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

Miguel A. Muñoz[1]

Juan Miguel Morales [1] Salvador Pineda [1]

University of Malaga [1]

October 2020

Cooking a data-driven decision

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

x	contextual info
y	uncertain parameter
z	decision

A receipt:

1. Select an uncertain parameter model
2. Train the model with a data set
3. Issue a point forecast for the unseen period $\tilde{t}$
4. Solve the deterministic optimization problem

$$\hat{y} = g(x; w)$$
$$\boxed{w^* = \phi(g(x_t; w), y_t)}$$
$$\hat{y}_{\tilde{t}} = g(x_{\tilde{t}}; w^*)$$
$$z_{\tilde{t}} = \arg\min_{z \in Z} f_0(z; \hat{y}_{\tilde{t}})$$

How to train the model?

Predictive approach

$$w^{\mathrm{FO}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} l^{\mathrm{FO}}(g^{\mathrm{FO}}(x_t; w), y_t)$$

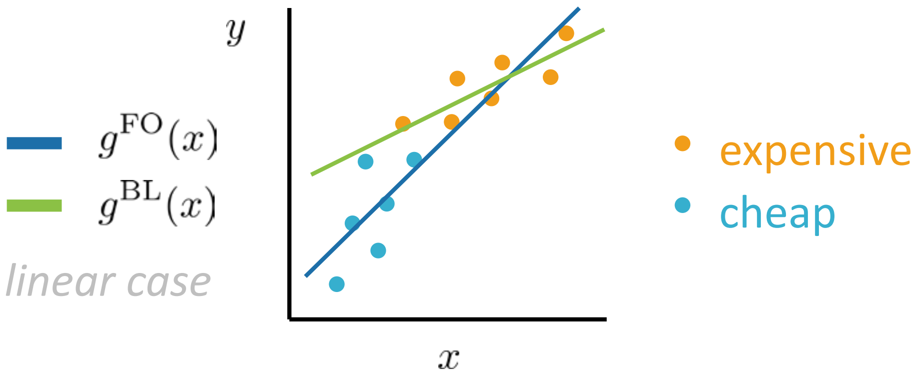
- ✓ Perform well in most tasks
- ✓ Simpler
- × Suboptimal for decision-making

Bilevel prescriptive approach

$$w^{\mathrm{BL}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} f_0(\hat{z}_t; y_t)$$

$$\text{s.t.} \quad \hat{z}_t = \arg\min_{z \in Z} f_0(z; g^{\mathrm{BL}}(x_t; w)) \ \ \forall t \in \mathcal{T}$$

- ✓ Yields better decisions
- × More complex
- × Issues for linear f_0



Reformulation of the lower level (I)

Lower Level

$$w^{\mathrm{BL}} = \arg\min_{w \in \mathbb{R}^q} \sum_{t \in \mathcal{T}} f_0(\hat{z}_t; y_t)$$

$$\text{s.t. } \hat{z}_t = \arg\min_z f_0(z; g^{\mathrm{BL}}(x_t; w)) \ \ \forall t \in \mathcal{T}$$

$$\left.\begin{array}{l} f_i(z_t) \le 0, \ \ \forall i \\ h_j(z_t) = 0, \ \ \forall j \end{array}\right\} \forall t \in \mathcal{T}$$

Feasible set Z

KKT optimality conditions

$$w^{\mathrm{BL}} = \arg\min_{w, \hat{z}_t, \lambda_{it}, \upsilon_{it}} \sum_{t \in \mathcal{T}} f_0(\hat{z}_t; y_t)$$

$$\text{s.t. } \nabla f_0(\hat{z}_t, g^{\mathrm{BL}}(x_t, w)) + \sum_{i=1}^{I} \lambda_{it} \nabla f_i(\hat{z}_t) + \sum_{j=1}^{J} \upsilon_{jt} \nabla h_j(\hat{z}_t) = 0, \ \ \forall t \in \mathcal{T}$$

(stationarity)

$$\left.\begin{array}{l} f_i(\hat{z}_t) \le 0, \ \ \forall i \\ h_j(\hat{z}_t) = 0, \ \ \forall j \\ \lambda_{it} \ge 0, \ \ \forall i \\ \lambda_{it} f_i(\hat{z}_t) = 0, \ \ \forall i \end{array}\right\} \forall t \in \mathcal{T}$$

(primal feasibility)

(dual feasibility)

(complementary slackness)

Reformulation of the lower level (II)

$$\lambda_{it} f_i(\hat{z}_t) = 0, \quad \forall i, \quad \forall t \in \mathcal{T}$$

(complementary slackness)

Approach 1: **Regularization**

$$-\sum_{\forall it} \lambda_{it} f_i(\hat{z}_t) \leq \epsilon \qquad \epsilon \to 0$$

Solver: CONOPT ↑ fast, ↓ non-optimal

Approach 2: **Big-M**

$$\left.\begin{aligned} &\lambda_{it} \leq u_{it} M^D, \\ &f_i(\hat{z}_t) \geq (u_{it} - 1) M^P, \\ &u_{it} \in \{0, 1\}, \end{aligned}\right\} \quad \forall i, \forall t \in \mathcal{T}$$

Under some conditions of f_0 and Z → **MIQP**

Solvers: CPLEX, Gurobi

↑ optimality, ↓ slow, big-M tunning

THANKS!

Miguel A. Muñoz
miguelangeljmd@uma.es
University of Malaga

Checkout more at:

M. A. Muñoz, S. Pineda, J. M. Morales, **A bilevel framework for decision-making under uncertainty with contextual information**. *arXiv preprint arXiv:2008.01500*, 2020.