
  - Only include $\bar{\omega}$ "scenarios".
- When optimized with respect to $\mathcal{Q}_L(x)$, compare to $\mathcal{Q}_U(x)$ (what we called $\mathbb{E}_\omega U(x, \omega)$).
- If $\mathcal{Q}_U(x) - \mathcal{Q}_L(x)$ is "sufficiently small". Stop.
- Otherwise, refine the partition (which improves the bounds), and repeat.

Bounds in the LShaped Method

- An example here — time permitting.

Next Time

Happy Spring Break!

- Monte Carlo Methods
- Stochastic Decomposition