Robust Optimization

Runyu Tang

March 2, 2023

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}, \boldsymbol{\xi}) \\ \text{s.t.} \quad & g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq \mathbf{0}, \quad \forall j \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- $\boldsymbol{x}$: the decision variable,
- $\boldsymbol{\xi}$: parameters,
- $g_j(\cdot)$: convex functions.

We have 4 potential options, each with a weight and a value. We want to choose a subset of options that maximizes the value while keeping the total weight below 500.

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 500 \\ & x_i \text{ integer }, \forall i. \end{aligned}$$

$\boldsymbol{x}^* = [4, 0, 0, 0]$

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \mathbb{E}^{\mathbb{P}} f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \\ \text{s.t.} \quad & g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq \mathbf{0}, \quad \forall j \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

- $\tilde{\boldsymbol{\xi}}$: random variables governed by $\mathbb{P}$.

We have a newsvendor who sell the newspaper at $p = \$7$ and buy in at $c = \$5$. We want to choose a quantity to sell that maximizes the expected profit. Assume the demand is D with a known distribution $\mathbb{P}$.

$$
\begin{aligned}
&\max_x \quad \mathbb{E}^{\mathbb{P}}\left[7\min\{x, D\} - 5x\right] \\
\Leftrightarrow &\min_x \quad \mathbb{E}^{\mathbb{P}}\left[-2x + 7(x - D)^+\right]
\end{aligned}
$$

$x^* = \inf\{y : F(y) < \frac{p-c}{p}\}$.

If the distribution $\mathbb{P} \sim N(\mu, \sigma^2)$, then $x^* = \phi^{-1}_{(\mu,\sigma)}(2/7) = \phi^{-1}(2/7)\sigma + \mu$.

Suppose we have N samples $\{\hat{\boldsymbol{\xi}}_i\}_{i=1}^N$ from $\mathbb{P}$, then we can approximate the expected value by:

$$\mathbb{E}^{\mathbb{P}} f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}) \approx \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}, \hat{\boldsymbol{\xi}}_i)$$

Then, the stochastic programming problem can be transformed into a deterministic problem:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}, \hat{\boldsymbol{\xi}}_i) \\ \text{s.t.} \quad & g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq \mathbf{0}, \quad \forall j \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

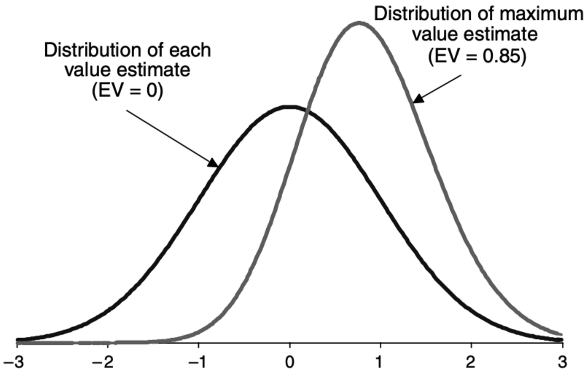
[Smith, Winkler 2006] (MS)

- Suppose that a decision maker is considering 3 alternatives whose true values are μ_1, μ_2, μ_3. Assume $\mu_1 = \mu_2 = \mu_3 \sim N(0, 1)$.
- The decision maker is uncertain about the true values of the alternatives and estimates them as $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$. The decision maker chooses the alternative with the highest estimated value.
- The expected disappointment will be 85% of the standard deviation of the value estimates, and will increase with the number of alternatives considered.

Order statistics/numerical simulation.

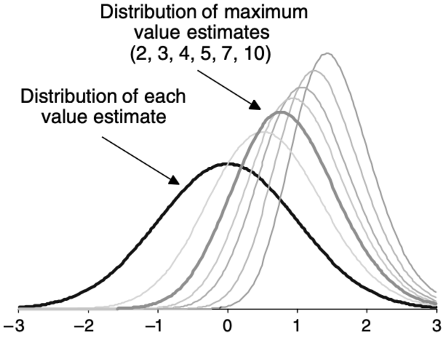
```
# python code
max(np.random.normal(0, 1, 3))
```

Figure 1 **The Distribution of the Maximum of Three Standard Normal Value Estimates**



The optimizer's curse

Figure 2 The Distribution of the Maximum of *n* Standard Normal Value Estimates

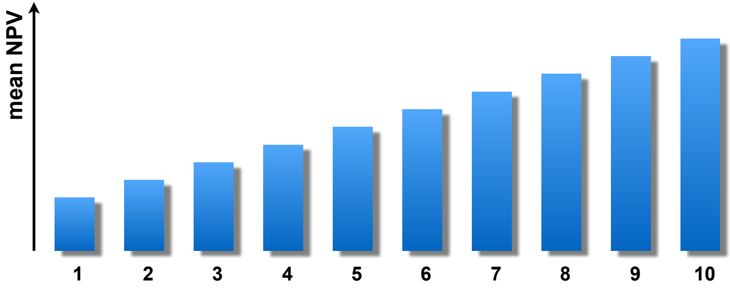


Number of alternatives	Expected disappointment
1	0.00
2	0.56
3	0.85
4	1.03
5	1.16
6	1.27
7	1.35
8	1.43
9	1.48
10	1.54

A decision maker who consistently chooses alternatives based on her estimated values should expect to be disappointed on average, even if the individual value estimates are conditionally unbiased.

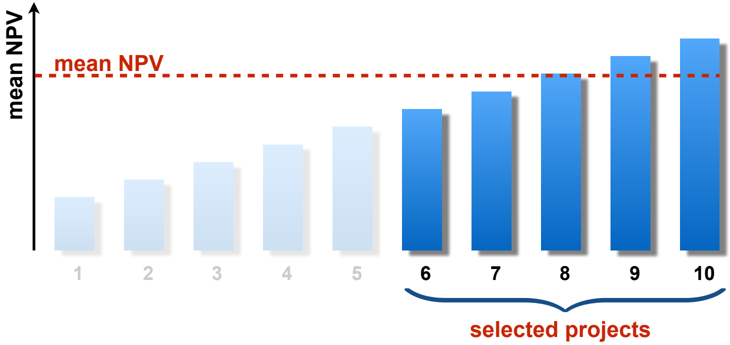
The Optimizer's Curse

Select 5 out of 10 projects!



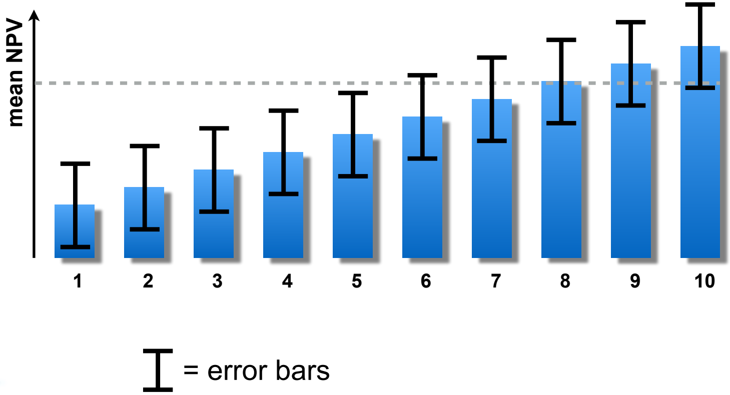
The Optimizer's Curse

Select 5 out of 10 projects!



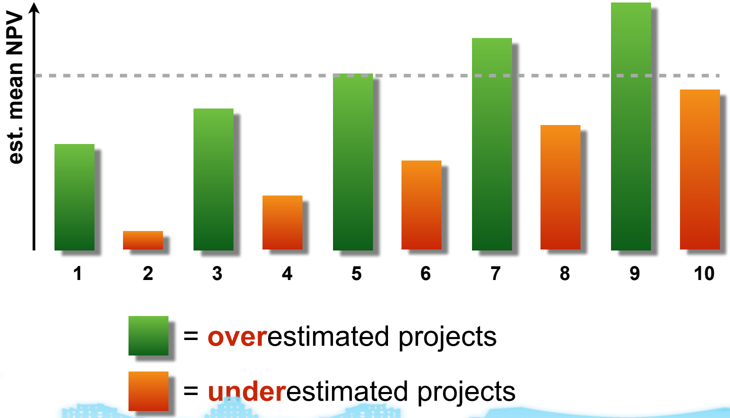
The Optimizer's Curse

Select 5 out of 10 projects!



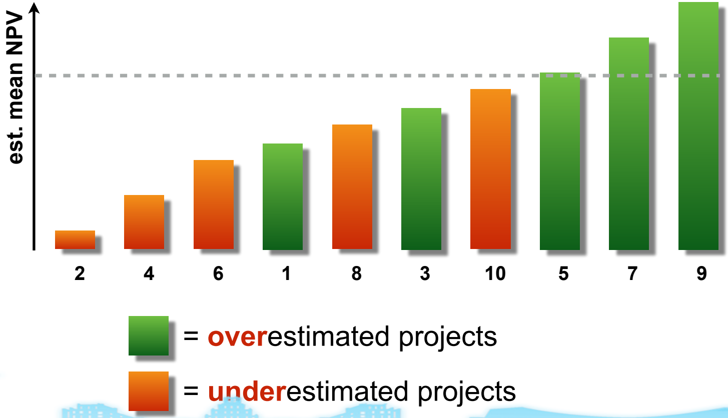
The Optimizer's Curse

Select 5 out of 10 projects!



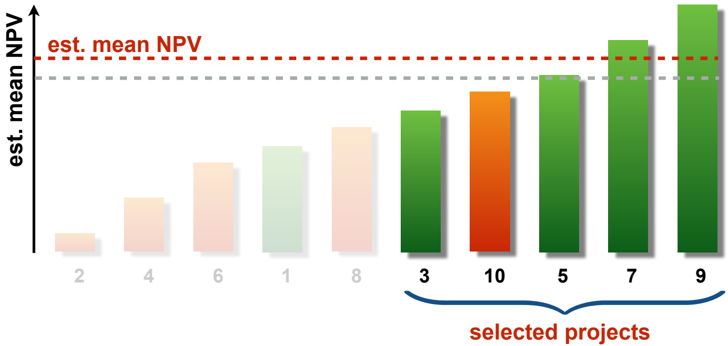
The Optimizer's Curse

Select 5 out of 10 projects!



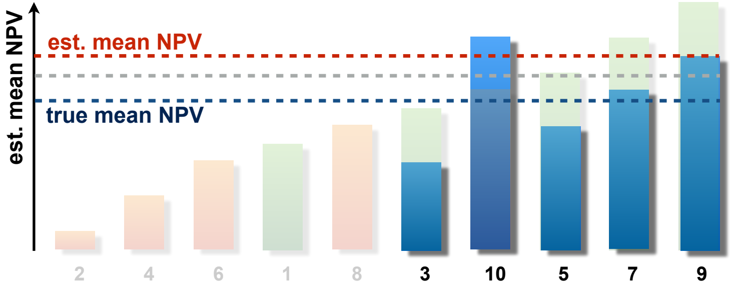
The Optimizer's Curse

Select 5 out of 10 projects!



The Optimizer's Curse

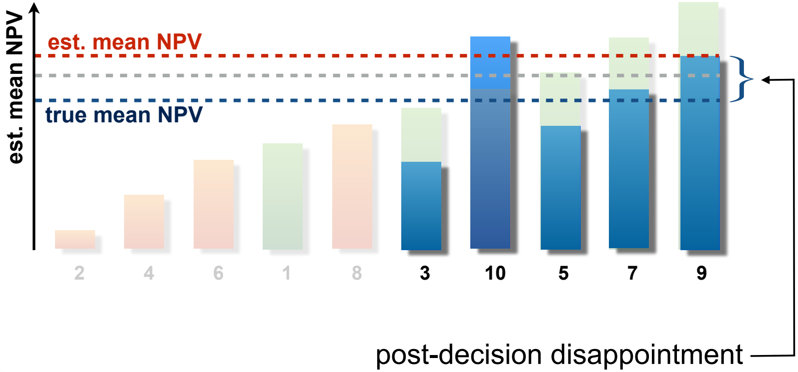
Select 5 out of 10 projects!



= **true** NPVs of selected projects

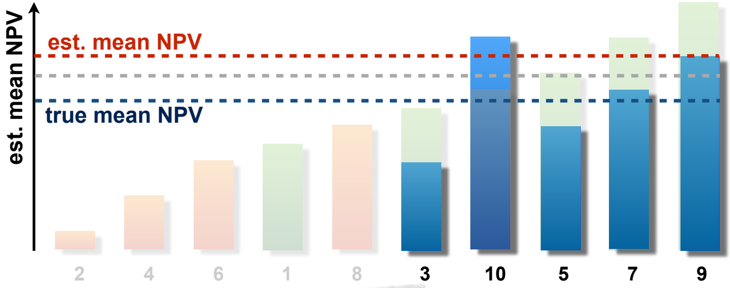
The Optimizer's Curse

Select 5 out of 10 projects!



The Optimizer's Curse

Select 5 out of 10 projects!



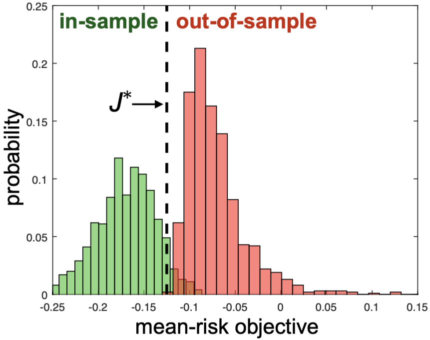
Even if input estimates are unbiased, the optimization results are biased!

Mean-risk portfolio problem

$$\min_{x \in \mathcal{X}} \left\{ \mathbb{E}^{\mathbb{P}}[-x^\top \xi] + \rho\, \mathbb{P}\text{-CVaR}_\alpha(-x^\top \xi) \right\}$$

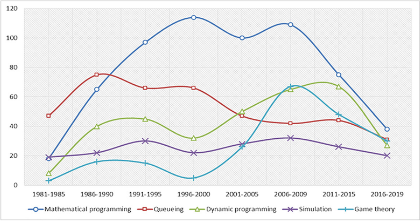
- 10 assets
- $\rho = 10$
- $\alpha = 20\%$
- $\xi_i = \psi + \zeta_i$ where $\psi \sim \mathcal{N}(0, 2\%)$ and $\zeta_i \sim \mathcal{N}(i \times 3\%, i \times 2.5\%)$

Performance of SAA solution



- 30 training samples
- in-sample: optimistic bias
- out-of-sample: pessimistic bias

Figure 3. (Color online) Number of Articles Applying the Top Five Research Methods over Time



Overall, we find that Operations Research welcomes diversified research methods and encourages the applications of multiple methods... Robust optimization has emerged as a popular research method. A total number of 60 articles have used this method since 2010, with a total citation of 1,089.

Angelito Calma , William Ho , Lusheng Shao , Huashan Li (2021) Operations Research: Topics, Impact, and Trends from 1952-2019. Operations Research 69(5):1487-1508.

Worst-case analysis

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \max_{\boldsymbol{\xi}} f(\boldsymbol{x}, \boldsymbol{\xi}) \\ \text{s.t.} \quad & g_j(\boldsymbol{x}, \boldsymbol{\xi}) \leq 0, \quad \forall j, \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}, \\ & \boldsymbol{x} \in \mathcal{X}, \end{aligned}$$

where $\boldsymbol{\Xi}$ is the uncertainty set

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 500 \\ & x_i \text{ integer }, \forall i. \end{aligned}$$

The numbers $50, 40, 60, 30, 120, 100, 180, 140$ can be not precise, and may belong to an uncertain set.
For example, $50 \in [45, 55]$, $40 \in [35, 45]$, $60 \in [55, 65]$, $30 \in [25, 35]$, $120 \in [115, 125]$, $100 \in [95, 105]$, $180 \in [175, 185]$, $140 \in [135, 145]$.

$$\begin{aligned} \max \quad & 45x_1 + 35x_2 + 55x_3 + 25x_4 \\ \text{s.t.} \quad & 125x_1 + 105x_2 + 185x_3 + 145x_4 \leq 500 \\ & x_i \text{ integer }, \forall i. \end{aligned}$$

$$\begin{aligned}
\min_{\boldsymbol{x}} \quad & \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}^{\mathbb{P}}\left[f(\boldsymbol{x},\boldsymbol{\xi})\right] \\
\text{s.t.} \quad & g_j(\boldsymbol{x},\boldsymbol{\xi}) \le 0, \quad \text{almost surely } \forall j, \mathbb{P}\in\mathcal{F}, \\
& \boldsymbol{x}\in\mathcal{X},
\end{aligned}$$

where $\mathcal{F}$ is the ambiguity set

$$\min_{x} \quad \max_{\mathbb{P}\in\mathcal{F}} \mathbb{E}^{\mathbb{P}} \left[-2x + 7(x - D)^{+}\right]$$

We may not know the distribution $\mathbb{P}$, and it may belong to an ambiguity set $\mathcal{F} = \{\mathbb{P}_1 \sim N(\mu_1, \sigma_1^2), \mathbb{P}_2 \sim N(\mu_2, \sigma_2^2)\}$.

Then we can compare the expected profit of two strategies: $x_1^* = F_1^{-1}(2/7)$ and $x_2^* = F_2^{-1}(2/7)$.

General Linear Optimization(LO) Problem:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} + d \\ \text{s.t.} \quad & A\boldsymbol{x} \leq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

Uncertain Linear Optimization problem

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} + d \\ \text{s.t.} \quad & A\boldsymbol{x} \leq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & (c, d, A, b) \in \mathcal{U} \end{aligned}$$

where $\mathcal{U}$ is a set of uncertain parameters.

[Soyster 1973] OR: "columnwise" uncertainty

Nominal problem:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} \\ \text{s.t.} \quad & \sum_j^n a_{ij} x_j \le b_i, i = 1, 2, \ldots, n \\ & \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u} \end{aligned}$$

Model uncertainty:
$\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.

Robust counterpart:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} \\ \text{s.t.} \quad & \sum_j^n a_{ij} x_j + \sum_j^n \hat{a}_{ij} |x_j| \le b_i, \forall i \\ & \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u} \end{aligned}$$

The solution is always feasible to the nomial problem, but too conservative.

[Ben-Tal and Nemirovski, 2000] MP: ellipsoid uncertainty

$$
\begin{aligned}
\max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} \\
\text{s.t.} \quad & \sum_j a_{ij} x_j + \sum_{j \in J_i} \hat{a}_{ij} y_{ij} + \Omega_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leqslant b_i \quad \forall i \\
& -y_{ij} \leqslant x_j - z_{ij} \leqslant y_{ij} \quad \forall i, j \in J_i \\
& \mathbf{1} \leqslant \mathbf{x} \leqslant \mathbf{u} \\
& \mathbf{y} \geqslant \mathbf{0}
\end{aligned}
$$

The probability that the i constraint is violated is at most $\exp(-\Omega_i^2/2)$.
But, it is **nonlinear**.

[Dimitris Bertsimas, Melvyn Sim, 2004] OR: The most cited OR paper!

Volume 52, Issue 1

January-February 2004
Pages 1-164

Article Information

Metrics

Information

Received: September 01, 2001
Accepted: December 01, 2002
Published Online: February 01, 2004

- The uncertainty: $\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.
- Speaking intuitively, it is unlikely that all of the a_{ij}, $j \in J_i$ will change. Our goal is to be protected against all cases that up to $\lfloor \Gamma_i \rfloor$ of these coefficients are allowed to change, and one coefficient a_{it} changes by $(\Gamma_i - \lfloor \Gamma_i \rfloor)\hat{a}_{it}$.
- $\Gamma_i = 0$, nominal LP; $\Gamma_i = J_i$, Soyster problem.

We have the following non-linear problem:

$$\begin{aligned} \max \quad & \mathbf{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \sum_j a_{ij} x_j + \max_{\{S_i \cup \{t_i\} | S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} y_j + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} y_t \right\} \leq b_i \quad \forall i \\ & -y_j \leqslant x_j \leqslant y_j \quad \forall j \\ & \mathbf{l} \leqslant \mathbf{x} \leqslant \mathbf{u} \\ & \mathbf{y} \geqslant \mathbf{0}. \end{aligned}$$

The protection function of the ith constraint,

$$\beta_i\left(\boldsymbol{x}^*, \Gamma_i\right) = \max_{\{S_i \cup \{t_i\} \mid S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} \left|x_j^*\right| + \left(\Gamma_i - \lfloor \Gamma_i \rfloor\right) \hat{a}_{it_i} \left|x_j^*\right| \right\},$$

$$\begin{aligned} = \max \quad & \sum_{j \in J_i} \hat{a}_{ij} \left|x_j^*\right| z_{ij} \\ \text{s.t.} \quad & \sum_{j \in J_i} z_{ij} \leqslant \Gamma_i \\ & 0 \leqslant z_{ij} \leqslant 1, \quad \forall j \in J_i. \end{aligned}$$

Dual problem

$$\begin{aligned} &\max_x \quad c^\top x \\ &\text{s.t.} \quad Ax \le b \\ \Leftrightarrow &\min_y \quad b^\top y \\ &\text{s.t.} \quad A^\top y \ge c \end{aligned}$$

$$\begin{aligned} &\max_z \quad \sum_{j \in J_i} \hat{a}_{ij} \left| x_j^* \right| z_{ij} \\ &\text{s.t.} \quad \sum_{j \in J_i} z_{ij} \leqslant \Gamma_i \qquad (\lambda_i) \\ &\qquad z_{ij} \leqslant 1, \quad \forall j \in J_i \qquad (p_{ij}). \\ \Leftrightarrow &\min_{\lambda, p} \quad \sum_i \lambda_i \Gamma_i + \sum_{ij} p_{ij} \\ &\text{s.t.} \quad \lambda_j + p_{ij} \ge \hat{a}_{ij} \left| x_j^* \right|, \quad \forall i, j \end{aligned}$$

The problem is equivalent to the following LP:

$$
\begin{aligned}
\max \quad & \mathbf{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \sum_j a_{ij} x_j + z_i \Gamma_i + \sum_{j \in J_i} p_{ij} \leqslant b_i \quad \forall i \\
& z_i + p_{ij} \geqslant \hat{a}_{ij} y_j \quad \forall i, j \in J_i \\
& -y_j \leqslant x_j \leqslant y_j \quad \forall j \\
& l_j \leqslant x_j \leqslant u_j \quad \forall j \\
& p_{ij} \geqslant 0 \quad \forall i, j \in J_i \\
& y_j \geqslant 0 \quad \forall j \\
& z_i \geqslant 0 \quad \forall i.
\end{aligned}
$$

Assumptions:

- $\eta_{ij} = (\tilde{a}_{ij} - \hat{a}_{ij})/\hat{a}_{ij} \in [-1, 1]$.
- η_{ij} are independent.

Probabilistic guarantee:

$$\Pr\left(\sum_j \tilde{a}_{ij} x_j^* > b_i\right) \leqslant \Pr\left(\sum_{j \in J_i} \gamma_{ij}\eta_{ij} \geqslant \Gamma_i\right) \leq \exp\left(-\frac{\Gamma_i^2}{2|J_i|}\right)$$

where

$$\gamma_{ij} = \begin{cases} 1, & \text{if } j \in S_i^* \\ \dfrac{\hat{a}_{ij}\left|x_j^*\right|}{\hat{a}_{ir^*}\left|x_{r^*}^*\right|}, & \text{if } j \in J_i \backslash S_i^* \end{cases} \quad \text{and } r^* = \arg\min_{r \in S_i^* \cup \{t_i^*\}} \hat{a}_{ir}\left|x_r^*\right|.$$

The bound can be further tightened. (See Theorem 3 in [Bertsimas and Sim, 2004])

The first inequality:

$$\Pr\left(\sum_j \tilde{a}_{ij}x_j^* > b_i\right)$$

$$= \Pr\left(\sum_j a_{ij}x_j^* + \sum_{j\in J_i} \eta_{ij}\hat{a}_{ij}x_j^* > b_i\right)$$

$$\leqslant \Pr\left(\sum_{j\in J_i} \eta_{ij}\hat{a}_{ij}|x_j^*| > \sum_{j\in S_i^*} \hat{a}_{ij}|x_j^*| + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\right)$$

$$= \Pr\left(\sum_{j\in J_i\setminus S_i^*} \eta_{ij}\hat{a}_{ij}|x_j^*| > \sum_{j\in S_i^*} \hat{a}_{ij}|x_j^*|(1-\eta_{ij}) + (\Gamma_i - \lfloor\Gamma_i\rfloor)\hat{a}_{it_i^*}|x_{t_i^*}^*|\right)$$

$$\leqslant \Pr\left(\sum_{j\in J_i\setminus S_i^*} \eta_{ij}\hat{a}_{ij}|x_j^*| > \hat{a}_{ir^*}|x_{r^*}^*| \cdot \left(\sum_{j\in S_i^*}(1-\eta_{ij}) + (\Gamma_i - \lfloor\Gamma_i\rfloor)\right)\right)$$

$$= \Pr\left(\sum_{j\in S_i^*} \eta_{ij} + \sum_{j\in J_i\setminus S_i^*} \frac{\hat{a}_{ij}|x_j^*|}{\hat{a}_{ir^*}|x_{r^*}^*|}\eta_{ij} > \Gamma_i\right)$$

$$= \Pr\left(\sum_{j\in J_i} \gamma_{ij}\eta_{ij} > \Gamma_i\right)$$

$$\leqslant \Pr\left(\sum_{j\in J_i} \gamma_{ij}\eta_{ij} \geqslant \Gamma_i\right).$$

The second inequality:

$$
\begin{aligned}
&\Pr\left(\sum_{j\in J_i}\gamma_{ij}\eta_{ij}\geqslant\Gamma_i\right)\\
&\leqslant\frac{E[\exp(\theta\sum_{j\in J_i}\gamma_{ij}\eta_{ij})]}{\exp(\theta\Gamma_i)} && (12)\\
&=\frac{\prod_{j\in J_i}E[\exp(\theta\gamma_{ij}\eta_{ij})]}{\exp(\theta\Gamma_i)} && (13)\\
&=\frac{\prod_{j\in J_i}2\int_0^1\sum_{k=0}^{\infty}((\theta\gamma_{ij}\eta)^{2k}/(2k)!)\,dF_{\eta_{ij}}(\eta)}{\exp(\theta\Gamma_i)} && (14)\\
&\leqslant\frac{\prod_{j\in J_i}\sum_{k=0}^{\infty}((\theta\gamma_{ij})^{2k})/(2k)!}{\exp(\theta\Gamma_i)}\leqslant\frac{\prod_{j\in J_i}\exp(\theta^2\gamma_{ij}^2/2)}{\exp(\theta\Gamma_i)}\\
&\leqslant\exp\left(|J_i|\frac{\theta^2}{2}-\theta\Gamma_i\right). && (15)
\end{aligned}
$$

Inequality (12) follows from Markov's inequality, Equations (13) and (14) follow from the independence and symmetric distribution assumption of the random variables η_{ij} . Inequality (15) follows from $\gamma_{ij}\leq 1$. Selecting $\theta=\Gamma_i/|J_i|$, we obtain the second inequality.

Model uncertainty: $\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.

Nominal problem:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} \\ \text{s.t.} \quad & \sum_{j}^{n} a_{ij} x_j \le b_i, i = 1, 2, \ldots, n \\ & \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u} \end{aligned}$$

Robust counterpart:

$$\begin{aligned} \max_{\boldsymbol{l} \le \boldsymbol{x} \le u} \quad & c^T \boldsymbol{x} \\ \text{s.t.} \quad & \max_{\tilde{a} \in \mathcal{U}} \sum_{j}^{n} a_{ij} x_j \le b_i, \forall i \\ \Rightarrow \max_{\boldsymbol{x}} \quad & c^T \boldsymbol{x} \\ \text{s.t.} \quad & \sum_{j}^{n} a_{ij} x_j + \sum_{j}^{n} \hat{a}_{ij} |x_j| \le b_i, \forall i \\ & - y_j \le x_j \le y_j, \forall j \\ & \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u} \end{aligned}$$

Protect Γ_i coefficients:

$$
\begin{aligned}
\max \quad & \mathbf{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \sum_j a_{ij} x_j + \max_{\{S_i \cup \{t_i\} | S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} y_j + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} y_t \right\} \leq b_i \quad \forall i \\
& -y_j \leqslant x_j \leqslant y_j \quad \forall j \\
& \mathbf{l} \leqslant \mathbf{x} \leqslant \mathbf{u} \\
& \mathbf{y} \geqslant \mathbf{0}.
\end{aligned}
$$

Reformulation:

$$
\begin{aligned}
\max \quad & \mathbf{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \sum_j a_{ij}x_j + z_i\Gamma_i + \sum_{j\in J_i} p_{ij} \leqslant b_i \quad \forall i \\
& z_i + p_{ij} \geqslant \hat{a}_{ij}y_j \quad \forall i, j \in J_i \\
& -y_j \leqslant x_j \leqslant y_j \quad \forall j \\
& l_j \leqslant x_j \leqslant u_j \quad \forall j \\
& p_{ij} \geqslant 0 \quad \forall i, j \in J_i \\
& y_j \geqslant 0 \quad \forall j \\
& z_i \geqslant 0 \quad \forall i.
\end{aligned}
$$

Model uncertainty:
$\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.

$$
\Pr\left(\sum_j \tilde{a}_{ij}x_j^* > b_i\right) \leq \exp\left(-\frac{\Gamma_i^2}{2|J_i|}\right)
$$

The derivative of the objective function value with respect to protection level Γ_i of the ith constraint is $-z_i^* q_i^*$,

Theorem (The Price of Robustness).

Let z^* and q^* be the optimal nondegenerate primal and dual solutions for the linear optimization problem (under nondegeneracy, the primal and dual optimal solutions are unique). Then, the derivative of the objective function value with respect to protection level Γ_i of the ith constraint is

$$-z_i^* q_i^*,$$

where z^* is the optimal primal variable corresponding to the protection level Γ_i and q_i^* is the optimal dual variable of the ith constraint.

$$\max_{\mathbf{x}} \quad \sum_{i}^{N} c_i x_i$$

$$\text{s.t.} \quad \sum_{i}^{N} w_i x_i \leq b,$$

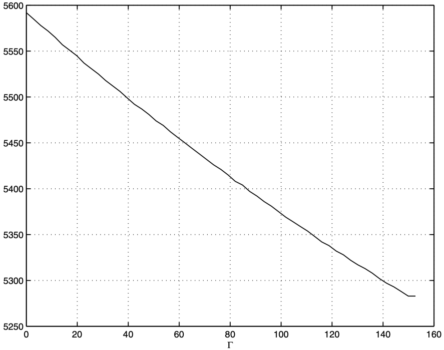
$$x_i \in \{0, 1\}, \quad i = 1, \cdots, N$$

model parameters:

- $N = 200, b = 4000$
- w_i randomly chosen from $\{20, 21, \cdots, 29\}$
- c_i randomly chosen from $\{16, 17, \cdots, 77\}$
- $\tilde{w}_i$ independently distributed and follow symmetric distribution $[w_i - \delta_i,\ w_i + \delta_i]$ and $\delta_i = 10\% w_i$

Optimal value if no uncertainty: 5592, Soyester's method: 5283.

Optimal value of the robust knapsack formulation as a function of Γ.



Optimal value of the robust knapsack formulation as a function of the probability bound of constraint violation given in Equation (18).

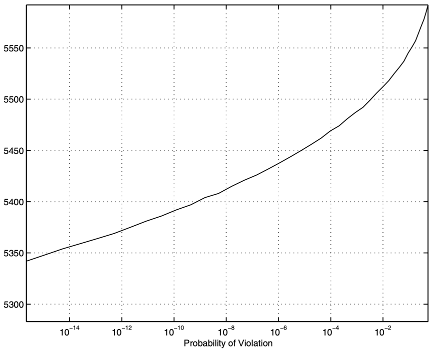


Table 2. Results of robust knapsack solutions.

Γ	Probability Bound	Optimal Value	Reduction (%)
2.8	4.49×10^{-1}	5,585	0.13
14.1	1.76×10^{-1}	5,557	0.63
25.5	4.19×10^{-2}	5,531	1.09
36.8	5.71×10^{-3}	5,506	1.54
48.1	4.35×10^{-4}	5,481	1.98
59.4	1.82×10^{-5}	5,456	2.43
70.7	4.13×10^{-7}	5,432	2.86
82.0	5.04×10^{-9}	5,408	3.29
93.3	3.30×10^{-11}	5,386	3.68
104.7	1.16×10^{-13}	5,364	4.08
116.0	2.22×10^{-16}	5,342	4.47

Insights:

- Our approach succeeds in reducing the price of robustness; that is, we do not heavily penalize the objective function value in order to protect ourselves against constraint violation.
- The proposed robust approach is computationally tractable in that the problem can be solved in reasonable computational times.

Uncertainty set: $U_i = \{\boldsymbol{a}_i | \boldsymbol{D}_i \boldsymbol{a}_i \leq \boldsymbol{d}_i\}$.

$$
\begin{aligned}
(\textbf{RO}) \max \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \max_{\boldsymbol{a}_i \in U_i} \boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i, \quad i = 1, \dots, m \\
& \boldsymbol{x} \geq 0
\end{aligned}
$$

$$
\begin{aligned}
(\textbf{RC}) \max \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{p}_i^\top d_i \leq b_i, \quad i = 1, \dots, m \\
& \boldsymbol{p}_i^\top \boldsymbol{D}_i = \boldsymbol{x}^\top, \quad i = 1, \dots, m \\
& \boldsymbol{x}, \boldsymbol{p}_i \geq 0
\end{aligned}
$$

Uncertainty set: $U_i = \{\boldsymbol{a}_i | \boldsymbol{a}_i = \bar{\boldsymbol{a}}_i + \Delta_i^\top \boldsymbol{u}_i, \|\boldsymbol{u}_i\| \leq \rho\}$.

$$
\begin{aligned}
(\textbf{RO}) \max \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \max_{\boldsymbol{a}_i \in U_i} \boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i, \quad i = 1, \ldots, m \\
& \boldsymbol{x} \geq 0
\end{aligned}
$$

$$
\begin{aligned}
(\textbf{RC}) \max \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \bar{\boldsymbol{a}}_i^\top \boldsymbol{x} + \rho \|\Delta_i \boldsymbol{x}\|_* \leq b_i, \quad i = 1, \ldots, m \\
& \boldsymbol{x} \geq 0
\end{aligned}
$$

Dual norm:

$$\|\boldsymbol{u}\|_* = \max\{\boldsymbol{u}^\top \boldsymbol{x} : \|\boldsymbol{x}\| \leq 1\}$$

For L_p−norm, if $\frac{1}{p} + \frac{1}{q} = 1$, then $\|\boldsymbol{u}\|_p$ and $\|\boldsymbol{u}\|_q$ are dual norm to each other.

Probabilistic Guarantee:
If u_i are independent, have zero means and $u_i \in [-1, 1]$.
Suppose $\boldsymbol{x}$ satisfies $\bar{\boldsymbol{a}}^\top \boldsymbol{x} + \rho\|\Delta_i \boldsymbol{x}\|_* \leq b$, then

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

We can select $\rho = \sqrt{2\log(\frac{1}{\epsilon})}$ to ensure the infeasibility probablity to be less than ϵ.

Proof sketch: $P(\tilde{\boldsymbol{a}}^\top \boldsymbol{x} > b) = P(\bar{\boldsymbol{a}}^\top \boldsymbol{x} + \boldsymbol{u}^\top \boldsymbol{\Delta x} > b) \leq P(-\rho\|\boldsymbol{\Delta x}\| + \boldsymbol{u}^\top \boldsymbol{\Delta x} > 0)$.
Note that when ξ_i are independent and zero mean in $[-1, 1]$, then
$P(w_0 + \boldsymbol{\xi}^\top \boldsymbol{w} > 0) \leq \exp(-\frac{w_0^2}{2\|\boldsymbol{w}\|^2})$. (In the same vein to page 35)

Many directions to work on:

- More general uncertain sets
- Chance constraints
- More complex function forms:
  - Mixed integer problem
  - Conic quadratic problem
  - Semidefinite problem
- Robust multi-stage optmization

To be uncertain is to be uncomfortable, but to be certain is to be ridiculous.

— *Chinese proverb*

Ben-Tal, A., Ghaoui, L.E., & Nemirovski, A. (2009). Robust Optimization. *Princeton Series in Applied Mathematics*.