IE 495 – Lecture 18

Stochastic Decomposition

Prof. Jeff Linderoth

March 26, 2003

Outline

- Review—Monte Carlo Methods
- “Interior” Sampling Methods
  - Stochastic Decomposition

Some Questions

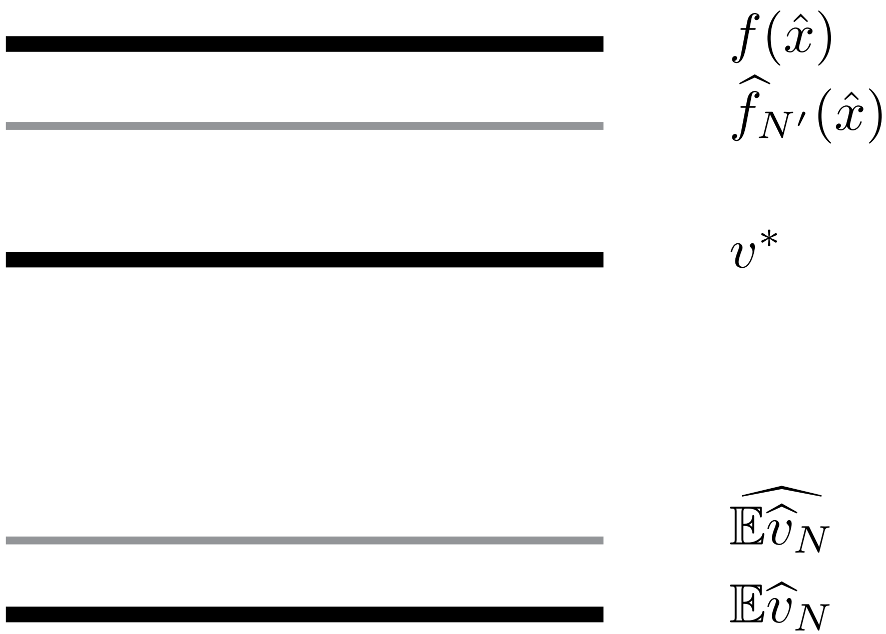
$$v^* = \min_{x \in S} \{ f(x) \equiv \mathbb{E}_P g(x; \xi) \}$$

- How can I use sampling to get a (statistical) lower bound on v^*?
- How can I use sampling to get a (statictical) upper bound on v^*?
- When do you think you will ever get your homeworks back?

Putting it all together

- $\widehat{f}_N(x)$ is the sample average function
  - Draw $\omega^1, \ldots \omega^N$ from P
  - $\widehat{f}_N(x) \equiv N^{-1} \sum_{j=1}^{N} g(x, \omega^j)$
  - $\mathbb{E}[\widehat{f}_N(x)] = f(x) \geq v^*$
- $\widehat{v}_N$ is the optimal solution value for the sample average function:
  - $\widehat{v}_N \equiv \min_{x \in S} \left\{ \widehat{f}_N(x) := N^{-1} \sum_{j=1}^{N} g(x, \omega^j) \right\}$
  - $\mathbb{E}[\widehat{v}_N] \leq v^*$
- Estimate $\mathbb{E}(\widehat{v}_N)$ as $\widehat{\mathbb{E}(\widehat{v}_N)} = L_{N,M} = M^{-1} \sum_{j=1}^{M} \widehat{v}_N^j$
  - Solve M stochastic LP's, each of sampled size N.

The Gap



$f(\hat{x})$
$\widehat{f}_{N'}(\hat{x})$
v^*
$\widehat{\mathbb{E}\widehat{v}_N}$
$\mathbb{E}\widehat{v}_N$

Sampling Methods

- "Interior" sampling methods.
- Sample during the course of the algorithm
  - ◇ LShaped Method (Dantzig and Infanger)
  - ◇ Stochastic Decomposition (Higle and Sen)
  - ◇ Stochastic Quasi-gradient methods (Ermoliev)
- "Exterior" sampling methods
  - ◇ Sample. Then solve problem approximating problem.
  - ◇ Can we get (statistical) bounds on key solution quantities?

Stochastic Decomposition

- Designed to solve 2-stage stochastic LP with fixed recourse
- Assume randomness only in objective function
- Assume relatively complete recourse and $\mathcal{Q}(x) \geq 0$.
- $X = \{\Re^n_+ | Ax = b\}$

$$\min_{x \in X} \mathcal{Q}(x) \equiv \mathbb{E}_P Q(x,\omega) = \int_\Omega Q(x,\omega) dP(\omega)$$

$$Q(x,\omega) = \min_{y \in \Re^p_+} \{q^T y : Wy = h(\omega) - T(\omega)x\}$$

Key Ideas–1

- Do sampling within the optimization
- Take one new sample per iteration
- A bit like "Using Bounds in the LShaped Method"
- There, we used Jensen's Inequality to give a lower bound.
  - ◇ There the lower bound was based on using a limited number of samples (namely the mean) and conditioning (to be in a region of Ω to get progressively tigther lower bounds).
- Here, we just randomly draw a sample, and we produce a lower bound that is valid in expectation only.

Key Ideas–2

- Method for “approximately” solving lots of LPs fast
- Cuts “phase out” as they get older. (Since they were based on less sampled information).

Bounds in LShaped Method—Review

- 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 $\mathcal{Q}_U(x)$, if we have one
- If $\mathcal{Q}_U(x) - \mathcal{Q}_L(x)$ is "sufficiently small". Stop.
- Otherwise, refine the partition (which improves the bounds), and repeat.

More Detail

Partition form of Jensen's inequality...

- Let $\mathcal{S} = \{\Omega^l, l = 1, 2, \ldots v\}$ be some partition of Ω:

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

Lower Bounding-Based LShaped Method

0. Let $l = 1, \Omega_l = \Omega, k = 1, x^k = 0$

1. For $j = 1, 2, \ldots l$ solve...

$$\pi_j = \arg \max_{\pi \in \Re^m} \{\pi^T (h(\bar{\omega}_j) - T(\bar{\omega}_j)x^k) | W^T \pi \leq q\}$$

$$\bar{\omega}_j = \mathbb{E}_P(\omega | \omega \in \Omega_j)$$

2. Create cut as...

$$\theta \geq \sum_{j=1}^{l} \pi_j^T \left[(h(\bar{\omega}_j) - T((\bar{\omega}_j)x^k) \right] - \sum_{j=1}^{l} \pi_j^T \left[\pi_j^T T((\bar{\omega}_j) \right] (x - x_k)$$

$$(p_j = \mathrm{P}(\omega \in \Omega_j).$$

Continuing

3. Let $k = k + 1$, add cut to master problem and solve

$$(x_k, \theta_k) = \arg\min_{x \in X} c^T x + \theta$$

subject to

$$\theta \geq \alpha_k + \beta_k x \quad \forall k = 1, 2, \ldots$$

4. $\mathcal{L}(x_k) = c^T x_k + \theta_x \leq \mathcal{Q}(x)$. If you are happy with x_k based on $\mathcal{L}(x_k)$, then quit, otherwise create a new partition of $\Omega = \{\Omega_1, \Omega_2, \ldots, \Omega_l\}$. Go to 1.

(I'll draw some pictures in class)...

Stochastic Decomposition

0. Let $k = 1$, $x_k = 0$, $V = \emptyset$

1a. Draw random sample ω_k, and solve...

$$\pi_k = \arg \max_{\pi \in \Re^m} \{\pi^T (h(\bar{\omega}_k) - T((\bar{\omega}_k)x^k) | W^T \pi \leq q\}$$

1b. $V = V \cup \pi^k$. For $j = 1, 2, \ldots k - 1$, solve

$$\pi^j = \arg \max_{\pi \in V} \{\pi^T ((h(\bar{\omega}_j) - T((\bar{\omega}_j)x^k)\}$$

Stochastic Decomposition

2a. Create cut as...

$$\theta \geq 1/k \sum_{j=1}^{k} \pi_j^T (h(\omega_j) - T(\omega_j)x_k)$$

- Call the cut $(\alpha_k + \beta_k^T x)$.

2b. For $j = 1, 2, \ldots, k-1$, *Phase Out* old cuts as

$$\alpha_k + \beta_k^T x = (k-1)/k(\alpha_{k-1} + \beta_{k-1}^T x).$$

Stochastic Decomposition

3. Solve Master Problem

$$(x_k, \theta_k) = \arg\min_{x \in X} c^T x + \theta$$

subject to

$$\theta \geq \alpha_k + \beta_k x \quad \forall k = 1, 2, \ldots$$

- Go to 1.

⋆ There is some *subsequence* of the $x^k \to x^*$

- Typically people use some sort of statistical based stopping criteria

LP Duality

- By duality

$$Q(x,\omega) = \max_{\pi \in \Re^m} \{\pi^T (h(\omega) - T(\omega)x) | \pi^T W \leq q$$

- So if $\hat{\pi}$ satisfies $\hat{\pi}^T W \leq q$, then

$$Q(x,\omega) \geq \hat{\pi}^T (h(\omega) - T(\omega)x) \quad \forall \omega \in \Omega, \forall x \in X$$

? (Why?)

⇒ The lower bounding function $\mathcal{L}(x)$ is (only) such that $\mathbb{E}_P \mathcal{L}(x) \leq \mathcal{Q}(x)$

Checkpoint! – What we've learned

- Modeling the Deterministic Equivalent of Stochastic (Linear) Programs
  - Two-stage
  - Multi-stage. (Modeling our favorite eight-syllable word).
- EVPI
- VSS
- Recourse Function

We Really Learned This Much?

- Two-Stage Stochastic LP With Recourse
  - ◇ Properties of the recourse function
    - Convex
    - Subdifferentiable
    - Lipschitz-Continuous
  - ◇ LShaped Method
    - Feasibility Cuts
    - Multicut methods
    - Regularized methods. Trust region and regularized decomposition.
    - Bunching and Trickling Down

WOW!

- ◇ Bounds
  - Jensen's Inequality
  - Edmunson-Madansky Inequality
- Monte Carlo Methods
  - ◇ Valid Statistical Lower and Upper Bounds on optimal objective function value
  - ◇ Variance Reduction (Latin Hypercube Sampling)
  - ◇ Convergence of Optimal Solutions

Math Stuff We've Learned

- KKT Conditions
- Convexity
- (Lipschitz)-Continuity
- Minkowski Sum
- Weak-Strong Long of Large Numbers
- Central Limit Theorem
- Lebesgue-Stiltjes Integral
- L'Hopital's Rule

Comouter Stuff We've Learned

- AMPL! Lots and lots of AMPL.
- Numerical Integration
- MPS format
- SMPS
- High Performance Computing
  - SMP
  - Message Passing
- Grid Computing
  - Condor

Survey Time

- A simple show of hands please...
- What Next?
  - Stochastic MIP
  - Probabilistic Constraints?
  - Multi-Stage Stochastic LP?

- Next Time...
  - Go over old homeworks
  - Get Homework "#5–#6"

Remainder of Course

- Start calling for people's projects to be done *soon!*
- Only 8 more classes left. Last 3–4 classes will be (interspersed) with project presentations $\Rightarrow$ you have roughly 2–3 weeks to finish.
- There will be 1.5 more assignments

$\Rightarrow$ There is *LOTS* of work that you have left to do for this course!