Stochastic Programming and Distributionally Robust Optimization for Unit Commitment

Mathis Azéma (CERMICS)

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

CERMICS Young Researchers' Days

2-4 April 2025

Contents

Fundamental results for LPs

Primal:

$$\min_{x \geq 0} \quad c^\top x \qquad = \qquad Ax \geq b$$

Dual formulation:

$$\max_{y \geq 0} \quad b^\top y \qquad A^\top y \leq c$$

$$\max_{k \in [K]} \quad b^\top y_k$$

Fundamental results for LPs

Primal:

$$\min_{x \geq 0} \quad c^\top x \quad = \quad Ax \geq b$$

Dual formulation:

$$\max_{y \geq 0} \quad b^\top y \quad A^\top y \leq c$$

$$\max_{k \in [K]} \quad b^\top y_k$$

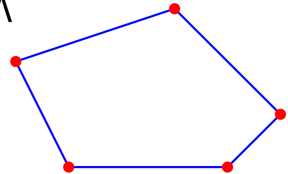
Fundamental results for LPs

Primal:

$$\min_{x \geq 0} \quad c^\top x \qquad =$$

$$Ax \geq b$$

Λ



Dual formulation:

$$\max_{y \geq 0} \quad b^\top y$$

$$A^\top y \leq c$$

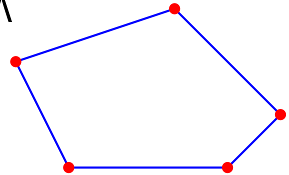
$$\max_{k \in [K]} \quad b^\top y_k$$

Fundamental results for LPs

Primal:

$$\min_{x \geq 0} \quad c^\top x \qquad =$$
$$Ax \geq b$$

Λ



Dual formulation:

$$\max_{y \geq 0} \quad b^\top y$$
$$A^\top y \leq c$$

$$\max_{k \in [K]} \quad b^\top y_k$$

Contents

Description

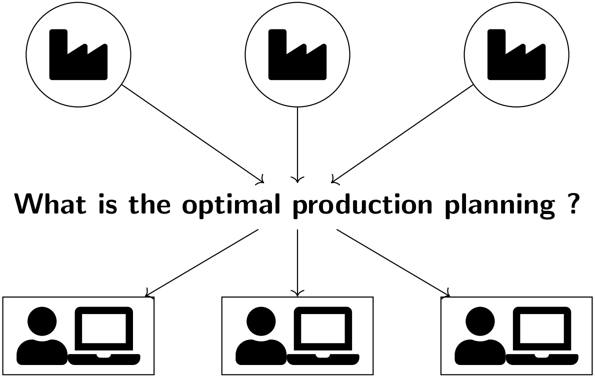


Figure: Unit Commitment

- $\mathcal{J}$: set of units
- T: time horizon (T=24 for a day)
- $\xi \in \mathbb{R}^T$: vector of demand

Description

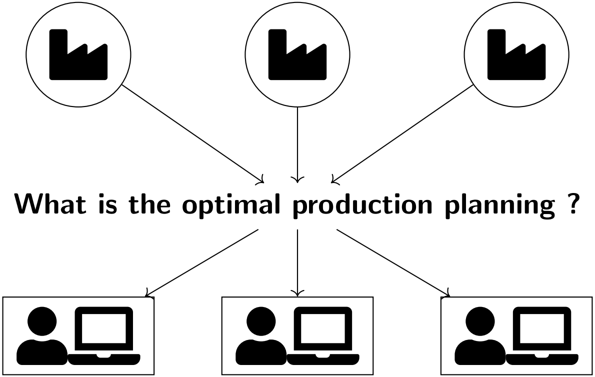


Figure: Unit Commitment

- $\mathcal{J}$: set of units
- T: time horizon (T=24 for a day)
- $\xi \in \mathbb{R}^T$: vector of demand

Variables and Constraints

Variables:

- $y_{j,t}$: Power generation of the unit j is on at time t.
- $x_{j,t}$: Binary variable equal to 1 if the unit j is on at time t.

Example of Constraints:

- Capacity constraints: $P_j^{min} x_{j,t} \leq y_{j,t} \leq P_j^{max} x_{j,t}$

x_t	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Variables and Constraints

Variables:

- $y_{j,t}$: Power generation of the unit j is on at time t.
- $x_{j,t}$: Binary variable equal to 1 if the unit j is on at time t.

Example of Constraints:

- Capacity constraints: $P_j^{min} x_{j,t} \leq y_{j,t} \leq P_j^{max} x_{j,t}$

x_t	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

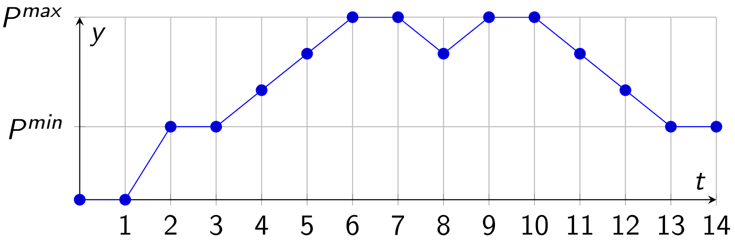
Variables and Constraints

Variables:

- $y_{j,t}$: Power generation of the unit j is on at time t.
- $x_{j,t}$: Binary variable equal to 1 if the unit j is on at time t.

Example of Constraints:

- Capacity constraints: $P_j^{min} x_{j,t} \leq y_{j,t} \leq P_j^{max} x_{j,t}$



x_t	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Mathematical formulation

$$
\begin{aligned}
\min_{x,y} \quad & c^\top x + b^\top y \\
s.t. \quad & Fx \geq f \\
& Ax + By \geq g \\
& \sum_{j \in \mathcal{J}} y_j = \xi
\end{aligned}
$$

But, in reality we have uncertainty in the demand, i.e. ξ is a random variable !

Mathematical formulation

$$
\begin{aligned}
\min_{x,y} \quad & c^\top x + b^\top y \\
s.t. \quad & Fx \geq f \\
& Ax + By \geq g \\
& \sum_{j \in \mathcal{J}} y_j = \boxed{\xi}
\end{aligned}
$$

But, in reality we have uncertainty in the demand, i.e. ξ is a random variable !

2-stage assumption

2-stage assumption:

- 1^{st} stage: commitment variables x (binary)
- 2^{nd} stage: production variables y (continuous)

The decisions x are taken without knowing the uncertainty ξ.

The decisions y are taken with full knowledge of the uncertainty ξ.

$$\begin{aligned} \min_{x,y} \quad & c^\top x + b^\top y^\xi \\ s.t. \quad & Fx \geq f \\ & Ax + By^\xi \geq g \\ & \sum_{j \in \mathcal{J}} y_j^\xi = \xi \end{aligned}$$

2-stage assumption

2-stage assumption:

- 1^{st} stage: commitment variables x (binary)
- 2^{nd} stage: production variables y (continuous)

The decisions x are taken without knowing the uncertainty ξ.

The decisions y are taken with full knowledge of the uncertainty ξ.

$$\begin{aligned} \min_{x,y} \quad & c^\top x + b^\top y^\xi \\ s.t. \quad & Fx \geq f \\ & Ax + By^\xi \geq g \\ & \sum_{j \in \mathcal{J}} y_j^\xi = \xi \end{aligned}$$

2-stage assumption

2-stage assumption:

- 1^{st} stage: commitment variables x (binary)
- 2^{nd} stage: production variables y (continuous)

The decisions x are taken without knowing the uncertainty ξ.

The decisions y are taken with full knowledge of the uncertainty ξ.

$$\begin{aligned} \min_{x} \quad & c^\top x + \min_{\substack{y: Ax + By^\xi \geq g \\ \sum_{j\in\mathcal{M}} y_j^\xi = \xi}} b^\top y^\xi \\ s.t. \quad & x \in X \end{aligned}$$

2-stage assumption

2-stage assumption:

- 1^{st} stage: commitment variables x (binary)
- 2^{nd} stage: production variables y (continuous)

The decisions x are taken without knowing the uncertainty ξ.
The decisions y are taken with full knowledge of the uncertainty ξ.

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

$Q(x,\xi)$: recourse value representing the optimal cost of the second stage, (i.e. value of the optimal planning knowing which units are on or off (x) and the demand (ξ))

Primal and dual formulation of Q

By strong duality in linear programs, Q has a primal and a dual formulation.

Primal formulation:

$$
\begin{aligned}
Q(x,\xi) = \min_{y} \quad & b^\top y \\
& Ax + By \geq d \\
& \sum_{j \in \mathcal{M}} y_j = \xi
\end{aligned}
$$

Primal and dual formulation of Q

By strong duality in linear programs, Q has a primal and a dual formulation.

Primal formulation:

$$Q(x,\xi) = \min_{y} \quad b^\top y$$
$$Ax + By \geq d$$
$$\sum_{j\in\mathcal{M}} y_j = \xi$$

Dual formulation:

$$Q(x,\xi) = \max_{\alpha,\beta,\gamma} \alpha^\top \xi + \beta^\top x + \gamma$$
$$(\alpha, \beta, \gamma) \in \Lambda$$

Primal and dual formulation of Q

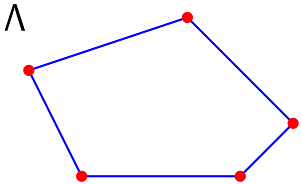
By strong duality in linear programs, Q has a primal and a dual formulation.

Primal formulation:

$$Q(x,\xi) = \min_{y} \quad b^\top y$$
$$Ax + By \geq d$$
$$\sum_{j \in \mathcal{M}} y_j = \xi$$

Dual formulation:

$$Q(x,\xi) = \max_{\alpha,\beta,\gamma} \alpha^\top \xi + \beta^\top x + \gamma$$
$$(\alpha,\beta,\gamma) \in \Lambda$$



Primal and dual formulation of Q

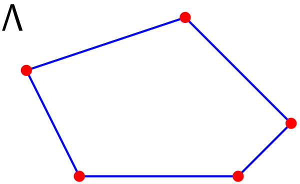
By strong duality in linear programs, Q has a primal and a dual formulation.

Primal formulation:

$$Q(x,\xi) = \min_{y} \quad b^\top y$$
$$Ax + By \geq d$$
$$\sum_{j \in \mathcal{M}} y_j = \xi$$

Dual formulation:

$$Q(x,\xi) = \max_{\alpha,\beta,\gamma} \ \alpha^\top \xi + \beta^\top x + \gamma$$
$$(\alpha, \beta, \gamma) \in \Lambda$$
$$= \max_{k \in [K]} \alpha_k^\top \xi + \beta_k^\top x + \gamma_k$$



What is the best decision?

EDF wants to take the best decisions x (i.e., determine which units are on tomorrow), we need a criteria !

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

This problem is not well-posed, as ξ is a random variable!

What is the best decision?

EDF wants to take the best decisions x (i.e., determine which units are on tomorrow), we need a criteria !

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

This problem is not well-posed, as ξ is a random variable!

$$\min_{x \in X} \quad \mathbb{E}_{\xi \sim \mathbb{P}}[c^\top x + Q(x,\xi)]$$

What is the best decision?

EDF wants to take the best decisions x (i.e., determine which units are on tomorrow), we need a criteria !

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

This problem is not well-posed, as ξ is a random variable!

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

What is the best decision?

EDF wants to take the best decisions x (i.e., determine which units are on tomorrow), we need a criteria !

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

This problem is not well-posed, as ξ is a random variable!

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Problem: This problem is well-posed, but the law of uncertainty $\mathbb{P}$ is unknown ! 😢

What is the best decision?

EDF wants to take the best decisions x (i.e., determine which units are on tomorrow), we need a criteria !

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

This problem is not well-posed, as ξ is a random variable!

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Problem: This problem is well-posed, but the law of uncertainty $\mathbb{P}$ is unknown ! 😢
But we have empirical data (scenarios) ξ^s forming the empirical law $\mathbb{P}_0$! 😄

Contents

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

We approximate $\mathbb{P}$ by $\mathbb{P}_0$

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x\in X} \quad c^\top x + \mathbb{E}_{\xi\sim\mathbb{P}_0}[Q(x,\xi)]$$

We use the primal formulation of $Q(x,\xi)$:

$$Q(x,\xi) = \min_{\substack{Ax+By\geq g \\ \sum_{j\in\mathcal{J}} y_j = \xi}} b^\top y$$

$$\min_{x\in X} c^\top x + \frac{1}{S}\sum_{s\in\mathcal{S}} Q(x,\xi^s)$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the primal formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \min_{\substack{Ax + By \geq g \\ \sum_{j \in \mathcal{J}} y_j = \xi}} b^\top y$$

$$\min_{x \in X} c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} \min_{\substack{Ax + By^s \geq g \\ \sum_{j \in \mathcal{J}} y_j^s = \xi^s}} b^\top y^s$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

We use the primal formulation of $Q(x,\xi)$:

$$Q(x,\xi) = \min_{\substack{Ax+By \geq g \\ \sum_{j \in \mathcal{J}} y_j = \xi}} b^\top y$$

$$\begin{aligned}
\min_{x,y} \quad & c^\top x + \frac{1}{S}\sum_{s \in \mathcal{S}} b^\top y^s \\
s.t. \quad & x \in X, \\
& Ax + By^s \geq g \qquad \forall s \in \mathcal{S} \\
& \sum_{j \in \mathcal{J}} y_j^s = \xi^s \qquad \forall s \in \mathcal{S}
\end{aligned}$$

Risk neutral: Extensive formulation (Primal formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

We use the primal formulation of $Q(x,\xi)$:

$$Q(x,\xi) = \min_{\substack{Ax+By \geq g \\ \sum_{j \in \mathcal{J}} y_j = \xi}} b^\top y$$

$$\begin{aligned} \min_{x,y} \quad & c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} b^\top y^s \\ s.t. \quad & x \in X, \\ & Ax + By^s \geq g \qquad \forall s \in \mathcal{S} \\ & \sum_{j \in \mathcal{J}} y_j^s = \xi^s \qquad \forall s \in \mathcal{S} \end{aligned}$$

$\Longrightarrow$ We have a single (very) large problem to solve.

Benders' Decomposition (Dual Formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the dual formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \max_{k \in [K]} \alpha_k^\top x + \beta_k^\top \xi + \gamma_k$$

Benders' Decomposition (Dual Formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the dual formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \max_{k \in [K]} \alpha_k^\top x + \beta_k^\top \xi + \gamma_k$$

$$\min_{x \in X} c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} Q(x, \xi^s)$$

Benders' Decomposition (Dual Formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the dual formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \max_{k \in [K]} \alpha_k^\top x + \beta_k^\top \xi + \gamma_k$$

$$\min_{x \in X} c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} z_s$$

$$z_s \geq Q(x, \xi^s)$$

Benders' Decomposition (Dual Formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the dual formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \max_{k \in [K]} \alpha_k^\top x + \beta_k^\top \xi + \gamma_k$$

$$\min_{z, x \in X} c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} z_s$$

$$z_s \geq \alpha_k^\top \xi^s + \beta_k^\top x + \gamma_k \qquad \forall k \in [K]$$

Benders' Decomposition (Dual Formulation)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

We use the dual formulation of $Q(x, \xi)$:

$$Q(x, \xi) = \max_{k \in [K]} \alpha_k^\top x + \beta_k^\top \xi + \gamma_k$$

$$\min_{z, x \in X} c^\top x + \frac{1}{S} \sum_{s \in \mathcal{S}} z_s$$

$$z_s \geq \alpha_k^\top \xi^s + \beta_k^\top x + \gamma_k \qquad \forall k \in [K]$$

$\implies$ **Problem:** the extreme points $(\alpha_k, \beta_k, \gamma_k)$ are both unknown and numerous.

$\implies$ Cut generating algorithm.

Benders' Decomposition

Master Problem:

$$\min_{x,z} \quad c^\top x + \frac{1}{S}\sum_{s\in\mathcal{S}} z_s$$
$$\text{s.t.} \quad x \in X,$$
$$z_s \geq (\alpha_\ell^s)^\top \xi^s + (\beta_\ell^s)^\top x + (\gamma_\ell^s), \quad \forall \ell \leq k$$

x^*

Subproblem for each scenario s:

$$\max_{\alpha,\beta,\gamma} \quad \alpha^\top \xi^s + \beta^\top x^* + \gamma$$
$$\text{s.t.} \quad (\alpha, \beta, \gamma) \in \Lambda$$

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

Convergence iff

$$z_s^* \geq (\alpha^*)^\top \xi^s + (\beta^*)^\top x^* + \gamma^*$$

Drawbacks

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Overfitting to the training scenarios. What does it mean?

Training phase

We solve by one of the previous algorithms with $\mathbb{P}_0 = 50$ scenarios:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Let x_0 be the optimal solution.

Test phase

Let $\hat{\mathbb{P}}_0$ another empirical distribution with 1000 scenarios ($=$ very close to $\mathbb{P}$)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \hat{\mathbb{P}}_0}[Q(x,\xi)]$$

$$x = x_0$$

Overfitting: x_0 performs well on the training data $\mathbb{P}_0$ but generalizes poorly to the test data $\hat{\mathbb{P}}_0$.

Drawbacks

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Overfitting to the training scenarios. What does it mean?

Training phase

We solve by one of the previous algorithms with $\mathbb{P}_0 = 50$ scenarios:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Let x_0 be the optimal solution.

Test phase

Let $\hat{\mathbb{P}}_0$ another empirical distribution with 1000 scenarios ($=$ very close to $\mathbb{P}$)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \hat{\mathbb{P}}_0}[Q(x,\xi)]$$

$$x = x_0$$

Overfitting: x_0 performs well on the training data $\mathbb{P}_0$ but generalizes poorly to the test data $\hat{\mathbb{P}}_0$.

Drawbacks

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Overfitting to the training scenarios. What does it mean?

Training phase

We solve by one of the previous algorithms with $\mathbb{P}_0 = 50$ scenarios:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Let x_0 be the optimal solution.

Test phase

Let $\hat{\mathbb{P}}_0$ another empirical distribution with 1000 scenarios ($=$ very close to $\mathbb{P}$)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \hat{\mathbb{P}}_0}[Q(x,\xi)]$$

$$x = x_0$$

Overfitting: x_0 performs well on the training data $\mathbb{P}_0$ but generalizes poorly to the test data $\hat{\mathbb{P}}_0$.

Drawbacks

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

Overfitting to the training scenarios. What does it mean?

Training phase

We solve by one of the previous algorithms with $\mathbb{P}_0 = 50$ scenarios:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

Let x_0 be the optimal solution.

Test phase

Let $\hat{\mathbb{P}}_0$ another empirical distribution with 1000 scenarios ($=$ very close to $\mathbb{P}$)

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \hat{\mathbb{P}}_0}[Q(x,\xi)]$$

$$x = x_0$$

Overfitting: x_0 performs well on the training data $\mathbb{P}_0$ but generalizes poorly to the test data $\hat{\mathbb{P}}_0$.

Contents

Introduction

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

SP problem:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

DRO problem:

$$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x,\xi)]$$

- If $\mathcal{P} = \{\mathbb{P}_0\}$, the DRO problem is the SP problem.
- If $\mathcal{P}$ is sufficiently large, it includes $\mathbb{P}$, but an excessively large $\mathcal{P}$ leads to poor decisions.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to be better than SP?**

Introduction

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

SP problem:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

DRO problem:

$$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x,\xi)]$$

- If $\mathcal{P} = \{\mathbb{P}_0\}$, the DRO problem is the SP problem.
- If $\mathcal{P}$ is sufficiently large, it includes $\mathbb{P}$, but an excessively large $\mathcal{P}$ leads to poor decisions.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to be better than SP?**

Introduction

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

SP problem:	**DRO problem:**
$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$	$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x,\xi)]$

- If $\mathcal{P} = \{\mathbb{P}_0\}$, the DRO problem is the SP problem.
- If $\mathcal{P}$ is sufficiently large, it includes $\mathbb{P}$, but an excessively large $\mathcal{P}$ leads to poor decisions.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to be better than SP?**

Introduction

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x, \xi)]$$

SP problem:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x, \xi)]$$

DRO problem:

$$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x, \xi)]$$

- If $\mathcal{P} = \{\mathbb{P}_0\}$, the DRO problem is the SP problem.
- If $\mathcal{P}$ is sufficiently large, it includes $\mathbb{P}$, but an excessively large $\mathcal{P}$ leads to poor decisions.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to be better than SP?**

Introduction

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}}[Q(x,\xi)]$$

SP problem:

$$\min_{x \in X} \quad c^\top x + \mathbb{E}_{\xi \sim \mathbb{P}_0}[Q(x,\xi)]$$

DRO problem:

$$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x,\xi)]$$

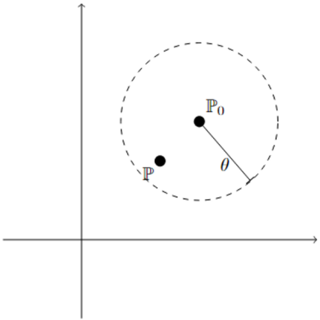
- If $\mathcal{P} = \{\mathbb{P}_0\}$, the DRO problem is the SP problem.
- If $\mathcal{P}$ is sufficiently large, it includes $\mathbb{P}$, but an excessively large $\mathcal{P}$ leads to poor decisions.

$\Longrightarrow$ **How can we select $\mathcal{P}$ to be better than SP?**

Choice of $\mathcal{P}$

$$\min_{x \in X} \quad c^\top x + \max_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_{\xi \sim \mathbb{Q}}[Q(x, \xi)]$$

Wasserstein distance $W_2(\mathbb{Q}, \mathbb{P})$ is a distance between the distributions $\mathbb{Q}$ and $\mathbb{P}$



$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}) \leq \theta\}$$

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

$$\begin{aligned} \min_{x} \quad & c^\top x + \max_{\substack{\mathbb{Q}:\, W_2(\mathbb{Q},\mathbb{P}_0) \leq \theta \quad [\lambda] \\ supp(\mathbb{Q}) \subset \mathbb{R}^T}} \mathbb{E}_{\mathbb{Q}}[Q(x,\xi)] \\ s.t. \quad & x \in X \end{aligned}$$

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

$$\begin{aligned} \min_{x, \lambda \geq 0} \quad & c^\top x + \lambda \theta^2 + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi \in \mathbb{R}^T} \left(Q(x, \xi) - \lambda \|\xi - \zeta_i\|_2^2 \right) \\ s.t. \quad & x \in X \end{aligned}$$

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \le \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

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

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

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

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

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

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \mid W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

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

[1] Gao and Kleywegt 2016.

DRO Unit Commitment problem

$$\mathcal{P} = \{\mathbb{Q} \,|\, W_2(\mathbb{Q}, \mathbb{P}_0) \leq \theta\}, \qquad \mathbb{P}_0 = \frac{1}{N} \sum_{i \in [N]} \delta_{\zeta_i}$$

Reformulation DRO Unit Commitment problem[1]:

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

$\Longrightarrow$ **Benders' algorithm**

[1] Gao and Kleywegt 2016.

Benders algorithm DRO

Master Problem:

$$\min_{x,w,\lambda,z} \quad c^\top x + w\theta^2 + \frac{1}{N}\sum_{i=1}^{N} z_i$$

$$\text{s.t.} \quad x \in X$$

$$\|(2, w-\lambda)\|_2 \leq w + \lambda$$

$$\boldsymbol{z_i \geq (\alpha_\ell^i)^\top \zeta^i + (\beta_\ell^i)^\top x + \gamma_\ell^i + \frac{\lambda\|\alpha_\ell^i\|_2^2}{4}, \quad \forall \ell \leq k, \forall i \in [N]}$$

$$\theta \in \mathbb{R}, \quad x_i \in \{0,1\}^{m_i \times T}$$

(x^*, λ^*)

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

Subproblem for each scenario i:

$$\max_{\alpha,\beta,\gamma} \quad \alpha^\top \zeta_i + \beta^\top x^* + \gamma + \frac{\lambda^*\|\alpha\|_2^2}{4}$$

$$\text{s.t.} \quad (\alpha, \beta, \gamma) \in \Lambda$$

Benders algorithm DRO

Master Problem:

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

(x^*, λ^*) → Subproblem; $(\alpha^*, \beta^*, \gamma^*)$ → Master Problem

Subproblem for each scenario i:

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

Convex maximization problem (NP-Hard)

Contents

Presentation

- Source: SMS++/ EDF.
- 5 instances with 10 units and 24 time steps.
- Source scenarios: `https://data.open-power-system-data.org/time_series/`
- Training on 25 scenarios and test on 1000 scenarios

Out-Of-Sample Costs over 1000 scenarios

Method	1	2	3	4	5
DRO	$1.88\,.10^6$	$1.32\,.10^6$	$1.81\,.10^6$	$1.82\,.10^6$	$1.89\,.10^6$
SP	$1.90\,.10^6$	$1.33\,.10^6$	$1.83\,.10^6$	$1.83\,.10^6$	$1.92\,.10^6$
Gap	1.24%	1.02%	1.01%	0.72%	1.48%

Table: Average cost over 1000 scenarios

$\Longrightarrow$ DRO leads to better decisions. But it requires slightly more time.

Contents

Conclusion

- Many optimization problems involve uncertain parameters.
- Presentation of the classical method to deal with uncertainty (Stochastic Programming)
- DRO is an innovative approach, yet it comes with challenges $\Longrightarrow$ An ideal topic for a PhD.😄