15.094J: Robust Modeling, Optimization, Computation

Lectures 2: The New Primitives: Uncertainty Sets

Outline

Linear Optimization

- **Nominal problem**

$$\begin{array}{rl} \text{maximize} & c_1x_1 + \ldots + c_nx_n \\ \text{subject to} & a_{11}x_1 + \ldots + a_{1n}x_n \leq b_1 \\ & \vdots \\ & a_{m1}x_1 + \ldots + a_{mn}x_n \leq b_m \\ & x_i \geq 0. \end{array}$$

- **Robust Problem**

$$\begin{array}{rll} \text{maximize} & c_1x_1 + \ldots + c_nx_n & \\ \text{subject to} & a_{11}x_1 + \ldots + a_{1n}x_n \leq b_1 & \\ & \vdots & \forall(a_{11}, \ldots, a_{mn}) \in \mathcal{U} \\ & a_{m1}x_1 + \ldots + a_{mn}x_n \leq b_m & \\ & x_i \geq 0. & \end{array}$$

- **What if $(c_1, \ldots, c_n, b_1, \ldots, b_m)$ are also uncertain?**

Robust modeling

- Replace probability distributions as primitives with *uncertainty sets*.
- Use *worst case* analysis, while bounding the power of nature - *Robust Optimization (RO).*
- Use *conclusions of probability theory*, and not its (Kolmogorov (1933)) axioms, to define uncertainty sets.

Constructing Uncertainty Sets

How do we construct uncertainty sets?

- We first suggest an approach based on the Central Limit Theorem by building on our previous example.
- Later, we'll see some more advanced approaches.

Constructing Uncertainty Sets: A First Try

Recall our previous example:

Project	1	2	3	4
Expected Cost	120	100	180	140
St Dev of Cost	12	10	18	14

- Although the cost of building a factory won't be exactly equal to its mean, we expect it to be close.
- Suggests Uncertainty set

$$-\Gamma \leq \frac{a_i - \overline{a}_i}{\sigma_i} \leq \Gamma.$$

Constructing Uncertainty Sets: A First Try (continued)

We might then solve:

$$\begin{array}{ll} \text{maximize} & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{subject to} & a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 \leq 500 \text{ for all } \mathbf{a} \in \mathcal{U} \\ & x_i \text{ integer.} \end{array}$$

where $\mathbf{a} \in \mathcal{U}$ means

$$-\Gamma \leq \frac{a_1 - 120}{12} \leq \Gamma$$
$$-\Gamma \leq \frac{a_2 - 100}{10} \leq \Gamma$$
$$\text{etc.}$$

- Notice that we insist the solution be feasible for *any* value of the costs. (In particular, the *worst* values.)
- By varying the value of Γ, (e.g. $\Gamma = 2, 3$), we control level of robustness.

A Criticism of our first set $\mathcal{U}$

- A fair criticism of this approach is that it is unlikely that all the costs are at their worst value at the same time.
- Suggests we should pick a better uncertainty set.
- Suggestions?

Central Limit Theorem

- X_i: iid with mean μ and standard deviation σ.
- Recall our old friend, the CLT

$$\frac{X_1 + \ldots + X_n - n\mu}{\sigma \cdot \sqrt{n}} \sim N(0, 1).$$

- Intuitively, the CLT tells us that sums of random variables tend to be close to their mean.
- How might we use this to improve our uncertainty set?

Constructing Uncertainty Sets: CLT

- Uncertainty set

$$-\Gamma\sqrt{n} \leq \sum_{i=1}^{n} \frac{a_i - \bar{a}_i}{\sigma_i} \leq \Gamma\sqrt{n}$$

- Sum of variations from mean values is limited.
- Γ still controls the degree of robustness.
- For $\Gamma = 2$, $\mathbb{P}[\mathbf{a} \in \mathcal{U}] \sim 0.95$.
- For $\Gamma = 3$, $\mathbb{P}[\mathbf{a} \in \mathcal{U}] \sim 0.997$.

Later, we'll see some numerical examples of how the above sets affect solution quality.

Modeling Correlation Information

- Factor model : $\{\tilde{z}_i\}_{i=1,\dots,n}$ depend on m factors $\tilde{\boldsymbol{f}} = \left(\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_m\right)$

$$\tilde{z}_i = \mathbf{a}_i' \cdot \tilde{\boldsymbol{f}} + \tilde{\epsilon}_i,$$

  $\{\tilde{\epsilon}_i\}$ are i.i.d.

-

$$\mathcal{U}^{\text{Corr}} = \left\{ (z_1, \dots, z_n) \;\middle|\; \begin{array}{ll} z_i = \displaystyle\sum_{j=1}^{m} a_{ij} f_j + \epsilon_i, & \forall i = 1, \dots, n, \\ -\Gamma_f \le \dfrac{\displaystyle\sum_{j=1}^{m} f_j - m \cdot \mu_f}{\sigma_f \cdot \sqrt{m}} \le \Gamma_f, & \\ -\Gamma_\epsilon \le \dfrac{\displaystyle\sum_{i=1}^{n} \epsilon_i - n \cdot \mu_\epsilon}{\sigma_\epsilon \cdot \sqrt{n}} \le \Gamma_\epsilon. & \end{array} \right\}.$$

Typical Sets: Incorporating Distributional Information

- Shannon (1948) introduced the idea of Typical Sets:
- Property (a): A typical set has probability nearly 1.
- Property (b): All elements of typical set are nearly equiprobable.
- Given pdf $f(\cdot)$,

$$\mathcal{U}^{f-\text{Typical}} = \left\{ (z_1, \ldots, z_n) \left| -\Gamma \leq \frac{\sum_{i=1}^{n} \log f(z_i) - n \cdot \mu_f}{\sigma_f \cdot \sqrt{n}} \leq \Gamma. \right. \right\},$$

$$\mu_f = \int_{-\infty}^{\infty} f(x) \log f(x) \, \mathrm{d}x,$$

$$\sigma_f = \int_{-\infty}^{\infty} f(x) \left(\log f(x) - \mu_{\log f} \right)^2 \mathrm{d}x.$$

Theorem

- (a) $\mathbb{P}\left[\tilde{\mathbf{z}} \in \mathcal{U}^{\text{f-Typical}}\right] \to g(\Gamma) = 2\Phi(\Gamma) - 1$, as $n \to \infty$.
- (b) The conditional pdf $h(\tilde{\mathbf{z}}) = f(\tilde{\mathbf{z}} | \tilde{\mathbf{z}} \in \mathcal{U}^{\text{f-Typical}})$ satisfies:

$$\left| \frac{1}{n} \log h(\tilde{\mathbf{z}}) - \mu_f \right| \le \epsilon_n,$$

  with $\epsilon_n \to 0$, as $n \to \infty$.
- $\tilde{u}_j = \log f(\tilde{z}_j)$, iid. Apply CLT: as $n \to \infty$,

$$\frac{\sum_{j=1}^{n} \tilde{u}_j - n\mu_f}{\sigma_f \cdot \sqrt{n}} \sim N(0,1),$$

- Let $\tilde{\mathbf{z}} \in \mathcal{U}^{\text{f-Typical}}$.

$$h(\tilde{\mathbf{z}}) = f(z_1) f(z_2) \dots f(z_n).$$

- Since $\tilde{\mathbf{z}} \in \mathcal{U}^{\text{f-Typical}}$, we have

$$\left| \frac{1}{n} \log h(\tilde{\mathbf{z}}) - \mu_f \right| = \left| \frac{1}{n} \sum_{j=1}^{n} \log f(z_j) - \mu_f \right| \le \frac{\Gamma \cdot \sigma_f}{\sqrt{n}} \to 0,$$

Typical Sets

- $\tilde{z}_i \sim N(0, \sigma)$

$$\mathcal{U}_\epsilon^{\mathrm{G}} = \left\{ \mathbf{z} \,\middle|\, -\Gamma_\epsilon^{\mathrm{G}} \le \|\mathbf{z}\|^2 - n\sigma^2 \le \Gamma_\epsilon^{\mathrm{G}} \right\}.$$

- $\tilde{z}_i \sim Exp(\lambda)$

$$\mathcal{U}_\epsilon^{\mathrm{E}} = \left\{ \mathbf{z} \,\middle|\, \frac{n}{\lambda} - \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}} \le \sum_{j=1}^n z_j \le \frac{n}{\lambda} + \frac{\sqrt{n}}{\lambda} \cdot \Gamma_\epsilon^{\mathrm{E}}, \ \mathbf{z} \ge \mathbf{0} \right\}.$$

- $\tilde{z}_i \sim U[a, b]$

$$\mathcal{U}_\epsilon^{U} = \left\{ \mathbf{z} \,\middle|\, \begin{array}{c} n\frac{a+b}{2} - \Gamma_\epsilon^U \sqrt{n} \le \sum_{j=1}^n z_j \le n\frac{a+b}{2} + \Gamma_\epsilon^U \sqrt{n}, \\ a \le z_j \le b, \ j = 1, \ldots, n, \end{array} \right\}..$$

- $\tilde{z}_i \sim \mathrm{Bin}(p)$

$$\mathcal{U}_\epsilon^{B} = \left\{ \mathbf{z} \,\middle|\, \begin{array}{c} np - \Gamma_\epsilon^B \sqrt{n} \le \sum_{j=1}^n z_j \le np + \Gamma_\epsilon^B \sqrt{n}, \\ z_j \in \{0, 1\}, \ j = 1, \ldots, n, \end{array} \right\}.$$

Insights

Recall our previous capacity expansion problem.

- Nominal problem

$$\begin{array}{rl} \text{maximize} & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{subject to} & 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 500 \\ & x_i \text{ integer.} \end{array}$$

- Nominal Solution: $x_1^* = 4$, $x_2^* = x_3^* = x_4^* = 0$.

Market	1	2	3	4
Expected Cost	120	100	180	140
St Dev of Cost	12	10	18	14

We will compare the solution of the nominal problem, the problem with our naive uncertainty set, and our CLT based uncertainty set in terms of their feasibility and optimality.

Uncertainty Sets

- **Robust 1:**

$$\mathcal{U} = \{-\Gamma \leq \frac{a_i - \bar{a}_i}{\sigma_i} \leq \Gamma\}$$

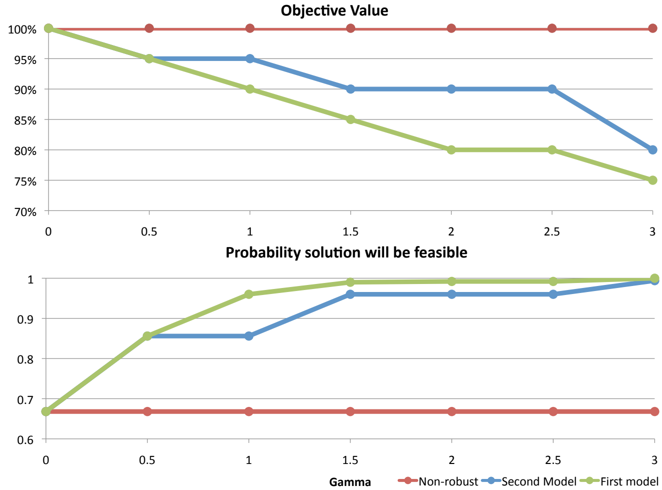
- **Robust 2:**

$$\mathcal{U} = \{-\Gamma\sqrt{n} \leq \sum_{i=1}^{n} \left| \frac{a_i - \bar{a}_i}{\sigma_i} \right| \leq \Gamma\sqrt{n}\}$$

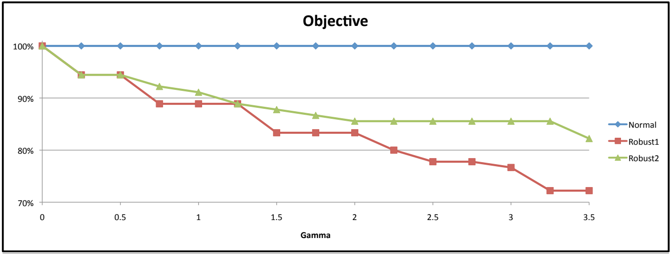
Robust Solution

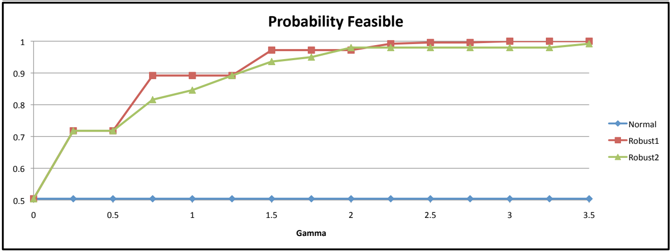
Problem	Γ	Objective	x_1	x_2	x_3	x_4	Unused Capital
Nominal	0	200	4	0	0	0	20
Robust 1	0.5	190	3	1	0	0	40
Robust 1	1	180	2	2	0	0	60
Robust 1	1.5	170	1	3	0	0	80
Robust 1	2	160	0	4	0	0	100
Robust 1	2.5	160	0	4	0	0	100
Robust 1	3	150	3	0	0	0	140
Robust 2	0.5	190	3	1	0	0	40
Robust 2	1	190	3	1	0	0	40
Robust 2	1.5	180	2	2	0	0	60
Robust 2	2	180	2	2	0	0	60
Robust 2	2.5	180	2	2	0	0	60
Robust 2	3	160	2	0	1	0	80

Tradeoff of robustness and optimality



10 variables





Modeling Demand

- We collected historical data: d_t. $t = 1, \dots, T$.
- D_t, is future demand for day $t = 1, \dots, n$.
- Compute $\mu = \dfrac{\sum_{t=1}^{T} d_t}{T}$.
- $\sigma^2 = \dfrac{\sum_{t=1}^{T}(d_t - \mu)^2}{T - 1}$.
- $U = \{(D_1, \dots, D_n) | \; -\Gamma \cdot \sigma \cdot \sqrt{n} \le \sum\limits_{t=1}^{n}(D_t - \mu) \le \Gamma \cdot \sigma \cdot \sqrt{n}, \; |D_t - \mu| \le \Gamma_1 \sigma\}$.

Insights

- Key intuition: Model uncertainty by conclusions of probability not its axioms.
- Often by sacrificing a bit of optimality, we can ensure feasibility for a large range of uncertainty.
- The price of robustness is often not great.