15.083J/6.859J Integer Optimization

Lecture 11-12: Robust Optimization
1 Papers

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

3 Motivation
- 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?

- Ben-Tal and Nemirovski (2000):
  In real-world applications of Linear Optimization (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

- Flexible adjustment of conservatism
- Nonlinear convex models
- Not extendable to discrete optimization

4 Goal

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

\[
\text{minimize } \mathbf{c'} \mathbf{x} \\
\text{subject to } \mathbf{A} \mathbf{x} \leq \mathbf{b} \\
\mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \\
x_i \in \mathbb{Z}, \quad i = 1, \ldots, k,
\]

WLOG data uncertainty affects only \( \mathbf{A} \) and \( \mathbf{c} \), but not the vector \( \mathbf{b} \).

- (Uncertainty for matrix \( \mathbf{A} \)): \( a_{ij}, \ j \in J_i \) 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 \( \mathbf{c} \)): \( c_j, \ j \in J_0 \) takes values in \([c_j, c_j + \hat{d}_j]\).

6 Robust MIP

- Consider an integer \( \Gamma_i \in [0, |J_i|], \ i = 0,1, \ldots, m. \)
- \( \Gamma_i \) 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 \( \Gamma_i \) of the \( a_{ij} \)'s are allowed to change.
- 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 $\Gamma_i$ change, then the robust solution will be feasible with very high probability.

6.1 Problem

\[
\begin{align*}
\text{minimize} & \quad c'x + \max_{\{s_0\} \in \mathbb{S}_0} \left\{ \sum_{j \in s_0} d_j |x_j| \right\} \\
\text{subject to} & \quad \sum_j a_{ij} x_j + \max_{\{s_i\} \in \mathbb{S}_i \mid \mathbb{S}_i \subseteq \mathbb{J}_i, |\mathbb{S}_i| \leq \Gamma_i} \left\{ \sum_{j \in \mathbb{S}_i} \hat{a}_{ij} |x_j| \right\} \leq b_i, \quad \forall i \\
& \quad l \leq x \leq u, \quad \forall i = 1, \ldots, k.
\end{align*}
\]

6.2 Theorem 1

The robust problem can be reformulated has an equivalent MIP:

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

6.3 Proof

Given a vector $x^*$, we define:

\[
\beta_i(x^*) = \max_{\{s_i\} \in \mathbb{S}_i \mid \mathbb{S}_i \subseteq \mathbb{J}_i, |\mathbb{S}_i| = \Gamma_i} \left\{ \sum_{j \in \mathbb{S}_i} \hat{a}_{ij} |x_j^*| \right\}.
\]

This equals to:

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

Dual:

\[
\begin{align*}
\beta_i(x^*) & = \min \sum_{j \in J_i} p_{ij} + \Gamma_i z_i \\
\text{s.t.} & \quad z_i + p_{ij} \geq \hat{a}_{ij} |x_j^*|, \quad \forall j \in J_i \\
& \quad p_{ij} \geq 0, \quad \forall j \in J_i \\
& \quad z_i \geq 0, \quad \forall i.
\end{align*}
\]
| $|J_i|$ | $\Gamma_i$ |
|------|-------|
| 5    | 5     |
| 10   | 8.3565 |
| 100  | 24.263 |
| 200  | 33.899 |

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

6.4 Size

- 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

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

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

$n = |J_i|$, $\nu = \frac{\Gamma_1 + n}{2}$ and $\mu = \nu - \lfloor \nu \rfloor$; bound is tight.
(b) As $n \to \infty$

$$\frac{1}{2^n}\left\{ (1 - \mu) \sum_{l=\lfloor \nu \rfloor}^{n} \binom{n}{l} + \mu \sum_{l=\lfloor \nu \rfloor + 1}^{n} \binom{n}{l} \right\} \sim 1 - \Phi\left(\frac{\Gamma_1 - 1}{\sqrt{n}}\right).$$

7 Experimental Results

7.1 Knapsack Problems

- maximize $\sum_{i \in N} c_i x_i$
- subject to $\sum_{i \in N} w_i x_i \leq b$
- $x \in \{0, 1\}^n$. 
<table>
<thead>
<tr>
<th>$\Gamma$</th>
<th>Violation Probability</th>
<th>Optimal Value</th>
<th>Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5</td>
<td>5592</td>
<td>0%</td>
</tr>
<tr>
<td>2.8</td>
<td>$4.49 \times 10^{-4}$</td>
<td>5585</td>
<td>0.13%</td>
</tr>
<tr>
<td>36.8</td>
<td>$5.71 \times 10^{-3}$</td>
<td>5506</td>
<td>1.54%</td>
</tr>
<tr>
<td>82.0</td>
<td>$5.04 \times 10^{-9}$</td>
<td>5408</td>
<td>3.29%</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>5283</td>
<td>5.50%</td>
</tr>
</tbody>
</table>

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

7.1.1 Data

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

Slide 20
8 Robust 0-1 Optimization

- Nominal combinatorial optimization:
  \[
  \begin{align*}
  &\text{minimize } \ c'x \\
  &\text{subject to } \ x \in X \subseteq \{0,1\}^n.
  \end{align*}
  \]

- Robust Counterpart:
  \[
  Z^* = \min c'x + \max_{\{S \subseteq J | |S| = \Gamma\}} \sum_{j \in S} d_j x_j
  \]
  \[
  \text{subject to } \ x \in X,
  \]

- WLOG \( d_1 \geq d_2 \geq \ldots \geq d_n \).

8.1 Remarks

- 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:
  \[
  \begin{align*}
  &\text{minimize } \max(c_1'x, c_2'x) \\
  &\text{subject to } \ x \in X.
  \end{align*}
  \]
  is NP-hard for the shortest path problem.

8.2 Approach

Primal: \( Z^* = \min_{x \in X} c'x + \max \sum_{j} d_j x_j u_j \)

\[
\begin{align*}
&\text{s.t. } \ 0 \leq u_j \leq 1, \ \forall \ j \\
&\text{s.t. } \ \sum_{j} u_j \leq \Gamma
\end{align*}
\]

Dual: \( Z^* = \min_{x \in X} c'x + \min \ \theta \Gamma + \sum_{j} y_j \)

\[
\begin{align*}
&\text{s.t. } \ y_j + \theta \geq d_j x_j, \ \forall \ j \\
&\text{s.t. } \ y_j, \theta \geq 0
\end{align*}
\]
8.3 Algorithm A

- Solution: \( y_j = \max(d_j x_j - \theta, 0) \)

\[
Z^* = \min_{x \in X, \theta \geq 0} \theta \Gamma + \sum_j (c_j x_j + \max(d_j x_j - \theta, 0))
\]

- Since \( X \subseteq \{0, 1\}^n \),
  \[
  \max(d_j x_j - \theta, 0) = \max(d_j - \theta, 0) x_j
  \]

\[
Z^* = \min_{x \in X, \theta \geq 0} \theta \Gamma + \sum_j (c_j + \max(d_j - \theta, 0)) x_j
\]

- \( d_1 \geq d_2 \geq \ldots \geq d_n \geq d_{n+1} = 0 \).
- For \( d_i \geq \theta \geq d_{i+1} \),
  \[
  \min_{x \in X, d_i \geq \theta \geq d_{i+1}} \theta \Gamma + \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{i} (d_j - \theta) x_j =
  \]
  \[
  d_i \Gamma + \min_{x \in X} \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{i} (d_j - d_i) x_j = Z_i
  \]

\[
Z^* = \min_{i=1, \ldots, n+1} d_i \Gamma + \min_{x \in X} \sum_{j=1}^{n} c_j x_j + \sum_{j=1}^{i} (d_j - d_i) x_j.
\]

8.4 Theorem 3

- 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

\[
\text{minimize } \sum_{i \in N} c_i x_i
\]

subject to \( \sum_{i \in N} x_i = k \)

\( x \in \{0, 1\}^n \).
\begin{array}{|c|c|c|c|c|}
\hline
\Gamma & Z(\Gamma) & \% \text{ change in } Z(\Gamma) & \sigma(\Gamma) & \% \text{ change in } \sigma(\Gamma) \\
\hline
0 & 8822 & 0 \% & 501.0 & 0.0 \% \\
10 & 8827 & 0.066 \% & 493.1 & -1.6 \% \\
20 & 8923 & 1.145 \% & 471.9 & -5.8 \% \\
30 & 9059 & 2.686 \% & 454.3 & -9.3 \% \\
40 & 9227 & 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 \% \\
\hline
\end{array}

\[ Z^*(\Gamma) = \min \ c' x + \max_{\{S|S \subseteq \{i,j\}, |S| = \Gamma\}} \sum_{j \in S} d_{ij} x_{ij} \]
subject to \[ \sum_{i \in N} x_i = k \]
\[ x \in [0, 1]^n. \]

9.1.1 Data
\begin{itemize}
\item \(|N| = 200;\)
\item \(k = 100;\)
\item \(c_j \sim U[50, 200]; d_j \sim U[20, 200];\)
\item For testing robustness, generate instances such that each cost component independently deviates with probability \(\rho = 0.2\) from the nominal value \(c_j\) to \(c_j + d_j.\)
\end{itemize}

9.1.2 Results

10 Robust Network Flows
\begin{itemize}
\item Nominal
\[ \min \ \sum_{(i,j) \in A} c_{ij} x_{ij} \]
s.t. \[ \sum_{(j \in S \in A)} x_{ij} - \sum_{(j \in S \in A)} x_{ji} = b_i \ \forall i \in N \]
\[ 0 \leq x_{ij} \leq u_{ij} \ \forall (i,j) \in A. \]
\item \(X\) set of feasible solutions flows.
\item Robust
\[ Z^* = \min \ c' x + \max_{\{S|S \subseteq A, |S| \leq \Gamma\}} \sum_{(i,j) \in S} d_{ij} x_{ij} \]
subject to \[ x \in X. \]
\end{itemize}
10.1 Reformulation

- \( Z^* = \min_{\theta \geq 0} Z(\theta) \),
- \( Z(\theta) = \Gamma \theta + \min \sum_{(i,j) \in A} p_{ij} \)
  subject to \( p_{ij} \geq d_{ij} x_{ij} - \theta \) \( \forall (i,j) \in A \)
  \( p_{ij} \geq 0 \) \( \forall (i,j) \in A \)
  \( x \in X \).

- Equivalently
  \[ Z(\theta) = \Gamma \theta + \min \sum_{(i,j) \in A} d_{ij} \max \left( x_{ij} - \frac{\theta}{d_{ij}}, 0 \right) \]
  subject to \( x \in X \).

10.2 Network Reformulation
Theorem: For fixed \( \theta \) we can solve the robust problem as a network flow problem

10.3 Complexity
- \( Z(\theta) \) is a convex function and for all \( \theta_1, \theta_2 \geq 0 \), we have
  \[ |Z(\theta_1) - Z(\theta_2)| \leq |A| |\theta_1 - \theta_2|. \]
- For any fixed \( \Gamma \leq |A| \) and every \( \epsilon > 0 \), we can find a solution \( \hat{x} \in X \) with robust objective value
  \[ \hat{Z} = c' \hat{x} + \max \sum_{(i,j) \in S} d_{ij} \hat{x}_{ij} \]
  such that
  \[ Z^* \leq \hat{Z} \leq (1 + \epsilon) Z^* \]
  by solving \( 2|\log_2(|A|/\epsilon)| + 3 \) network flow problems, where \( \overline{\theta} = \max\{u_{ij} d_{ij} : (i,j) \in A\} \).
11 Experimental Results

12 Conclusions

- 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.
- 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