PowerTech Belgrade 2023
LEADING INNOVATIONS FOR RESILIENT & CARBON-NEUTRAL POWER SYSTEMS
25-29 JUNE, 2023, BELGRADE, SERBIA
IEEE PES Power & Energy Society®
IEEE

A bilevel framework for decision-making under uncertainty with contextual information

Salvador Pineda

(joint work with J. M. Morales and M. A. Muñoz)

OASYS Research Group
University of Málaga (Spain)

OASYS & Málaga

- **O**ptimization and **A**nalytics for **S**ustainable energ**Y** **S**ystems
- Over 300 sundays per year (known as Costa del Sol)

Framework

- At 10 am we have to decide how much ice cream to make (decision z)
- At 10 am we do not know the demand in the afternoon (uncertain parameter y)

Framework

Decision-making under uncertainty:

$$z^* = \arg\min_{z \in Z} \ \mathbb{E}[f_0(z; y)]$$

where

- z are the decisions variables
- y are uncertain parameters, $y \sim Y$
- f_0 is the objective function
- Z is the feasible set
- Stochastic programming, robust optimization, etc

Framework

- We can use some available information such as the temperature at 10 am (contextual data x)
- Obviously there is a relationship between the morning temperature (x) and the ice cream demand in the afternoon (y).
- We would like to use such relation to make better decisions about ice cream quantity

Framework

Decision-making under uncertainty with contextual information:

$$z^* = \arg\min_{z \in Z} \mathbb{E}[f_0(z; y)|x]$$

where

- z are the decisions variables
- y are uncertain parameters, $y \sim Y$
- x are contextual features, $x \sim X$
- f_0 is the objective function
- Z is the feasible set
- Input: available data $S = \{(x_t, y_t, z_t^*), \forall t \in \mathcal{T}\}$ (training)
- Output: optimal decision for a new period $z_{\tilde{t}}^*$ with known $x_{\tilde{t}}$ (test)

Approaches

Problem

$$z^* = \arg\min_{z \in Z} \ \mathbb{E}[f_0(z; y)|x]$$

Data

$$S = \{(x_t, y_t, z_t^*), \forall t \in \mathcal{T}\}$$

Approaches

Problem	Data
$z^* = \arg\min_{z \in Z} \ \mathbb{E}[f_0(z; y) \mid x]$	$S = \{(x_t, y_t, z_t^*), \forall t \in \mathcal{T}\}$

- Forecasting approach (FO)

learns the relation between y and x ignoring f_0 and Z

Approaches

Problem

$$z^* = \arg\min_{z \in Z} \; \mathbb{E}[f_0(z; y)|x]$$

Data

$$S = \{(x_t, y_t, z_t^*), \forall t \in \mathcal{T}\}$$

- Forecasting approach (FO)

  learns the relation between y and x ignoring f_0 and Z

- Decision rule approach (DR)

  learns the relation between z^* and x

Approaches

Problem

$$z^* = \arg\min_{z \in Z} \ \mathbb{E}[f_0(z; y)|x]$$

Data

$$S = \{(x_t, y_t, z_t^*), \forall t \in \mathcal{T}\}$$

- Forecasting approach (FO)

  learns the relation between y and x ignoring f_0 and Z

- Decision rule approach (DR)

  learns the relation between z^* and x

- Bilevel approach (BL)

  learns the relation between y and x taking into account f_0 and Z

Strategic producer

- Linear inverse demand function $p = \alpha - \beta q$ (Uncertain)
- Quadratic cost function $c_2 q^2 + c_1 q$
- Bounds on produced quantity $\underline{q} \leq q \leq \overline{q}$
- A strategic producer maximizes profits by solving

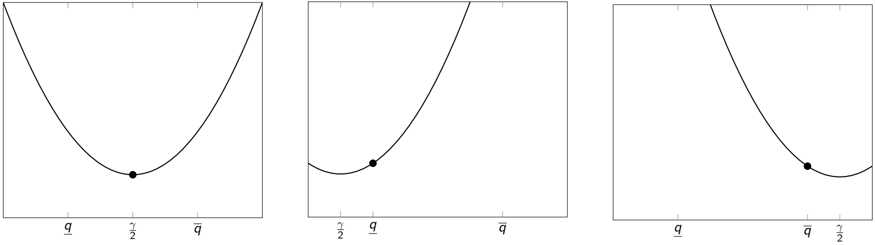
$$q^* = \arg \min_{\underline{q} \leq q \leq \overline{q}} q^2 - \gamma q$$

- Parameter γ is usually unknown
- Historical data set $S = \{(x_t, \gamma_t, q_t^*), \forall t \in \mathcal{T}\}$

Strategic producer

The solution of the quadratic optimization problem is the following

$$\min_{\underline{q} \leq q \leq \overline{q}} q^2 - \gamma q$$

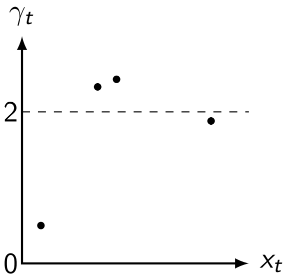


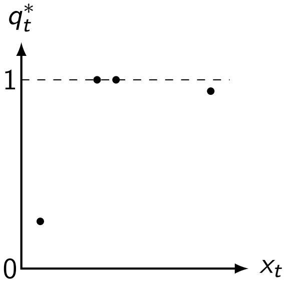
$q^* = \frac{\gamma}{2}$ $\qquad$ $q^* = \underline{q}$ $\qquad$ $q^* = \overline{q}$

Strategic producer

t	x_t	γ_t	$\gamma_t/2$	q_t^*
1	1	0.50	0.25	0.25
2	4	2.33	1.16	1.00
3	5	2.43	1.21	1.00
4	10	1.88	0.94	0.94

$$0 \leq q \leq 1$$

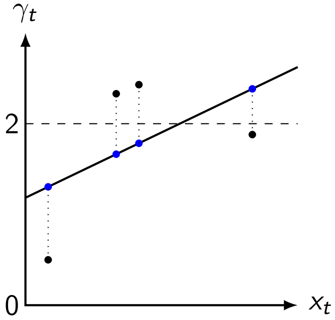




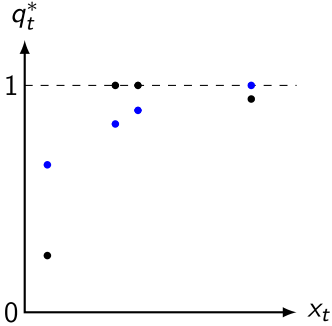
Income = 21.16 (100%)

Strategic producer

Forecasting approach (FO): $\hat{\gamma}_t = 1.184 + 0.12x_t$



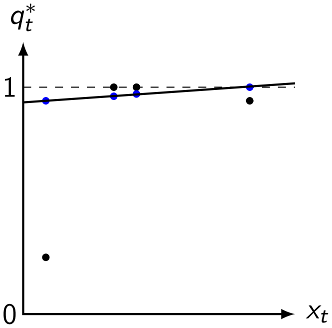
Forecast error $= 0.665$



Income $= 20.14 \quad (95.2\%)$

Strategic producer

Decision-rule approach (DR): $\hat{q}_t = 0.933 + 0.007x_t$



Income $= 20.02 \quad (94.6\%)$

Bilevel programming

- John and Peter go to the movies every week.
- First John decides the theater, and then Peter decides the movie.
- John prefers action over terror, and terror over romantic.
- Peter prefers romantic over terror, and terror over action.

Theater A
The Matrix
Notting Hill

Theater B
The Exorcist
The Shining

Bilevel programming

Bilevel optimization is a special kind of optimization where one problem is embedded (nested) within another and can be formulated as follows

$$
\begin{aligned}
\min_{x\ (theater)} \quad & F_0(x,y) \quad (John) \\
\text{s.t.} \quad & F_i(x,y) \leq 0, \qquad i = 1,\dots,I \\
& H_j(x,y) = 0, \qquad j = 1,\dots,J \\
& \min_{y\ (movie)} \quad f_0(x,y) \quad (Peter) \\
& \text{s.t.} \quad f_k(x,y) \leq 0, \qquad k = 1,\dots,K \\
& \qquad\;\; h_l(x,y) = 0, \qquad l = 1,\dots,L
\end{aligned}
$$

Bilevel approach (BL)

1) Function to learn the relation between y and x so that $\hat{y} = g^{\mathrm{BL}}(x; w)$.

Bilevel approach (BL)

1) Function to learn the relation between y and x so that $\hat{y} = g^{\mathrm{BL}}(x; w)$.
2) Adjust parameters w by solving the following bilevel problem

$$
\begin{aligned}
w^{\mathrm{BL}} = \arg\min_{w} \sum_{t \in \mathcal{T}} & f_0(\hat{z}_t; y_t) \quad (\textit{Forecaster}) \\
\text{s.t.} \;\; \hat{z}_t = & \arg\min_{z \in Z} f_0(z; g^{\mathrm{BL}}(x_t; w)) \;\; \forall t \in \mathcal{T} \;\; (\textit{Decision maker})
\end{aligned}
$$

Bilevel approach (BL)

1) Function to learn the relation between y and x so that $\hat{y} = g^{\mathrm{BL}}(x; w)$.
2) Adjust parameters w by solving the following bilevel problem

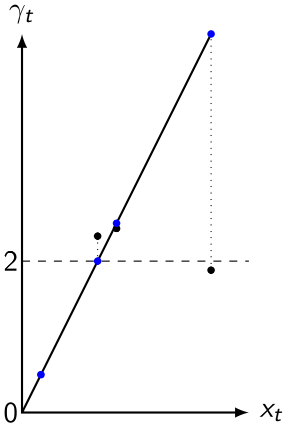
$$
\begin{aligned}
w^{\mathrm{BL}} = \arg\min_{w} \sum_{t \in \mathcal{T}} f_0(\hat{z}_t; y_t) \quad & (\textit{Forecaster}) \\
\text{s.t.} \;\; \hat{z}_t = \arg\min_{z \in Z} f_0(z; g^{\mathrm{BL}}(x_t; w)) \;\; \forall t \in \mathcal{T} \;\; & (\textit{Decision maker})
\end{aligned}
$$

3) Determine the optimal decision of a unseen time period $\tilde{t}$ as

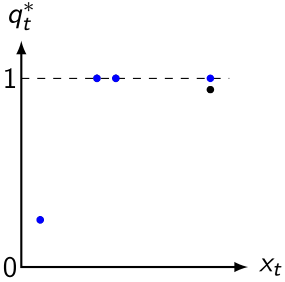
$$
z_{\tilde{t}}^{\mathrm{BL}} = \arg\min_{z \in Z} \;\; f_0(z; g^{\mathrm{BL}}(x_{\tilde{t}}; w^{\mathrm{BL}}))
$$

Strategic producer

Bilevel approach (BL): $\hat{\gamma}_t = 0.000 + 0.5x_t$



Forecast error $= 0.745$



Income $= 21.13 \quad (99.9\%)$

Strategic producer

	Forecast error	Income
Forecasting (FO)	0.665	95.2%
Decision-rule (DR)	-	94.6%
Bilevel (BL)	0.745	99.9%

- FO minimizes forecast error, but yields suboptimal decisions
- DR simplifies decision-making, but also yields suboptimal decisions
- BL uses contextual information to prescribe the best decisions

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information
- For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information
- For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- Bilevel problems are hard to solve (single-level reformulation)

Bilevel approach (BL)

- 👍 Natural framework to prescribe decisions using contextual information
- 👎 For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- 👎 Bilevel problems are hard to solve (single-level reformulation)
  - If the lower-level is not convex, KKT are not sufficient for optimality

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information
- For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- Bilevel problems are hard to solve (single-level reformulation)
  - If the lower-level is not convex, KKT are not sufficient for optimality
  - If the lower-level is convex, it can be replaced by its KKT

Bilevel approach (BL)

- 👍 Natural framework to prescribe decisions using contextual information
- 👎 For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- 👎 Bilevel problems are hard to solve (single-level reformulation)
  - If the lower-level is not convex, KKT are not sufficient for optimality
  - If the lower-level is convex, it can be replaced by its KKT
    - 👎 The obtained single-level problem is non-convex and violates CQ

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information
- For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- Bilevel problems are hard to solve (single-level reformulation)
  - If the lower-level is not convex, KKT are not sufficient for optimality
  - If the lower-level is convex, it can be replaced by its KKT
    - The obtained single-level problem is non-convex and violates CQ
    - Iterative regularization approach (local optimal solutions)

Bilevel approach (BL)

- Natural framework to prescribe decisions using contextual information
- For some problems, non-anticipativity can be violated if the value of w yields decisions $\hat{z}_t$ that anticipate y_t (optimistic solution)
- Bilevel problems are hard to solve (single-level reformulation)
  - If the lower-level is not convex, KKT are not sufficient for optimality
  - If the lower-level is convex, it can be replaced by its KKT
    - The obtained single-level problem is non-convex and violates CQ
    - Iterative regularization approach (local optimal solutions)
    - Big-M approach (global optimality, but computationally costly)

Case study

- Real data from Iberian electricity market
- Wind and solar power forecasts are used as contextual information
- Three technologies: base (nuclear), medium (carbon) and peak (gas)
- 43 sets of 200 hours (160 hours as training and 40 hours as test)

Case study

	Base	Medium	Peak
Forecasting (FO)	96.0%	77.3%	41.6%
Decision-rule (DR)	94.6%	62.6%	18.9%
Bilevel (BL)	96.3%	80.0%	58.7%

- All methods provide similar incomes for the base unit
- Market uncertainty affects the operation of medium and peak units
- The BL approach obtains the highest incomes for all technologies
- DR approach lead to a significant number of infeasible cases

Conclusions

- Forcasting approach (FO)
  - learns the relation between y and x ignoring f_0 and Z
  - wide variety of learning techniques can be applied
  - obtained decisions may be suboptimal

Conclusions

- Forecasting approach (FO)
  - learns the relation between y and x ignoring f_0 and Z
  - wide variety of learning techniques can be applied
  - obtained decisions may be suboptimal
- Decision rule approach (DR)
  - learns the relation between z^* and x
  - decisions are quickly obtained wihtout solving an optimization problem
  - obtained decisions may be suboptimal and infeasible

Conclusions

- Forecasting approach (FO)
  - learns the relation between y and x ignoring f_0 and Z
  - wide variety of learning techniques can be applied
  - obtained decisions may be suboptimal
- Decision rule approach (DR)
  - learns the relation between z^* and x
  - decisions are quickly obtained wihtout solving an optimization problem
  - obtained decisions may be suboptimal and infeasible
- Bilevel approach (BL)
  - learns the relation between y and x taking into account f_0 and Z
  - best possible decisions using available contextual information
  - bilevel problem can be only solved under certain assumptions

Thanks for the attention!! Questions??

Omega 108 (2022) 102575

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

A bilevel framework for decision-making under uncertainty with contextual information☆

M.A. Muñoz, S. Pineda*, J.M. Morales

OASYS Group, University of Malaga, Malaga, Spain

OASYS
Optimization and Analytics for Sustainable energY Systems

More info: oasys.uma.es

Email: spineda@uma.es