Optimization Under Uncertainty

(but really, just Robust Optimization)

Lecture 18

December 2, 2024

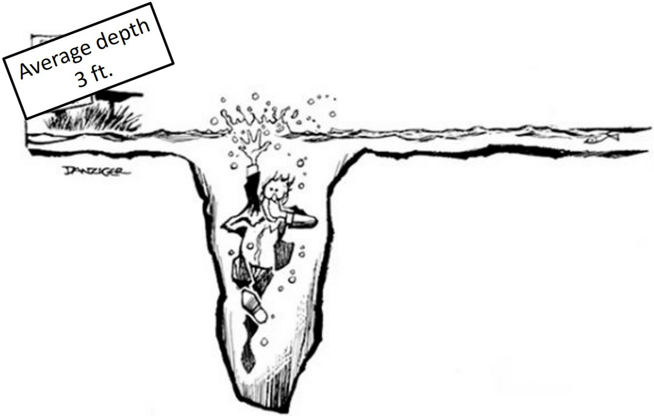
Quick Announcements

- Homework 5 due on Tuesday (Dec 3)
- Office Hours this week - extended schedule (Ed Announcement coming up)
- Final exam topics
- Any questions?

Outline for Today

1. Introduction
   - Some Motivating Examples
   - A History Detour
   - Pros and Cons of Probabilistic Models
2. Robust Optimization
   - Basic Premises
   - Modeling with Basic Uncertainty Sets
   - Reformulating and Solving Robust Models
   - Extensions
   - Some Applications
   - Calibrating Uncertainty Sets
   - Distributionally Robust Optimization
   - Connections with Other Areas
3. Dynamic Robust Optimization
   - Properly Writing a Robust DP
   - An Inventory Example
   - Tractable Approximations with Decision Rules
   - Some Practical Issues
   - Bellman Optimality
   - An Application in Monitoring

Optimization based on nominal values can lead to severe issues...



Taken from *"Flaw of averages"* Sam Savage (2009, 2012)

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Consider a **real-world scheduling problem** problem (PILOT4) in NETLIB Library
  - One of the constraints is the following linear constraint $\bar{\mathbf{a}}^\top \mathbf{x} \geqslant b$:

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

  - Coefficients like *8.598819* are estimated and potentially inaccurate

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Consider a **real-world scheduling problem** problem (PILOT4) in NETLIB Library
  - One of the constraints is the following linear constraint $\bar{\boldsymbol{a}}^T\boldsymbol{x} \geqslant b$:

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

  - Coefficients like *8.598819* are estimated and potentially inaccurate

- What if these coefficients are just 0.1% inaccurate?
  - i.e., suppose the true $\boldsymbol{a}$ is not $\bar{\boldsymbol{a}}$, but $|a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|$?

- Will the optimal solution to the problem still be feasible?

- How can we test?

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Original constraint: $\bar{\mathbf{a}}^\top \boldsymbol{x} \geqslant b$, optimal solution $\boldsymbol{x}^\star$
- Suppose true $\mathbf{a}$ satisfies $|a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|, \forall\, i$

- How to determine if the constraint is violated?

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Original constraint: $\bar{\boldsymbol{a}}^\top \boldsymbol{x} \geqslant b$, optimal solution $\boldsymbol{x}^\star$
- Suppose true $\boldsymbol{a}$ satisfies $|a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|$, $\forall\, i$

- How to determine if the constraint is violated?

$$\begin{aligned} \min_{\boldsymbol{a}} \;& \boldsymbol{a}^\top \boldsymbol{x}^\star - b \\ \text{s.t.} \;& |a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|, \;\forall\, i \end{aligned}$$

  - For PILOT4, this comes to $-128.8 \approx -4.5b$, so 450% violation!

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Original constraint: $\bar{\mathbf{a}}^{\mathsf{T}}\mathbf{x} \geqslant b$, optimal solution $\mathbf{x}^\star$
- Suppose true $\mathbf{a}$ satisfies $|a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|$, $\forall\, i$
- How to determine if the constraint is violated?

$$\begin{aligned} \min_{\mathbf{a}}\ & \mathbf{a}^{\mathsf{T}}\mathbf{x}^\star - b \\ \text{s.t.}\ & |a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|,\ \forall\, i \end{aligned}$$

  - For PILOT4, this comes to $-128.8 \approx -4.5b$, so 450% violation!

- OK, but perhaps we're too conservative?
  - Suppose $a_i = \bar{a}_i + \varepsilon_i|\bar{a}_i|$, where $\varepsilon_i \sim \text{Uniform}[-0.001, 0.001]$
  - Using Monte-Carlo simulation with 1,000 samples:
    - ★ $\mathbb{P}(\text{infeasible}) = 50\%$, $\mathbb{P}(\text{violation} > 150\%) = 18\%$, $\mathbb{E}[\text{violation}] = 125\%$

How Robust Are Optimal Solutions? (Ben-Tal & Nemirovski)

- Original constraint: $\bar{a}^\top x \geqslant b$, optimal solution $x^\star$
- Suppose true a satisfies $|a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|, \forall\, i$

- How to determine if the constraint is violated?

$$\begin{aligned} \min_{a}\ & a^\top x^\star - b \\ \text{s.t.}\ & |a_i - \bar{a}_i| \leqslant 0.001|\bar{a}_i|, \forall\, i \end{aligned}$$

  - For PILOT4, this comes to $-128.8 \approx -4.5b$, so 450% violation!

- OK, but perhaps we're too conservative?
  - Suppose $a_i = \bar{a}_i + \varepsilon_i|\bar{a}_i|$, where $\varepsilon_i \sim \text{Uniform}[-0.001, 0.001]$
  - Using Monte-Carlo simulation with 1,000 samples:
    - ★ $\mathbb{P}(\text{infeasible}) = 50\%$, $\mathbb{P}(\text{violation} > 150\%) = 18\%$, $\mathbb{E}[\text{violation}] = 125\%$

- Disturbing that nominal solutions are likely highly infeasible
- Turns out to be the case for many **NETLIB** problems
- We should **capture uncertainty more explicitly** apriori!

Decisions Under Uncertainty

- Decision Maker (DM) must chose x, without knowing z

- DM incurs a **cost** $C(x, z)$

Decisions Under Uncertainty

- Decision Maker (DM) must chose x, without knowing z
- DM incurs a **cost** $C(x, z)$
- How to model z? How to properly formalize the decision problem?

Decisions Under Uncertainty

- Decision Maker (DM) must chose x, without knowing z
- DM incurs a **cost** $C(x, z)$
- How to model z? How to properly formalize the decision problem?
- "Standard" probabilistic model:
  - There is a unique probability distribution $\mathbb{P}$ for z
  - DM considers an objective: $\min_x \ \mathbb{E}_{z\sim\mathbb{P}}\big[C(x, z)\big]$

Classical Probabilistic Model: DM knows $\mathbb{P}$, solves $\min_{x} \ \mathbb{E}_{z\sim\mathbb{P}}\left[C(x,z)\right]$

Classical Probabilistic Model: DM knows $\mathbb{P}$, solves $\min_x \ \mathbb{E}_{z\sim\mathbb{P}}\big[C(x, z)\big]$

- What if there are constraints?
  $f_i\,(x, z) \geqslant 0,\ \forall\, i \in I$
- Where is $\mathbb{P}$ coming from?
- When is this reasonable?
- What if $\mathbb{P}$ is not the actual distribution?
- What if $\mathbb{P}$ is not exogenous?

Classical Probabilistic Model: DM knows $\mathbb{P}$, solves $\min_x \ \mathbb{E}_{z\sim\mathbb{P}}\big[C(x, z)\big]$

- What if there are constraints?
  $f_i\,(x, z) \geqslant 0,\ \forall\, i \in I$
- Where is $\mathbb{P}$ coming from?
- When is this reasonable?
- What if $\mathbb{P}$ is not the actual distribution?
- What if $\mathbb{P}$ is not exogenous?

- expectation constraint:
  $\mathbb{E}_{\mathbb{P}}[\, f_i\big(x, z\big)] \geqslant 0,\ \forall\, i$
- chance constraint:
  individual: $\mathbb{P}[\, f_i\big(x, z\big) \geqslant 0\,] \geqslant 1 - \epsilon,\ \forall\, i$
  joint: $\mathbb{P}[\, f_i\big(x, z\big) \geqslant 0,\ \forall\, i] \geqslant 1 - \epsilon$
- robust (a.s.) constraint: $F\big(x, z\big) \geqslant 0, \forall\, z$
- easy to check / enforce?

Classical Probabilistic Model: DM knows $\mathbb{P}$, solves $\min_x \ \mathbb{E}_{z\sim\mathbb{P}}\big[C(x, z)\big]$

- What if there are constraints? $f_i\,(x, z) \geqslant 0, \ \forall\, i \in I$
- Where is $\mathbb{P}$ coming from?
- When is this reasonable?
- What if $\mathbb{P}$ is not the actual distribution?
- What if $\mathbb{P}$ is not exogenous?

- expectation constraint: $\mathbb{E}_{\mathbb{P}}[\,f_i\big(x, z\big)] \geqslant 0, \ \forall\, i$
- chance constraint:
  - individual: $\mathbb{P}[\,f_i\big(x, z\big) \geqslant 0\,] \geqslant 1 - \epsilon, \ \forall\, i$
  - joint: $\mathbb{P}[\,f_i\big(x, z\big) \geqslant 0, \ \forall\, i] \geqslant 1 - \epsilon$
- robust (a.s.) constraint: $F\big(x, z\big) \geqslant 0, \forall\, z$
- easy to check / enforce?

- Theory unable to analyze **complex, real-world** dynamics
  - poor data, changing environments (future $\neq$ past), many agents, ...
- Framework not geared towards **computing decisions**
  - Limited computational tractability, particularly in higher dimensions
- With $C = -u(\cdot)$ (u utility function), unclear if this is a good behavioral model

An Alternative Model of Uncertainty

- Let's admit **explicitly** that our model of reality is **incorrect**
- From **classical view**: *"we know distribution* $\mathbb{P}$ *for* z*, and solve:* $\min_{x} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

  to **robust view**: *"we only know that* $\mathbb{P} \in \mathcal{P}$*, and solve:* $\min_{x} \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

An Alternative Model of Uncertainty

- Let's admit **explicitly** that our model of reality is **incorrect**
- From **classical view**: *"we know distribution* $\mathbb{P}$ *for* z*, and solve:* $\min_{x} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

  to **robust view**: *"we only know that* $\mathbb{P} \in \mathcal{P}$*, and solve:* $\min_{x} \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

Long history of **robust decision-making** and **model misspecification**:

- **Economics:**
  - Frank Knight (1921) - risk vs. Knightian uncertainty, Abraham Wald (1939), John von Neumann (1944) zero-sum games
  - Savage (1951): minimax regret, Scarf (1958): robust Newsvendor model
  - Schmeidler, Gilboa (1980s): axiomatic frameworks, Ben-Haim (1980s): info-gap theory
  - Hansen & Sargent (2008): *"Robustness"* - robust control in macroeconomics
  - Bergemann & Morris (2012): *"Robust mechanism design"* book, Carroll (2015), ...
- **Engineering and robust control:** Bertsekas (1970s), Doyle (1980s), etc.
- **Computer science:** complexity analysis; adversarial training (modern!)
- **Statistics:** M-estimators Huber (1981)
- **Operations Research:**
  - Early work by Soyster (1973), Libura (1980), Bard (1984), Kouvelis (1997)
  - **Robust Optimization**: Ben-Tal, Nemirovski, El-Ghaoui ('90s), Bertsimas, Sim ('00s)
  - Two books: Ben-Tal, El-Ghaoui, Nemirovski (2009), Bertsimas, den Hertog (2020)
  - Many tutorials!

An Alternative Model of Uncertainty

- Let's admit **explicitly** that our model of reality is **incorrect**
- From **classical view**: *"we know distribution* $\mathbb{P}$ *for* z*, and solve:* $\min_{x} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

  to **robust view**: *"we only know that* $\mathbb{P} \in \mathcal{P}$*, and solve:* $\min_{x} \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}\big[C(x, z)\big]$*"*

Why robust optimization? (in my view)

1. Very sensible
2. Modest modeling requirements
3. Modest in its premise: *"always under-promises, and over-delivers"*
4. Tractable: quickly becoming "technology"
5. Very sensible results: can rationalize simple rules in complex problems

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array}}$$

“Classical” Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{array}{l} \inf_{x} \sup_{z \in \mathcal{U}} \; C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \; \forall i \in I \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
   - Other options possible, based on notions of **regret**

- Conservative?

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \quad \begin{aligned} &\inf_{x} \sup_{z \in \mathcal{U}} \; C(x, z) \\ &\text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \; \forall i \in I \end{aligned}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
    - Other options possible, based on notions of **regret**

- Conservative?
    - **Not necessarily!**
    - $\mathcal{U}$ directly trades off robustness and conservatism, and is ultimately a **modeling choice**

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{aligned} &\inf_{x} \; \sup_{z \in \mathcal{U}} \; C(x, z) \\ &\text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \; \forall i \in I \end{aligned}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
   - Other options possible, based on notions of **regret**

- Conservative?
  - **Not necessarily!**
  - $\mathcal{U}$ directly trades off robustness and conservatism, and is ultimately a **modeling choice**

- Is there a probabilistic interpretation?

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(P) \quad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
   - Other options possible, based on notions of **regret**

- Conservative?
  - **Not necessarily!**
  - $\mathcal{U}$ directly trades off robustness and conservatism, and is ultimately a **modeling choice**

- Is there a probabilistic interpretation?
  - Objective $= \sup\limits_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{z \sim \mathbb{P}}[C(x, z)]$ where $\mathcal{P}$ is the set of all measures with support $\mathcal{U}$
  - So we are assuming that the only information about $\mathbb{P}$ is the support $\mathcal{U}$

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{aligned} &\inf_{x} \sup_{z \in \mathcal{U}} C(x, z) \\ &\text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{aligned}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z

What is the optimal value of the following robust LP?

$$\begin{aligned} \min_{x} \max_{a \in \mathcal{U}} \quad & -(x_1 + x_2) \\ \text{such that} \quad & x_1 \leqslant a_1 \\ & x_2 \leqslant a_2 \\ & x_1 + x_2 \leqslant 1. \end{aligned}$$

where $\mathcal{U} = \{(a_1, a_2) \in [0,1]^2 : a_1 + a_2 \leqslant 1\}$

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} \; C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \; \forall i \in I \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z

What is the optimal value of the following robust LP?

$$\begin{array}{ll} \min\limits_{x} \max\limits_{a \in \mathcal{U}} & -(x_1 + x_2) \\ \text{such that} & x_1 \leqslant a_1 \\ & x_2 \leqslant a_2 \qquad\qquad \text{where } \mathcal{U} = \left\{(a_1, a_2) \in [0,1]^2 \; : \; a_1 + a_2 \leqslant 1\right\} \\ & x_1 + x_2 \leqslant 1. \end{array}$$

Optimal value 0. In RO, **each constraint must be satisfied separately, robustly.**

“Classical” Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{aligned} &\inf_{x} \sup_{z \in \mathcal{U}} \; C(x, z) \\ &\text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{aligned}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
2. Each constraint is “hard”: must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{\begin{array}{ll} (\mathrm{P}) & \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array} \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem

$$\boxed{f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}} \quad \Leftrightarrow \quad \boxed{\sup_{z \in \mathcal{U}} f_i(x, z) \leqslant 0}$$

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$
\boxed{
(\mathrm{P}) \qquad
\begin{array}{l}
\inf\limits_{x} \sup\limits_{z \in \mathcal{U}} \; C(x, z) \\
\text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \ \forall i \in I
\end{array}
}
$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem
4. Without loss, we can consider a problem where z only appears in constraints

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \quad \begin{aligned} & \inf_{x} \sup_{z \in \mathcal{U}} C(x, z) \\ & \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{aligned}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem
4. Without loss, we can consider a problem where z only appears in constraints

(P) is equivalent to the following problem:

$$\begin{aligned} & \inf_{x,t} t \\ & \text{s.t. } t \geqslant C(x, z), \forall z \in \mathcal{U} \\ & \qquad f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{aligned}$$

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{\begin{array}{ll} (P) & \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array} \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem
4. Without loss, we can consider a problem where z only appears in constraints

   (P) is equivalent to the following problem:

$$\begin{aligned} &\inf_{x,t} \ t \\ &\text{s.t. } t \geqslant C(x, z), \forall z \in \mathcal{U} \\ &\qquad f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{aligned}$$

   Many RO models are in this *epigraph reformulation*, and focus on constraints

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \qquad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} \; C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \; \forall i \in I \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} \; C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem
4. Without loss, we can consider a problem where z only appears in constraints
5. DM only responsible for objective and constraints when $z \in \mathcal{U}$
    - If $z \notin \mathcal{U}$ actually occurs, all bets are off
    - Can extend framework to ensure **gradual** degradation of performance: Globalized robust counterparts (Ben-Tal & Nemirovski)

"Classical" Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$\boxed{(\mathrm{P}) \quad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} \; C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array}}$$

Remarks.

1. Objective: worst-case performance $\sup_{z \in \mathcal{U}} C(x, z)$
2. Each constraint is "hard": must be satisfied *robustly*, for any realization of z
3. Each constraint can be re-written as an optimization problem
4. Without loss, we can consider a problem where z only appears in constraints
5. DM only responsible for objective and constraints when $z \in \mathcal{U}$
6. Robust model seems to lead to a **difficult** optimization problem
   - For any given x, checking constraints/solving the "adversary" problem may be tough
   - We must also solve our original problem of finding x!

“Classical” Robust Optimization (RO)

- Robust Optimization: the values of z belong to an **uncertainty set** $\mathcal{U}$
- DM reformulates the original optimization problem as:

$$(\mathrm{P}) \qquad \begin{array}{l} \inf\limits_{x} \sup\limits_{z \in \mathcal{U}} C(x, z) \\ \text{s.t. } f_i(x, z) \leqslant 0, \forall z \in \mathcal{U}, \forall i \in I \end{array}$$

1. How to model $\mathcal{U}$
2. How to formulate and solve the **robust counterpart**
3. Why is this useful, in theory and in practice

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some “safety buffers” for constraint like #372:

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

- Consider a **linear constraint** in $\boldsymbol{x}$ with coefficients that depend **linearly** on $\boldsymbol{z}$

$$\boxed{(\bar{\mathbf{a}} + P\boldsymbol{z})^{\mathsf{T}}\boldsymbol{x} \leqslant b, \ \forall \boldsymbol{z} \in \mathcal{U}}$$

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some "safety buffers" for constraint like #372:

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

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- P is a known matrix; z is primitive uncertainty

- **Q:** Why this more general form?

  A: For modeling flexibility:
  - Suppose the same physical quantity (i.e., coefficient) appears in multiple constraints
  - Can capture "correlations", e.g., with a factor model

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some "safety buffers" for constraint like #372:

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

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$
\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}
$$

- How about a **box** uncertainty set? For some confidence level ρ:

$$
\mathcal{U}_{\text{box}} := \{z : -\rho \leqslant z_i \leqslant \rho\} = \{z : \|z\|_\infty \leqslant \rho\}
$$

"Too conservative?"

- In PILOT4, **robust** solution is within 1% of x^* for objective
- Recall that x^* would violate this constraint by 450%
- Sometimes not much is sacrificed for robustness!

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some "safety buffers" for constraint like #372:

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

- Consider a **linear constraint** in x with coefficients that depend **linearly** on $\textcolor{red}{z}$

$$
\boxed{(\bar{a} + P\textcolor{red}{z})^{T} x \leqslant b, \ \textcolor{red}{\forall z \in \mathcal{U}}}
$$

- How about a **box** uncertainty set? For some confidence level ρ:

$$
\mathcal{U}_{\text{box}} := \{z : -\rho \leqslant z_i \leqslant \rho\} = \{z : \|z\|_\infty \leqslant \rho\}
$$

- Or maybe an **ellipsoid** would be less conservative:

$$
\mathcal{U}_{\text{ellipsoid}} := \{z : \|z\|_2 \leqslant \rho\}
$$

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some "safety buffers" for constraint like #372:

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

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- How about a **box** uncertainty set? For some confidence level ρ:

$$\mathcal{U}_{\text{box}} := \{z : -\rho \leqslant z_i \leqslant \rho\} = \{z : \|z\|_\infty \leqslant \rho\}$$

- Or maybe an **ellipsoid** would be less conservative:

$$\mathcal{U}_{\text{ellipsoid}} := \{z : \|z\|_2 \leqslant \rho\}$$

- Or what if we gave "nature" a **budget** on how many coefficients it could change:

$$\mathcal{U}_{\text{budget}} := \{z : \|z\|_\infty \leqslant \rho, \ \|z\|_1 \leqslant \Gamma\rho\}$$

Intuition for Some Basic Uncertainty Sets

- Recall PILOT4; how to build some "safety buffers" for constraint like #372:

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

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- How about a **box** uncertainty set? For some confidence level ρ:

$$\mathcal{U}_{\text{box}} := \{z : -\rho \leqslant z_i \leqslant \rho\} = \{z : \|z\|_\infty \leqslant \rho\}$$

- Or maybe an **ellipsoid** would be less conservative:

$$\mathcal{U}_{\text{ellipsoid}} := \{z : \|z\|_2 \leqslant \rho\}$$

- Or what if we gave "nature" a **budget** on how many coefficients it could change:

$$\mathcal{U}_{\text{budget}} := \{z : \|z\|_\infty \leqslant \rho, \ \|z\|_1 \leqslant \Gamma\rho\}$$

- How to formulate the robust counterpart? How to set ρ, Γ? How to use in practice?

Formulating the Robust Counterpart (RC) for Box Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\max_{z : \|z\|_\infty \leqslant \rho} (\bar{a} + Pz)^T x \leqslant b,$$

Formulating the Robust Counterpart (RC) for Box Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\max_{z:\|z\|_\infty \leqslant \rho} (\bar{a} + Pz)^T x \leqslant b,$$

or

$$\bar{a}^T x + \max_{z:\|z\|_\infty \leqslant \rho} (P^T x)^T z \leqslant b,$$

Formulating the Robust Counterpart (RC) for Box Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^\top x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\max_{z:\|z\|_\infty \leqslant \rho} (\bar{a} + Pz)^\top x \leqslant b,$$

or

$$\bar{a}^\top x + \max_{z:\|z\|_\infty \leqslant \rho} (P^\top x)^\top z \leqslant b,$$

or

$$\bar{a}^\top x + \max_{z:|z_i| \leqslant \rho} \sum_i (P^\top x)_i z_i \leqslant b,$$

Formulating the Robust Counterpart (RC) for Box Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\max_{z:\|z\|_\infty \leqslant \rho} (\bar{a} + Pz)^T x \leqslant b,$$

or

$$\bar{a}^T x + \max_{z:\|z\|_\infty \leqslant \rho} (P^T x)^T z \leqslant b,$$

or

$$\bar{a}^T x + \max_{z:|z_i| \leqslant \rho} \sum_i (P^T x)_i \, z_i \leqslant b,$$

or

$$\bar{a}^T x + \rho \sum_i |(P^T x)_i| \leqslant b,$$

Formulating the Robust Counterpart (RC) for Box Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\max_{z : \|z\|_\infty \leqslant \rho} (\bar{a} + Pz)^T x \leqslant b,$$

or

$$\bar{a}^T x + \max_{z : \|z\|_\infty \leqslant \rho} (P^T x)^T z \leqslant b,$$

or

$$\bar{a}^T x + \max_{z : |z_i| \leqslant \rho} \sum_i (P^T x)_i \, z_i \leqslant b,$$

or

$$\bar{a}^T x + \rho \sum_i |(P^T x)_i| \leqslant b,$$

or

$$\boxed{\bar{a}^T x + \rho \|P^T x\|_1 \leqslant b.}$$

Formulating the Robust Counterpart (RC) for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

Formulating the Robust Counterpart (RC) for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^{\top} x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^{\top} x + \max_{z : Dz \leqslant d} (P^{\top} x)^{\top} z \leqslant b. \tag{1}$$

Formulating the Robust Counterpart (RC) for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{\mathbf{a}} + \mathrm{P}\mathbf{z})^{\mathsf{T}}\mathbf{x} \leqslant \mathrm{b},\ \forall\, \mathbf{z} \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{\mathbf{z} : \mathrm{D}\mathbf{z} \leqslant \mathbf{d}\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^{\mathsf{T}}x + \max_{z\,:\,\mathrm{D}z \leqslant d} (\mathrm{P}^{\mathsf{T}}x)^{\mathsf{T}}z \leqslant b. \qquad (1)$$

By strong LP duality, when the left-hand-side in (1) is finite, we must have:

$$\max\{(\mathrm{P}^{\mathsf{T}}x)^{\mathsf{T}}z \;:\; \mathrm{D}z \leqslant d\} = \min\{d^{\mathsf{T}}y : \mathrm{D}^{\mathsf{T}}y = \mathrm{P}^{\mathsf{T}}x,\ y \geqslant 0\}.$$

Formulating the Robust Counterpart (RC) for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^{\mathsf{T}} x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^{\mathsf{T}} x + \max_{z : Dz \leqslant d} (P^{\mathsf{T}} x)^{\mathsf{T}} z \leqslant b. \tag{1}$$

By strong LP duality, when the left-hand-side in (1) is finite, we must have:

$$\max\{(P^{\mathsf{T}} x)^{\mathsf{T}} z \ : \ Dz \leqslant d\} = \min\{d^{\mathsf{T}} y : D^{\mathsf{T}} y = P^{\mathsf{T}} x, \ y \geqslant 0\}.$$

Hence (1) is equivalent to

$$\bar{a}^{\mathsf{T}} x + \min_{y}\{d^{\mathsf{T}} y \ : \ D^{\mathsf{T}} y = P^{\mathsf{T}} x, \ y \geqslant 0\} \leqslant b,$$

Formulating the Robust Counterpart (RC) for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^T x + \max_{z : Dz \leqslant d} (P^T x)^T z \leqslant b. \tag{1}$$

By strong LP duality, when the left-hand-side in (1) is finite, we must have:

$$\max\{(P^T x)^T z \ : \ Dz \leqslant d\} = \min\{d^T y : D^T y = P^T x, \ y \geqslant 0\}.$$

Hence (1) is equivalent to

$$\bar{a}^T x + \min_y \{d^T y \ : \ D^T y = P^T x, \ y \geqslant 0\} \leqslant b,$$

or

$$\boxed{\exists\, y \ : \ \bar{a}^T x + d^T y \leqslant b, \quad D^T y = P^T x, \quad y \geqslant 0.}$$

Formulating the Robust Counterpart for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}} \tag{2}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\boxed{\exists y : \bar{a}^T x + d^T y \leqslant b, \quad D^T y = P^T x, \quad y \geqslant 0.}$$

Remarks.

- To formulate the RC for (2), we must introduce a set of **auxiliary decision variables** y
  - these are **decision variables**, chosen together with x
- *How many* **auxiliary variables** *are needed to derive the RC for* (2)*?*
- *How many* **constraints** *are needed to derive the RC for* (2)*?*
- *Suppose we were solving* $\min_x\{c^T x : Ax \leqslant b\}$, *with* $A \in \mathbb{R}^{m \times n}$ *being uncertain. Under* $\mathcal{U}_{\text{polyhedral}}$ *and* $D \in \mathbb{R}^{p \times q}$, *what kind of problem is the RC of this LO, and how large is it?*

Formulating the Robust Counterpart for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}} \tag{2}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\boxed{\exists y : \bar{a}^T x + d^T y \leqslant b, \quad D^T y = P^T x, \quad y \geqslant 0.}$$

Remarks.

- To formulate the RC for (2), we must introduce a set of **auxiliary decision variables** y
  - these are **decision variables**, chosen together with x
- *How many* **auxiliary variables** *are needed to derive the RC for* (2)*?*
  - \# rows of D, i.e., as many as the constraints defining $\mathcal{U}_{\text{polyhedral}}$
- *How many* **constraints** *are needed to derive the RC for* (2)*?*
  - $1 + (\#\text{columns of } D) + (\#\text{rows of } D)$
- *Suppose we were solving* $\min_x\{c^T x : Ax \leqslant b\}$, *with* $A \in \mathbb{R}^{m \times n}$ *being uncertain. Under* $\mathcal{U}_{\text{polyhedral}}$ *and* $D \in \mathbb{R}^{p \times q}$, *what kind of problem is the RC of this LO, and how large is it?*

Formulating the Robust Counterpart for Polyhedral Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}} \tag{2}$$

- For $\mathcal{U}_{\text{polyhedral}} = \{z : Dz \leqslant d\}$, satisfying the constraint robustly is equivalent to:

$$\boxed{\exists\, y : \bar{a}^T x + d^T y \leqslant b, \quad D^T y = P^T x, \quad y \geqslant 0.}$$

Remarks.

- To formulate the RC for (2), we must introduce a set of **auxiliary decision variables** y
  - these are **decision variables**, chosen together with x
- *How many* **auxiliary variables** *are needed to derive the RC for* (2)*?*
  - $\#$ rows of D, i.e., as many as the constraints defining $\mathcal{U}_{\text{polyhedral}}$
- *How many* **constraints** *are needed to derive the RC for* (2)*?*
  - $1 + (\#\text{columns of } D) + (\#\text{rows of } D)$
- *Suppose we were solving* $\min_x \{c^T x : Ax \leqslant b\}$, *with* $A \in \mathbb{R}^{m \times n}$ *being uncertain. Under* $\mathcal{U}_{\text{polyhedral}}$ *and* $D \in \mathbb{R}^{p \times q}$, *what kind of problem is the RC of this LO, and how large is it?*
  - the RC of a **linear optimization** with $\mathcal{U}_{\text{polyhedral}}$ is still a linear optimization
  - $n + m \cdot p$ variables, $m \cdot (1 + p + q)$ constraints

Formulating the Robust Counterpart (RC) for Ellipsoidal Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^{T}x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{ellipsoid}} = \{z : \|z\|_2 \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

Formulating the Robust Counterpart (RC) for Ellipsoidal Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^\top x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{ellipsoid}} = \{z : \|z\|_2 \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^\top x + \max_{z:\|z\|_2 \leqslant \rho} (P^\top x)^\top z \leqslant b.$$

Formulating the Robust Counterpart (RC) for Ellipsoidal Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^\mathsf{T} x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{ellipsoid}} = \{z : \|z\|_2 \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^\mathsf{T} x + \max_{z:\|z\|_2 \leqslant \rho} (P^\mathsf{T} x)^\mathsf{T} z \leqslant b.$$

Intermezzo: $\max\ \{q^\mathsf{T} z : \|z\|_2 \leqslant \rho\}$ *or* $\max\ \{q^\mathsf{T} z : z^\mathsf{T} z \leqslant \rho^2\}$

Lagrange: $z = q/\lambda$, *and* $\lambda = \|q\|_2/\rho$.

Optimal objective value: $\frac{q^\mathsf{T} q}{\lambda} = \rho\|q\|_2$.

Formulating the Robust Counterpart (RC) for Ellipsoidal Uncertainty Set

- Consider a **linear constraint** in x with coefficients that depend **linearly** on z

$$\boxed{(\bar{a} + Pz)^T x \leqslant b, \ \forall z \in \mathcal{U}}$$

- For $\mathcal{U}_{\text{ellipsoid}} = \{z : \|z\|_2 \leqslant \rho\}$, satisfying the constraint robustly is equivalent to:

$$\bar{a}^T x + \max_{z:\|z\|_2 \leqslant \rho} (P^T x)^T z \leqslant b.$$

Intermezzo: $\max \{q^T z : \|z\|_2 \leqslant \rho\}$ *or* $\max \{q^T z : z^T z \leqslant \rho^2\}$

Lagrange: $z = q/\lambda$, *and* $\lambda = \|q\|_2/\rho$.

Optimal objective value: $\frac{q^T q}{\lambda} = \rho\|q\|_2$.

Hence robust counterpart (RC) is:

$$\boxed{\bar{a}^T x + \rho\|P^T x\|_2 \leqslant b.}$$

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{\mathbf{a}} + P\mathbf{z})^\top \mathbf{x} \leqslant b, \ \forall\, \mathbf{z} \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	**Robust Counterpart**	**Tractability**
Box	$\|\mathbf{z}\|_\infty \leqslant \rho$	$\bar{a}^\top x + \rho\|P^\top x\|_1 \leqslant b$	LO
Ellipsoidal	$\|\mathbf{z}\|_2 \leqslant \rho$	$\bar{a}^\top x + \rho\|P^\top x\|_2 \leqslant b$	CQO
Polyhedral	$D\mathbf{z} \leqslant d$	$\exists y \ : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \|\mathbf{z}\|_\infty \leqslant \rho \\ \|\mathbf{z}\|_1 \leqslant \Gamma \end{cases}$	$\exists y \ : \ \bar{a}^\top x + \rho\|y\|_1 + \Gamma\|P^\top x - y\|_\infty \leqslant b$	LO

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{\mathbf{a}} + P\mathbf{z})^\top \mathbf{x} \leqslant b, \ \forall \mathbf{z} \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	Robust Counterpart	Tractability
Box	$\|\mathbf{z}\|_\infty \leqslant \rho$	$\bar{a}^\top x + \rho \|P^\top x\|_1 \leqslant b$	LO
Ellipsoidal	$\|\mathbf{z}\|_2 \leqslant \rho$	$\bar{a}^\top x + \rho \|P^\top x\|_2 \leqslant b$	CQO
Polyhedral	$D\mathbf{z} \leqslant d$	$\exists y : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \|\mathbf{z}\|_\infty \leqslant \rho \\ \|\mathbf{z}\|_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^\top x + \rho \|y\|_1 + \Gamma \|P^\top x - y\|_\infty \leqslant b$	LO

- Problems above can be handled by large-scale modern solvers: CPLEX, Gurobi, etc.
- Some software now also handling automatic problem re-formulation
- If some of the decisions $\mathbf{x}$ are integer, problems above become MI-LO/CQO
- Already a lot of mileage in many practical problems:
  logistics and supply chain management, radiation therapy, scheduling, ...

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{\mathbf{a}} + P\mathbf{z})^\top \mathbf{x} \leqslant b,\ \forall \mathbf{z} \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	**Robust Counterpart**	**Tractability**
Box	$\\|z\\|_\infty \leqslant \rho$	$\bar{a}^\top x + \rho \\|P^\top x\\|_1 \leqslant b$	LO
Ellipsoidal	$\\|z\\|_2 \leqslant \rho$	$\bar{a}^\top x + \rho \\|P^\top x\\|_2 \leqslant b$	CQO
Polyhedral	$Dz \leqslant d$	$\exists y : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \\|z\\|_\infty \leqslant \rho \\ \\|z\\|_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^\top x + \rho \\|y\\|_1 + \Gamma \\|P^\top x - y\\|_\infty \leqslant b$	LO

- **Uncertainty in the right-hand side:** $(\bar{\mathbf{a}} + P\mathbf{z})^\top \mathbf{x} \leqslant b + \mathbf{p}^\top \mathbf{z},\ \forall \mathbf{z} \in \mathcal{U}$

  $\Leftrightarrow \bar{\mathbf{a}}^\top \mathbf{x} + (P^\top \mathbf{x} - p)^\top \mathbf{z} \leqslant b,\ \forall \mathbf{z} \in \mathcal{U}$, so can use base model

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{\mathbf{a}} + P\boldsymbol{z})^\top \boldsymbol{x} \leqslant b,\ \forall \boldsymbol{z} \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	**Robust Counterpart**	**Tractability**
Box	$\lVert \boldsymbol{z} \rVert_\infty \leqslant \rho$	$\bar{a}^\top x + \rho \lVert P^\top x \rVert_1 \leqslant b$	LO
Ellipsoidal	$\lVert \boldsymbol{z} \rVert_2 \leqslant \rho$	$\bar{a}^\top x + \rho \lVert P^\top x \rVert_2 \leqslant b$	CQO
Polyhedral	$D\boldsymbol{z} \leqslant d$	$\exists y : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \lVert \boldsymbol{z} \rVert_\infty \leqslant \rho \\ \lVert \boldsymbol{z} \rVert_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^\top x + \rho \lVert y \rVert_1 + \Gamma \lVert P^\top x - y \rVert_\infty \leqslant b$	LO

- **Uncertainty in the right-hand side:** $(\bar{a} + P\boldsymbol{z})^\top \boldsymbol{x} \leqslant b + \mathbf{p}^\top \boldsymbol{z},\ \forall \boldsymbol{z} \in \mathcal{U}$
  $\Leftrightarrow \bar{\mathbf{a}}^\top \boldsymbol{x} + (P^\top \boldsymbol{x} - p)^\top \boldsymbol{z} \leqslant b,\ \forall \boldsymbol{z} \in \mathcal{U}$, so can use base model
- **General convex uncertainty set:** $\mathcal{U} = \{\boldsymbol{z} : h_k(\boldsymbol{z}) \leqslant 0,\ k \in K\}$, $h_k(\cdot)$ convex?
  RC is $\exists \{\boldsymbol{w}_k, u_k\}_{k \in K} : \bar{\mathbf{a}}^\top \boldsymbol{x} + \sum_k \mathbf{u}_k h_k^\star(\boldsymbol{w}_k / u_k) \leqslant b,\ \sum_k \boldsymbol{w}^k = P^\top x,\ \mathbf{u} \geqslant 0$. $h_k^\star$ is convex conjugate of h_k

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{\mathbf{a}} + P\mathbf{z})^\top \mathbf{x} \leqslant b, \ \forall \mathbf{z} \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	Robust Counterpart	Tractability
Box	$\\|\mathbf{z}\\|_\infty \leqslant \rho$	$\bar{a}^\top x + \rho \\|P^\top x\\|_1 \leqslant b$	LO
Ellipsoidal	$\\|\mathbf{z}\\|_2 \leqslant \rho$	$\bar{a}^\top x + \rho \\|P^\top x\\|_2 \leqslant b$	CQO
Polyhedral	$D\mathbf{z} \leqslant d$	$\exists y : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \\|\mathbf{z}\\|_\infty \leqslant \rho \\ \\|\mathbf{z}\\|_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^\top x + \rho \\|y\\|_1 + \Gamma \\|P^\top x - y\\|_\infty \leqslant b$	LO

- **Uncertainty in the right-hand side:** $(\bar{a} + P\mathbf{z})^\top \mathbf{x} \leqslant b + \mathbf{p}^\top \mathbf{z}, \ \forall \mathbf{z} \in \mathcal{U}$
  $\Leftrightarrow \bar{\mathbf{a}}^\top \mathbf{x} + (P^\top \mathbf{x} - p)^\top \mathbf{z} \leqslant b, \ \forall \mathbf{z} \in \mathcal{U}$, so can use base model
- **General convex uncertainty set:** $\mathcal{U} = \{\mathbf{z} : h_k(\mathbf{z}) \leqslant 0, \ k \in K\}$, $h_k(\cdot)$ convex?
  RC is $\exists \{\mathbf{w}_k, u_k\}_{k \in K} : \bar{\mathbf{a}}^\top \mathbf{x} + \sum_k \mathbf{u}_k h_k^\star(\mathbf{w}_k / u_k) \leqslant b, \ \sum_k \mathbf{w}^k = P^\top x, \ \mathbf{u} \geqslant 0.$ $h_k^\star$ is convex conjugate of h_k
- **Constraint LHS general in x, linear in $\mathbf{z}$:** $(P\mathbf{z})^\top g(x) \leqslant b, \ \forall \mathbf{z} \in \mathcal{U}$
  To calculate RC, take $\bar{\mathbf{a}} = 0$ and replace $\mathbf{x}$ with $g(\mathbf{x})$ in our base-case model

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{a} + Pz)^\top x \leqslant b, \ \forall z \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	Robust Counterpart	Tractability
Box	$\|z\|_\infty \leqslant \rho$	$\bar{a}^\top x + \rho \|P^\top x\|_1 \leqslant b$	LO
Ellipsoidal	$\|z\|_2 \leqslant \rho$	$\bar{a}^\top x + \rho \|P^\top x\|_2 \leqslant b$	CQO
Polyhedral	$Dz \leqslant d$	$\exists y : \begin{cases} \bar{a}^\top x + d^\top y \leqslant b \\ D^\top y = P^\top x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \|z\|_\infty \leqslant \rho \\ \|z\|_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^\top x + \rho \|y\|_1 + \Gamma \|P^\top x - y\|_\infty \leqslant b$	LO

- **Uncertainty in the right-hand side:** $(\bar{a} + Pz)^\top x \leqslant b + p^\top z, \ \forall z \in \mathcal{U}$
  $\Leftrightarrow \bar{a}^\top x + (P^\top x - p)^\top z \leqslant b, \ \forall z \in \mathcal{U}$, so can use base model
- **General convex uncertainty set:** $\mathcal{U} = \{z : h_k(z) \leqslant 0, \ k \in K\}$, $h_k(\cdot)$ convex?
  RC is $\exists \{w_k, u_k\}_{k \in K} : \bar{a}^\top x + \sum_k u_k h_k^\star(w_k/u_k) \leqslant b, \ \sum_k w^k = P^\top x, \ u \geqslant 0$. $h_k^\star$ is convex conjugate of h_k
- **Constraint LHS general in x, linear in z:** $(Pz)^\top g(x) \leqslant b, \ \forall z \in \mathcal{U}$
  To calculate RC, take $\bar{a} = 0$ and replace x with $g(x)$ in our base-case model
- **Constraint LHS linear in $x \geqslant 0$, concave in z:** $x^\top g(\bar{a} + Pz) \leqslant b, \ \forall z \in \mathcal{U}, \ g_i(y)$ concave
  $\Leftrightarrow d^\top x \leqslant b, \ \forall (z, d) \in \mathcal{U}^+ := \{(z, d) \mid \exists a : a = \bar{a} + Pz, \ d \leqslant f(a), \ z \in \mathcal{U}\}$; now **linear** in (z, d), and $\mathcal{U}^+$ convex

RC for **Linear Optimization** Problems with Classical Sets

The robust counterpart for $\boxed{(\bar{a} + P z)^T x \leqslant b, \ \forall z \in \mathcal{U}}$ is:

U-set	$\mathcal{U}$	Robust Counterpart	Tractability
Box	$\|z\|_\infty \leqslant \rho$	$\bar{a}^T x + \rho \|P^T x\|_1 \leqslant b$	LO
Ellipsoidal	$\|z\|_2 \leqslant \rho$	$\bar{a}^T x + \rho \|P^T x\|_2 \leqslant b$	CQO
Polyhedral	$Dz \leqslant d$	$\exists y : \begin{cases} \bar{a}^T x + d^T y \leqslant b \\ D^T y = P^T x \\ y \geqslant 0 \end{cases}$	LO
Budget	$\begin{cases} \|z\|_\infty \leqslant \rho \\ \|z\|_1 \leqslant \Gamma \end{cases}$	$\exists y : \bar{a}^T x + \rho \|y\|_1 + \Gamma \|P^T x - y\|_\infty \leqslant b$	LO

- **Uncertainty in the right-hand side:** $(\bar{a} + Pz)^T x \leqslant b + p^T z, \ \forall z \in \mathcal{U}$
  $\Leftrightarrow \bar{a}^T x + (P^T x - p)^T z \leqslant b, \ \forall z \in \mathcal{U}$, so can use base model
- **General convex uncertainty set:** $\mathcal{U} = \{z : h_k(z) \leqslant 0, \ k \in K\}$, $h_k(\cdot)$ convex?
  RC is $\exists \{w_k, u_k\}_{k \in K} : \bar{a}^T x + \sum_k u_k h_k^\star(w_k / u_k) \leqslant b, \ \sum_k w^k = P^T x, \ u \geqslant 0$. $h_k^\star$ is convex conjugate of h_k
- **Constraint LHS general in x, linear in z:** $(Pz)^T g(x) \leqslant b, \ \forall z \in \mathcal{U}$
  To calculate RC, take $\bar{a} = 0$ and replace x with $g(x)$ in our base-case model
- **Constraint LHS linear in $x \geqslant 0$, concave in z:** $x^T g(\bar{a} + Pz) \leqslant b, \ \forall z \in \mathcal{U}$, $g_i(y)$ concave
  $\Leftrightarrow d^T x \leqslant b, \ \forall (z, d) \in \mathcal{U}^+ := \{(z, d) \mid \exists a : a = \bar{a} + Pz, \ d \leqslant f(a), \ z \in \mathcal{U}\}$; now **linear** in (z, d), and $\mathcal{U}^+$ convex
- **Constraint LHS convex in x and convex in z:** $f(x, z) \leqslant b$, f jointly convex
  Tractable if f has "easy" piece-wise description: $f(x, z) = \max_{k \in K} f_k(x, z)$, where f_k are cases that "worked"

Extensions

- **Uncertainty in the right-hand side:** $(\bar{a} + P\boldsymbol{z})^\top \boldsymbol{x} \leqslant b + \mathbf{p}^\top \boldsymbol{z},\ \forall\, \boldsymbol{z} \in \mathcal{U}$
  $\Leftrightarrow \bar{\mathbf{a}}^\top \boldsymbol{x} + (P^\top \boldsymbol{x} - p)^\top \boldsymbol{z} \leqslant b,\ \forall\, \boldsymbol{z} \in \mathcal{U}$, so can use base model

Extensions

- **Uncertainty in the right-hand side:** $(\bar{a} + P\boldsymbol{z})^\top \boldsymbol{x} \leqslant b + \mathbf{p}^\top \boldsymbol{z}, \ \forall\, \boldsymbol{z} \in \mathcal{U}$

  $\Leftrightarrow \bar{\mathbf{a}}^\top \boldsymbol{x} + (P^\top \boldsymbol{x} - p)^\top \boldsymbol{z} \leqslant b, \ \forall\, \boldsymbol{z} \in \mathcal{U}$, so can use base model

- **General convex uncertainty set:** $\mathcal{U} = \{\boldsymbol{z} : h_k(\boldsymbol{z}) \leqslant 0, \ k \in K\}$, $h_k(\cdot)$ convex?

  $\Leftrightarrow \ \exists\, \{\boldsymbol{w}_k, u_k\}_{k \in K} : \bar{\mathbf{a}}^\top \boldsymbol{x} + \sum_k \mathbf{u}_k h_k^\star(\boldsymbol{w}_k / u_k) \leqslant b, \ \sum_k \boldsymbol{w}^k = P^\top x, \ \mathbf{u} \geqslant 0.$

  $h_k^\star$ is convex conjugate of h_k

Extensions

- **Uncertainty in the right-hand side:** $(\bar{a} + Pz)^T x \leqslant b + p^T z,\ \forall z \in \mathcal{U}$

  $\Leftrightarrow \bar{a}^T x + (P^T x - p)^T z \leqslant b,\ \forall z \in \mathcal{U}$, so can use base model

- **General convex uncertainty set:** $\mathcal{U} = \{z : h_k(z) \leqslant 0,\ k \in K\}$, $h_k(\cdot)$ convex?

  $\Leftrightarrow \quad \exists \{w_k, u_k\}_{k \in K} : \bar{a}^T x + \sum_k u_k h_k^\star(w_k/u_k) \leqslant b,\ \sum_k w^k = P^T x,\ u \geqslant 0.$

  $h_k^\star$ is convex conjugate of h_k

- **LHS general in x, linear in z:** $(Pz)^T g(x) \leqslant b,\ \forall z \in \mathcal{U}$

  To calculate RC, take $\bar{a} = 0$ and replace x with $g(x)$ in our base-case model

- **Uncertainty in the right-hand side:** $(\bar{a} + Pz)^T x \leqslant b + p^T z,\ \forall z \in \mathcal{U}$
  $\Leftrightarrow \bar{a}^T x + (P^T x - p)^T z \leqslant b,\ \forall z \in \mathcal{U}$, so can use base model
- **General convex uncertainty set:** $\mathcal{U} = \{z : h_k(z) \leqslant 0,\ k \in K\}$, $h_k(\cdot)$ convex?
  $\Leftrightarrow \quad \exists \{w_k, u_k\}_{k \in K} : \bar{a}^T x + \sum_k u_k h_k^\star(w_k/u_k) \leqslant b,\ \sum_k w^k = P^T x,\ u \geqslant 0.$
  $h_k^\star$ is convex conjugate of h_k
- **LHS general in x, linear in z:** $(Pz)^T g(x) \leqslant b,\ \forall z \in \mathcal{U}$
  To calculate RC, take $\bar{a} = 0$ and replace x with $g(x)$ in our base-case model
- **LHS linear in $x \geqslant 0$, concave in z:** $x^T g(\bar{a} + Pz) \leqslant b,\ \forall z \in \mathcal{U}$, g concave
  $\Leftrightarrow d^T x \leqslant b,\ \forall (z, d) \in \mathcal{U}^+ := \left\{(z, d) \mid \exists a : a = \bar{a} + Pz,\ \ d \leqslant f(a),\ \ z \in \mathcal{U}\right\}$
  now **linear** in (z, d), and $\mathcal{U}^+$ convex

Extensions

- **Uncertainty in the right-hand side:** $(\bar{\mathbf{a}} + P\mathbf{z})^T \mathbf{x} \leqslant b + \mathbf{p}^T \mathbf{z},\ \forall\, \mathbf{z} \in \mathcal{U}$

  $\Leftrightarrow\ \bar{\mathbf{a}}^T \mathbf{x} + (P^T \mathbf{x} - p)^T \mathbf{z} \leqslant b,\ \forall\, \mathbf{z} \in \mathcal{U}$, so can use base model

- **General convex uncertainty set:** $\mathcal{U} = \{\mathbf{z} : h_k(\mathbf{z}) \leqslant 0,\ k \in K\}$, $h_k(\cdot)$ convex?

  $\Leftrightarrow\ \ \exists\, \{\mathbf{w}_k, u_k\}_{k \in K} \ :\ \bar{\mathbf{a}}^T \mathbf{x} + \sum_k u_k h_k^\star(\mathbf{w}_k / u_k) \leqslant b,\ \sum_k \mathbf{w}^k = P^T x,\ \mathbf{u} \geqslant 0.$

  $h_k^\star$ is convex conjugate of h_k

- **LHS general in x, linear in z:** $(P\mathbf{z})^T g(\mathbf{x}) \leqslant b,\ \forall \mathbf{z} \in \mathcal{U}$

  To calculate RC, take $\bar{\mathbf{a}} = 0$ and replace $\mathbf{x}$ with $g(\mathbf{x})$ in our base-case model

- **LHS linear in $x \geqslant 0$, concave in z:** $x^T g(\bar{a} + Pz) \leqslant b,\ \forall\, z \in \mathcal{U}$, g concave

  $\Leftrightarrow\ d^T x \leqslant b,\ \forall (z, d) \in \mathcal{U}^+ := \left\{(z, d) \mid \exists a : a = \bar{a} + Pz,\ \ d \leqslant f(a),\ \ z \in \mathcal{U}\right\}$

  now **linear** in (z, d), and $\mathcal{U}^+$ convex

- **LHS convex in x and convex in z:** $f(\mathbf{x}, \mathbf{z}) \leqslant b$, f jointly convex

  Tractable if f has "easy" piece-wise description: $f(\mathbf{x}, \mathbf{z}) = \max_{k \in K} f_k(\mathbf{x}, \mathbf{z})$, where f_k are cases that "worked"

Used in many applications

- **inventory management** e.g., [Ben-Tal et al., 2005, Bertsimas and Thiele, 2006, Bienstock and Özbay, 2008, ...]
- **facility location and transportation** [Baron et al., 2011, ...]
- **scheduling** [Lin et al., 2004, Yamashita et al., 2007, Mittal et al., 2014, ...]
- **revenue management** [Perakis and Roels, 2010, Adida and Perakis, 2006, ...]
- **project management** [Wiesemann et al., 2012, Ben-Tal et al., 2009, ...]
- **energy generation and distribution** [Zhao et al., 2013, Lorca and Sun, 2015, ...]
- **portfolio optimization** [Goldfarb and Iyengar, 2003, Tütüncü and Koenig, 2004, Ceria and Stubbs, 2006, Pinar and Tütüncü, 2005, Bertsimas and Pachamanova, 2008, ...]
- **healthcare** [Borfeld et al., 2008, Hanne et al., 2009, Chen et al., 2011, I., Trichakis, Yoon (2018), ...]
- **humanitarian** [Uichano 2017, den Hertog et al., 2019, ...]

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

Step 1. Start with a **deterministic** model formulation:

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

Step 1. Start with a **deterministic** model formulation:

$$
\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i + K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \qquad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \qquad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I \in \{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

Parameters:
$\mathcal{T}$: discrete planning horizon, indexed by τ
$\mathcal{F}$: potential facility locations, indexed by i
$\mathcal{N}$: demand node locations, indexed by j
η: unit price of goods
c_i: cost per unit of production at facility i
C_i: cost per unit of capacity for facility i
K_i: cost of opening a facility at location i
d_{ij}: cost of shipping units from location i to j
$D_{j\tau}$: demand in period τ at location j.

Decision variables:
$X_{ij\tau}$: fraction of demand j in period τ satisfied by i
$P_{i\tau}$: quantity produced at facility i in period τ
I_i: whether facility i is open (0/1)
Z_i: capacity of facility i if open.

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

Step 1. Start with a **deterministic** model formulation:

$$\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i + K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \qquad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \qquad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}$$

Step 2. Identify all uncertain parameters and model the uncertainty set $\mathcal{U}$.

Baron et al. 2011 captured uncertain demands:

$$\mathcal{U} = \left\{ D\in\mathbb{R}^{|\mathcal{N}|\cdot|\mathcal{T}|} \;\middle|\; \sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}}\left(\frac{D_{jt}-\bar{D}_{jt}}{\epsilon_t \bar{D}_{jt}}\right)^2 \leqslant \rho^2 \right\},$$

$\{\bar{D}_{jt}\}_{j\in\mathcal{N};t\in\mathcal{T}}$ are "nominal" demands, ϵ_t is allowed deviation (%), ρ is the size of the ellipsoid.

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

Step 1. Start with a **deterministic** model formulation:

$$
\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i + K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \quad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \quad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \quad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \quad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

Step 2. Identify all uncertain parameters and model the uncertainty set $\mathcal{U}$.

Baron et al. 2011 captured uncertain demands:

$$
\mathcal{U} = \left\{ D \in \mathbb{R}^{|\mathcal{N}|\cdot|\mathcal{T}|} \;\middle|\; \sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}} \left(\frac{D_{jt} - \bar{D}_{jt}}{\epsilon_t \bar{D}_{jt}}\right)^2 \leqslant \rho^2 \right\},
$$

Equivalently, can write $D_{jt} = \bar{D}_{jt}(1 + \varepsilon_t \cdot z_{jt})$, where $z \in \mathcal{U} = \{z \in \mathbb{R}^{|\mathcal{N}|\cdot|\mathcal{T}|} : \|z\|_2 \leqslant \rho\}$

A Quick Example: A Facility Location Problem (Baron et al. 2011)

Need to decide where to open facilities, how much capacity to install, and how to assign customer demands over a future planning horizon, in order to maximize profits.

Step 1. Start with a **deterministic** model formulation:

$$
\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i + K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \qquad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \qquad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \qquad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I \in \{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

Step 2. Identify all uncertain parameters and model the uncertainty set $\mathcal{U}$.

Baron et al. 2011 captured uncertain demands:

$$
\mathcal{U} = \left\{ D \in \mathbb{R}^{|\mathcal{N}|\cdot|\mathcal{T}|} \;\middle|\; \sum_{j\in\mathcal{N}}\sum_{t\in\mathcal{T}} \left(\frac{D_{jt} - \bar{D}_{jt}}{\epsilon_t \bar{D}_{jt}}\right)^2 \leqslant \rho^2 \right\},
$$

Step 3. Derive robust counterpart for the problem. Here, this will be a Conic Quadratic program.

Caution when Building Robust Models:
Equivalent deterministic/nominal models may lead to different robust counterparts!

Caution when Building Robust Models:
Equivalent deterministic/nominal models may lead to different robust counterparts!

Our initial **deterministic** model:

$$
\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i - K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \quad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \quad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \quad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ Z_i \leqslant M\cdot I_i, \quad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \quad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

An equivalent **deterministic** model, with decisions for quantities:

$$
\begin{aligned}
\max_{Q,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})Y_{ij\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i - K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} Y_{ij\tau} \leqslant D_{j\tau}, \quad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} Y_{ij\tau} \leqslant P_{i\tau}, \quad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& Y_{ij\tau} \geqslant 0, \quad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ Z_i \leqslant M\cdot I_i, \quad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \quad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

Caution when Building Robust Models:
Equivalent deterministic/nominal models may lead to different robust counterparts!

Our initial **deterministic** model:

$$
\begin{aligned}
\max_{X,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})X_{ij\tau}D_{j\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i - K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} X_{ij\tau} \leqslant 1, \quad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} X_{ij\tau}D_{j\tau} \leqslant P_{i\tau}, \quad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& X_{ij\tau} \geqslant 0, \quad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \quad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

An equivalent **deterministic** model, with decisions for quantities:

$$
\begin{aligned}
\max_{Q,I,Z,P} \quad & \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}}\sum_{j\in\mathcal{N}}(\eta - d_{ij})Y_{ij\tau} - \sum_{\tau\in\mathcal{T}}\sum_{i\in\mathcal{F}} c_i P_{i\tau} - \sum_{i\in\mathcal{F}}(C_i Z_i - K_i I_i) \\
\text{s.t.} \quad & \sum_{i\in\mathcal{F}} Y_{ij\tau} \leqslant D_{j\tau}, \quad j\in\mathcal{N},\ \tau\in\mathcal{T}, \\
& \sum_{j\in\mathcal{N}} Y_{ij\tau} \leqslant P_{i\tau}, \quad i\in\mathcal{F},\ \tau\in\mathcal{T}, \\
& Y_{ij\tau} \geqslant 0, \quad i\in\mathcal{F},\ j\in\mathcal{N},\ \tau\in\mathcal{T} \\
& P_{i\tau} \leqslant Z_i,\ \ Z_i \leqslant M\cdot I_i, \quad i\in\mathcal{F},\ \tau\in\mathcal{T} \\
& I\in\{0,1\}^{|\mathcal{F}|}, \qquad \text{(where } M \text{ is a large enough constant.)}
\end{aligned}
$$

Are the **robust counterparts** of the two formulations **equivalent**? Which do you think will be **more conservative**?

HINT: In which model are future shipping decisions more "flexible," e.g., allowed to depend on realized demands?

Are Robust Solutions **Pareto-Efficient**?

$$\max_{x \in \mathcal{X}} \quad \min_{u \in \mathcal{U}} u^\top x$$

- Feasible set of solutions $\mathcal{X} = \left\{x \in \mathbb{R}^n : Ax \leqslant b\right\}$
- Uncertainty set of objective coefficients $\mathcal{U} = \{u \in \mathbb{R}^n : Du \geqslant d\}$

$$\max_{x \in \mathcal{X}} \quad \min_{u \in \mathcal{U}} u^\top x$$

- Feasible set of solutions $\mathcal{X} = \left\{x \in \mathbb{R}^n : Ax \leqslant b\right\}$
- Uncertainty set of objective coefficients $\mathcal{U} = \{u \in \mathbb{R}^n : Du \geqslant d\}$

- Classical RO framework results in
  - Optimal value $J^\star_{\mathsf{RO}}$
  - Set of robustly optimal solutions

$$X^{\mathsf{RO}} = \left\{x \in \mathcal{X} : \exists\, y \geqslant 0 \text{ such that } D^\top y = x, \quad y^\top d \geqslant J^\star_{\mathsf{RO}}\right\}$$

Set of **Robustly Optimal** Solutions

- $X^{\text{RO}} = \left\{x \in \mathcal{X} : \exists\, y \geqslant 0 \text{ such that } D^\top y = x, \quad y^\top d \geqslant J^\star_{\text{RO}}\right\}$
- $x \in X^{\text{RO}}$ guarantees that no other solution exists with higher **worst-case** objective value $u^\top x$

Set of **Robustly Optimal** Solutions

- $X^{\mathrm{RO}} = \left\{x \in \mathcal{X} : \exists\, y \geqslant 0 \text{ such that } D^{\top}y = x, \quad y^{\top}d \geqslant J^{\star}_{\mathrm{RO}}\right\}$
- $x \in X^{\mathrm{RO}}$ guarantees that no other solution exists with higher **worst-case** objective value $u^{\top}x$
- What if an uncertainty scenario materializes that does not correspond to the worst-case?
- Are there any guarantees that no other solution $\bar{x}$ exists that, apart from protecting us from worst-case scenarios, also performs better overall, under all possible uncertainty realizations?

Pareto Robustly Optimal solutions (Iancu & Trichakis 2014)

$$\max_{x \in \mathcal{X}} \quad \min_{u \in \mathcal{U}} u^\top x \tag{3}$$

Definition

A solution x is called a **Pareto Robustly Optimal (PRO) solution** for Problem (3) if

(a) it is robustly optimal, i.e., $x \in X^{\mathrm{RO}}$, and

(b) there is no $\bar{x} \in \mathcal{X}$ such that

$$u^\top \bar{x} \geqslant u^\top x, \quad \forall u \in \mathcal{U}, \quad \text{and}$$
$$\bar{u}^\top \bar{x} > \bar{u}^\top x, \text{ for some } \bar{u} \in \mathcal{U}.$$

Pareto Robustly Optimal solutions (Iancu & Trichakis 2014)

$$\max_{x \in \mathcal{X}} \quad \min_{u \in \mathcal{U}} u^\top x \tag{3}$$

Definition

A solution x is called a **Pareto Robustly Optimal (PRO) solution** for Problem (3) if

(a) it is robustly optimal, i.e., $x \in X^{\text{RO}}$, and

(b) there is no $\bar{x} \in \mathcal{X}$ such that

$$\begin{aligned} u^\top \bar{x} &\geqslant u^\top x, \quad \forall u \in \mathcal{U}, \quad \text{and} \\ \bar{u}^\top \bar{x} &> \bar{u}^\top x, \ \text{for some } \bar{u} \in \mathcal{U}. \end{aligned}$$

- $X^{\text{PRO}} \subseteq X^{\text{RO}}$: set of all PRO solutions

Some questions

- Given a RO solution, is it also PRO?
- How can one find a PRO solution?
- Can we optimize over $\mathcal{X}^{\text{PRO}}$?
- Can we characterize $\mathcal{X}^{\text{PRO}}$?
  - Is it non-empty?
  - Is it convex?
  - When is $\mathcal{X}^{\text{PRO}} = \mathcal{X}^{\text{RO}}$?
- How does the notion generalize in other RO formulations?

Finding PRO solutions

Theorem

Given a solution $x \in X^{\mathrm{RO}}$ *and an arbitrary point* $\bar{p} \in \mathrm{ri}(\mathcal{U})$, *consider the following linear optimization problem:*

$$\begin{array}{ll} \text{maximize} & \bar{p}^\top y \\ \text{subject to} & y \in \mathcal{U}^* \\ & x + y \in \mathcal{X}. \end{array}$$

Then, either

- $\mathcal{U}^* \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : y^\top u \geqslant 0, \forall\, u \in \mathcal{U}\}$ is the dual of $\mathcal{U}$

Finding PRO solutions

Theorem

Given a solution $x \in X^{\mathrm{RO}}$ *and an arbitrary point* $\bar{p} \in \mathrm{ri}(\mathcal{U})$*, consider the following linear optimization problem:*

$$\begin{array}{ll} \text{maximize} & \bar{p}^\top y \\ \text{subject to} & y \in \mathcal{U}^* \\ & x + y \in \mathcal{X}. \end{array}$$

Then, either

- *the optimal value is zero and* $x \in X^{\mathrm{PRO}}$*, or*

- $\mathcal{U}^* \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : y^\top u \geqslant 0, \forall\, u \in \mathcal{U}\}$ is the dual of $\mathcal{U}$

Finding PRO solutions

Theorem

Given a solution $x \in X^{\text{RO}}$ *and an arbitrary point* $\bar{p} \in \text{ri}(\mathcal{U})$, *consider the following linear optimization problem:*

$$\begin{array}{ll} \text{maximize} & \bar{p}^\top y \\ \text{subject to} & y \in \mathcal{U}^* \\ & x + y \in \mathcal{X}. \end{array}$$

Then, either

- *the optimal value is zero and* $x \in X^{\text{PRO}}$, *or*
- *the optimal value is strictly positive and* $\bar{x} = x + y^\star \in X^{\text{PRO}}$, *for any optimal* $y^\star$.

- $\mathcal{U}^* \stackrel{\text{def}}{=} \{y \in \mathbb{R}^n : y^\top u \geqslant 0, \forall\, u \in \mathcal{U}\}$ is the dual of $\mathcal{U}$

Remarks

- Finding a point $\bar{u} \in \mathrm{ri}(\mathcal{U})$ can be done efficiently using LP techniques
- Testing whether $x \in X^{\mathrm{RO}}$ is no harder than solving the classical RO problem in this setting
- Finding a PRO solution $x \in X^{\mathrm{PRO}}$ is no harder than solving the classical RO problem in this setting

Corollaries

- If $\bar{u} \in \text{ri}(\mathcal{U})$, all optimal solutions to the problem below are PRO:

$$\begin{array}{ll} \text{maximize} & \bar{u}^\top x \\ \text{subject to} & x \in X^{\text{RO}} \end{array}$$

- If $0 \in \text{ri}(\mathcal{U})$, then $X^{\text{PRO}} = X^{\text{RO}}$

- If $\bar{u} \in \text{ri}(\mathcal{U})$, then $X^{\text{PRO}} = X^{\text{RO}}$ if and only if the optimal value of this LP is zero:

$$\begin{array}{ll} \text{maximize} & \bar{u}^\top y \\ \text{subject to} & x \in X^{\text{RO}} \\ & y \in \mathcal{U}^* \\ & x + y \in \mathcal{X} \end{array}$$

Optimizing over / Understanding X^{PRO}

- Secondary objective r: can we solve

$$\begin{array}{ll} \text{maximize} & r^\top x \\ \text{subject to} & x \in X^{\mathsf{PRO}}? \end{array}$$

- Interesting case: $X^{\mathsf{RO}} \neq X^{\mathsf{PRO}}$

- Secondary objective r: can we solve

$$\begin{array}{ll} \text{maximize} & r^\top x \\ \text{subject to} & x \in X^{\text{PRO}}? \end{array}$$

Proposition

X^{PRO} *is not necessarily convex.*

- $\mathcal{X} = \{x \in \mathbb{R}^4_+ : x_1 \leqslant 1,\ x_2 + x_3 \leqslant 6,\ x_3 + x_4 \leqslant 5,\ x_2 + x_4 \leqslant 5\}$
- $\mathcal{U} = \text{conv}\Big(\{e_i,\ i \in \{1, \ldots, 4\}\}\Big)$
- $J^\star_{\text{RO}} = 1$, and $X^{\text{RO}} = \{x \in X : x \geqslant \mathbf{1}\}$
- $x^1 = \begin{bmatrix}1 & 2 & 4 & 1\end{bmatrix}^\text{T}$, $x^2 = \begin{bmatrix}1 & 4 & 2 & 1\end{bmatrix}^\text{T} \in X^{\text{PRO}}$
- $0.5\,x^1 + 0.5\,x^2$ is Pareto dominated by $\begin{bmatrix}1 & 3 & 3 & 2\end{bmatrix}^\text{T} \in X^{\text{RO}}$.

Optimizing over / Understanding X^{PRO}

- Secondary objective r: can we solve

$$\begin{array}{ll} \text{maximize} & r^\top x \\ \text{subject to} & x \in X^{\text{PRO}} ? \end{array}$$

Proposition

If $X^{\text{RO}} \neq X^{\text{PRO}}$, *then* $X^{\text{PRO}} \cap \text{ri}(X^{\text{RO}}) = \emptyset$.

- Whether solution to nominal RO is PRO depends on algorithm used for solving LP
- Simplex **better for RO problems** than interior point methods

What Are The Gains?

Example (Portfolio)

- $n+1$ assets, with returns r_i
- $r_i = \mu_i + \sigma_i \zeta_i$, $i = 1, \ldots, n$, $r_{n+1} = \mu_{n+1}$
- ζ unknown, $\mathcal{U} = \{\zeta \in \mathbb{R}^n : -\mathbf{1} \leqslant \zeta \leqslant \mathbf{1},\ \mathbf{1}^T\zeta = 0\}$
- Objective: select weights x to maximize worst-case portfolio return

What Are The Gains?

Example (Portfolio)

- $n + 1$ assets, with returns r_i
- $r_i = \mu_i + \sigma_i \zeta_i$, $i = 1, \dots, n$, $r_{n+1} = \mu_{n+1}$
- ζ unknown, $\mathcal{U} = \{\zeta \in \mathbb{R}^n : -\mathbf{1} \leqslant \zeta \leqslant \mathbf{1}, \mathbf{1}^T \zeta = 0\}$
- Objective: select weights x to maximize worst-case portfolio return

What Are The Gains?

Example (Portfolio)

- $n+1$ assets, with returns r_i
- $r_i = \mu_i + \sigma_i \zeta_i$, $i = 1, \ldots, n$, $r_{n+1} = \mu_{n+1}$
- ζ unknown, $\mathcal{U} = \{\zeta \in \mathbb{R}^n : -\mathbf{1} \leqslant \zeta \leqslant \mathbf{1},\ \mathbf{1}^T \zeta = 0\}$
- Objective: select weights x to maximize worst-case portfolio return

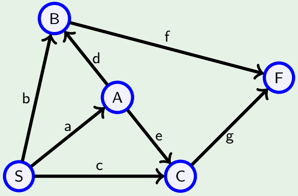
Example (Inventory)

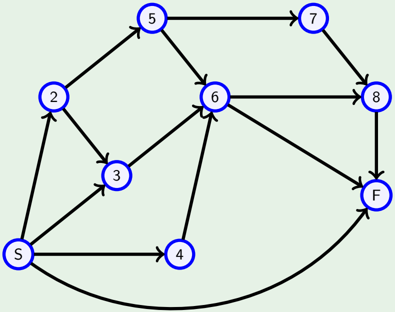
- One warehouse, N retailers where uncertain demand is realized
- Transportation, holding costs and profit margins differ for each retailer
- Demand driven by market factors $d_i = d_i^0 + q_i^T z, \quad i = 1, \ldots, N$
- Market factors z are uncertain

$$z \in \mathcal{U} = \{z \in \mathbb{R}^N : -b \cdot \mathbf{1} \leqslant z \leqslant b \cdot \mathbf{1},\ -B \leqslant \mathbf{1}^T z \leqslant B\}$$

Example (Project management)

- A PERT diagram given by directed, acyclic graph $G = (\mathcal{N}, \mathcal{E})$
- $\mathcal{N}$ are project events, $\mathcal{E}$ are project activities / tasks





Example (Project management)

- A PERT diagram given by directed, acyclic graph $G = (\mathcal{N}, \mathcal{E})$
- $\mathcal{N}$ are project events, $\mathcal{E}$ are project activities / tasks
- Task $e \in \mathcal{E}$ has uncertain duration $\tau_e = \tau_e^0 + \delta_e$

$$\delta \in \mathcal{U} := \left\{\delta \in \mathbb{R}_+^{|\mathcal{E}|} : \delta \leqslant b \cdot \mathbf{1}, \quad \mathbf{1}^\top \delta_e \leqslant B\right\}$$

- Task $e \in \mathcal{E}$ can be expedited by allocating a budgeted resource x_e

$$\begin{aligned} \tau_e &= \tau_e^0 + \delta_e - x_e \\ \mathbf{1}^\top x &\leqslant C \end{aligned}$$

- Goal: find resource allocation x to minimize worst-case completion time

Results – finance and inventory examples (10K instances)

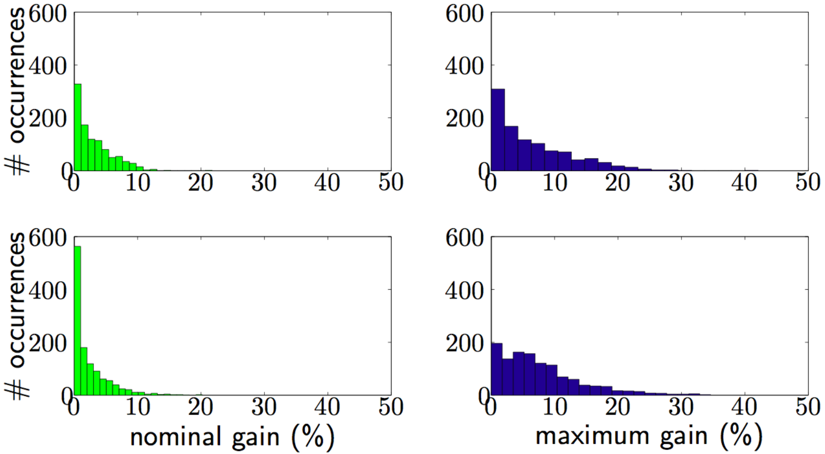
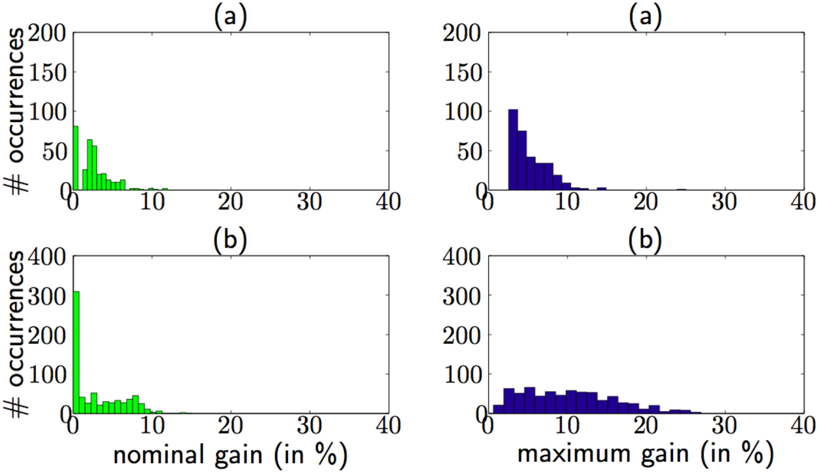


Figure: TOP: portfolio example. BOTTOM: inventory example. LEFT: performance gains in nominal scenario. RIGHT: maximal performance gains.



Careful To Avoid Naïve Inefficiencies In Robust Models!

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\big[(\bar{a} + Pz)^T x \leqslant b\big]$ is "large", $\forall\, \mathbb{P}$

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\big[(\bar{a} + Pz)^T x \leqslant b\big]$ is "large", $\forall\, \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big]$ is "large", $\forall\, \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

- *if x feasible for RC with $\mathcal{U}_{\text{box}} \stackrel{\text{def}}{=} \{z : \|z\|_\infty \leqslant 1\}$, $\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big] = 1$.*
  (no need for independence or 0-mean)

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\big[(\bar{a} + Pz)^T x \leqslant b\big]$ is "large", $\forall\, \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

- *if x feasible for RC with* $\mathcal{U}_{\text{box}} \stackrel{\text{def}}{=} \{z : \|z\|_\infty \leqslant 1\}$, $\mathbb{P}\big[(\bar{a} + Pz)^T x \leqslant b\big] = 1$.
  (no need for independence or 0-mean)
- *if x feasible for RC with* $\mathcal{U}_{\text{ellipsoid}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}\}$,

$$\mathbb{P}\big[(\bar{a} + Pz)^T x \leqslant b\big] \geqslant 1 - \epsilon.$$

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:
Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right]$ is "large", $\forall\, \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

- *if x feasible for RC with* $\mathcal{U}_{\text{box}} \stackrel{\text{def}}{=} \{z : \|z\|_\infty \leqslant 1\}$, $\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] = 1$.
  (no need for independence or 0-mean)
- *if x feasible for RC with* $\mathcal{U}_{\text{ellipsoid}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}\}$,
  $$\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] \geqslant 1 - \epsilon.$$
- *if x feasible for RC with* $\mathcal{U}_{\text{ellipsoid-box}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}, \|z\|_\infty \leqslant 1\}$,
  $$\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] \geqslant 1 - \epsilon.$$

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big]$ is "large", $\forall\, \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

- *if x feasible for RC with $\mathcal{U}_{\text{box}} \stackrel{\text{def}}{=} \{z : \|z\|_\infty \leqslant 1\}$, $\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big] = 1$.*
  (no need for independence or 0-mean)
- *if x feasible for RC with $\mathcal{U}_{\text{ellipsoid}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}\}$,*
  $$\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big] \geqslant 1 - \epsilon.$$
- *if x feasible for RC with $\mathcal{U}_{\text{ellipsoid-box}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}, \|z\|_\infty \leqslant 1\}$,*
  $$\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big] \geqslant 1 - \epsilon.$$
- *if x feasible for RC with $\mathcal{U}_{\text{budget}} = \{z \in \mathbb{R}^L : \|z\|_\infty \leqslant 1,\ \|z\|_1 \leqslant \Gamma = \sqrt{2\ln(1/\epsilon)}\sqrt{L}\}$,*
  $$\mathbb{P}\big[(\bar{a} + Pz)^\top x \leqslant b\big] \geqslant 1 - \epsilon.$$

How to Calibrate Uncertainty Sets?

We know how to derive RC for "structured" uncertainty sets. **But how to pick ρ, Γ?**

Let's take a **probabilistic** view for a moment:

Suppose z_i are really random, and we seek ρ, Γ to ensure $\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right]$ is "large", $\forall \mathbb{P}$

Theorem (High Probability of Constraint Satisfaction)

Suppose z_i are **independent r.v. with mean 0 and support on** $[-1, 1]$. *Then:*

- *if x feasible for RC with $\mathcal{U}_{\text{box}} \stackrel{\text{def}}{=} \{z : \|z\|_\infty \leqslant 1\}$, $\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] = 1$.*
  (no need for independence or 0-mean)
- *if x feasible for RC with $\mathcal{U}_{\text{ellipsoid}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}\}$,*
$$\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] \geqslant 1 - \epsilon.$$
- *if x feasible for RC with $\mathcal{U}_{\text{ellipsoid-box}} \stackrel{\text{def}}{=} \{z : \|z\|_2 \leqslant \rho \stackrel{\text{def}}{=} \sqrt{2\ln(1/\epsilon)}, \|z\|_\infty \leqslant 1\}$,*
$$\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] \geqslant 1 - \epsilon.$$
- *if x feasible for RC with $\mathcal{U}_{\text{budget}} = \{z \in \mathbb{R}^L : \|z\|_\infty \leqslant 1, \|z\|_1 \leqslant \Gamma = \sqrt{2\ln(1/\epsilon)}\sqrt{L}\}$,*
$$\mathbb{P}\left[(\bar{a} + Pz)^T x \leqslant b\right] \geqslant 1 - \epsilon.$$

- Some probabilistic information allows controlling conservatism: **very useful in applications**
- The budget Γ depends on the dimension of z (L), whereas ρ does not!
- Proof based on concentration inequalities

Another Quick Example: A Portfolio Problem (Ben-Tal and Nemirovski)

- 200 risky assets; asset # 200 is cash, with yearly return $r_{200} = 5\%$ and zero risk
- Yearly returns r_i are **independent r.v.** with values in $[\mu_i - \sigma_i, \mu_i + \sigma_i]$ and means μ_i:

$$\mu_i = 1.05 + 0.3\frac{(200-i)}{199}, \quad \sigma_i = 0.05 + 0.6\frac{(200-i)}{199}, \quad i = 1, \ldots, 199.$$

- Goal: distribute \$1 so as to maximize worst-case value-at-risk at level $\epsilon = 0.5\%$:

$$\max_{x,t}\left\{t \;:\; \mathbb{P}\left[\sum_{i=1}^{199} r_i x_i + r_{200}x_{200} \geqslant t\right] \geqslant 1-\epsilon,\ \forall\,\mathbb{P},\ \sum_{i=1}^{200} x_i = 1,\ x \geqslant 0\right\},$$

Another Quick Example: A Portfolio Problem (Ben-Tal and Nemirovski)

- 200 risky assets; asset # 200 is cash, with yearly return $r_{200} = 5\%$ and zero risk
- Yearly returns r_i are **independent r.v.** with values in $[\mu_i - \sigma_i, \mu_i + \sigma_i]$ and means μ_i:

$$\mu_i = 1.05 + 0.3\frac{(200-i)}{199}, \quad \sigma_i = 0.05 + 0.6\frac{(200-i)}{199}, \quad i = 1, \ldots, 199.$$

- Goal: distribute \$1 so as to maximize worst-case value-at-risk at level $\epsilon = 0.5\%$:

$$\max_{x,t}\left\{t \ : \ \mathbb{P}\left[\sum_{i=1}^{199} r_i x_i + r_{200}x_{200} \geqslant t\right] \geqslant 1-\epsilon, \ \forall\,\mathbb{P}, \ \sum_{i=1}^{200} x_i = 1, \ x \geqslant 0\right\},$$

- With $z_i \stackrel{\text{def}}{=} (r_i - \mu_i)/\sigma_i$, let's consider 3 uncertainty sets:
  1. $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant 1\}$
  2. $\mathcal{U}_{\text{ellipsoid-box}} = \{z : \|z\|_\infty \leqslant 1, \|z\|_2 \leqslant \rho\}$, with $\rho = \sqrt{2\ln(1/\epsilon)} = 3.255$
  3. $\mathcal{U}_{\text{budget}} = \{z : \|z\|_\infty \leqslant 1, \|z\|_1 \leqslant \Gamma\}$ with $\Gamma = \sqrt{2\ln(1/\epsilon)}\sqrt{199} = 45.921$.

Another Quick Example: A Portfolio Problem (Ben-Tal and Nemirovski)

- 200 risky assets; asset # 200 is cash, with yearly return $r_{200} = 5\%$ and zero risk
- Yearly returns r_i are **independent r.v.** with values in $[\mu_i - \sigma_i, \mu_i + \sigma_i]$ and means μ_i:

$$\mu_i = 1.05 + 0.3\frac{(200-i)}{199}, \quad \sigma_i = 0.05 + 0.6\frac{(200-i)}{199}, \quad i = 1, \ldots, 199.$$

- Goal: distribute \$1 so as to maximize worst-case value-at-risk at level $\epsilon = 0.5\%$:

$$\max_{x,t}\left\{t \ : \ \mathbb{P}\left[\sum_{i=1}^{199} r_i x_i + r_{200}x_{200} \geqslant t\right] \geqslant 1-\epsilon, \ \forall\, \mathbb{P}, \ \sum_{i=1}^{200} x_i = 1, \ x \geqslant 0\right\},$$

- With $z_i \stackrel{\text{def}}{=} (r_i - \mu_i)/\sigma_i$, let's consider 3 uncertainty sets:
  1. $\mathcal{U}_{\text{box}} = \{z : \|z\|_\infty \leqslant 1\}$
  2. $\mathcal{U}_{\text{ellipsoid-box}} = \{z : \|z\|_\infty \leqslant 1, \|z\|_2 \leqslant \rho\}$, with $\rho = \sqrt{2\ln(1/\epsilon)} = 3.255$
  3. $\mathcal{U}_{\text{budget}} = \{z : \|z\|_\infty \leqslant 1, \|z\|_1 \leqslant \Gamma\}$ with $\Gamma = \sqrt{2\ln(1/\epsilon)}\sqrt{199} = 45.921$.
- Results:
  - $\mathcal{U}_{\text{box}}$: worst-case returns $r_i = \mu_i - \sigma_i$ yield less than risk-free return of 5%, so optimal to keep all money in cash; robust optimal return 1.05, risk 0
  - $\mathcal{U}_{\text{ellipsoid-box}}$: robust optimal value is 1.12, risk 0.5%
  - $\mathcal{U}_{\text{budget}}$: robust optimal value is 1.10, risk 0.5%
- $\mathcal{U}_{\text{box}}$ can be quite conservative, a tiny bit of risk can go a long way...

Using Concentration Results to Model Uncertainty Sets

- Dimitris Bertsimas: let's use the **implications** of the **Central Limit Theorem**

  "All epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, non-random regularity" (Gnedenko and Kolmogorov, 1954)

Using Concentration Results to Model Uncertainty Sets

- Dimitris Bertsimas: let's use the **implications** of the **Central Limit Theorem**

  "All epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, non-random regularity" (Gnedenko and Kolmogorov, 1954)

- Suppose we have uncertainties $\{X_i\}_{i=1}^{n}$, each with mean μ, standard deviation σ

$$\mathcal{U}_{\mathrm{CLT}} \stackrel{\text{def}}{=} \left\{(x_1, \ldots, x_n) : \left|\sum_{i=1}^{n} x_i - n\mu\right| \leqslant \Gamma\sigma\sqrt{n}\right\}.$$

  - If n large and CLT premises hold, $\Gamma = 2$ (3) would give 95% (99%) coverage

Using Concentration Results to Model Uncertainty Sets

- Dimitris Bertsimas: let's use the **implications** of the **Central Limit Theorem**

  "All epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, non-random regularity" (Gnedenko and Kolmogorov, 1954)

- Suppose we have uncertainties $\{X_i\}_{i=1}^n$, each with mean μ, standard deviation σ

$$\mathcal{U}_{\text{CLT}} \stackrel{\text{def}}{=} \left\{ (x_1, \ldots, x_n) \,:\, \left| \sum_{i=1}^{n} x_i - n\mu \right| \leqslant \Gamma\sigma\sqrt{n} \right\}.$$

  - If n large and CLT premises hold, $\Gamma = 2$ (3) would give 95% (99%) coverage

- Many extensions possible
  - Modeling correlations through a factor model:
    $\mathcal{U}_{\text{corr}} \stackrel{\text{def}}{=} \left\{ \boldsymbol{x} \,:\, \boldsymbol{x} = P\boldsymbol{z} + \boldsymbol{\epsilon},\ \left|\sum_{i=1}^{m} z_i - m\mu_y\right| \leqslant \Gamma\sigma_z\sqrt{m},\ \left|\sum_{i=1}^{n} \epsilon_i\right| \leqslant \Gamma\sigma_\epsilon\sqrt{n}, \right\}.$
  - Using stable laws to model heavy-tailed cases where variance is undefined:
    $\mathcal{U}_{\text{HT}} \stackrel{\text{def}}{=} \left\{ (x_1, \ldots, x_n) \,:\, \left|\sum_{i=1}^{n} x_i - n\mu\right| \leqslant \Gamma n^{1/\alpha} \right\}.$
  - Constructing typical sets: if H_f is the (Shannon) entropy of f,
    (i) $\mathbb{P}[\tilde{z} \in \mathcal{U}_{\text{typical}}] \to 1$, (ii) $\left| \frac{1}{n} \log f\left(\tilde{z} | \tilde{z} \in \mathcal{U}_{\text{typical}}\right) + H_f \right| \leqslant \varepsilon_n$

Using Concentration Results to Model Uncertainty Sets

- Dimitris Bertsimas: let's use the **implications** of the **Central Limit Theorem**

  "All epistemological value of the theory of probability is based on this: that large-scale random phenomena in their collective action create strict, non-random regularity" (Gnedenko and Kolmogorov, 1954)

- Suppose we have uncertainties $\{X_i\}_{i=1}^n$, each with mean μ, standard deviation σ

$$\mathcal{U}_{\mathrm{CLT}} \stackrel{\text{def}}{=} \left\{(x_1, \ldots, x_n) \,:\, \left|\sum_{i=1}^{n} x_i - n\mu\right| \leqslant \Gamma\sigma\sqrt{n}\right\}.$$

  - If n large and CLT premises hold, $\Gamma = 2$ (3) would give 95% (99%) coverage

- Many extensions possible
  - Modeling correlations through a factor model:
    $\mathcal{U}_{\mathrm{corr}} \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \,:\, \boldsymbol{x} = P\boldsymbol{z} + \boldsymbol{\epsilon},\ \left|\sum_{i=1}^{m} z_i - m\mu_y\right| \leqslant \Gamma\sigma_z\sqrt{m},\ \left|\sum_{i=1}^{n} \epsilon_i\right| \leqslant \Gamma\sigma_\epsilon\sqrt{n},\right\}.$
  - Using stable laws to model heavy-tailed cases where variance is undefined:
    $\mathcal{U}_{\mathrm{HT}} \stackrel{\text{def}}{=} \left\{(x_1, \ldots, x_n) \,:\, \left|\sum_{i=1}^{n} x_i - n\mu\right| \leqslant \Gamma n^{1/\alpha}\right\}.$
  - Constructing typical sets: if H_f is the (Shannon) entropy of f,
    (i) $\mathbb{P}[\tilde{z} \in \mathcal{U}_{\text{typical}}] \to 1$, (ii) $\left|\frac{1}{n}\log f\left(\tilde{z} \middle| \tilde{z} \in \mathcal{U}_{\text{typical}}\right) + H_f\right| \leqslant \varepsilon_n$

- Bertsimas & Bandi used these to derive **robust equivalents** for several classical queueing theory and information theory results

Using Hypothesis Tests to Model Uncertainty Sets

Another powerful idea: derive **data-driven** uncertainty sets from **hypothesis tests**

From Bertsimas, Gupta, Kallus (2017):

Table 1 Summary of data-driven uncertainty sets proposed in this paper. SOC, EC and LMI denote second-order cone representable sets, exponential cone representable sets, and linear matrix inequalities, respectively

Assumptions on $\mathbb{P}^*$	Hypothesis test	Geometric description	Eqs.	Inner problem
Discrete support	χ^2-test	SOC	(13, 15)	
Discrete support	G-test	Polyhedral*	(13, 16)	
Independent marginals	KS Test	Polyhedral*	(21)	Line search
Independent marginals	K Test	Polyhedral*	(76)	Line search
Independent marginals	CvM Test	SOC*	(76, 69)	
Independent marginals	W Test	SOC*	(76, 70)	
Independent marginals	AD Test	EC	(76, 71)	
Independent marginals	Chen et al. [23]	SOC	(27)	Closed-form
None	Marginal Samples	Box	(31)	Closed-form
None	Linear Convex Ordering	Polyhedron	(34)	
None	Shawe-Taylor and Cristianini [46]	SOC	(39)	Closed-form
None	Delage and Ye [25]	LMI	(41)	

The additional "*" notation indicates a set of the above type with one additional, relative entropy constraint. KS, K, CvM, W, and AD denote the Kolmogorov–Smirnov, Kuiper, Cramer-von Mises, Watson and Anderson-Darling goodness of fit tests, respectively. In some cases, we can identify a worst-case realization of $\mathbf{u}$ in (1) for bi-affine f and a candidate $\mathbf{x}$ with a specialized algorithm. In these cases, the column "Inner Problem" roughly describes this algorithm

Distributionally Robust Modeling

- We saw that embedding some more probabilistic information can help

Distributionally Robust Modeling

- We saw that embedding some more probabilistic information can help
- Let's change the paradigm slightly:
  - we know the probability distribution for $\tilde{z}$ belongs to an *ambiguity set*: $\mathbb{P} \in \mathcal{P}$
  - we model $\mathcal{P}$, and are interested in robust expected constraint satisfaction:

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[f(x, \tilde{z})] \leqslant b$$

- Now, the adversary is choosing $\mathbb{P}$, instead of z
  - **Key advantage:** $\mathbb{E}_{\mathbb{P}}[f(x, \tilde{z})]$ as an expression of $\mathbb{P}$ is **always linear**, so much of our earlier machinery (e.g., convex duality) can be applied if the set $\mathcal{P}$ is "well-behaved"

Distributionally Robust Modeling

- We saw that embedding some more probabilistic information can help
- Let's change the paradigm slightly:
  - we know the probability distribution for $\tilde{z}$ belongs to an *ambiguity set*: $\mathbb{P} \in \mathcal{P}$
  - we model $\mathcal{P}$, and are interested in robust expected constraint satisfaction:

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[f(x, \tilde{z})] \leqslant b$$

- Now, the adversary is choosing $\mathbb{P}$, instead of z
  - **Key advantage:** $\mathbb{E}_{\mathbb{P}}[f(x, \tilde{z})]$ as an expression of $\mathbb{P}$ is **always linear**, so much of our earlier machinery (e.g., convex duality) can be applied if the set $\mathcal{P}$ is "well-behaved"
- Very old idea, dating to the 1950s
  - Scarf (1958) studied a Newsvendor model with mean and variance of demand known
  - Zackova (1966) studied stochastic LPs with knowledge of mean and support
- Recent tutorial by Kuhn, Shafiee, Wiesemann (2024) very nice summary of state-of-the-art; can model:
  - known (**bounds** on) moments, e.g., means, covariance matrix, higher order
  - information about various spread statistics, e.g., absolute mean spread ($E[X|X > \theta] - E[X|X < \theta]$), mean absolute deviation ($E[|X - m|]$), etc.
  - known (**bounds on**) quantiles, e.g., median
  - multiple confidence regions
  - distance from a nominal distribution (e.g., the empirical one)

Connections with Risk Measures

- A lot of effort in math finance since 2000 to characterize the "right" risk measures
- X is a random **loss**; a **risk measure** $\mu(X)$ should capture the "riskiness" in X
  - used to determine capital requirements for banks, so $\mu(X)$ should be "money"

Connections with Risk Measures

- A lot of effort in math finance since 2000 to characterize the "right" risk measures
- X is a random **loss**; a **risk measure** $\mu(X)$ should capture the "riskiness" in X
  - used to determine capital requirements for banks, so $\mu(X)$ should be "money"

Desirable Axioms (Artzner et al 1999)

- **[P1] Monotonicity**: If $X \leqslant Y$, then $\mu(X) \leqslant \mu(Y)$.
- **[P2] Influence of cash**: If $m \in \mathbb{R}$, then $\mu(X + m) = \mu(X) + m$.
- **[P3] Diversification**: $\mu\big(\lambda X + (1-\lambda)Y\big) \leqslant \lambda\, \mu(X) + (1-\lambda)\mu(Y)$, for $\lambda \in [0, 1]$.
- **[P4] Leverage**: If $\lambda \geqslant 0$, then $\mu\big(\lambda X\big) = \lambda\, \mu(X)$.

Connections with Risk Measures

- A lot of effort in math finance since 2000 to characterize the "right" risk measures
- X is a random **loss**; a **risk measure** $\mu(X)$ should capture the "riskiness" in X
  - used to determine capital requirements for banks, so $\mu(X)$ should be "money"

Desirable Axioms (Artzner et al 1999)

- **[P1] Monotonicity**: If $X \leqslant Y$, then $\mu(X) \leqslant \mu(Y)$.
- **[P2] Influence of cash**: If $m \in \mathbb{R}$, then $\mu(X+m) = \mu(X) + m$.
- **[P3] Diversification**: $\mu\big(\lambda X + (1-\lambda)Y\big) \leqslant \lambda\, \mu(X) + (1-\lambda)\mu(Y)$, for $\lambda \in [0, 1]$.
- **[P4] Leverage**: If $\lambda \geqslant 0$, then $\mu\big(\lambda X\big) = \lambda\, \mu(X)$.

  - E.g., variance does not satisfy **[P1]**, Value-at-Risk (VaR) does not satisfy **P3**
  - Conditional-Value-at-Risk (CVaR) satisfies [P1-P4]

Connections with Risk Measures

- A lot of effort in math finance since 2000 to characterize the "right" risk measures
- X is a random **loss**; a **risk measure** $\mu(X)$ should capture the "riskiness" in X
  - used to determine capital requirements for banks, so $\mu(X)$ should be "money"

Desirable Axioms (Artzner et al 1999)

- **[P1] Monotonicity**: If $X \leqslant Y$, then $\mu(X) \leqslant \mu(Y)$.
- **[P2] Influence of cash**: If $m \in \mathbb{R}$, then $\mu(X + m) = \mu(X) + m$.
- **[P3] Diversification**: $\mu\big(\lambda X + (1-\lambda)Y\big) \leqslant \lambda\, \mu(X) + (1-\lambda)\mu(Y)$, for $\lambda \in [0, 1]$.
- **[P4] Leverage**: If $\lambda \geqslant 0$, then $\mu\big(\lambda X\big) = \lambda\, \mu(X)$.

  - E.g., variance does not satisfy **[P1]**, Value-at-Risk (VaR) does not satisfy **P3**
  - Conditional-Value-at-Risk (CVaR) satisfies [P1-P4]

THM. μ satisfies **[P1-3]** iff there exists a convex function ϕ and

$$\mu(X) = \sup_{\mathbb{P}} \big[\, \mathbb{E}_{\mathbb{P}}[X] - \phi(\mathbb{P}) \,\big]$$

THM. μ satisfies **[P1-4]** iff there exists a convex set of measures $\mathcal{P}$ and

$$\mu(X) = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[X]$$

- Can construct uncertainty sets using risk measures (Bertsimas, Brown, Sim)
- Similar results also in decision theory under ambiguity (Gilboa, Schmeidler, etc.)

References I

E. Adida and G. Perakis. A robust optimization approach to dynamic pricing and inventory control with no backorders. *Mathematical Programming*, 107: 97–129, 2006.

O. Baron, J. Milner, and H. Naseraldin. Facility location: A robust optimization approach. *Production and Operations Management*, 20(5):772–785, 2011.

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2): 351–376, 2004.

A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial. Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management*, 7(3):248–271, 2005.

A. Ben-Tal, L. El-Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, 2009.

D. Bertsimas and D. Pachamanova. Robust multiperiod portfolio management in the presence of transaction costs. *Computers and Operations Research*, 35(1):3–17, 2008.

D. Bertsimas and A. Thiele. A robust optimization approach to inventory theory. *Operations Research*, 54(1):150–168, 2006.

D. Bienstock and N. Özbay. Computing robust basestock levels. *Discrete Optimization*, 5(2):389 – 414, 2008.

S. Ceria and R. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2): 109–127, July 2006.

F. de Ruiter, A. Ben-Tal, R. Brekelmans, and D. den Hertog. Adjustable robust optimization with decision rules based on inexact revealed data. Center discussion paper series no. 2014-003, CentER, 2014.

D. Goldfarb and G. Iyengar. Robust Portfolio Selection Problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

X. Lin, S. Janak, and C. Floudas. A new robust optimization approach for scheduling under uncertainty: I. bounded uncertainty. *Computers and Chemical Engineering*, 28:1069–1085, 2004.

A. Lorca and A. Sun. Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind. *IEEE Transactions on Power Systems*, 30(4):1702–1713, 2015.

S. Mittal, A. Schulz, and S. Stiller. Robust appointment scheduling. In *APPROX*, 2014.

G. Perakis and G. Roels. Robust controls for network revenue management. *Manufacturing & Service Operations Management*, 12(1):56–76, 2010.

M. Pinar and R. Tütüncü. Robust profit opportunities in risky financial portfolios. *Operations Research Letters*, 33(4):331 – 340, 2005.

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming*. MPS / SIAM Series on Optimization. SIAM, 2009.

R. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187, November 2004.

W. Wiesemann, D. Kuhn, and B. Rustem. Robust resource allocations in temporal networks. *Mathematical Programming*, 135:437–471, 2012. ISSN 0025-5610.

D. Yamashita, V. Armentano, and M. Laguna. Robust optimization models for project scheduling with resource availability cost. *Journal of Scheduling*, 10 (1):67–76, 2007.

C. Zhao, J.Wang, J. Watson, and Y.Guan. Multi-stage robust unit commitment considering wind and demand response uncertainties. *IEEE Transactions on Power Systems*, 28(3):2708–2717, 2013.