Comparison between Robust Optimization, Stochastic Programming and Distributionally Robust Optimization for Unit Commitment

Mathis Azéma (CERMICS)

Supervisors: Vincent Leclère, Wim Van Ackooij (EDF R&D)

PGMO Days

November 19th, 2024

Contents

Contents

Deterministic UC Problem

- $\mathcal{M}$: set of units
- T: time horizon
- $d \in \mathbb{R}^T$: demand vector
- x_i: commitment variables (binary).
- y_i: production variables (continuous).

$$
\begin{aligned}
\min_{x_i, y_i} \quad & \sum_{i \in \mathcal{M}} c_i^\top x_i + \sum_{i \in \mathcal{M}} b_i^\top y_i \\
s.t. \quad & F_i x_i \geq f_i && \forall i \in \mathcal{M} \\
& H_i y_i \geq h_i && \forall i \in \mathcal{M} \\
& A_i x_i + B_i y_i \geq g_i && \forall i \in \mathcal{M} \\
& \sum_{i \in \mathcal{M}} y_i = d \\
& x_i \in \{0,1\}^{m_i \times T}, \quad y_i \in \mathbb{R}_+^T
\end{aligned}
$$

Assumptions

- Uncertainty on the demand.
- 2-stage assumption:
  - 1^{st} stage: commitment variables (binary)
  - 2^{nd} stage: production variables (continuous)

$$\begin{aligned} \min_{x_i, y_i} \quad & c^\top x + b^\top y \\ s.t. \quad & x \in X \\ & y_i \in Y_i(x_i) \qquad \forall i \in \mathcal{M} \\ & \sum_{i \in \mathcal{M}} y_i = \boxed{d}\,? \end{aligned}$$

Assumptions

- Uncertainty on the demand.
- 2-stage assumption:
  - 1^{st} stage: commitment variables (binary)
  - 2^{nd} stage: production variables (continuous)

$$\begin{aligned} \min_{x_i, y_i} \quad & c^\top x + b^\top y \\ s.t. \quad & x \in X \\ & y_i \in Y_i(x_i) \qquad \forall i \in \mathcal{M} \\ & \sum_{i \in \mathcal{M}} y_i = d \end{aligned}$$

Assumptions

- Uncertainty on the demand.
- 2-stage assumption:
    - 1^{st} stage: commitment variables (binary)
    - 2^{nd} stage: production variables (continuous)

$$\begin{aligned} \min_{x_i} \quad & c^\top x + \min_{\substack{y_i : y_i \in Y_i(x_i) \\ \sum_{i \in \mathcal{M}} y_i = d}} b^\top y \\ s.t. \quad & x \in X \end{aligned}$$

Assumptions

- Uncertainty on the demand.
- 2-stage assumption:
  - 1^{st} stage: commitment variables (binary)
  - 2^{nd} stage: production variables (continuous)

$$\begin{aligned} \min_{x_i} \quad & c^\top x + Q(x,d) \\ s.t. \quad & x \in X \end{aligned}$$

$Q(x,d)$: recourse function representing the optimal cost of the second stage, considering which units are on or off (first stage) and the demand d

$$Q(x,d) = \min_{y} \quad b^\top y \qquad = \max_{\alpha,\beta,\gamma} \alpha^\top d + \beta^\top x + \gamma$$

$$y_i \in Y_i(x_i) \qquad (\alpha,\beta,\gamma) \in \Lambda$$

$$\sum_{i \in \mathcal{M}} y_i = d$$

Contents

Contents

Principle

Robust problem: minimize the worst-case

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{d \in \mathcal{D}} Q(x,d) & \\ s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\ & x_i \in \{0,1\}^{m_i \times T} & \end{aligned}$$

Budget Uncertainty Set:

$$\mathcal{D} = \left\{ d = (d_t)_{t \in [T]} \quad \middle| \quad d_t = \hat{d}_t + \gamma_t \Delta_t \quad \sum |\gamma_t| \leq \Gamma \quad \gamma_t \in [-1,1] \right\}$$

$\implies$ Objective: Find a tractable reformulation to solve the min-max-min problem

Principle

Robust problem: minimize the worst-case

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{d \in \mathcal{D}} Q(x,d) & \\ s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\ & x_i \in \{0,1\}^{m_i \times T} & \end{aligned}$$

Budget Uncertainty Set:

$$\mathcal{D} = \left\{ d = (d_t)_{t\in[T]} \quad \middle| \quad d_t = \hat{d}_t + \gamma_t \Delta_t \quad \sum |\gamma_t| \leq \Gamma \quad \gamma_t \in [-1,1] \right\}$$

$\Longrightarrow$ Objective: Find a tractable reformulation to solve the min-max-min problem

Principle

Robust problem: minimize the worst-case

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{d \in \mathcal{D}} Q(x,d) & \\ s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\ & x_i \in \{0,1\}^{m_i \times T} & \end{aligned}$$

Budget Uncertainty Set:

$$\mathcal{D} = \left\{ d = (d_t)_{t \in [T]} \quad \middle| \quad d_t = \hat{d}_t + \gamma_t \Delta_t \quad \sum |\gamma_t| \leq \Gamma \quad \gamma_t \in [-1,1] \right\}$$

$\implies$ Objective: Find a tractable reformulation to solve the min-max-min problem

Steps for reformulating $\max_{d\in\mathcal{D}} Q(x,d)$

1. If Γ is an integer, then:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D})} Q(x,d)$$

$$ext(\mathcal{D}) = \left\{ d = (d_t)_{t\in[T]} \,\middle|\, d_t = \bar{d}_t + \gamma_t \Delta_t \quad \sum |\gamma_t| \leq \Gamma \quad \gamma_t \in \{-1,0,1\} \right\}$$

2. Dualize the recourse function:

$$Q(x,d) = \max \left(\alpha^\top d + \beta^\top x + \gamma \quad s.t. \quad (\alpha,\beta,\gamma) \in \Lambda \right)$$

3. Transform the bilinear worst-case problem into a MILP:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D}),(\alpha,\beta,\gamma)\in\Lambda} \alpha^\top d + \beta^\top x + \gamma$$

Contribution: Adaptation of the method developed by Billionnet et al. [2016][1] which assumes $\alpha \geq 0$ (stemming from $\sum y_i \geq d$ instead of $\sum y_i = d$)

---

[1] Billionnet, Costa, and Poirion 2016.

Steps for reformulating $\max_{d\in\mathcal{D}} Q(x,d)$

1. If Γ is an integer, then:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D})} Q(x,d)$$

$$ext(\mathcal{D}) = \Big\{ d = (d_t)_{t\in[T]} \,|\, d_t = \bar{d}_t + \gamma_t\Delta_t \quad \sum|\gamma_t| \leq \Gamma \quad \gamma_t \in \{-1,0,1\} \Big\}$$

2. Dualize the recourse function:

$$Q(x,d) = \max\Big(\alpha^\top d + \beta^\top x + \gamma \quad s.t. \quad (\alpha,\beta,\gamma)\in\Lambda\Big)$$

3. Transform the bilinear worst-case problem into a MILP:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D}),(\alpha,\beta,\gamma)\in\Lambda} \alpha^\top d + \beta^\top x + \gamma$$

Contribution: Adaptation of the method developed by Billionnet et al. [2016][1] which assumes $\alpha \geq 0$ (stemming from $\sum y_i \geq d$ instead of $\sum y_i = d$)

---

[1] Billionnet, Costa, and Poirion 2016.

Steps for reformulating $\max\limits_{d\in\mathcal{D}} Q(x,d)$

1. If Γ is an integer, then:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D})} Q(x,d)$$

$$ext(\mathcal{D}) = \left\{ d = (d_t)_{t\in[T]} \,|\, d_t = \bar{d}_t + \gamma_t \Delta_t \quad \sum |\gamma_t| \leq \Gamma \quad \gamma_t \in \{-1, 0, 1\} \right\}$$

2. Dualize the recourse function:

$$Q(x,d) = \max \left(\alpha^\top d + \beta^\top x + \gamma \quad s.t. \quad (\alpha,\beta,\gamma) \in \Lambda \right)$$

3. Transform the bilinear worst-case problem into a MILP:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D}),(\alpha,\beta,\gamma)\in\Lambda} \alpha^\top d + \beta^\top x + \gamma$$

Contribution: Adaptation of the method developed by Billionnet et al. [2016][1] which assumes $\alpha \geq 0$ (stemming from $\sum y_i \geq d$ instead of $\sum y_i = d$)

---

[1] Billionnet, Costa, and Poirion 2016.

Benders Algorithm

Initial Robust Problem:

$$\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\
& \theta \geq \max_{d \in \mathcal{D}} Q(x, d) \\
& x_i \in \{0, 1\}^{m_i \times T}
\end{aligned}$$

Benders Algorithm

If Γ is an integer, then:

$$\max_{d\in\mathcal{D}} Q(x,d) = \max_{d\in ext(\mathcal{D})} Q(x,d)$$

$$\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i && \forall i \in \mathcal{M} \\
& \theta \geq Q(x,d) && \forall d \in ext(\mathcal{D}) \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}$$

Benders Algorithm

Dualize the recourse function:

$$Q(x,d) = \max \left(\alpha^\top d + \beta^\top x + \gamma \quad s.t. \quad (\alpha, \beta, \gamma) \in \Lambda \right)$$

$$\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\
& \theta \geq \max \left(\alpha^\top d + \beta^\top x + \gamma \quad s.t. \quad (\alpha, \beta, \gamma) \in \Lambda \right) & \forall d \in ext(\mathcal{D}) \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}$$

Benders Algorithm

Decomposition over the extreme points $(\alpha_k, \beta_k, \gamma_k)$ of the dual polytope Λ:

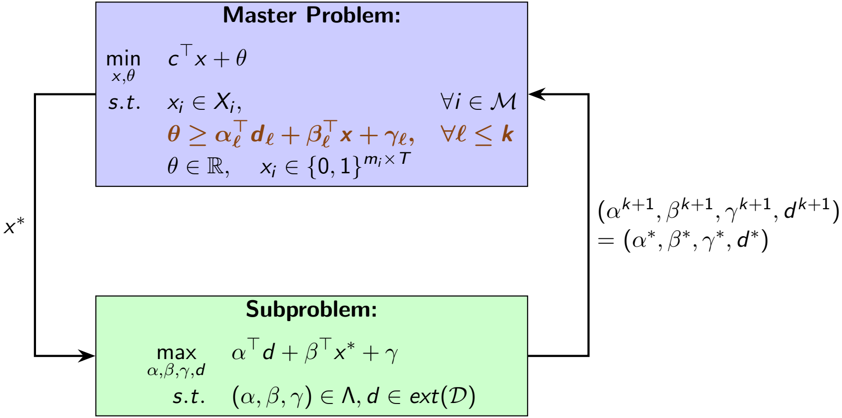
$$\begin{aligned}\min_{x,\theta} \quad & c^\top x + \theta \\ s.t. \quad & x_i \in X_i && \forall i \in \mathcal{M} \\ & \theta \geq \max_{k \in [K]} \left(\alpha_k^\top d + \beta_k^\top x + \gamma_k \right) && \forall d \in ext(\mathcal{D}) \\ & x_i \in \{0,1\}^{m_i \times T}\end{aligned}$$

Benders Algorithm

Linearization:

$$\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i && \forall i \in \mathcal{M} \\
& \theta \geq \left(\alpha_k^\top d + \beta_k^\top x + \gamma_k \right) && \forall d \in ext(\mathcal{D}) \forall k \in [K] \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}$$

Benders Algorithm



Master Problem:

$$\begin{aligned} \min_{x,\theta} \quad & c^\top x + \theta \\ s.t. \quad & x_i \in X_i, && \forall i \in \mathcal{M} \\ & \boldsymbol{\theta \geq \alpha_\ell^\top d_\ell + \beta_\ell^\top x + \gamma_\ell}, && \boldsymbol{\forall \ell \leq k} \\ & \theta \in \mathbb{R}, \quad x_i \in \{0,1\}^{m_i \times T} \end{aligned}$$

x^*

$(\alpha^{k+1}, \beta^{k+1}, \gamma^{k+1}, d^{k+1})$
$= (\alpha^*, \beta^*, \gamma^*, d^*)$

Subproblem:

$$\begin{aligned} \max_{\alpha,\beta,\gamma,d} \quad & \alpha^\top d + \beta^\top x^* + \gamma \\ s.t. \quad & (\alpha, \beta, \gamma) \in \Lambda, d \in ext(\mathcal{D}) \end{aligned}$$

CCG Algorithm

Initial Robust Problem:

$$
\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i \qquad\qquad \forall i \in \mathcal{M} \\
& \theta \geq \max_{d \in \mathcal{D}} Q(x, d) \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}
$$

CCG Algorithm

If Γ is an integer, then:

$$\max_{d \in \mathcal{D}} Q(x,d) = \max_{d \in ext(\mathcal{D})} Q(x,d)$$

$$\begin{aligned} \min_{x,\theta} \quad & c^\top x + \theta \\ s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\ & \theta \geq Q(x,d) & \forall d \in ext(\mathcal{D}) \\ & x_i \in \{0,1\}^{m_i \times T} \end{aligned}$$

CCG Algorithm

Primal form of the recourse function:

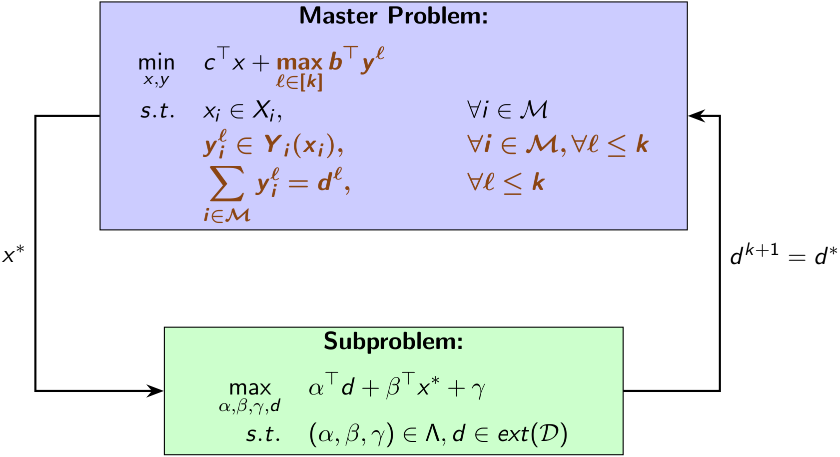
$$
\begin{aligned}
Q(x,d) = \min_{y} \quad & b^\top y \\
& y_i \in Y_i(x_i) \qquad \forall i \in \mathcal{M} \\
& \sum_{i \in \mathcal{M}} y_i = d
\end{aligned}
$$

$$
\begin{aligned}
\min_{x,\theta} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i & \forall i \in \mathcal{M} \\
& \theta \geq \min_{\substack{y_i^d \in Y_i(x_i) \\ \sum_{i\in\mathcal{M}} y_i^d = d}} \quad b^\top y^d & \forall d \in ext(\mathcal{D}) \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}
$$

CCG Algorithm

$$
\begin{aligned}
\min_{x,y} \quad & c^\top x + \theta \\
s.t. \quad & x_i \in X_i && \forall i \in \mathcal{M} \\
& \theta \geq b^\top y^d && \forall d \in ext(\mathcal{D}) \\
& y_i^d \in Y_i(x_i) && \forall d \in ext(\mathcal{D}), \forall i \in \mathcal{M} \\
& \sum_{i \in \mathcal{M}} y_i^d = d && \forall d \in ext(\mathcal{D}) \\
& x_i \in \{0,1\}^{m_i \times T}
\end{aligned}
$$

CCG Algorithm



Master Problem:

$$
\begin{aligned}
\min_{x,y} \quad & c^\top x + \max_{\ell \in [k]} \boldsymbol{b}^\top \boldsymbol{y}^\ell & \\
s.t. \quad & x_i \in X_i, & \forall i \in \mathcal{M} \\
& \boldsymbol{y_i^\ell \in Y_i(x_i)}, & \boldsymbol{\forall i \in \mathcal{M}, \forall \ell \leq k} \\
& \boldsymbol{\sum_{i \in \mathcal{M}} y_i^\ell = d^\ell}, & \boldsymbol{\forall \ell \leq k}
\end{aligned}
$$

x^*

$d^{k+1} = d^*$

Subproblem:

$$
\begin{aligned}
\max_{\alpha,\beta,\gamma,d} \quad & \alpha^\top d + \beta^\top x^* + \gamma \\
s.t. \quad & (\alpha, \beta, \gamma) \in \Lambda, d \in ext(\mathcal{D})
\end{aligned}
$$

Contents

Different Risk Measures

Risk neutral: $\mathbb{E}[X]$

$$\min\ c^\top x + \mathbb{E}[Q(x,d)]$$
$$s.t.\quad x_i \in X_i, \quad \forall i \in \mathcal{M}$$

Risk Averse (TVAR): $\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\min\ c^\top x + z + \frac{1}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)]$$
$$s.t.\quad x_i \in X_i, \quad \forall i \in \mathcal{M}$$

MTVaR: $\beta\mathbb{E}[X] + (1-\beta)\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\min\ c^\top x + \beta\mathbb{E}[Q(x,d)] + (1-\beta)z + \frac{1-\beta}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)]$$
$$s.t.\quad x_i \in X_i, \quad \forall i \in \mathcal{M}$$

Different Risk Measures

Risk neutral: $\mathbb{E}[X]$

$$\begin{aligned} &\min\ c^\top x + \mathbb{E}[Q(x,d)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Risk Averse (TVAR): $\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\begin{aligned} &\min\ c^\top x + z + \frac{1}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

MTVaR: $\beta\mathbb{E}[X] + (1-\beta)\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\begin{aligned} &\min\ c^\top x + \beta\mathbb{E}[Q(x,d)] + (1-\beta)z + \frac{1-\beta}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Different Risk Measures

Risk neutral: $\mathbb{E}[X]$

$$\begin{aligned} &\min\ c^\top x + \mathbb{E}[Q(x,d)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Risk Averse (TVAR): $\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\begin{aligned} &\min\ c^\top x + z + \frac{1}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

MTVaR: $\beta\mathbb{E}[X] + (1-\beta)\mathbb{E}[X|X \geq VaR_\alpha(X)]$

$$\begin{aligned} &\min\ c^\top x + \beta\mathbb{E}[Q(x,d)] + (1-\beta)z + \frac{1-\beta}{1-\alpha}\mathbb{E}[max(Q(x,d) - z, 0)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Algorithms

Extensive formulation:

$$\begin{aligned} &\min\ c^\top x + \mathbb{E}[Q(x,d)] \\ &s.t. \quad x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Algorithms

Extensive formulation:

$$\begin{aligned} \min \; & c^\top x + \sum_{s\in\mathcal{S}} \pi_s Q(x, d^s) \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Algorithms

Extensive formulation:

$$\begin{aligned} \min \; & c^\top x + \sum_{s \in \mathcal{S}} \pi_s \min_{\substack{y_i^s \in Y_i(x_i) \\ \sum_{i \in \mathcal{M}} y_i^s = d^s}} b^\top y^s \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

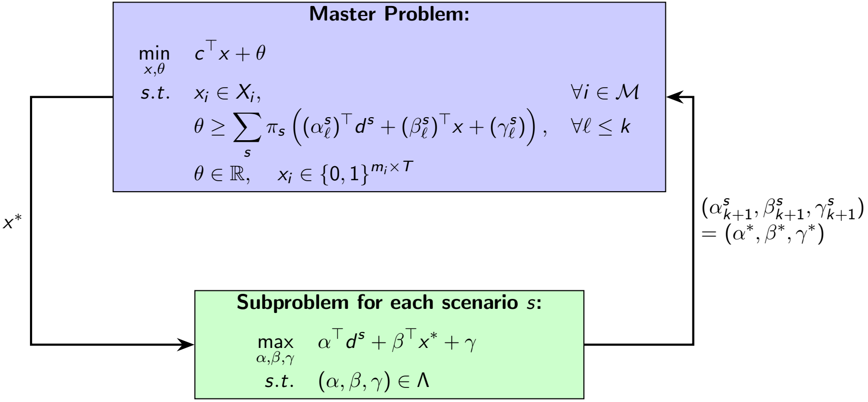
Algorithms

Extensive formulation:

$$\begin{aligned} \min \quad & c^\top x + \sum_{s\in\mathcal{S}} \pi_s b^\top y^s \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \\ & y_i^s \in Y_i(x_i), \quad \forall i \in \mathcal{M}, \forall s \in \mathcal{S} \\ & \sum_{i\in\mathcal{M}} y_i^s = d^s, \quad \forall s \in \mathcal{S} \end{aligned}$$

Algorithms

L-shaped (Benders):



Master Problem:

$$\begin{aligned} \min_{x,\theta} \quad & c^\top x + \theta \\ s.t. \quad & x_i \in X_i, && \forall i \in \mathcal{M} \\ & \theta \geq \sum_s \pi_s \left((\alpha_\ell^s)^\top d^s + (\beta_\ell^s)^\top x + (\gamma_\ell^s) \right), && \forall \ell \leq k \\ & \theta \in \mathbb{R}, \quad x_i \in \{0,1\}^{m_i \times T} \end{aligned}$$

x^*

$(\alpha_{k+1}^s, \beta_{k+1}^s, \gamma_{k+1}^s)$
$= (\alpha^*, \beta^*, \gamma^*)$

Subproblem for each scenario s:

$$\begin{aligned} \max_{\alpha,\beta,\gamma} \quad & \alpha^\top d^s + \beta^\top x^* + \gamma \\ s.t. \quad & (\alpha, \beta, \gamma) \in \Lambda \end{aligned}$$

Contents

Contents

Introduction

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\xi\sim\mathbb{Q}}[Q(x,\xi)] \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Link between RO and SP:

- If $\mathcal{P} = \{\mathbb{P}_0\}$, then the DRO problem is equal to the risk-neutral one.
- If $\mathcal{P}$ includes all distributions supported on $\mathcal{U}$, then the DRO problem is equivalent to the robust optimization (RO) problem with $\mathcal{U}$ as the uncertainty set.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to benefit from the advantages of both RO and SP?**

Introduction

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\xi\sim\mathbb{Q}}[Q(x,\xi)] \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Link between RO and SP:

- If $\mathcal{P} = \{\mathbb{P}_0\}$, then the DRO problem is equal to the risk-neutral one.
- If $\mathcal{P}$ includes all distributions supported on $\mathcal{U}$, then the DRO problem is equivalent to the robust optimization (RO) problem with $\mathcal{U}$ as the uncertainty set.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to benefit from the advantages of both RO and SP?**

Introduction

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\xi\sim\mathbb{Q}}[Q(x,\xi)] \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Link between RO and SP:

- If $\mathcal{P} = \{\mathbb{P}_0\}$, then the DRO problem is equal to the risk-neutral one.
- If $\mathcal{P}$ includes all distributions supported on $\mathcal{U}$, then the DRO problem is equivalent to the robust optimization (RO) problem with $\mathcal{U}$ as the uncertainty set.

$\Longrightarrow$ How can we select $\mathcal{P}$ to benefit from the advantages of both RO and SP?

Introduction

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\xi\sim\mathbb{Q}}[Q(x,\xi)] \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Link between RO and SP:

- If $\mathcal{P} = \{\mathbb{P}_0\}$, then the DRO problem is equal to the risk-neutral one.
- If $\mathcal{P}$ includes all distributions supported on $\mathcal{U}$, then the DRO problem is equivalent to the robust optimization (RO) problem with $\mathcal{U}$ as the uncertainty set.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to benefit from the advantages of both RO and SP?**

Choice of $\mathcal{P}$

Idea: $\mathcal{P}$ has to include distributions "close" to the empirical one.

- **Moments-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \,|\, |\mathbb{E}[\mathbb{Q}] - \mathbb{E}[\mathbb{P}_0]| \leq \theta\}$$

- **ϕ-divergence-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \,|\, D_\phi(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

Problem: $D_\phi(\mathbb{Q}, \mathbb{P}_0) = \mathbb{E}_\mathbb{P}\left[\phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\right)\right]$ is not a metric and only considers distributions with the same support as $\mathbb{P}_0$.

- **Wasserstein distance-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \in \mathcal{P}(\mathcal{U}) \,|\, W_p(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

Choice of $\mathcal{P}$

Idea: $\mathcal{P}$ has to include distributions "close" to the empirical one.

- **Moments-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \mid |\mathbb{E}[\mathbb{Q}] - \mathbb{E}[\mathbb{P}_0]| \leq \theta\}$$

- **ϕ-divergence-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \mid D_\phi(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

  Problem: $D_\phi(\mathbb{Q}, \mathbb{P}_0) = \mathbb{E}_\mathbb{P}\left[\phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\right)\right]$ is not a metric and only considers distributions with the same support as $\mathbb{P}_0$.

- **Wasserstein distance-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \in \mathcal{P}(\mathcal{U}) \mid W_p(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

Choice of $\mathcal{P}$

Idea: $\mathcal{P}$ has to include distributions “close” to the empirical one.

- **Moments-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \mid |\mathbb{E}[\mathbb{Q}] - \mathbb{E}[\mathbb{P}_0]| \leq \theta\}$$

- ϕ**-divergence-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \mid D_\phi(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

  Problem: $D_\phi(\mathbb{Q}, \mathbb{P}_0) = \mathbb{E}_\mathbb{P}\left[\phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\right)\right]$ is not a metric and only considers distributions with the same support as $\mathbb{P}_0$.

- **Wasserstein distance-based ambiguity sets**

$$\mathcal{P} = \{\mathbb{Q} \in \mathcal{P}(\mathcal{U}) \mid W_p(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}$$

Contents

ϕ-divergence

DRO problem can be reformulated as:

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\mathbb{Q}: D_\phi(\mathbb{Q}, \mathbb{P}_0) \leq \theta} \mathbb{E}_{\mathbb{Q}}[Q(x, \xi)] \\ s.t. \quad & x \in X \end{aligned}$$

Kullback-Leibler divergence:

$$\phi_{KL}(x) = x \log(x) - x + 1, \qquad \phi^*_{KL}(y) = e^y - 1$$

$$D_{\phi_{KL}}(\mathbb{Q}, \mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]$$

ϕ-divergence

DRO problem can be reformulated as:

$$\begin{aligned} \min_{x,\beta,\alpha\geq 0} \quad & c^\top x + \beta + \theta \cdot \alpha + \alpha \frac{1}{N} \sum_{i=1}^{N} \phi^* \left(\frac{Q(x,\zeta_i) - \beta}{\alpha} \right) \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Kullback-Leibler divergence:

$$\phi_{KL}(x) = x \log(x) - x + 1, \qquad \phi_{KL}^*(y) = e^y - 1$$

$$D_{\phi_{KL}}(\mathbb{Q},\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]$$

ϕ-divergence

DRO problem can be reformulated as:

$$\begin{aligned} \min_{x,\beta,\alpha\geq 0} \quad & c^\top x + \beta + \theta \cdot \alpha + \alpha \frac{1}{N} \sum_{i=1}^{N} \phi^* \left(\frac{Q(x,\zeta_i) - \beta}{\alpha} \right) \\ s.t. \quad & x_i \in X_i, \quad \forall i \in \mathcal{M} \end{aligned}$$

Kullback-Leibler divergence:

$$\phi_{KL}(x) = x \log(x) - x + 1, \qquad \phi^*_{KL}(y) = e^y - 1$$

$$D_{\phi_{KL}}(\mathbb{Q}, \mathbb{P}) = \mathbb{E}_{\mathbb{Q}} \left[\log \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right]$$

KL-divergence

$$\begin{aligned}\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x+\beta+(\theta-1)\cdot\alpha+\frac{\alpha}{N}\sum_{i=1}^{N} z_i \\ s.t. \quad & z_i \geq e^{\frac{Q(x,\zeta_i)-\beta}{\alpha}} && \forall i\in[N] \\ & x\in X\end{aligned}$$

Tangents approximation:

$$\begin{aligned}\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x+\beta+(\theta-1)\cdot\alpha+\frac{1}{N}\sum_{i=1}^{N} z_i \\ s.t. \quad & z_i \geq \alpha e^{z_r}+e^{z_r}\left(Q(x,\zeta_i)-\beta-\alpha z_r\right) && \forall i\in[N], \quad \forall r\in\mathcal{R} \\ & x\in X\end{aligned}$$

$\Longrightarrow$ Solve with an extensive or L-shaped formulation.

KL-divergence

$$
\begin{aligned}
\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x + \beta + (\theta - 1)\cdot\alpha + \frac{\alpha}{N}\sum_{i=1}^{N} z_i \\
s.t. \quad & z_i \geq e^{\frac{Q(x,\zeta_i)-\beta}{\alpha}} && \forall i \in [N] \\
& x \in X
\end{aligned}
$$

Tangents approximation:

$$
\begin{aligned}
\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x + \beta + (\theta - 1)\cdot\alpha + \frac{1}{N}\sum_{i=1}^{N} z_i \\
s.t. \quad & z_i \geq \alpha e^{z_r} + e^{z_r}\left(Q(x,\zeta_i) - \beta - \alpha z_r\right) && \forall i \in [N], \quad \forall r \in \mathcal{R} \\
& x \in X
\end{aligned}
$$

$\Longrightarrow$ Solve with an extensive or L-shaped formulation.

KL-divergence

$$
\begin{aligned}
\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x + \beta + (\theta-1)\cdot\alpha + \frac{\alpha}{N}\sum_{i=1}^{N} z_i \\
s.t. \quad & z_i \geq e^{\frac{Q(x,\zeta_i)-\beta}{\alpha}} & \forall i \in [N] \\
& x \in X
\end{aligned}
$$

Tangents approximation:

$$
\begin{aligned}
\min_{x,\beta,\alpha\geq 0,z\in\mathbb{R}^N_+} \quad & c^\top x + \beta + (\theta-1)\cdot\alpha + \frac{1}{N}\sum_{i=1}^{N} z_i \\
s.t. \quad & z_i \geq \alpha e^{z_r} + e^{z_r}\left(Q(x,\zeta_i) - \beta - \alpha z_r\right) & \forall i \in [N], \quad \forall r \in \mathcal{R} \\
& x \in X
\end{aligned}
$$

$\implies$ Solve with an extensive or L-shaped formulation.

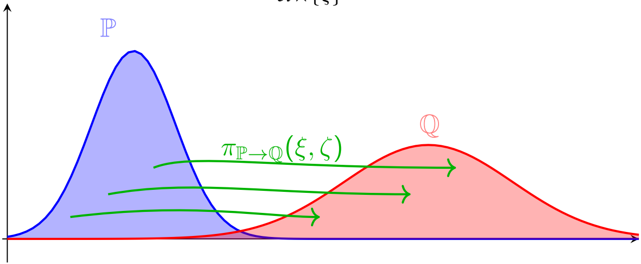
Contents

Wassertein-distance definition

$$(W_p(\mathbb{Q},\mathbb{P}))^p = \inf_{\pi \in \mathcal{P}(\mathcal{U},\mathcal{U})} \int_{\mathcal{U}\times\mathcal{U}} \|\xi-\zeta\|^p d\pi(\xi,\zeta)$$

$$s.t. \quad \int_{\{\xi\}\times\mathcal{U}} d\pi(\xi,\zeta) = \mathbb{Q}(\xi) \qquad \forall \xi \in \mathcal{U}$$

$$\int_{\mathcal{U}\times\{\zeta\}} d\pi(\xi,\zeta) = \mathbb{P}(\zeta) \qquad \forall \zeta \in \mathcal{U}$$



$$\mathcal{P} = \{\mathbb{Q} \,|\, W_p(\mathbb{Q},\mathbb{P}) \le \theta, \quad supp(\mathbb{Q}) \subset \mathcal{U}\}$$

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x} \quad c^\top x + \max_{\substack{\mathbb{Q}:\, W_p(\mathbb{Q},\mathbb{P}_0)\leq\theta \quad [\lambda] \\ supp(\mathbb{Q})\subset\mathcal{U}}} \mathbb{E}_{\mathbb{Q}}[Q(x,\xi)]$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1-**norm**: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top\zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\Longrightarrow$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \sup_{\xi\in\mathcal{U}} \left(Q(x,\xi) - \lambda\|\xi - \zeta_i\|^p\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1**-norm**: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty \theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\implies$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \max_{\xi\in\mathcal{U},k\in[K]} \left(\alpha_k^\top \xi + \beta_k^\top x + \gamma_k - \lambda\|\xi-\zeta_i\|^p\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1-norm: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\Longrightarrow$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \max_{\xi\in\mathcal{U}}(\alpha_k^\top \xi - \lambda\|\xi - \zeta_i\|^p)\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1-norm: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\Longrightarrow$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \max_{\xi\in\mathbb{R}^T}(\alpha_k^\top \xi - \lambda\|\xi - \zeta_i\|^p)\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1-**norm**: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\implies$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \max_{\xi\in\mathbb{R}^T}(\alpha_k^\top \xi - \lambda\|\xi - \zeta_i\|^p)\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1**-norm**: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\implies$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

Choice of L^2 Wassertein distance: Problem with L^1 norm

General reformulation[2]:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^p + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \max_{\xi\in\mathbb{R}^T}(\alpha_k^\top \xi - \lambda\|\xi - \zeta_i\|^p)\right)$$

$$s.t. \quad x \in X$$

In our study, no information is given in the support of the distributions: $\mathcal{U} = \mathbb{R}^T$

L^1**-norm**: with $\lambda_\infty := \max_{k\in[K]} \|\alpha_k\|_\infty$

$$\min_{x} \quad c^\top x + \overset{\text{Cste}}{\boxed{\lambda_\infty\theta}} + \frac{1}{N}\sum_{i=1}^{N} \underbrace{\max_{k\in[K]} \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k\right)}_{Q(x,\zeta_i)}$$

$$s.t. \quad x \in X$$

$\Longrightarrow$ No link between x and θ !

---

[2] Gao and Kleywegt 2016.

L^2-norm reformulation

DRO problem with norm L^2 is equivalent to:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^2 + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \max_{\xi\in\mathbb{R}^T}(\alpha_k^\top \xi - \lambda\|\xi - \zeta_i\|_2^2)\right)$$

$$s.t. \quad x \in X$$

Note: We remove the ξ variables without losing the link between x and θ.

$$\begin{aligned} \min_{x,\lambda\geq 0,z\geq 0,w\geq 0} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i \\ s.t. \quad & x \in X \\ & \|(2, w-\lambda)\|_2 \leq w + \lambda \\ & z_i \geq \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k + \frac{\lambda\|\alpha_k\|_2^2}{4}\right) \quad \forall k \in [K] \end{aligned}$$

$\Longrightarrow$ Benders' algorithm

L^2-norm reformulation

DRO problem with norm L^2 is equivalent to:

$$\begin{aligned} \min_{x,\lambda\geq 0} \quad & c^\top x + \lambda\theta^2 + \frac{1}{N}\sum_{i=1}^{N}\max_{k\in[K]}\left(\beta_k^\top x + \gamma_k + \alpha_k^\top \zeta_i + \frac{\|\alpha_k\|_2^2}{4\lambda}\right) \\ s.t. \quad & x \in X \end{aligned}$$

Note: We remove the ξ variables without losing the link between x and θ.

$$\begin{aligned} \min_{x,\lambda\geq 0,z\geq 0,w\geq 0} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i \\ s.t. \quad & x \in X \\ & \|(2, w-\lambda)\|_2 \leq w + \lambda \\ & z_i \geq \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k + \frac{\lambda\|\alpha_k\|_2^2}{4}\right) \quad \forall k \in [K] \end{aligned}$$

$\Longrightarrow$ Benders' algorithm

L^2-norm reformulation

DRO problem with norm L^2 is equivalent to:

$$\begin{aligned}\min_{x,\lambda\geq 0} \quad & c^\top x + \lambda\theta^2 + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \alpha_k^\top \zeta_i + \frac{\|\alpha_k\|_2^2}{4\lambda}\right)\\ s.t. \quad & x \in X\end{aligned}$$

Note: We remove the ξ variables without losing the link between x and θ.

$$\begin{aligned}\min_{x,\lambda\geq 0,z\geq 0,w\geq 0} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i\\ s.t. \quad & x \in X\\ & \|(2, w-\lambda)\|_2 \leq w + \lambda\\ & z_i \geq \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k + \frac{\lambda\|\alpha_k\|_2^2}{4}\right) \quad \forall k \in [K]\end{aligned}$$

$\Longrightarrow$ Benders' algorithm

L^2-norm reformulation

DRO problem with norm L^2 is equivalent to:

$$\begin{aligned}\min_{x,\lambda\geq 0} \quad & c^\top x + \lambda\theta^2 + \frac{1}{N}\sum_{i=1}^{N}\max_{k\in[K]}\left(\beta_k^\top x + \gamma_k + \alpha_k^\top \zeta_i + \frac{\|\alpha_k\|_2^2}{4\lambda}\right)\\ s.t. \quad & x \in X\end{aligned}$$

Note: We remove the ξ variables without losing the link between x and θ.

$$\begin{aligned}\min_{x,\lambda\geq 0,z\geq 0,w\geq 0} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i\\ s.t. \quad & x \in X\\ & \|(2, w-\lambda)\|_2 \leq w + \lambda\\ & z_i \geq \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k + \frac{\lambda\|\alpha_k\|_2^2}{4}\right) \quad \forall k \in [K]\end{aligned}$$

$\Longrightarrow$ Benders' algorithm

L^2-norm reformulation

DRO problem with norm L^2 is equivalent to:

$$\min_{x,\lambda\geq 0} \quad c^\top x + \lambda\theta^2 + \frac{1}{N}\sum_{i=1}^{N} \max_{k\in[K]} \left(\beta_k^\top x + \gamma_k + \alpha_k^\top \zeta_i + \frac{\|\alpha_k\|_2^2}{4\lambda}\right)$$

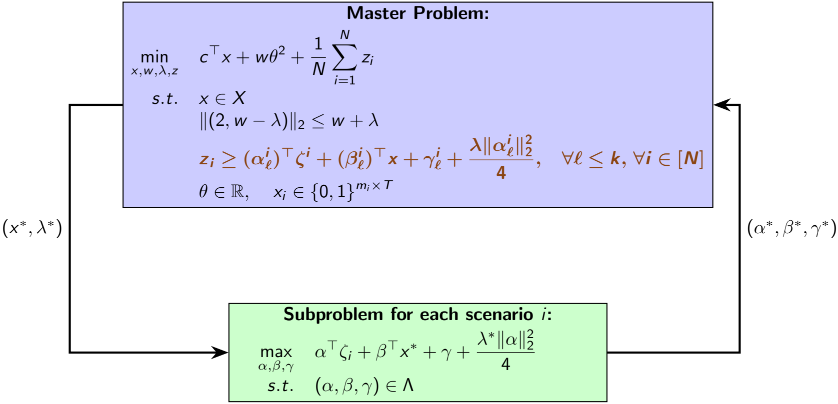
$$s.t. \quad x \in X$$

Note: We remove the ξ variables without losing the link between x and θ.

$$\begin{aligned} \min_{x,\lambda\geq 0, z\geq 0, w\geq 0} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i \\ s.t. \quad & x \in X \\ & \|(2, w-\lambda)\|_2 \leq w + \lambda \\ & z_i \geq \left(\alpha_k^\top \zeta_i + \beta_k^\top x + \gamma_k + \frac{\lambda\|\alpha_k\|_2^2}{4}\right) \quad \forall k \in [K] \end{aligned}$$

$\implies$ Benders' algorithm

Benders algorithm DRO



Master Problem:

$$
\begin{aligned}
\min_{x,w,\lambda,z} \quad & c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i \\
s.t. \quad & x \in X \\
& \|(2, w-\lambda)\|_2 \leq w + \lambda \\
& \boldsymbol{z_i \geq (\alpha_\ell^i)^\top \zeta^i + (\beta_\ell^i)^\top x + \gamma_\ell^i + \frac{\lambda \|\alpha_\ell^i\|_2^2}{4}}, \quad \forall \ell \leq \boldsymbol{k}, \forall \boldsymbol{i} \in [\boldsymbol{N}] \\
& \theta \in \mathbb{R}, \quad x_i \in \{0,1\}^{m_i \times T}
\end{aligned}
$$

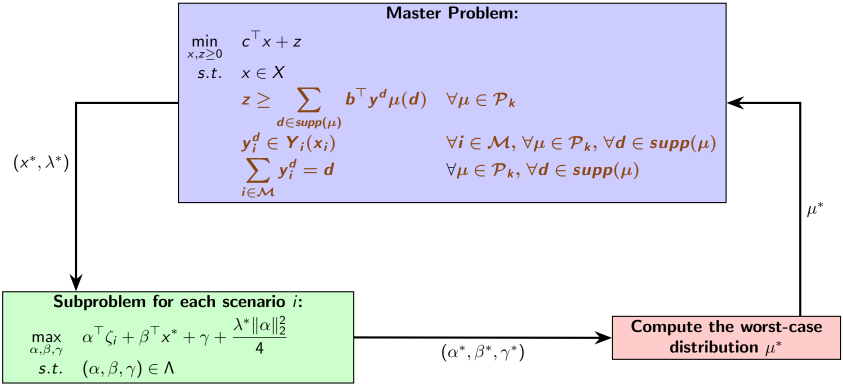
(x^*, λ^*)

$(\alpha^*, \beta^*, \gamma^*)$

Subproblem for each scenario i**:**

$$
\begin{aligned}
\max_{\alpha,\beta,\gamma} \quad & \alpha^\top \zeta_i + \beta^\top x^* + \gamma + \frac{\lambda^* \|\alpha\|_2^2}{4} \\
s.t. \quad & (\alpha,\beta,\gamma) \in \Lambda
\end{aligned}
$$

Another algorithm using worstcase distributions (CCG)



Master Problem:

$$\begin{aligned}
\min_{x,z\geq 0} \quad & c^\top x + z \\
s.t. \quad & x \in X \\
& \boldsymbol{z \geq \sum_{d\in supp(\mu)} b^\top y^d \mu(d)} && \boldsymbol{\forall \mu \in \mathcal{P}_k} \\
& \boldsymbol{y_i^d \in Y_i(x_i)} && \boldsymbol{\forall i \in \mathcal{M}, \forall \mu \in \mathcal{P}_k, \forall d \in supp(\mu)} \\
& \boldsymbol{\sum_{i\in\mathcal{M}} y_i^d = d} && \boldsymbol{\forall \mu \in \mathcal{P}_k, \forall d \in supp(\mu)}
\end{aligned}$$

Subproblem for each scenario *i*:

$$\begin{aligned}
\max_{\alpha,\beta,\gamma} \quad & \alpha^\top \zeta_i + \beta^\top x^* + \gamma + \frac{\lambda^* \|\alpha\|_2^2}{4} \\
s.t. \quad & (\alpha, \beta, \gamma) \in \Lambda
\end{aligned}$$

Compute the worst-case distribution μ^*

Contents

Presentation

- Consideration of instances with thermal units only (source: SMS++/ EDF).
- Analysis centered on small-scale instances (10 units) to ensure convergence across all approaches.
- Source scenarios: `https://data.open-power-system-data.org/time_series/`
- Out-of-sample tests on 300 scenarios

Out-Of-Sample Costs

Method	**10/1/700**	**10/2/700**	**10/3/700**	**10/4/700**	**10/5/700**
`TVaR`	0.04%	0.34%	**0.00%**	0.09%	0.13%
`MTVaR`	0.04%	0.34%	**0.00%**	0.09%	0.62%
`KL`	0.04%	0.78%	0.08%	0.01%	0.55%
`DW`	0.04%	**0.00%**	**0.00%**	0.01%	**0.00%**
`RO`	**0.00%**	**0.00%**	**0.00%**	**0.00%**	0.04%
`N:50`	1.24%	1.02%	1.01%	0.72%	1.48%
`Det.`	1.75%	0.51%	0.21%	0.95%	2.50%
Min cost	**1880117**	**1315961**	**1809417**	**1819979**	**1888259**

Table: Gap between average cost out-of-sample and best cost found among all the approaches

$\Longrightarrow$ `RO` and `DW` are the best but `RO` is much easier to solve.

Costs distributions

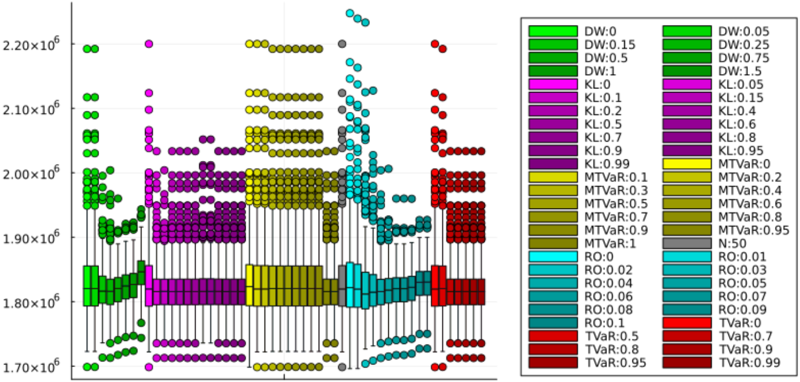
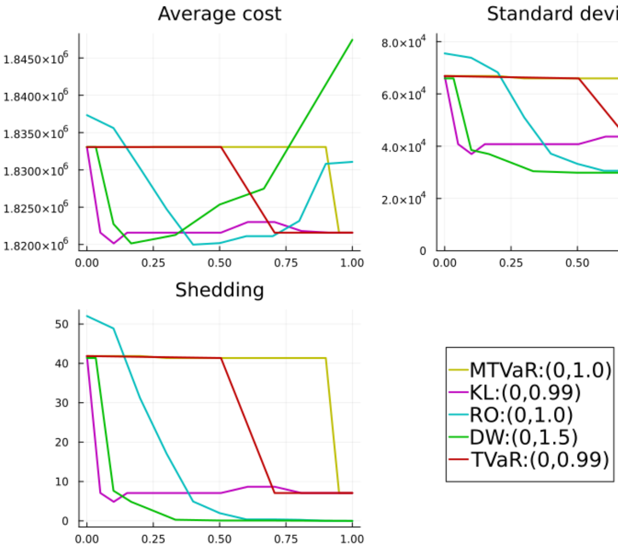


Figure: Boxplots of the out-of-sample costs of the instance 10/4/700

Parameters Sensitivity



Contents

Conclusion

Contributions:

- Analysis of traditional and novel methods for UC under uncertainty.
- Adaptation of a RO method.
- In-depth exploration of DRO using ϕ-divergence and Wasserstein distance.
- Development of specific algorithms to the L^2 Wasserstein distance.
- Numerical performance evaluation.

Future Work:

- Investigate the concave oracle problem for the L^2 Wasserstein distance to design specific algorithms (e.g., KKT-based reformulations).
- Integrate regularization techniques in Benders' algorithms.
- Study the case $\mathcal{U}$ is bounded and includes specific information.
- Comparison with Chance-Constrained models.
- Computational experiments to other instances and scenarios.

Conclusion

Contributions:

- Analysis of traditional and novel methods for UC under uncertainty.
- Adaptation of a RO method.
- In-depth exploration of DRO using ϕ-divergence and Wasserstein distance.
- Development of specific algorithms to the L^2 Wasserstein distance.
- Numerical performance evaluation.

Future Work:

- Investigate the concave oracle problem for the L^2 Wasserstein distance to design specific algorithms (e.g., KKT-based reformulations).
- Integrate regularization techniques in Benders' algorithms.
- Study the case $\mathcal{U}$ is bounded and includes specific information.
- Comparison with Chance-Constrained models.
- Computational experiments to other instances and scenarios.