15.094J: Robust Modeling, Optimization, Computation

Lecture 4: RLO: Probabilistic Guarantees

February 2015

Outline

1. Guarantees for independent uncertainty
2. Guarantees for non-independent distributions
3. Philosophical Remarks

Objectives Today

- Probabilistic Guarantees for RLO
- Insights in selecting parameters

Row-wise Ellipsoidal uncertainty

- RO:

$$\begin{array}{rl} \max & c'x \\ \text{s.t.} & \max_{a_i \in \mathcal{U}_i} a_i'x \leq b_i. \\ & x \geq \mathbf{0}. \end{array}$$

- $\mathcal{U}_i = \{a_i |\ a_i = \overline{a}_i + \boldsymbol{\Delta}_i' u_i,\ ||u_i||_2 \leq \rho\}$, $\boldsymbol{\Delta}_i :\ k_i \times n,\ u_i : k_i \times 1.$
- RC:

$$\begin{array}{rl} \max & c'x \\ \text{s.t.} & \overline{a}_i'x + \rho||\boldsymbol{\Delta}_i x||_2 \leq b_i, \qquad i = 1, \ldots, m. \\ & x \geq \mathbf{0}. \end{array}$$

Probabilistic Guarantee

- Suppose u_i are independent, have zero mean and have support in $[-1, 1]$.
- Suppose that x satisfies $\bar{a}'x + \rho||\mathbf{\Delta}x|| \leq b$.
- Then

$$P(\tilde{a}'x > b) \leq e^{-\rho^2/2}.$$

- **Remarkable property**: Independent of the distributions of u (we do not even require identical distributions).
- How to select ρ: Suppose our tolerance for infeasibility is ϵ, that is $P(\tilde{a}'x > b) \leq \epsilon$.
- Use $\epsilon = e^{-\frac{\rho^2}{2}}$, select $\rho = \sqrt{2\log\left(\frac{1}{\epsilon}\right)}$.

ϵ	ρ
10^{-6}	5.25
10^{-5}	4.79
10^{-4}	4.29
10^{-3}	3.71
10^{-2}	3.03
10^{-1}	2.14

Proof from First Principles

- Let $X(\boldsymbol{\xi}) = w_0 + \sum_{i=1}^{k} w_i \xi_i$, where ξ_i are independent with zero mean and with support in $[-1, 1]$.
- Let $w = (w_1, \ldots, w_k)'$. We will first show that

$$P(X(\boldsymbol{\xi}) > 0) = P\left(w_0 + \sum_{i=1}^{k} w_i \xi_i > 0 \right) \leq exp\left(-\frac{w_0^2}{2||w||^2} \right).$$

$$P(X(\boldsymbol{\xi}) > 0) = \int \chi(X(\boldsymbol{\xi}))\, dP(\boldsymbol{\xi}), \quad \chi(s) = \begin{cases} 0, & s \leq 0, \\ 1, & s > 0 \end{cases}$$

- Note that $\chi(s) \leq \gamma(s) = e^s$.
- Let $\alpha > 0$. Note also that $\chi(s) = \chi(\alpha \cdot s) \leq \gamma(\alpha \cdot s)$.
-

$$P(X(\boldsymbol{\xi}) > 0) \leq E[exp(\alpha w_0 + \sum_{i=1}^{k} \alpha w_i \xi_i)] = exp(\alpha w_0) \prod_{i=1}^{k} E[exp(\alpha w_i \xi_i)].$$

Proof continued

- For every random variable ξ with zero mean and support in $[-1, 1]$

$$E[e^{t\xi}] \leq e^{t^2/2}.$$

- Let $f(s) = e^{ts} - \frac{e^t - e^{-t}}{2} s$.
- $f(s)$ convex in s. Maximum in $[-1, 1]$ is at endpoint.
- $\max_{|s| \leq 1} f(s) = f(1) = f(-1) = \frac{e^t + e^{-t}}{2}$.

$$\begin{aligned} E[e^{t\xi}] &= \int f(s)\, dP(s) \quad \text{[zero mean]} \\ &\leq \max_{|s|\leq 1} f(s) \\ &= \frac{e^t + e^{-t}}{2} \\ &\leq e^{t^2/2} \quad \text{[Taylor series].} \end{aligned}$$

Proof continued

- For all $\alpha > 0$:

$$\begin{aligned} P(X(\boldsymbol{\xi}) > 0) &\leq exp(\alpha w_0) \prod_{i=1}^{k} E[exp(\alpha w_i \xi_i)] \\ &\leq exp\left(\alpha w_0 + \frac{\alpha^2}{2} \sum_{i=1}^{k} w_i^2\right). \end{aligned}$$

- Select α to minimize the upper bound.
-

$$P(X(\boldsymbol{\xi}) > 0) \leq \min_{\alpha > 0} exp\left(\alpha w_0 + \frac{\alpha^2}{2} ||w||^2\right).$$

- $\alpha^* = -w_0/||w||^2$.
- $P(X(\boldsymbol{\xi}) > 0) = P\left(w_0 + \sum_{i=1}^{k} w_i \xi_i > 0\right) \leq exp\left(-\frac{w_0^2}{2||w||^2}\right)$.

Proof of the key guarantee

- Suppose that x satisfies $\overline{a}'x + \rho||\boldsymbol{\Delta}x|| \leq b$.
- Then

$$P(\tilde{a}'x > b) = P(\overline{a}'x + u'\boldsymbol{\Delta}x > b) \leq P(u'\boldsymbol{\Delta}x > \rho||\boldsymbol{\Delta}x||).$$

- Select $w_0 = -\rho||\boldsymbol{\Delta}x||$ and $w = \boldsymbol{\Delta}x$, we obtain

$$P(\tilde{a}'x > b) \leq e^{-\frac{\rho^2}{2}}.$$

Guarantees for non-independent distributions

- RO:

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \tilde{A}x \leq b \\ & x \in P \\ & \forall \tilde{A} \in \mathcal{U} = \left\{ \tilde{A} \mid ||M(\text{vec}(\tilde{A}) - \text{vec}(\overline{A}))|| \leq \Delta \right\}. \end{aligned}$$

- RC:

$$\begin{aligned} \max \quad & c'x \\ \text{s.t.} \quad & \overline{a}_i x + \Delta ||M^{-1} x_i||^* \leq b_i, \quad i = 1, \ldots, m \\ & x \in P, \end{aligned}$$

- $\text{vec}(\tilde{A}) \sim (\text{vec}(\overline{A}), \boldsymbol{\Sigma})$.
- Let $M = \boldsymbol{\Sigma}^{-\frac{1}{2}}$.

Probabilistic Guarantees

- $$P\left(\tilde{a}_i' x^* > b_i\right) \leq \frac{1}{1 + \Delta^2 \left(\frac{\|\boldsymbol{\Sigma}^{\frac{1}{2}} x_i^*\|^*}{\|\boldsymbol{\Sigma}^{\frac{1}{2}} x_i^*\|_2} \right)^2}.$$
- If L_p norm used in $\mathcal{U}$, then
$$P\left(\tilde{a}_i' x^* > b_i\right) \leq \frac{1}{1 + \Delta^2 \min\left\{ \frac{1}{p^2}, \frac{1}{n} \right\}}.$$
- If L_2 used in $\mathcal{U}$, then
$$P\left(\tilde{a}_i' x^* > b_i\right) \leq \frac{1}{1 + \Delta^2}.$$
- Remark: Arbitrary Dependence structure.
- How to select Δ?

Proof

- Optimal robust solution x_i^* satisfies

$$(\text{vec}(\overline{A}))' x_i^* + \Delta ||\boldsymbol{\Sigma}^{\frac{1}{2}} x_i^*||^* \leq b_i,$$

- Thus

$$P\left((\text{vec}(\tilde{A}))' x_i^* > b_i \right) \leq P\left((\text{vec}(\tilde{A}))' x_i^* \geq (\text{vec}(\overline{A}))' x_i^* + ||\boldsymbol{\Sigma}^{\frac{1}{2}} x_i^*||^* \right).$$

- Bertsimas and Popescu: if S is a convex set, and $\tilde{X} \sim (\overline{X}, \boldsymbol{\Sigma})$, then

$$P\left(\tilde{X} \in S \right) \leq \frac{1}{1 + d^2},$$

where

$$d^2 = \inf_{\tilde{X} \in S} \left(\tilde{X} - \overline{X} \right)' \boldsymbol{\Sigma}^{-1} \left(\tilde{X} - \overline{X} \right).$$

Proof continued

- $S_i = \left\{ \text{vec}(\tilde{A}) \mid (\text{vec}(\tilde{A}))' x_i \geq (\text{vec}(\overline{A}))' x_i + \Delta ||\mathbf{\Sigma}^{\frac{1}{2}} x_i||^* \right\}$.
- $d_i^2 = \inf_{\text{vec}(\tilde{A}) \in S_i} \left(\text{vec}(\tilde{A}) - \text{vec}(\overline{A}) \right)' \mathbf{\Sigma}^{-1} \left(\text{vec}(\tilde{A}) - \text{vec}(\overline{A}) \right)$.
- Optimal solution (KKT):

$$\text{vec}(\overline{A}) + \Delta \left(\frac{||\mathbf{\Sigma}^{\frac{1}{2}} x_i||^*}{||\mathbf{\Sigma}^{\frac{1}{2}} x_i||_2} \right)^2 \mathbf{\Sigma} x_i,$$

- $$d^2 = \Delta^2 \left(\frac{||\mathbf{\Sigma}^{\frac{1}{2}} x_i||^*}{||\mathbf{\Sigma}^{\frac{1}{2}} x_i||_2} \right)^2 .$$

On the interplay of probability and optimization

- Use Probability theorems to select parameters.
- Use optimization ideas to find best possible results in probability.
- In exercise we will explore other bounds.
- Use RO to solve problems under uncertainty computationally.