Why won't local algorithms work for bilevel games?

Based on a work with J. Dutta and D. Thirumulanathan.

Parin Chaipunya

KMUTT

└Mathematics @ Faculty of Science
└The Joint Graduate School of Energy and Environments

Areas of research:

- Multi-agent optimization: Bilevel programs, Games
- Optimization modeling for energy and environmental applications
- Nonsmooth geometry in optimization

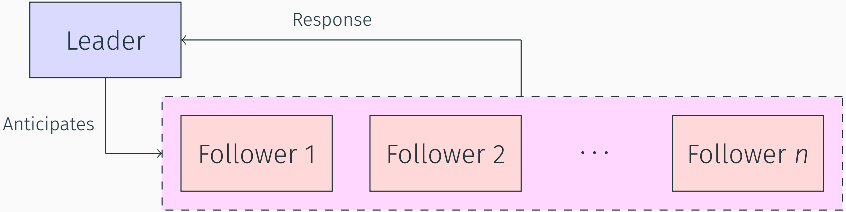
AMC2025, Chiang Mai
August 3–7, 2025.

Scan to get the slides.

Excerpt

Schematic.

We consider a bilevel game of **von Stackelberg** with 1 **leader** and n **followers**.



Decisions and outcomes

Firstly, the **leader** makes the decision (Anticipation).

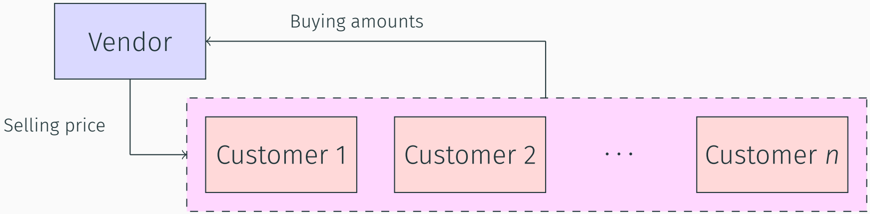
Secondly, the **followers** simultaneously and independently make the decisions.

The **outcome** of each player is **collectively affected** by the decision of **every players**.

Excerpt

Schematic.

One could think of the following situation.



Decisions and outcomes

Firstly, the **vendor** sets the selling price .

Secondly, the **customers** and independently decides the buying amounts.

The **outcome** of each player is **collectively affected** by the decision of **every players** and revealed after all parties made their decisions.

Ingredients

Players

Players	$\mathtt{a} \in \mathtt{A} := \{\mathtt{l}, \mathtt{f}_1, \mathtt{f}_2, \ldots, \mathtt{f}_n\}$
Leader	$\mathtt{l}$
Followers	$\mathtt{f} \in \mathtt{F} := \{\mathtt{f}_1, \ldots, \mathtt{f}_n\}$

Decision and Outcome of player a

Decision	$u^{\mathtt{a}} \in U^{\mathtt{a}} := \mathbb{R}^{n_{\mathtt{a}}}$	
Outcome	$\Theta^{\mathtt{a}} : U^{\mathtt{a}} \times U^{-\mathtt{a}} \to \mathbb{R}$	$U^{-\mathtt{a}} := \prod_{\mathtt{b} \in \mathtt{A} \setminus \{\mathtt{a}\}} U^{\mathtt{b}}$
Constraint functions	$g_j^{\mathtt{f}} : U^{\mathtt{l}} \times U^{\mathtt{f}} \to \mathbb{R}$	$j = 1, \ldots, p^{\mathtt{f}}$

Excerpt

Problem formulation The bilevel game that we consider takes the **optimistic** form.

$$\begin{cases} \min\limits_{u^{\mathtt{l}}, u^{\mathsf{F}}} & \Theta^{\mathtt{l}}(u^{\mathtt{l}}, u^{\mathsf{F}}) \\ s.t. & u^{\mathtt{l}} \in Y^{\mathtt{l}} \subset U^{\mathtt{l}} \\ & \forall \mathsf{f} \in \mathsf{F} : \\ & \quad \lfloor\, u^{\mathsf{f}} \in \arg\min\limits_{\hat{u}^{\mathsf{f}}} \left\{ \Theta^{\mathsf{f}}(u^{\mathtt{l}}, \hat{u}^{\mathsf{f}}) \mid g_j^{\mathsf{f}}(u^{\mathtt{l}}, \hat{u}^{\mathsf{f}}) \leq 0 \ (\forall j = 1, \ldots, p^{\mathsf{f}}) \right\} \end{cases}$$

Aim of this talk: To discuss **systematically**...

- why local algorithms are not suitable for this class of problem, and
- under which conditions that they work.

Talk anatomy

How to deal with a bilevel game (Analytically and Numerically) ?

Why only local solutions ?

What happens at the solutions ?

Further results

How to deal with a bilevel game (Analytically and Numerically) ?

Common difficulties

One usually faces the following issues when considering a bilevel game.

- The ill-posedness of the problem formulation. (Resolved by specifying the response nature.)
- The nested optimization problems is not tractable. (Resolved by reformulation techniques.)
- The poor geometric properties. (Not resolved, required special attention.)

Reformulation technique I: Using value functions

To each follower $\mathsf{f} \in \mathsf{F}$, define a value function $V^{\mathsf{f}} : U^{\mathsf{l}} \to [-\infty, +\infty]$ by

$$V^{\mathsf{f}}(u^{\mathsf{l}}) := \underbrace{\inf_{\hat{u}^{\mathsf{f}}} \left\{ \Theta^{\mathsf{f}}(u^{\mathsf{l}}, \hat{u}^{\mathsf{f}}) \mid g_j^{\mathsf{f}}(u^{\mathsf{l}}, \hat{u}^{\mathsf{f}}) \leq 0 \ (\forall j = 1, \ldots, p^{\mathsf{f}}) \right\}.}_{\text{Best outcome of } \mathsf{f}\text{, corresponded to } u^{\mathsf{l}}.}$$

Then

$$u^{\mathsf{f}} \in \arg\min_{\hat{u}^{\mathsf{f}}} \left\{ \Theta^{\mathsf{f}}(u^{\mathsf{l}}, \hat{u}^{\mathsf{f}}) \mid g_j^{\mathsf{f}}(u^{\mathsf{l}}, \hat{u}^{\mathsf{f}}) \leq 0 \ (\forall j = 1, \ldots, p^{\mathsf{f}}) \right\}$$

$$\Updownarrow$$

$$\Theta^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}}) - V^{\mathsf{f}}(u^{\mathsf{l}}) \leq 0 \quad \text{and} \quad g_j^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}}) \leq 0 \quad (\forall j = 1, \ldots, p^{\mathsf{f}}).$$

Reformulation technique I: Using value functions

Thus the original bilevel game reduces to

$$\begin{cases} \min\limits_{u^{\mathrm{l}}, u^{\mathrm{F}}} & \Theta^{\mathrm{l}}(u^{\mathrm{l}}, u^{\mathrm{F}}) \\ s.t. & u^{\mathrm{l}} \in Y^{\mathrm{l}} \subset U^{\mathrm{l}} \\ & \forall \mathrm{f} \in \mathrm{F}: \\ & \quad \left| \begin{array}{l} \Theta^{\mathrm{f}}(u^{\mathrm{l}}, u^{\mathrm{f}}) - V^{\mathrm{f}}(u^{\mathrm{l}}) \leq 0 \\ g_j^{\mathrm{f}}(u^{\mathrm{l}}, u^{\mathrm{f}}) \leq 0 \qquad (\forall j = 1, \ldots, p^{\mathrm{f}}). \end{array} \right. \end{cases}$$

Pros.

- **Unconditionally equivalent** to the original problem.
- **Theoretically tractable** using nonsmooth analysis. (Clarke or limiting subdifferentials, calmness, etc.)

Cons.

- A value function is **not computable** in practice.
- **BCQ always fails**, which means the problem is **unstable**.[a]

[a] **C, Dutta, Thirumulanathan**, 2025.

Reformulation technique II: KKT reformulations

Now recall that the KKT conditions for

$$u^{\mathtt{f}} \in \arg\min_{\hat{u}^{\mathtt{f}}} \left\{ \Theta^{\mathtt{f}}(u^{\mathtt{l}}, \hat{u}^{\mathtt{f}}) \mid g_j^{\mathtt{f}}(u^{\mathtt{l}}, \hat{u}^{\mathtt{f}}) \leq 0 \ (\forall j = 1, \ldots, p^{\mathtt{f}}) \right\} \qquad \text{(Optimality)}$$

reads

$$(u^{\mathtt{f}}, \lambda^{\mathtt{f}}) \in U^{\mathtt{f}} \times \mathbb{R}_+^{p^{\mathtt{f}}} \quad \text{solves} \begin{cases} \nabla_{u^{\mathtt{f}}} \Theta^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}}) + \sum_{j=1}^{p^{\mathtt{f}}} \lambda_j^{\mathtt{f}} \nabla_{u^{\mathtt{f}}} g_j^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}}) = 0 \\ \lambda_j^{\mathtt{f}} g_j^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}}) = 0 \quad (\forall j = 1, \ldots, p^{\mathtt{f}}) \\ g_j^{\mathtt{f}}(u^{\mathtt{l}}, \hat{u}^{\mathtt{f}}) \leq 0 \ (\forall j = 1, \ldots, p^{\mathtt{f}}) \end{cases} \qquad \text{(Stationarity)}$$

Note that $u^{\mathtt{f}}$ is optimal implies an existence of $\lambda^{\mathtt{f}} \geq 0$ such that $(u^{\mathtt{f}}, \lambda^{\mathtt{f}})$ is stationary, if a **CQ** holds.

The converse holds under convexity assumptions.

At the least, we require that each follower's problem is convex with Slater's CQ for each $u^{\mathtt{l}} \in Y^{\mathtt{l}}$.

Reformulation technique II: KKT reformulations

The bilevel game is reformulated into the following **mathematical program with conplementarity constraints** (MPCC):

$$
\begin{cases}
\min_{u^{\mathrm{l}}, u^{\mathrm{F}}, \lambda^{\mathrm{F}}} & \Theta^{\mathrm{l}}(u^{\mathrm{l}}, u^{\mathrm{F}}) \\
s.t. & u^{\mathrm{l}} \in Y^{\mathrm{l}} \subset U^{\mathrm{l}} \\
 & \forall \mathrm{f} \in \mathrm{F} : \\
 & \quad \left| \begin{array}{l} \nabla_{u^{\mathrm{f}}} \Theta^{\mathrm{f}}(u^{\mathrm{l}}, u^{\mathrm{f}}) + \sum_{j=1}^{p^{\mathrm{f}}} \lambda_j^{\mathrm{f}} \nabla_{u^{\mathrm{f}}} g_j^{\mathrm{f}}(u^{\mathrm{l}}, u^{\mathrm{f}}) = 0 \\ \lambda_j^{\mathrm{f}} g_j^{\mathrm{f}}(u^{\mathrm{l}}, u^{\mathrm{f}}) = 0 \quad (\forall j = 1, \ldots, p^{\mathrm{f}}) \\ g_j^{\mathrm{f}}(u^{\mathrm{l}}, \hat{u}^{\mathrm{f}}) \leq 0 \quad (\forall j = 1, \ldots, p^{\mathrm{f}}) \end{array} \right.
\end{cases}
$$

Pros.

- Explicitly defined.
- More numerically tractable, but still tricky.

Cons.

- **Equivalence issues!**

Why only local solutions ?

Let us illustrate typical poor geometric properties of a bilevel game in the following example.

Example 1
Consider a problem with one leader, $\mathtt{l}$, and on follower, $\mathtt{f}$. Let us write $u_{\mathtt{l}} \equiv x$ and $u_{\mathtt{f}} \equiv y$.

The bilevel game under consideration reads

$$\begin{cases} \min_{x,y} & x + y \\ s.t. & -1 \leq x \leq 1 \\ & y - 0.25x \leq 0.5 \\ & y \in \mathcal{S}(x) := \arg\min_{\hat{y}} \left\{ -\hat{y} \;\middle|\; \begin{array}{l} 0 \leq \hat{y} \leq 1 \\ \hat{y} - x \leq 1 \\ x + \hat{y} \leq 1 \end{array} \right\}. \end{cases}$$

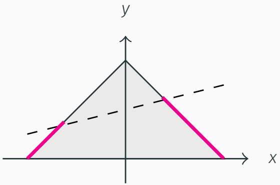


Figure 1:
The disconnected feasible region.

The solution is at $(\bar{x}, \bar{y}) = (-1, 0)$.

The poor geometry = poor algorithms

$$
\begin{aligned}
\text{Nonconvex problem} = \text{No global optimality condition} \\
&\implies \text{Speaking of global optima does not make sense.} \\
&\implies \text{Algorithms would arrive at a local solution (if at all).}
\end{aligned}
$$

What happens at the solutions ?

Bilevel game vs MPCC reformulation

When each follower's problem is convex with Slater CQ, one may equivalently replace the followers' problems with their KKT conditions.

Even then, the original bilevel game and its MPCC reformulation are not equivalent.

In this final section, we **present counterexamples** to illustrate the non-equivalence and also **provide conditions under which one could obtain implications between them**.

One shall see that the issue occurs due to the extra dimensions of the multipliers.

Neccessity of the MPCC reformulation

First, we show that the MPCC reformulation is a relaxation of the original bilevel game.

Theorem 2

Let $(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathsf{F}})$ ***be a local solution of the bilevel game*** *and each follower* $\mathsf{f} \in \mathsf{F}$*, the maps* $u^{\mathsf{f}} \mapsto \Theta^{\mathsf{f}}(u^{\mathtt{l}}, u^{\mathsf{f}}, u^{-\mathsf{f}})$ ***and*** $u^{\mathsf{f}} \mapsto g_j^{\mathsf{f}}(u^{\mathtt{l}}, u^{\mathsf{f}})$ ***are convex with Slater's CQ satisfied*** *at* $u^{\mathtt{l}} = \bar{u}^{\mathtt{l}}$.

Then ***there exist*** $\bar{\lambda}^{\mathsf{F}} \geq 0$ ***such that*** $(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathsf{F}}, \bar{\lambda}^{\mathsf{F}})$ ***is a local solution of the MPCC reformulation***.

Remark
We have a counterexample showing that Slater's CQ is required otherwise the necessary condition fails. (Kindly refer to the full paper.)

Insufficiency of the MPCC reformulation

We argue that the MPCC reformulation is a **pure relaxation** in the following example.

Example Consider again the two-follower setting with the same notation, and with $U^{\mathtt{l}} = U^{\mathtt{f}_i} = \mathbb{R}$ for $i = 1, 2.$ The follower $\mathtt{f}_i$'s problem is

$$\begin{cases} \min_{y_i} & -y_1 - y_2 \\ \text{s.t.} & x + y_i \leq 1 \\ & y_i - x \leq 1 \end{cases}$$

Solving for the primal-dual optimal solution, we obtain

$$\bar{y}_i = \begin{cases} 1 + x & \text{if } x \leq 0 \\ 1 - x & \text{if } x > 0 \end{cases} \qquad \bar{\lambda}_i = \begin{cases} (0, 1) & \text{if } x < 0 \\ (1, 0) & \text{if } x > 0 \\ \text{conv}\{(0, 1), (1, 0)\} & \text{if } x = 0. \end{cases}$$

Insufficiency of the MPCC reformulation

Example (cont.) Suppose that the leader's problem reads

$$\begin{cases} \min & (x-1)^2 + (y_1-1)^2 + (y_2-1)^2 \\ \text{s.t.} & (y_1, y_2) \in \mathcal{S}(x) \end{cases}$$

This bilevel game has a **unique global solution** $(\bar{x}, \bar{y}_1, \bar{y}_2) = (\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$ and **no other local solution**.

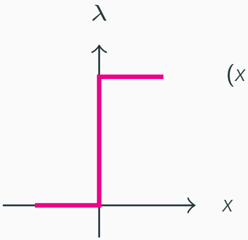
If the followers' problems are replaced with their KKT conditions, the resulting MPCC has another local solution $(\tilde{x}, \tilde{y}_1, \tilde{y}_2, \tilde{\lambda}_1, \tilde{\lambda}_2) = (0, 1, 1, (0,1), (0,1))$.

The reason for this failure is: **there exists a multiplier $((1,0),(1,0))$ at which $(0,1,1)$ is not locally optimal.**

Insufficiency is the burden of the hidden dimension!

Let us illustrate the source of the problem graphically.

The top view does not show it all.



$$(x, \lambda) \in \mathbb{R} \times \mathbb{R}_+$$

$$\text{with } \lambda \in \begin{cases} \{0\}, & x < 0 \\ \{1\}, & x > 0 \\ [0, 1], & x = 0 \end{cases}$$

We need to see the hidden dimension too!

Theorem 3

Suppose for all $\mathsf{f} \in \mathsf{F}$ *that* ***the maps*** $u^{\mathsf{f}} \mapsto \Theta^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}}, u^{-\mathsf{f}})$ ***and*** $u^{\mathsf{f}} \mapsto g_j^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}})$ ***are convex with Slater's CQ*** *satisfied at* $u^{\mathsf{l}} = \bar{u}^{\mathsf{l}}$.

Let $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}}) \in U^{\mathsf{l}} \times U^{\mathsf{F}}$ *and suppose that the vector* $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}}, \bar{\lambda}^{\mathsf{F}})$ ***is a local solution of the MPCC reformulation for any*** $\bar{\lambda}^{\mathsf{F}} \in \Lambda^{\mathsf{F}}(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}})$**.**

Then $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}})$ ***is a local solution of the original bilevel game****.*

The sufficiency follows from the assumption that the optaimality at every mulitpliers.
This is difficult to verify!

A more practical condition under CRCQ.

Definition 4
The follower's problem for the follower $\mathtt{f} \in \mathtt{F}$ satisfies the **constant rank constraint qualification** (briefly, **CRCQ**), at $(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathtt{f}})$ if there exists a neighborhood $V^{\mathtt{f}}$ of $(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathtt{f}})$ such that for each active index subset

$$I^{\mathtt{f}} \subset \mathcal{J}^{\mathtt{f}}(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathtt{f}}) := \left\{ j \mid g_j^{\mathtt{f}}(\bar{u}^{\mathtt{l}}, \bar{u}^{\mathtt{f}}) = 0 \right\},$$

the rank of the matrix $(\nabla_{u^{\mathtt{f}}} g_j^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}}))_{j \in I^{\mathtt{f}}}$ is a constant for $(u^{\mathtt{l}}, u^{\mathtt{f}}) \in V^{\mathtt{f}}$, where the columns of $(\nabla_{u^{\mathtt{f}}} g_j^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}}))_{j \in I^{\mathtt{f}}}$ are given by the vectors $\nabla_{u^{\mathtt{f}}} g_j^{\mathtt{f}}(u^{\mathtt{l}}, u^{\mathtt{f}})$'s in the increasing order of indices in the set $I^{\mathtt{f}}$.

A more practical condition under CRCQ.

Theorem 5
Suppose for all $\mathsf{f} \in \mathsf{F}$ *that **the maps*** $u^{\mathsf{f}} \mapsto \Theta^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}}, u^{-\mathsf{f}})$ ***and*** $u^{\mathsf{f}} \mapsto g^{\mathsf{f}}(u^{\mathsf{l}}, u^{\mathsf{f}})$ ***are convex with Slater's CQ*** *at* $u^{\mathsf{l}} = \bar{u}^{\mathsf{l}}$.

Let $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}}, \bar{\lambda}^{\mathsf{F}})$ ***be a local minimum of the MPCC reformulation for all vertex points*** $\bar{\lambda}^{\mathsf{F}}$ ***of*** $\Lambda^{\mathsf{F}}(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}})$ *and that the follower's problem satisfies the* ***CRCQ*** *at* $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{f}})$ *for all* $\mathsf{f} \in \mathsf{F}$.

Then $(\bar{u}^{\mathsf{l}}, \bar{u}^{\mathsf{F}})$ ***is a local minimum of the original bilevel game.***

Implications on numerical methods.

- Since the original bilevel games are not adapted with any numerical approach, **any algorithm has to be applied to the MPCC reformulation**.
- Since a bilevel game (and its MPCC reformulation) is typically nonconvex, **any local algorithm**[1] **stops at a local solution of the MPCC problem**.
- Since a local solution of the original bilevel game and its MPCC reformulation are different, **the solution of the original problem is not guaranteed by any local algorithm**.

What works ?

- The **SOS-1 method**[2] branches on the complementarity conditions, and it is guaranteed to stop at a global solution (under mild assumptions). A heavy trade-off is the **exponential complexity** it has.

[1] gradient-based, ADMM, augmented Lagrangian, penalization, etc.
[2] See **Sanguansuttigul, Chayawatto, C** (2024) and **Khe, C, Bangviwat** (2025) for the real implementation in Smart grid applications.

Further results

Further results

For the **bilevel games with shared constraints**, additional consideration is required. We refer the interested audience to our full paper.

Full paper

Advertisement

`optdefi` : A LaTeX package that helps typesetting optimization problems.

Acknowledgement

- The hardworking committees and organizers of AMC2025 and at CMU.
- Center of Excellence in Theoretical and Computational Science (TaCS-CoE) at KMUTT.

☻ Thank you for your attendance.

Mer de Glace, 2025
parin.cha@kmutt.ac.th