Robust optimization

Michaël Poss

Optimization under uncertainty

Stochastic programming

$$\begin{aligned} \min \quad & \mathbb{E}_{\boldsymbol{\xi} \in \Xi}[Q(\boldsymbol{x}, \boldsymbol{\xi})] \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \in X \end{aligned}$$

$Q(\boldsymbol{x}, \boldsymbol{\xi})$: cost of solution $\boldsymbol{x}$ under uncertain realization $\boldsymbol{\xi}$

- Requires knowledge of probability distribution
- Assumes decision-makers are risk-neutral
- Works well if decision-making process is repeated sufficiently many times

Optimization under uncertainty

Risk-averse stochastic programming

$$\begin{aligned} \min \quad & \mathrm{CVaR}_{\boldsymbol{\xi} \in \Xi}[Q(\boldsymbol{x}, \boldsymbol{\xi})] \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \in X \end{aligned}$$

$Q(\boldsymbol{x}, \boldsymbol{\xi})$: cost of solution $\boldsymbol{x}$ under uncertain realization $\boldsymbol{\xi}$

- Requires knowledge of probability distribution
- Assumes decision-makers are risk-averse
- May take deviations from the expected value into account
- But what the heck is CVaR (conditional value at risk)?

Optimization under uncertainty

Distributionally robust optimization

$$\begin{aligned} \min \quad \max_{\mathbb{P}\in\mathcal{P}} \quad & \mathbb{E}^{\mathbb{P}}_{\boldsymbol{\xi}\in\Xi}[Q(\boldsymbol{x},\boldsymbol{\xi})] \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \in X \end{aligned}$$

$Q(\boldsymbol{x},\boldsymbol{\xi})$: cost of solution $\boldsymbol{x}$ under uncertain realization $\boldsymbol{\xi}$
$\mathcal{P}$: a family of probability distributions

- Can be used without exact knowledge of probability distribution
- Requires knowledge of some attributes of the distribution (moments, support etc.)
- Especially appropriate if some data about the uncertain parameters is available (data-driven optimization)

Optimization under uncertainty

Chance constrained programming

$$\begin{aligned} \min \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \mathbb{P}\{\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\} \geq \alpha \\ & \boldsymbol{x} \in X \end{aligned}$$

- Requires knowledge of probability distribution
- Is appropriate when we want to avoid undesirable events or enforce that a desirable event happens with high proability
- Is often used in contexts where recourse is not possible

Optimization under uncertainty

Robust optimization

$$
\begin{aligned}
\min \quad \max_{\boldsymbol{\xi} \in \Xi} \quad & \boldsymbol{c}(\boldsymbol{\xi})^\top \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}(\boldsymbol{\xi})\boldsymbol{x} \leq \boldsymbol{b}(\boldsymbol{\xi}) \qquad\qquad \forall \boldsymbol{\xi} \in \Xi \\
& \boldsymbol{x} \in X
\end{aligned}
$$

- Assumes no knowledge of probability distribution
- Only an "uncertainty set" of possible values for uncertain parameters
- Optimizes with respect to the worst-case
- Appropriate when decision-making involves high risks or adversarial participants

Motivating example: numerical error

PILOT4 from NETLIB library

- Constraint 372

$$
\begin{aligned}
a^T x \equiv & -15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417x_{829} \\
& -1.526049x_{830} - 0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} \\
& -0.19004x_{852} - 2.757176x_{853} - 12.290832x_{854} + 717.562256x_{855} \\
& -0.057865x_{856} - 3.785417x_{857} - 78.30661x_{858} - 122.163055x_{859} \\
& -6.46609x_{860} - 0.48371x_{861} - 0.615264x_{862} - 1.353783x_{863} \\
& -84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712592x_{870} \\
& -0.401597x_{871} + x_{880} - 0.946049x_{898} - 0.946049x_{916} \\
\geq & \; b \equiv 23.387405
\end{aligned}
$$

- Optimal "deterministic" solution

$$
\begin{aligned}
x^*_{826} &= 255.6112787181108 & x^*_{827} &= 6240.488912232100 \\
x^*_{828} &= 3624.613324098961 & x^*_{829} &= 18.20205065283259 \\
x^*_{849} &= 174397.0389573037 & x^*_{870} &= 14250.00176680900 \\
x^*_{871} &= 25910.00731692178 & x^*_{880} &= 104958.3199274139
\end{aligned}
$$

- Can the coefficients be known with such a high accuracy?
- Define the relative constraint violation as:

$$V = \frac{b - a^T x^*}{b} \times 100\%$$

- Assume 0.1%-accurate approximation $\Longrightarrow$ worst V is about 450% !!!
- Considering instead

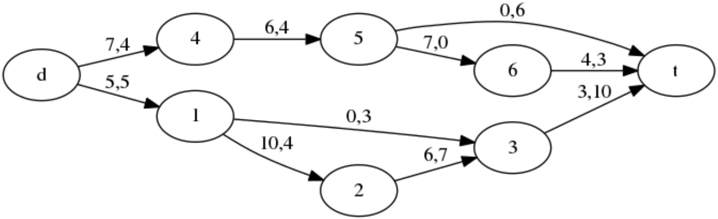
$$\tilde{a} = (1 + \xi_j) a_j$$

where ξ_j aree iid in $[-0.001, 0.001]$ yields:

Prob$\{V > 0\}$	Prob$\{V > 150\%\}$	Mean(V)
0.50	0.18	125%

An example: Shortest path

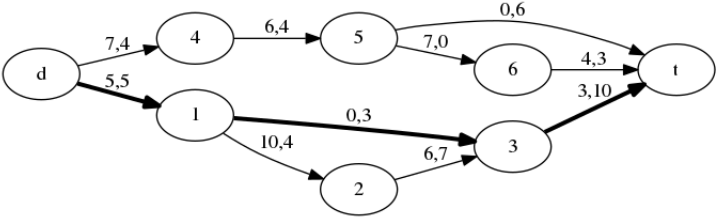
- Consider a shortest path problem with arc cost uncertainty.
- Two cost scenarios with probability 0.5 for each arc.



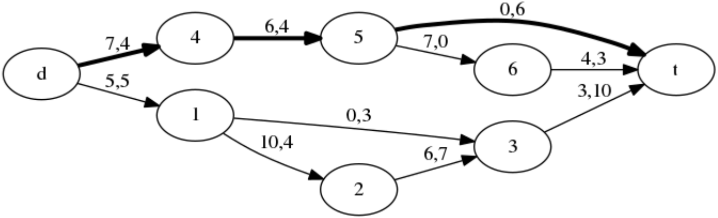
- Minimize the expected value of the cost of going from d to t.
- Minimize the worst case value of the cost of going from d to t.

An example: Shortest path

An example: Shortest path



An example: Shortest path



Let's formalize

- Consider the following MILP where the parameters $\boldsymbol{c}$, $\boldsymbol{A}$, and $\boldsymbol{b}$ can be uncertain:

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\
& \boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}
\end{aligned}
$$

- In robust optimization, an uncertain mixed integer linear programming problem is defined as a collection

$$
\left\{ \min_{\boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}} \{ \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \} \middle| (\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b}) \in \mathcal{U} \right\}
$$

  of problems of a common structure where the data $(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$ varies in a given uncertainty set $\mathcal{U}$.
- We will assume that $\mathcal{U}$ is a compact set.

Fundamental assumptions

1. All decision variables represent “here and now” decisions; the problem should be solved before the actual uncertain data “reveals itself”, *i.e.*, no recourse.
2. The decision maker is fully responsible for consequences of the decisions to be made when, and only when, the actual uncertain data is within the prespecified uncertainty set $\mathcal{U}$.
3. The constraints are “hard”, *i.e.*, we cannot tolerate the violations of constraints, even small ones, when the data is in $\mathcal{U}$.

Feasibility?/Optimality?

- A *robust feasible solution* satisfies the constraints $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ for all realizations of $\boldsymbol{A}$ and $\boldsymbol{b}$ in the uncertainty set.
- A *robust optimal solution* is a robust feasible solution that achieves the minimum "worst-case" objective value, defined as

$$\sup\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \mid \boldsymbol{c} \in \mathcal{U}\}.$$

  for a given $\boldsymbol{x}$.
- Then the robust counterpart of the LP is written as:

$$\min_{\boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}} \left\{ \sup_{\boldsymbol{c} \in \mathcal{U}} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \;\middle|\; \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \quad \forall (\boldsymbol{A}, \boldsymbol{b}) \in \mathcal{U} \right\}$$

- Since $\mathcal{U}$ is assumed to be compact, equivalently written as:

$$\min_{\boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}} \left\{ \max_{\boldsymbol{c} \in \mathcal{U}} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \;\middle|\; \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \quad \forall (\boldsymbol{A}, \boldsymbol{b}) \in \mathcal{U} \right\}$$

Key observations

Proposition

An uncertain linear optimization problem can always be reformulated as an uncertain linear optimization problem with a deterministic objective and right-hand side.

A standard robust optimization model

- Only constraint uncertainty
- $\mathcal{U}_i$ for $i = 1, \ldots, m$ for each constraint closed and convex
- $\mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_m$

$$
\text{(RLP)} \qquad \begin{array}{lll} \min & \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} & \\ \text{s.t.} & \boldsymbol{a}_i^{\mathrm{T}} \boldsymbol{x} \geq \boldsymbol{b}_i & \forall \boldsymbol{a}_i \in \mathcal{U}_i, \forall i = 1, \ldots, m \\ & \boldsymbol{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2} & \end{array}
$$

Budgeted uncertainty set

- One interesting polyhedral construction is due to Bertsimas and Sim.
- Let for each a_{ij}, $\bar{a}_{ij}$ be its mean value and $\hat{a}_{ij}$ be its maximum deviation.
- Let Γ be a given parameter.
- We propose the polyhedron

$$\mathcal{U}_i^{\Gamma} = \left\{ \boldsymbol{a}_i \in \mathbb{R}^n \middle| \sum_{j=1}^{n} \left| \frac{a_{ij} - \bar{a}_{ij}}{\hat{a}_{ij}} \right| \leq \Gamma, \bar{a}_{ij} - \hat{a}_{ij} \leq a_{ij} \leq \bar{a}_{ij} + \hat{a}_{ij}, \forall j = 1, \ldots, n \right\}.$$

Budgeted uncertainty set

- Γ restricts the number of parameters that deviate from their mean value simultaneously.
- It models the level of "conservatism" of a decision maker.

Relation to chance constraints

- Consider a tolerance value $\epsilon \in]0, 1[$.
- Let random variables $\tilde{a}_i$ iid uniform in $[\bar{a}_i - \hat{a}_i, \bar{a}_i + \hat{a}_i]$
- Consider the following chance constraint

$$\mathbb{P}\left\{\sum_{i=1}^{n} \tilde{a}_i x_i \leq b_i\right\} \geq 1 - \epsilon.$$

- A well-known theorem states that if $\boldsymbol{x}$ satisfies the robust constraint

$$\sum_{i=1}^{n} a_i x_i \leq b \qquad \forall \boldsymbol{a} \in \mathcal{U}^{\Gamma}$$

then $\boldsymbol{x}$ satisfies the chance constraint with $\epsilon \leq e^{-\frac{\Gamma^2}{2n}}$ [Bertsimas and Sim, 2004].

- This immediately yields a safe approximation of chance constraints. It suffices to fix the tolerance ϵ and solve for Γ.