15.093 Optimization Methods

Lecture 8: Robust Optimization

1 Papers

SLIDE 1

- B. and Sim, The Price of Robustness, Operations Research, 2003.
- B. and Sim, Robust Discrete optimization, Mathematical Programming, 2003.

2 Structure

SLIDE 2

- Motivation
- Data Uncertainty
- Robust Mixed Integer Programming
- Robust 0-1 Programming
- Robust Approximation Algorithms
- Robust Network Flows
- Experimental Results
- Summary and Conclusions

3 Motivation

SLIDE 3

- The classical paradigm in optimization is to develop a model that assumes that the input data is precisely known and equal to some nominal values. This approach, however, does not take into account the influence of data uncertainties on the quality and feasibility of the model.
- Can we design solution approaches that are immune to data uncertainty, that is they are robust?

SLIDE 4

- Ben-Tal and Nemirovski (2000):

> In real-world applications of Linear Programming (Net Lib library), one cannot ignore the possibility that a small uncertainty in the data can make the usual optimal solution completely meaningless from a practical viewpoint.

3.1 Literature

SLIDE 5

- Ellipsoidal uncertainty; Robust convex optimization Ben-Tal and Nemirovski (1997), El-Ghaoui et. al (1996)
- Flexible adjustment of conservativism
- Nonlinear convex models
- Not extendable to discrete optimization

4 Goal

SLIDE 6

Develop an approach to address data uncertainty for optimization problems that:

- It allows to control the degree of conservatism of the solution;
- It is computationally tractable both practically and theoretically.

5 Data Uncertainty

SLIDE 7

$$\begin{array}{rll} \text{minimize} & \boldsymbol{c}'\boldsymbol{x} & \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} & \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} & \\ & x_i \in \mathcal{Z}, & i = 1, \ldots, k, \end{array}$$

WLOG data uncertainty affects only $\boldsymbol{A}$ and $\boldsymbol{c}$, but not the vector $\boldsymbol{b}$. SLIDE 8

- **(Uncertainty for matrix $\boldsymbol{A}$)**: $a_{ij}, \ j \in J_i$ is independent, symmetric and bounded random variable (but with unknown distribution) $\tilde{a}_{ij}, j \in J_i$ that takes values in $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.
- **(Uncertainty for cost vector $\boldsymbol{c}$)**: $c_j, \ j \in J_0$ takes values in $[c_j, c_j + d_j]$.

6 Robust MIP

SLIDE 9

- Consider an integer $\Gamma_i \in [0, |J_i|]$, $i = 0, 1, \ldots, m$.
- Γ_i adjusts the robustness of the proposed method against the level of conservativeness of the solution.
- Speaking intuitively, it is unlikely that all of the $a_{ij}, \ j \in J_i$ will change. We want to be protected against all cases that up to Γ_i of the a_{ij}'s are allowed to change.

SLIDE 10

- Nature will be restricted in its behavior, in that only a subset of the coefficients will change in order to adversely affect the solution.

- We will guarantee that if nature behaves like this then the robust solution will be feasible deterministically. Even if more than Γ_i change, then the robust solution will be feasible with very high probability.

6.1 Problem

SLIDE 11

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}'\boldsymbol{x} + \max_{\{S_0|\ S_0 \subseteq J_0, |S_0| \leq \Gamma_0\}} \left\{ \sum_{j \in S_0} d_j |x_j| \right\} \\ \text{subject to} \quad & \sum_j a_{ij} x_j + \max_{\{S_i|\ S_i \subseteq J_i, |S_i| \leq \Gamma_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} |x_j| \right\} \leq b_i, \quad \forall i \\ & \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u} \\ & x_i \in \mathcal{Z}, \qquad \forall i = 1, \dots k. \end{aligned}$$

6.2 Theorem 1

SLIDE 12

The robust problem can be reformulated has an equivalent MIP:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}'\boldsymbol{x} + z_0 \Gamma_0 + \textstyle\sum_{j \in J_0} p_{0j} & \\ \text{subject to} \quad & \sum_j a_{ij} x_j + z_i \Gamma_i + \sum_{j \in J_i} p_{ij} \leq b_i & \forall i \\ & z_0 + p_{0j} \geq d_j y_j & \forall j \in J_0 \\ & z_i + p_{ij} \geq \hat{a}_{ij} y_j & \forall i \neq 0, j \in J_i \\ & p_{ij}, y_j, z_i \geq 0 & \forall i, j \in J_i \\ & -y_j \leq x_j \leq y_j & \forall j \\ & l_j \leq x_j \leq u_j & \forall j \\ & x_i \in \mathcal{Z} & i = 1, \dots, k. \end{aligned}$$

6.3 Proof

SLIDE 13

Given a vector $\boldsymbol{x}^*$, we define:

$$\beta_i(\boldsymbol{x}^*) = \max_{\{S_i|\ S_i \subseteq J_i, |S_i| = \Gamma_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} |x_j^*| \right\}.$$

This equals to:

$$\begin{aligned} \beta_i(\boldsymbol{x}^*) = \max \quad & \sum_{j \in J_i} \hat{a}_{ij} |x_j^*| z_{ij} \\ \text{s.t.} \quad & \sum_{j \in J_i} z_{ij} \leq \Gamma_i \\ & 0 \leq z_{ij} \leq 1 \qquad \forall i, j \in J_i. \end{aligned}$$

SLIDE 14

Dual:

$$\begin{aligned} \beta_i(\boldsymbol{x}^*) = \min \quad & \sum_{j \in J_i} p_{ij} + \Gamma_i z_i & \\ \text{s.t.} \quad & z_i + p_{ij} \geq \hat{a}_{ij} |x_j^*| & \forall j \in J_i \\ & p_{ij} \geq 0 & \forall j \in J_i \\ & z_i \geq 0 & \forall i. \end{aligned}$$

$\lvert J_i \rvert$	Γ_i
5	5
10	8.3565
100	24.263
200	33.899

Table 1: Choice of Γ_i as a function of $|J_i|$ so that the probability of constraint violation is less than 1%.

6.4 Size

SLIDE 15

- Original Problem has n variables and m constraints
- Robust counterpart has $2n + m + l$ variables, where $l = \sum_{i=0}^{m} |J_i|$ is the number of uncertain coefficients, and $2n + m + l$ constraints.

6.5 Probabilistic Guarantee

6.5.1 Theorem 2

SLIDE 16

Let $\boldsymbol{x}^*$ be an optimal solution of robust MIP.
(a) If $\boldsymbol{A}$ is subject to the model of data uncertainty U:

$$\Pr\left(\sum_j \tilde{a}_{ij} x_j^* > b_i\right) \le \frac{1}{2^n}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\},$$

$n = |J_i|$, $\nu = \frac{\Gamma_i + n}{2}$ and $\mu = \nu - \lfloor\nu\rfloor$; bound is tight.
(b) As $n \to \infty$

$$\frac{1}{2^n}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{n}{l} + \mu \sum_{l=\lfloor\nu\rfloor+1}^{n}\binom{n}{l}\right\} \sim 1 - \Phi\left(\frac{\Gamma_i - 1}{\sqrt{n}}\right).$$

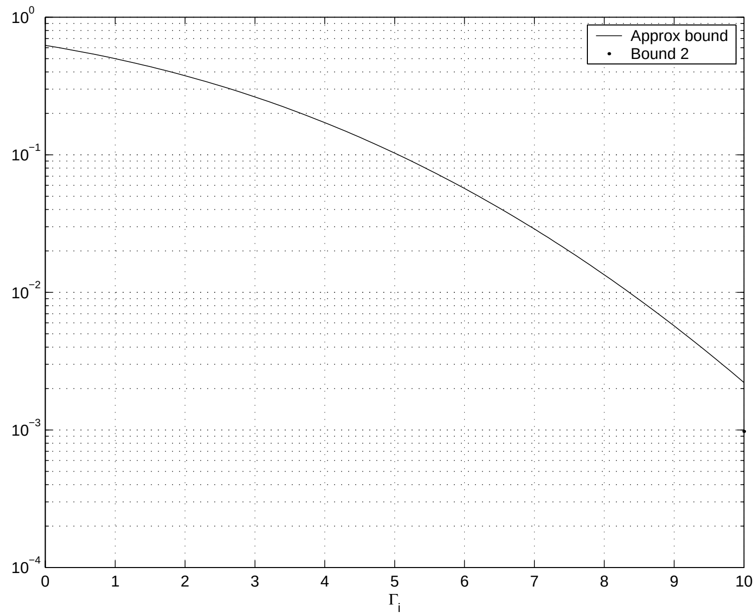
SLIDE 17
SLIDE 18

7 Experimental Results

7.1 Knapsack Problems

SLIDE 19

- $$\begin{aligned} \text{maximize} \quad & \sum_{i\in N} c_i x_i \\ \text{subject to} \quad & \sum_{i \in N} w_i x_i \le b \\ & \boldsymbol{x} \in \{0,1\}^n. \end{aligned}$$



Γ	Violation Probability	Optimal Value	Reduction
0	0.5	5592	0%
2.8	4.49×10^{-1}	5585	0.13%
36.8	5.71×10^{-3}	5506	1.54%
82.0	5.04×10^{-9}	5408	3.29%
200	0	5283	5.50%

- $\tilde{w}_i$ are independently distributed and follow symmetric distributions in $[w_i - \delta_i, w_i + \delta_i]$;
- $\boldsymbol{c}$ is not subject to data uncertainty.

7.1.1 Data

Slide 20

- $|N| = 200$, $b = 4000$,
- w_i randomly chosen from $\{20, 21, \ldots, 29\}$.
- c_i randomly chosen from $\{16, 17, \ldots, 77\}$.
- $\delta_i = 0.1 w_i$.

7.1.2 Results

SLIDE 21

8 Robust 0-1 Optimization

SLIDE 22

- Nominal combinatorial optimization:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x} \in X \subset \{0,1\}^n. \end{aligned}$$

- Robust Counterpart:

$$\begin{aligned} Z^* = \quad \text{minimize} \quad & \boldsymbol{c}'\boldsymbol{x} + \max_{\{S|\ S\subseteq J, |S|=\Gamma\}} \sum_{j\in S} d_j x_j \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned}$$

- WLOG $d_1 \geq d_2 \geq \ldots \geq d_n$.

8.1 Remarks

SLIDE 23

- Examples: the shortest path, the minimum spanning tree, the minimum assignment, the traveling salesman, the vehicle routing and matroid intersection problems.
- Other approaches to robustness are hard. Scenario based uncertainty:

$$\begin{aligned} \text{minimize} \quad & \max(\boldsymbol{c_1}'\boldsymbol{x}, \boldsymbol{c_2}'\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in X. \end{aligned}$$

  is NP-hard for the shortest path problem.

8.2 Approach

SLIDE 24

$$\begin{aligned} \text{Primal}: Z^* = \min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + \max \quad & \sum_j d_j x_j u_j \\ \text{s.t.} \quad & 0 \leq u_j \leq 1, \qquad \forall\, j \\ & \sum_j u_j \leq \Gamma \end{aligned}$$

$$\begin{aligned} \text{Dual}: Z^* = \min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + \min \quad & \theta\Gamma + \sum_j y_j \\ \text{s.t.} \quad & y_j + \theta \geq d_j x_j, \qquad \forall\, j \\ & y_j, \theta \geq 0 \end{aligned}$$

8.3 Algorithm A

SLIDE 25

- Solution: $y_j = \max(d_j x_j - \theta, 0)$
-
$$Z^* = \min_{\boldsymbol{x} \in X, \theta \geq 0} \theta \Gamma + \sum_j \left(c_j x_j + \max(d_j x_j - \theta, 0) \right)$$
- Since $X \subset \{0,1\}^n$,
$$\max(d_j x_j - \theta, 0) = \max(d_j - \theta, 0)\ x_j$$
-
$$Z^* = \min_{\boldsymbol{x} \in X, \theta \geq 0} \theta \Gamma + \sum_j \left(c_j + \max(d_j - \theta, 0) \right) x_j$$

SLIDE 26

- $d_1 \geq d_2 \geq \ldots \geq d_n \geq d_{n+1} = 0$.
- For $d_l \geq \theta \geq d_{l+1}$,
$$\min_{\boldsymbol{x} \in X, d_l \geq \theta \geq d_{l+1}} \theta \Gamma + \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{l} (d_j - \theta) x_j =$$
$$d_l \Gamma + \min_{\boldsymbol{x} \in X} \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{l} (d_j - d_l) x_j = Z_l$$
-
$$Z^* = \min_{l=1,\ldots,n+1} d_l \Gamma + \min_{\boldsymbol{x} \in X} \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{l} (d_j - d_l) x_j.$$

8.4 Theorem 3

SLIDE 27

- Algorithm A correctly solves the robust 0-1 optimization problem.
- It requires at most $|J| + 1$ solutions of nominal problems. Thus, If the nominal problem is polynomially time solvable, then the robust 0-1 counterpart is also polynomially solvable.
- Robust minimum spanning tree, minimum assignment, minimum matching, shortest path and matroid intersection, are polynomially solvable.

9 Experimental Results

9.1 Robust Sorting

SLIDE 28

$$\begin{array}{ll} \text{minimize} & \sum_{i \in N} c_i x_i \\ \text{subject to} & \sum_{i \in N} x_i = k \\ & \boldsymbol{x} \in \{0,1\}^n. \end{array}$$

Γ	$\bar{Z}(\Gamma)$	% change in $\bar{Z}(\Gamma)$	$\sigma(\Gamma)$	% change in $\sigma(\Gamma)$
0	8822	0 %	501.0	0.0 %
10	8827	0.056 %	493.1	-1.6 %
20	8923	1.145 %	471.9	-5.8 %
30	9059	2.686 %	454.3	-9.3 %
40	9627	9.125 %	396.3	-20.9 %
50	10049	13.91 %	371.6	-25.8 %
60	10146	15.00 %	365.7	-27.0 %
70	10355	17.38 %	352.9	-29.6 %
80	10619	20.37 %	342.5	-31.6 %
100	10619	20.37 %	340.1	-32.1 %

$$
\begin{aligned}
Z^*(\Gamma) = \quad & \text{minimize} \quad \boldsymbol{c}'\boldsymbol{x} + \max_{\{S|\ S\subseteq J, |S|=\Gamma\}} \sum_{j\in S} d_j x_j \\
& \text{subject to} \quad \sum_{i\in N} x_i = k \\
& \qquad\qquad\quad \boldsymbol{x} \in \{0,1\}^n.
\end{aligned}
$$

9.1.1 Data

SLIDE 29

- $|N| = 200$;
- $k = 100$;
- $c_j \sim U[50,200]$; $d_j \sim U[20,200]$;
- For testing robustness, generate instances such that each cost component independently deviates with probability $\rho = 0.2$ from the nominal value c_j to $c_j + d_j$.

9.1.2 Results

SLIDE 30

10 Robust Network Flows

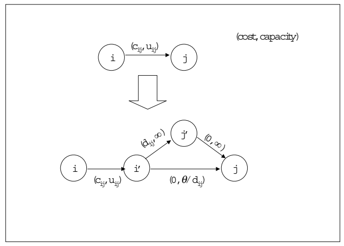
SLIDE 31

- Nominal

$$
\begin{aligned}
\min \quad & \sum_{(i,j)\in\mathcal{A}} c_{ij} x_{ij} \\
\text{s.t.} \quad & \sum_{\{j:(i,j)\in\mathcal{A}\}} x_{ij} - \sum_{\{j:(j,i)\in\mathcal{A}\}} x_{ji} = b_i \quad && \forall i \in \mathcal{N} \\
& 0 \le x_{ij} \le u_{ij} && \forall (i,j) \in \mathcal{A}.
\end{aligned}
$$

- X set of feasible solutions flows.
- Robust

$$
\begin{aligned}
Z^* = \min \quad & \boldsymbol{c}'\boldsymbol{x} + \max_{\{S|\ S\subseteq\mathcal{A}, |S|\le\Gamma\}} \sum_{(i,j)\in S} d_{ij} x_{ij} \\
\text{subject to} \quad & \boldsymbol{x} \in X.
\end{aligned}
$$



10.1 Reformulation

SLIDE 32

- $$Z^* = \min_{\theta \geq 0} Z(\theta),$$

$$\begin{aligned} Z(\theta) = \Gamma\theta + \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sum_{(i,j)\in\mathcal{A}} p_{ij} \\ \text{subject to} \quad & p_{ij} \geq d_{ij}x_{ij} - \theta \quad \forall (i,j) \in \mathcal{A} \\ & p_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{A} \\ & \boldsymbol{x} \in X. \end{aligned}$$

- Equivalently

$$\begin{aligned} Z(\theta) = \Gamma\theta + \min \quad & \boldsymbol{c}'\boldsymbol{x} + \sum_{(i,j)\in\mathcal{A}} d_{ij} \max\left(x_{ij} - \frac{\theta}{d_{ij}}, 0\right) \\ \text{subject to} \quad & \boldsymbol{x} \in X. \end{aligned}$$

10.2 Network Reformulation

SLIDE 33

Theorem: For fixed θ we can solve the robust problem as a network flow problem

10.3 Complexity

SLIDE 34

- $Z(\theta)$ is a convex function and for all $\theta_1, \theta_2 \geq 0$, we have

$$|Z(\theta_1) - Z(\theta_2)| \leq |\mathcal{A}||\theta_1 - \theta_2|.$$

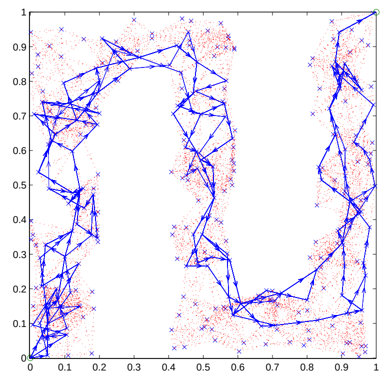
- For any fixed $\Gamma \leq |\mathcal{A}|$ and every $\epsilon > 0$, we can find a solution $\hat{\boldsymbol{x}} \in X$ with robust objective value

$$\hat{Z} = \boldsymbol{c}'\hat{\boldsymbol{x}} + \max_{\{S|\ S\subseteq\mathcal{A}, |S|\leq\Gamma\}} \sum_{(i,j)\in S} d_{ij}\hat{x}_{ij}$$

such that

$$Z^* \leq \hat{Z} \leq (1+\epsilon)Z^*$$

by solving $2\lceil \log_2(|\mathcal{A}|\overline{\theta}/\epsilon) \rceil + 3$ network flow problems, where $\overline{\theta} = \max\{u_{ij}d_{ij} : (i,j) \in \mathcal{A}\}$.



11 Experimental Results

11.1 Shortest Path

Slide 35

Slide 36

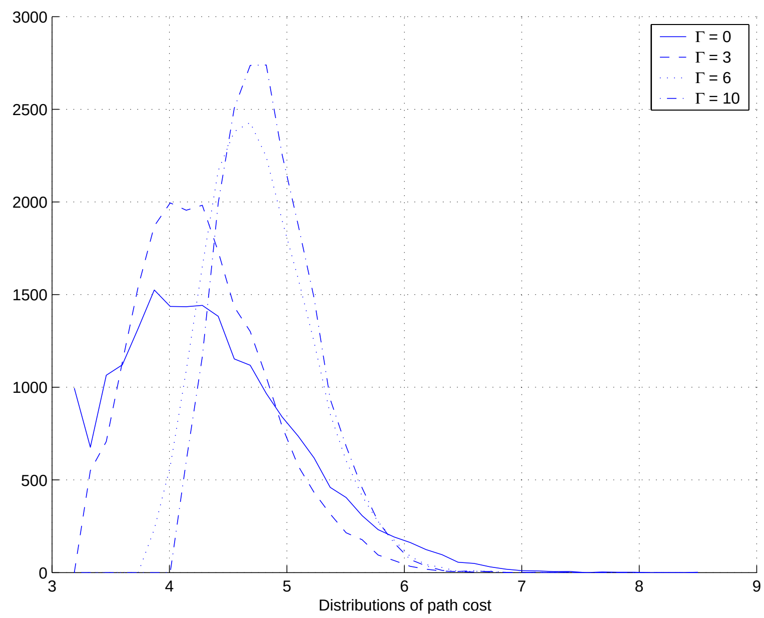
12 Conclusions

Slide 37

- Robust counterpart of a MIP remains a MIP, of comparable size.
- Approach permits flexibility of adjusting the level of conservatism in terms of probabilistic bound of constraint violation
- For polynomial solvable 0-1 optimization problems with cost uncertainty, the robust counterpart is polynomial solvable.

Slide 38

- Robust network flows are solvable as a series of nominal network flow problems.
- Robust optimization is tractable for stochastic optimization problems without the curse of dimensionality


3000
2500
2000
1500
1000
500
0
3
4
5
6
7
8
9
Distributions of path cost
Γ = 0
Γ = 3
Γ = 6
Γ = 10