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

SEIO 2022

S. Pineda, M. A. Muñoz, J.M. Morales

OASYS group, University of Málaga (Spain)

June 8, 2022

Framework

Optimization problem:

$$z^* = \arg\min_{z \in Z} \; f(z; y)$$

where

- z are the decisions variables
- y are parameters
- f is the objective function
- Z is the feasible set

Framework

Stochastic optimization problem:

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

where

- z are the decisions variables
- y are uncertain parameters following a probability distribution $y \sim Y$
- f is the objective function
- Z is the feasible set
- Solved with stochastic programming or robust optimization

Framework

Conditional stochastic optimization problem:

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

where

- z are the decisions variables
- y are uncertain parameters, $y \sim Y$
- x are contextual features with explanatory power about y, $x \sim X$
- f is the objective function
- Z is the known 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

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

Data

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

Approaches

Problem

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

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 and Z

Approaches

Problem

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

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 and Z
- Decision rule approach (DR)
  - learns the relation between z^* and x

Approaches

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

- Forecasting approach (FO)
  - learns the relation between y and x ignoring f 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 and Z

Application: strategic producer

- Linear relation between price (p) and quantity (q) $p = \alpha - \beta q$
- Quadratic cost function $c_2 q^2 + c_1 q$
- Quantity bounds $\underline{q} \leqslant q \leqslant \overline{q}$
- Producer's profit

$$\Pi(q) = pq - c_2 q^2 - c_1 q = (\alpha - \beta q)q - c_2 q^2 - c_1 q = -\beta' q^2 + \alpha' q$$

- A strategic producer maximizes profits by solving

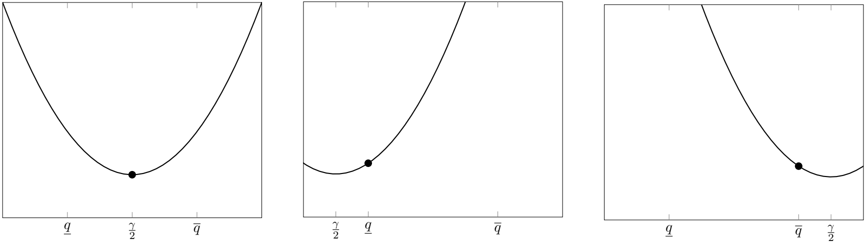
$$q^* = \arg \max_{\underline{q} \leqslant q \leqslant \overline{q}} -\beta' q^2 + \alpha' q = \arg \min_{\underline{q} \leqslant q \leqslant \overline{q}} q^2 - \gamma q$$

- Parameter γ is usually unknown and the producer has available a set of historical measures $S = \{(x_t, \alpha'_t, \beta'_t), \forall t \in \mathcal{T}\}$, with $x_t \in \mathbb{R}^q$

Application: strategic producer

The solution of the quadratic optimization problem is the following

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

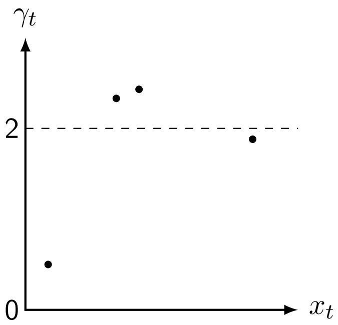


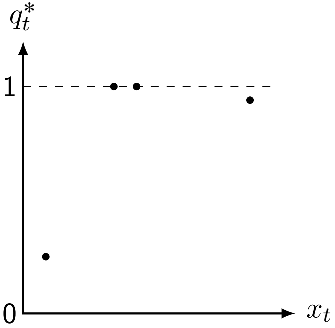
$q^* = \frac{\gamma}{2}$ $\quad$ $q^* = \underline{q}$ $\quad$ $q^* = \overline{q}$

Illustrative example

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

$$0 \leqslant q \leqslant 1$$





Income = 21.16 € (100%)

Approaches

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

- Forecasting approach (FO)
  - learns the relation between y and x ignoring f 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 and Z

Forecasting approach (FO)

- We assume a linear function so that $\hat{\gamma}_t = w^T x_t$ with $w \in \mathbb{R}^q$
- We choose the squared error as the loss function $l^{\mathsf{FO}}(\gamma, \hat{\gamma}) = (\gamma - \hat{\gamma})^2$
- We compute w by solving the quadratic problem

$$w^{\mathsf{FO}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} (\gamma_t - w^T x_t)^2$$

- We estimate $\hat{\gamma}$ for a new time period $\tilde{t}$ as

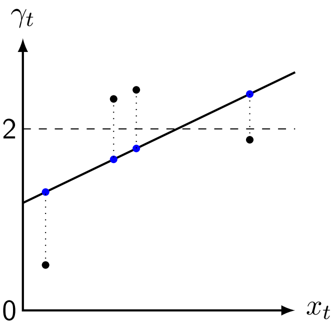
$$\hat{\gamma}_{\tilde{t}} = (w^{\mathsf{FO}})^T x_{\tilde{t}}$$

- We determine the produced quantity for a new time period $\tilde{t}$ as

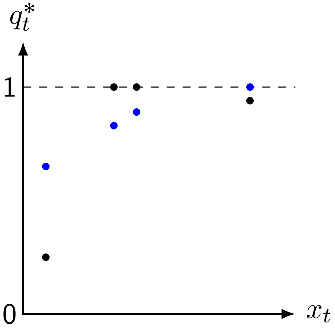
$$q^*_{\tilde{t}} = \arg\min_{\underline{q} \leqslant q_{\tilde{t}} \leqslant \overline{q}} q^2_{\tilde{t}} - \hat{\gamma}_{\tilde{t}} q_{\tilde{t}}$$

Illustrative example

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



RMSE = 0.665



Income = 20.14 € (95.2%)

Approaches

Problem

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

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 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 and Z

Decision-rule approach (DR)

- We assume a linear function so that $q_t^* = w^T x_t$
- We compute w by solving the quadratic problem

$$w^{\mathsf{DR}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} \beta_t'(q_t)^2 - \alpha_t' q_t$$

$$\text{s.t. } \underline{q} \leqslant q_t \leqslant \overline{q}, \quad \forall t \in \mathcal{T}$$

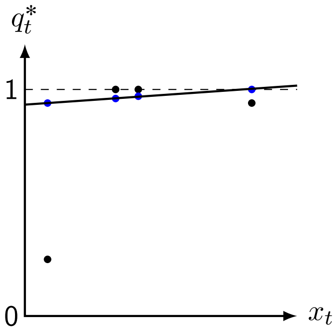
$$q_t = w^T x_t, \quad \forall t \in \mathcal{T}$$

- We determine the produced quantity for a new time period $\tilde{t}$ as

$$q_{\tilde{t}}^* = (w^{\mathsf{DR}})^T x_{\tilde{t}}$$

Illustrative example

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



Income $=$ 20.02 € (94.6%)

Approaches

Problem

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

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 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 and Z

Bilevel approach (BL)

- We assume a linear function so that $\hat{\gamma}_t = w^T x_t$ with $w \in \mathbb{R}^q$
- We compute w by solving the bilevel

$$w^{\mathsf{BL}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} \beta'_t \hat{q}_t^2 - \alpha'_t \hat{q}_t$$

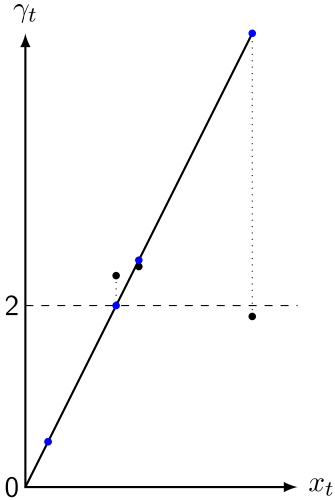
$$\text{s.t. } \hat{q}_t = \arg\min_{\underline{q} \leqslant q_t \leqslant \overline{q}} q_t^2 - w^T x_t q_t, \ \ \forall t \in \mathcal{T}$$

- We determine the produced quantity for a new time period $\tilde{t}$ as

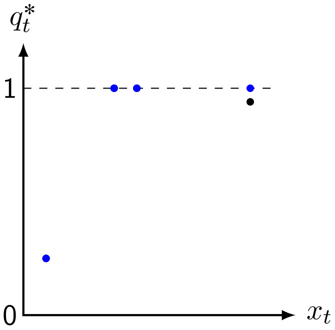
$$q^*_{\tilde{t}} = \arg\min_{\underline{q} \leqslant q_{\tilde{t}} \leqslant \overline{q}} q_{\tilde{t}}^2 - (w^{\mathsf{BL}})^T x_{\tilde{t}} q_{\tilde{t}}$$

Illustrative example

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



RMSE = 0.745



Income = 21.13 € (99.9%)

Illustrative example

	RMSE	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 relates uncertainty and context to derive the best decisions

Case study

- Real data from Iberian electricity market is used to approximate the inverse demand function
- Wind and solar power forecasts is used as contextual information
- Three different generation technologies: base (nuclear), medium (carbon) and peak (gas)
- 43 sets of 200 hours (160 hours as training and 40 hours as test) are used to compute average results

Case study

	Base	Medium	Peak
Relative income FO	96.0%	77.3%	41.6%
Relative income DR	94.6%	62.6%	18.9%
Relative income BL	96.3%	80.0%	58.7%
Infeasible cases DR	4.9%	1.7%	0.1 %

- All methods provide similar incomes for the base unit since it is at full capacity most of the time
- The uncertainty of the inverse demand function significantly affects the operation of medium and peak units
- The proposed BL approach obtains the highest incomes for the three generating technologies
- DR approach lead to a significant number of infeasible cases

Conclusions

- Forecasting approach (FO)
  - learns the relation between y and x ignoring f 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 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