Multilevel Optimization: Basics, an Application to the European Gas Market, and an Open Research Problem

Yasmine Beck, Daniel Bienstock, Holger Heitsch, René Henrion,
Thomas Kleinert, **Martin Schmidt**, Johannes Thürauf

@schmaidt

June 23, 2022 — SIAM UKIE National Student Chapter Conference

The Team

Smart People

The Mascot

Overview

What is Bilevel Optimization Anyway?

A Real-World Application:
The European Gas Market
with Chance Constraints

An Open Problem:
Continuous & Nonconvex Lower Levels

What is Bilevel Optimization Anyway?

Bilevel Optimization in a Nutshell

"Usual" optimization models

- single decision maker
- one set of variables and constraints
- one objective function

Bilevel Optimization in a Nutshell

"Usual" optimization models

- single decision maker
- one set of variables and constraints
- one objective function

Bilevel optimization

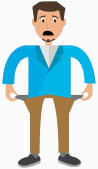
- two decision makers
- both interact in a hierarchical way

Hierarchical Decision Making

Leader: Alice x

decides first

anticipates follower (Bob)



Follower: Bob y

decides second (of course)

A Bit More Formal

Upper-level problem

$$\begin{aligned} \text{“}\min_x\text{”} \quad & F(x, y) \\ \text{s.t.} \quad & G(x, y) \geq 0 \end{aligned}$$

A Bit More Formal

Upper-level problem

$$\begin{aligned} \text{``}\min_{x}\text{''} \quad & F(x,y) \\ \text{s.t.} \quad & G(x,y) \geq 0, \quad y \in \mathcal{S}(x) \end{aligned}$$

Upper-level problem

$$\begin{aligned} \text{“}\min_{x}\text{”} \quad & F(x, y) \\ \text{s.t.} \quad & G(x, y) \geq 0, \quad y \in \mathcal{S}(x) \end{aligned}$$

Lower-level problem

$$\begin{aligned} \min_{y} \quad & f(x, y) \\ \text{s.t.} \quad & g(x, y) \geq 0 \end{aligned}$$

A Bit More Formal

Upper-level problem

$$\begin{aligned} \text{“}\min_{x}\text{”} \quad & F(x, y) \\ \text{s.t.} \quad & G(x, y) \geq 0, \quad y \in \mathcal{S}(x) \end{aligned}$$

Lower-level problem

$$\begin{aligned} \min_{y} \quad & f(x, y) \\ \text{s.t.} \quad & g(x, y) \geq 0 \end{aligned}$$

Different solution concepts: optimistic vs. pessimistic

The Linear-Linear Case

$$\min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \quad c^\top x + d^\top y \quad \text{s.t.} \quad Ax + By \geq a, \quad y \in \mathcal{S}(x)$$

The Linear-Linear Case

$$\min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \quad c^\top x + d^\top y \quad \text{s.t.} \quad Ax + By \geq a, \quad y \in \mathcal{S}(x)$$

$\mathcal{S}(x)$ denotes the set of optimal solutions of the x-parameterized linear problem

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad Dy \geq b - Cx$$

The Linear-Linear Case

$$\min_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \quad c^\top x + d^\top y \quad \text{s.t.} \quad Ax + By \geq a, \quad y \in \mathcal{S}(x)$$

$\mathcal{S}(x)$ denotes the set of optimal solutions of the x-parameterized linear problem

$$\min_y \quad f^\top y \quad \text{s.t.} \quad Dy \geq b - Cx$$

- Strongly NP-hard problem in general (Hansen, Jaumard, Savard 1992)
- Optimistic variant (Dempe 2002)

How to solve these problems: The KKT reformulation

The lower-level problem is an LP:

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad Dy \geq b - Cx$$

How to solve these problems: The KKT reformulation

The lower-level problem is an LP:

$$\min_y \quad f^\top y \quad \text{s.t.} \quad Dy \geq b - Cx$$

The KKT conditions

$$\begin{aligned} Cx + Dy &\geq b \\ \lambda \geq 0,\ D^\top \lambda &= f \\ \lambda^\top (Cx + Dy - b) &= 0 \end{aligned}$$

are both necessary and sufficient

How to solve these problems: The KKT reformulation

The lower-level problem is an LP:

$$\min_y \quad f^\top y \quad \text{s.t.} \quad Dy \geq b - Cx$$

The KKT conditions

$$\begin{aligned} Cx + Dy &\geq b \\ \lambda \geq 0,\ D^\top \lambda &= f \\ \lambda^\top (Cx + Dy - b) &= 0 \end{aligned}$$

are both necessary and sufficient

Single-level reformulation

$$\begin{aligned} \min_{x,y,\lambda} \quad & c^\top x + d^\top y \\ \text{s.t.} \quad & Ax + By \geq a, \quad Cx + Dy \geq b \\ & \lambda \in \Omega_D := \{\lambda \geq 0 : D^\top \lambda = f\} \\ & \lambda^\top (Cx + Dy - b) = 0 \end{aligned}$$

KKT Reformulation

$$
\begin{aligned}
\min_{x,y,\lambda} \quad & c^\top x + d^\top y \\
\text{s.t.} \quad & Ax + By \geq a, \quad Cx + Dy \geq b \\
& \lambda \in \Omega_D := \{\lambda \geq 0 \colon D^\top \lambda = f\} \\
& \lambda^\top (Cx + Dy - b) = 0
\end{aligned}
$$

KKT Reformulation

$$\begin{aligned} \min_{x,y,\lambda} \quad & c^\top x + d^\top y \\ \text{s.t.} \quad & Ax + By \geq a, \quad Cx + Dy \geq b \\ & \lambda \in \Omega_D := \{\lambda \geq 0 : D^\top \lambda = f\} \\ & \lambda^\top (Cx + Dy - b) = 0 \end{aligned}$$

- Be careful if the dual multipliers are not unique (Dempe, Dutta 2012)
- Otherwise, all is nice ...
- ... except for the nasty KKT complementarity conditions

$$\lambda^\top (Cx + Dy - b) = 0$$

How to deal with KKT complementarity conditions

$$\lambda^\top (Cx + Dy - b) = 0$$

How to deal with KKT complementarity conditions

$$\lambda^\top (Cx + Dy - b) = 0$$

That's a disjunction

$$\lambda_i = 0 \quad \vee \quad (Cx + Dy - b)_i = 0, \quad i \in \{1, \dots, \ell\}$$

Introduce a binary variable and some big-Ms ...

$$Cx + Dy - b \leq M_{\mathrm{P}}(1 - u)$$
$$\lambda \leq M_{\mathrm{D}} u$$
$$u \in \{0, 1\}^\ell$$

Mixed-Integer Linear Reformulation

$$
\begin{aligned}
\min_{x,y,\lambda,u} \quad & c^\top x + d^\top y \\
\text{s.t.} \quad & Ax + By \geq a, \quad Cx + Dy \geq b \\
& \lambda \in \Omega_D := \{\lambda \geq 0 \colon D^\top \lambda = f\} \\
& Cx + Dy - b \leq M_{\mathrm{P}}(1 - u) \\
& \lambda \leq M_{\mathrm{D}} u \\
& u \in \{0, 1\}^\ell
\end{aligned}
$$

Mixed-Integer Linear Reformulation

$$
\begin{aligned}
\min_{x,y,\lambda,u} \quad & c^\top x + d^\top y \\
\text{s.t.} \quad & Ax + By \geq a, \quad Cx + Dy \geq b \\
& \lambda \in \Omega_D := \{\lambda \geq 0 \colon D^\top \lambda = f\} \\
& Cx + Dy - b \leq M_{\mathrm{P}}(1 - u) \\
& \lambda \leq M_{\mathrm{D}} u \\
& u \in \{0, 1\}^\ell
\end{aligned}
$$

But how to choose the nasty big-*M*s?

Pitfall #1: Heuristics do not work!

Solving Linear Bilevel Problems Using Big-Ms: Not All That Glitters Is Gold

Salvador Pineda and Juan Miguel Morales

***Abstract*—The most common procedure to solve a linear bilevel problem in the PES community is, by far, to transform it into an equivalent single-level problem by replacing the lower level with its KKT optimality conditions. Then, the complementarity conditions are reformulated using additional binary variables and large enough constants (big-Ms) to cast the single-level problem as a mixed-integer linear program that can be solved using optimization software. In most cases, such large constants are tuned by trial and error. We show, through a counterexample, that this widely used trial-and-error approach may lead to highly suboptimal solutions. Then, further research is required to properly select big-M values to solve linear bilevel problems.**

***Index Terms*—Bilevel programming, optimality conditions, mathematical program with equilibrium constraints (MPEC).**

in [5]. Dealing with the solution to this variant goes beyond the purposes of this letter and thus, we assume $d_i = 0$. This assumption is common in several applications of linear bilevel programming in the PES technical literature. For example, in long-term planning models formulated as bilevel problems [6], [7], [8], [9], the upper-level problem determines investment decisions to maximize investor's profit, while the lower-level problem yields the dispatch quantities to minimize operating cost. In most cases, upper-level constraints model maximum available capacities to be installed and/or budget limitations, but do not include lower-level dispatch variables.

Since the lower-level optimization problem is linear, it can be replaced with its KKT optimality conditions as follows:

Home > Operations Research > Vol. 68, No. 6 >

Technical Note—There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization

Thomas Kleinert, Martine Labbé, Fr¨ank Plein, Martin Schmidt

Published Online: 30 Jun 2020 | https://doi.org/10.1287/opre.2019.1944

Abstract

One of the most frequently used approaches to solve linear bilevel optimization problems consists in replacing the lower-level problem with its Karush–Kuhn–Tucker (KKT) conditions and by reformulating the KKT complementarity conditions using techniques from mixed-integer linear optimization. The latter step requires to determine some big-*M* constant in order to bound the lower level's dual feasible set such that no bilevel-optimal solution is cut off. In practice, heuristics are often used to find a big-*M* although it is known that these approaches may fail. In this paper, we consider the hardness of two proxies for the above mentioned concept of a bilevel-correct big-*M*. First, we prove that verifying that a given big-*M* does not cut off any feasible vertex of the lower level's dual polyhedron cannot be done in polynomial time unless P = NP. Second, we show that verifying that a given big-*M* does not cut off any optimal point of the lower level's dual problem (for any point in the projection of the high-point relaxation onto the leader's decision space) is as hard as solving the original bilevel problem.

A Real-World Application: The European Gas Market with Chance Constraints

Level 1 TSO announces technical capacities and booking price floors

Level 2 Traders book, i.e., sign mid- to long-term capacity contracts

Level 3 Traders nominate at a day-ahead market

Level 4 TSO cost-optimally transports the given nominations

Level 1 TSO announces technical capacities and booking price floors

Level 2 Traders book, i.e., sign mid- to long-term capacity contracts

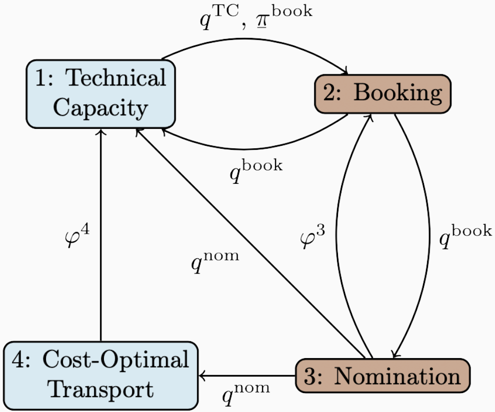
Level 3 Traders nominate at a day-ahead market

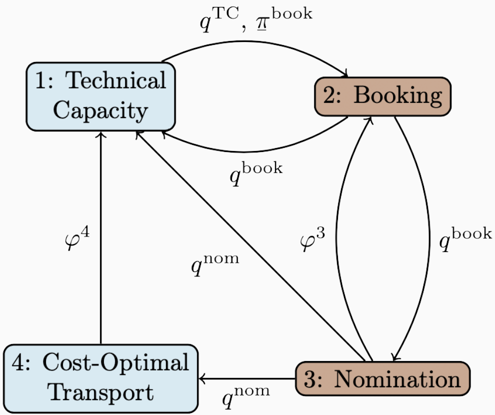
Level 4 TSO cost-optimally transports the given nominations

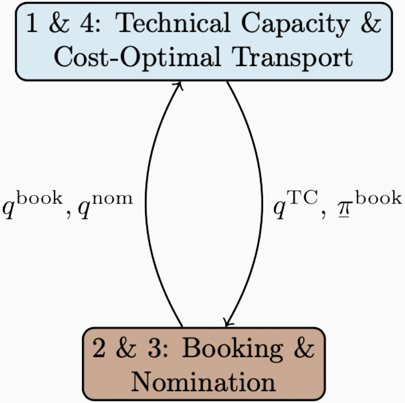
Grimm, Schewe, S., Zöttl (2019)

- Four-level modeling of the European entry-exit gas market
- Identification of assumptions that allow to simplify the model
- Perfect competition $\rightarrow$ reduction to a bilevel model

The European Entry-Exit Gas Market




1: Technical Capacity
2: Booking
3: Nomination
4: Cost-Optimal Transport
q^{TC}, $\underline{\pi}^{\mathrm{book}}$
q^{book}
φ^4
q^{nom}
φ^3
q^{book}
q^{nom}


1 & 4: Technical Capacity & Cost-Optimal Transport
2 & 3: Booking & Nomination
q^{book}, q^{nom}
q^{TC}, $\underline{\pi}^{\mathrm{book}}$

$$
\begin{aligned}
\max_{q^{\mathrm{TC}},\underline{\pi}^{\mathrm{book}},\pi,q} \quad & \varphi^{\mathrm{u}}(q^{\mathrm{nom}},q) = \sum_{t\in T}\left(\sum_{i\in\mathcal{P}_-}\int_0^{q_{i,t}^{\mathrm{nom}}} P_{i,t}(s)\,\mathrm{d}s - \sum_{i\in\mathcal{P}_+} c_i^{\mathrm{var}} q_{i,t}^{\mathrm{nom}}\right) - \sum_{t\in T}\sum_{a\in A} c^{\mathrm{trans}}(q_{a,t}) \\
\text{s.t.} \quad & 0 \le q_u^{\mathrm{TC}},\ 0 \le \underline{\pi}_u^{\mathrm{book}} \quad \text{for all } u \in V_+ \cup V_- \\
& \sum_{u\in V_+\cup V_-}\sum_{i\in\mathcal{P}_u} \underline{\pi}_u^{\mathrm{book}} q_i^{\mathrm{book}} = \sum_{t\in T}\sum_{a\in A} c^{\mathrm{trans}}(q_{a,t}) \\
& (\pi,q) \in \mathcal{F}(q^{\mathrm{nom}}) \\
& (q^{\mathrm{book}},q^{\mathrm{nom}}) \in \operatorname*{arg\,max}\{\text{ lower-level problem }\}
\end{aligned}
$$

$$\begin{aligned}
\max_{q^{\text{book}}, q^{\text{nom}}} \quad & \sum_{t \in T} \left(\sum_{i \in \mathcal{P}_-} \int_0^{q_{i,t}^{\text{nom}}} P_{i,t}(s)\, \mathrm{ds} - \sum_{i \in \mathcal{P}_+} c_i^{\text{var}} q_{i,t}^{\text{nom}} \right) - \sum_{u \in V_+ \cup V_-} \sum_{i \in \mathcal{P}_u} \underline{\pi}_u^{\text{book}} q_i^{\text{book}} \\
\text{s.t.} \quad & \sum_{i \in \mathcal{P}_u} q_i^{\text{book}} \leq q_u^{\text{TC}} \quad \text{for all } u \in V_+ \cup V_- \\
& 0 \leq q_{i,t}^{\text{nom}} \leq q_i^{\text{book}} \quad \text{for all } i \in \mathcal{P}_- \cup \mathcal{P}_+,\ t \in T \\
& \sum_{i \in \mathcal{P}_-} q_{i,t}^{\text{nom}} - \sum_{i \in \mathcal{P}_+} q_{i,t}^{\text{nom}} = 0 \quad \text{for all } t \in T
\end{aligned}$$

- In reality, exit players $i \in \mathcal{P}_-$ nominate quantities $q_{i,t}^{\text{nom}}$ without exactly knowing the actual load $\xi_{i,t}$
- Load vector $\xi = (\xi_{i,t})_{i \in \mathcal{P}_-, t \in T}$ with log-concave cumulative distribution function
- In particular: $\xi \sim \mathcal{N}(m, \Sigma)$
- Modeling assumption: the TSO imposes a fee μ on the exit players $i \in \mathcal{P}_-$ to ensure that the realized loads are covered up to a specified safety level $p \in [0, 1]$
- Joint (over all times and exit players) probabilistic constraint

$$\mathbb{P}\left(\xi_{i,t} \leq q_{i,t}^{\text{nom}} \text{ for all } i \in \mathcal{P}_-,\ t \in T\right) \geq p$$

- Log-concavity of the Gaussian distribution function implies that the log-transformed probabilistic load coverage constraint

$$h(q_-^{\text{nom}}) := \log p - \log \mathbb{P}\left(\xi_{i,t} \leq q_{i,t}^{\text{nom}} \text{ for all } i \in \mathcal{P}_-,\ t \in T\right) \leq 0$$

  is convex

$$
\begin{aligned}
\min_{x,y} \quad & F(x, y) \\
\text{s.t.} \quad & G(x, y) \leq 0 \\
& x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y} \\
& y \in \mathcal{S}(x)
\end{aligned}
$$

Bilevel Optimization

$$\begin{aligned} \min_{x,y} \quad & F(x, y) \\ \text{s.t.} \quad & G(x, y) \leq 0 \\ & x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y} \\ & y \in \mathcal{S}(x) \end{aligned}$$

$\mathcal{S}(x)$ is the solution set of the convex lower-level problem

$$\mathcal{S}(x) = \arg\min_{y} \left\{ f(x, y) \colon g(x, y) \leq 0, \; y \in \mathbb{R}^{n_y} \right\}$$

A "small" extension

$$\mathcal{S}(x) = \arg\min_{y} \left\{ f(x,y) \colon g(x,y) \le 0,\ b(y) \le 0,\ y \in \mathbb{R}^{n_y} \right\}$$

Black-Box Constraint in the Lower Level

A "small" extension

$$\mathcal{S}(x) = \arg\min_{y} \left\{ f(x,y) \colon g(x,y) \leq 0,\; b(y) \leq 0,\; y \in \mathbb{R}^{n_y} \right\}$$

Assumption

The black-box function b is convex and for all $(x,y) \in \{(x,y) \colon G(x,y) \leq 0,\ g(x,y) \leq 0\}$, ...

1. we can evaluate the function $b(y)$,
2. we can evaluate the gradient $\nabla b(y)$,
3. the gradient is bounded, i.e., $\|\nabla b(y)\| \leq K$ for a fixed $K \in \mathbb{R}$.

Some Notation & Single-Level Reformulation

- Shared constraint set

$$\Omega := \{(x, y) \colon G(x, y) \leq 0,\ g(x, y) \leq 0,\ b(y) \leq 0\}$$

Some Notation & Single-Level Reformulation

- Shared constraint set

$$\Omega := \{(x, y)\colon G(x, y) \le 0,\ g(x, y) \le 0,\ b(y) \le 0\}$$

- Projection onto the decision space of the leader

$$\Omega_{\mathrm{u}} := \{x\colon \exists y \text{ with } (x, y) \in \Omega\}$$

Some Notation & Single-Level Reformulation

- Shared constraint set

$$\Omega := \{(x,y) \colon G(x,y) \leq 0,\ g(x,y) \leq 0,\ b(y) \leq 0\}$$

- Projection onto the decision space of the leader

$$\Omega_{\mathrm{u}} := \{x \colon \exists y \text{ with } (x,y) \in \Omega\}$$

- Optimal value function of the lower level

$$\varphi(x) = \min_{y} \{f(x,y) \colon g(x,y),\ b(y) \leq 0,\ y \in \mathbb{R}^{n_y}\}$$

Some Notation & Single-Level Reformulation

- Shared constraint set

$$\Omega := \{(x,y)\colon G(x,y) \leq 0,\ g(x,y) \leq 0,\ b(y) \leq 0\}$$

- Projection onto the decision space of the leader

$$\Omega_{\mathrm{u}} := \{x\colon \exists y \text{ with } (x,y) \in \Omega\}$$

- Optimal value function of the lower level

$$\varphi(x) = \min_{y} \{f(x,y)\colon g(x,y),\ b(y) \leq 0,\ y \in \mathbb{R}^{n_y}\}$$

- Single-level reformulation

$$\begin{aligned} \min_{x,y} \quad & F(x,y) \\ \text{s.t.} \quad & G(x,y) \leq 0, \quad g(x,y) \leq 0, \quad b(y) \leq 0 \\ & f(x,y) \leq \varphi(x) \\ & x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y} \end{aligned}$$

Obstacles and Pitfalls

- Main challenge: black-box constraint $b(y) \leq 0$
- Not given explicitly $\rightarrow$ optimality conditions (KKT) are not given explicitly as well

Obstacles and Pitfalls

- Main challenge: black-box constraint $b(y) \leq 0$
- Not given explicitly $\rightarrow$ optimality conditions (KKT) are not given explicitly as well
- Possible remedies
  - Cutting plane techniques (Kelley 1960)
  - Outer approximation (Duran, Grossmann 1986; Fletcher, Leyffer 1994)

Obstacles and Pitfalls

- Main challenge: black-box constraint $b(y) \leq 0$
- Not given explicitly $\rightarrow$ optimality conditions (KKT) are not given explicitly as well
- Possible remedies
  - Cutting plane techniques (Kelley 1960)
  - Outer approximation (Duran, Grossmann 1986; Fletcher, Leyffer 1994)
- But: $b(y) \leq 0$ can only by satisfied up to a prescribed tolerance
- Specifying the quality of solutions via ε-δ-optimality
  - Global optimization (Locatelli, Schoen 2013)
  - Bilevel optimization (Mitsos, Lemonidis, Barton 2008)

Definition

For $\delta = (\delta_G, \delta_g, \delta_b, \delta_f) \in \mathbb{R}^{m_\text{u}+m_\ell+2}_{\geq 0}$, a point $(\bar{x}, \bar{y}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ is called δ-feasible for the bilevel problem if $G(\bar{x}, \bar{y}) \leq \delta_G$, $g(\bar{x}, \bar{y}) \leq \delta_g$, $b(y) \leq \delta_b$, and $f(x, y) \leq \varphi(x) + \delta_f$ hold. Moreover, for $\varepsilon \geq 0$, a point $(x^*, y^*) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ is called ε-δ-optimal for the bilevel problem, if it is δ-feasible and if $F(x^*, y^*) \leq F^* + \varepsilon$ holds, with F^* denoting the optimal objective function value of the bilevel problem.

Definition
For $\delta = (\delta_G, \delta_g, \delta_b, \delta_f) \in \mathbb{R}^{m_u+m_\ell+2}_{\geq 0}$, a point $(\bar{x}, \bar{y}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ is called δ-feasible for the bilevel problem if $G(\bar{x}, \bar{y}) \leq \delta_G$, $g(\bar{x}, \bar{y}) \leq \delta_g$, $b(y) \leq \delta_b$, and $f(x, y) \leq \varphi(x) + \delta_f$ hold. Moreover, for $\varepsilon \geq 0$, a point $(x^*, y^*) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ is called ε-δ-optimal for the bilevel problem, if it is δ-feasible and if $F(x^*, y^*) \leq F^* + \varepsilon$ holds, with F^* denoting the optimal objective function value of the bilevel problem.

- A δ-feasible point $(\bar{x}, \bar{y})$ is δ_f-(δ_g, δ_b)-optimal for the lower level with fixed $x = \bar{x}$
- Assume f and g pose no challenges $\rightarrow$ choose $\delta_f = \delta_g = 0$
- Assume F and G pose no challenges $\rightarrow$ can we obtain 0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

- Consider the relaxed lower-level problem

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; b(y) \leq \delta_b$$

- Denote the optimal value function by $\underline{\varphi}(x)$
- Relaxation property yields $\underline{\varphi}(x) \leq \varphi(x)$ for all feasible $x \in \Omega_{\text{u}}$

0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

- Consider the relaxed lower-level problem

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; b(y) \leq \delta_b$$

- Denote the optimal value function by $\underline{\varphi}(x)$
- Relaxation property yields $\underline{\varphi}(x) \leq \varphi(x)$ for all feasible $x \in \Omega_{\text{u}}$
- First-relax-then-reformulate leads to a single-level problem with $f(x, y) \leq \underline{\varphi}(x)$
- If $\underline{\varphi}(x) < \varphi(x)$ holds for any $x \in \Omega_{\text{u}}$, this single-level reformulation is not a relaxation of the original single-level reformulation

0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

- Consider the relaxed lower-level problem

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; b(y) \leq \delta_b$$

- Denote the optimal value function by $\underline{\varphi}(x)$
- Relaxation property yields $\underline{\varphi}(x) \leq \varphi(x)$ for all feasible $x \in \Omega_{\text{u}}$
- First-relax-then-reformulate leads to a single-level problem with $f(x, y) \leq \underline{\varphi}(x)$
- If $\underline{\varphi}(x) < \varphi(x)$ holds for any $x \in \Omega_{\text{u}}$, this single-level reformulation is not a relaxation of the original single-level reformulation
- It is not clear whether and how ε-δ-optimality can be guaranteed

0-δ-optimal solutions with $\delta = (0, 0, \delta_b, 0)$?

- Consider the relaxed lower-level problem

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; b(y) \leq \delta_b$$

- Denote the optimal value function by $\underline{\varphi}(x)$
- Relaxation property yields $\underline{\varphi}(x) \leq \varphi(x)$ for all feasible $x \in \Omega_\text{u}$
- First-relax-then-reformulate leads to a single-level problem with $f(x, y) \leq \underline{\varphi}(x)$
- If $\underline{\varphi}(x) < \varphi(x)$ holds for any $x \in \Omega_\text{u}$, this single-level reformulation is not a relaxation of the original single-level reformulation
- It is not clear whether and how ε-δ-optimality can be guaranteed

Can we hope for the δ-feasible points with $\delta = (0, 0, \delta_b, 0)$?

A "First-Relax-Then-Reformulate" Approach

- Block-box constraint $b(y) \geq 0$ is convex
- Construct a sequence of linear outer approximations $(E^r, e^r)_{r \in \mathbb{N}}$ of the black-box constraint $b(y) \leq 0$ with the property

$$\{y \in \mathbb{R}^{n_y} : b(y) \leq 0\} \subseteq \{y \in \mathbb{R}^{n_y} : E^{r+1} y \leq e^{r+1}\} \subseteq \{y \in \mathbb{R}^{n_y} : E^r y \leq e^r\}$$

A "First-Relax-Then-Reformulate" Approach

- Block-box constraint $b(y) \geq 0$ is convex
- Construct a sequence of linear outer approximations $(E^r, e^r)_{r \in \mathbb{N}}$ of the black-box constraint $b(y) \leq 0$ with the property

$$\{y \in \mathbb{R}^{n_y} : b(y) \leq 0\} \subseteq \{y \in \mathbb{R}^{n_y} : E^{r+1} y \leq e^{r+1}\} \subseteq \{y \in \mathbb{R}^{n_y} : E^r y \leq e^r\}$$

- For a given upper-level solution $\bar{x} \in \Omega_\mathrm{u}$ and $r \in \mathbb{N}$, the adapted lower-level problem reads

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; E^r y \leq e^r$$

- This is a relaxation of the original lower-level problem

A "First-Relax-Then-Reformulate" Approach

- Block-box constraint $b(y) \geq 0$ is convex
- Construct a sequence of linear outer approximations $(E^r, e^r)_{r \in \mathbb{N}}$ of the black-box constraint $b(y) \leq 0$ with the property

$$\{y \in \mathbb{R}^{n_y} : b(y) \leq 0\} \subseteq \{y \in \mathbb{R}^{n_y} : E^{r+1} y \leq e^{r+1}\} \subseteq \{y \in \mathbb{R}^{n_y} : E^r y \leq e^r\}$$

- For a given upper-level solution $\bar{x} \in \Omega_{\mathrm{u}}$ and $r \in \mathbb{N}$, the adapted lower-level problem reads

$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; E^r y \leq e^r$$

- This is a relaxation of the original lower-level problem
- $\underline{\varphi}^r(x)$: optimal value function
- Assumption: Slater's constraint qualification holds

A "First-Relax-Then-Reformulate" Approach

- Block-box constraint $b(y) \geq 0$ is convex
- Construct a sequence of linear outer approximations $(E^r, e^r)_{r \in \mathbb{N}}$ of the black-box constraint $b(y) \leq 0$ with the property

$$\{y \in \mathbb{R}^{n_y} : b(y) \leq 0\} \subseteq \{y \in \mathbb{R}^{n_y} : E^{r+1} y \leq e^{r+1}\} \subseteq \{y \in \mathbb{R}^{n_y} : E^r y \leq e^r\}$$

- For a given upper-level solution $\bar{x} \in \Omega_{\mathrm{u}}$ and $r \in \mathbb{N}$, the adapted lower-level problem reads

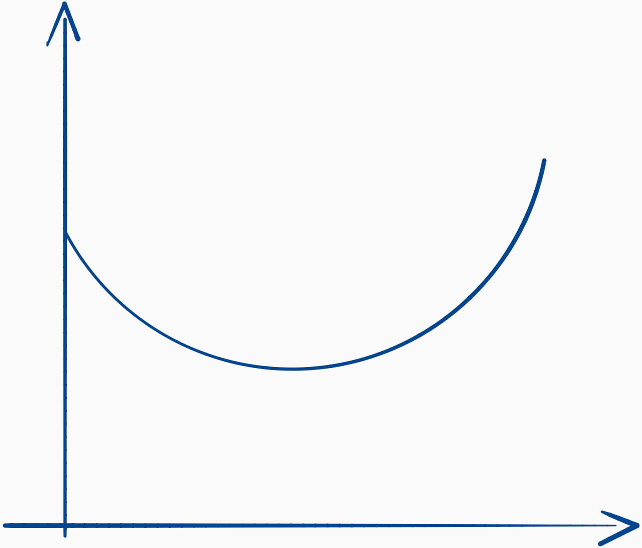
$$\min_{y \in \mathbb{R}^{n_y}} \quad f(\bar{x}, y) \quad \text{s.t.} \quad g(\bar{x}, y) \leq 0, \; E^r y \leq e^r$$

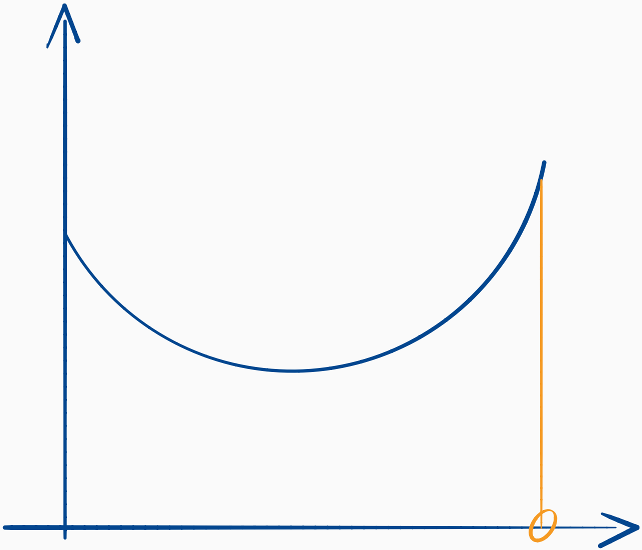
- This is a relaxation of the original lower-level problem
- $\underline{\varphi}^r(x)$: optimal value function
- Assumption: Slater's constraint qualification holds

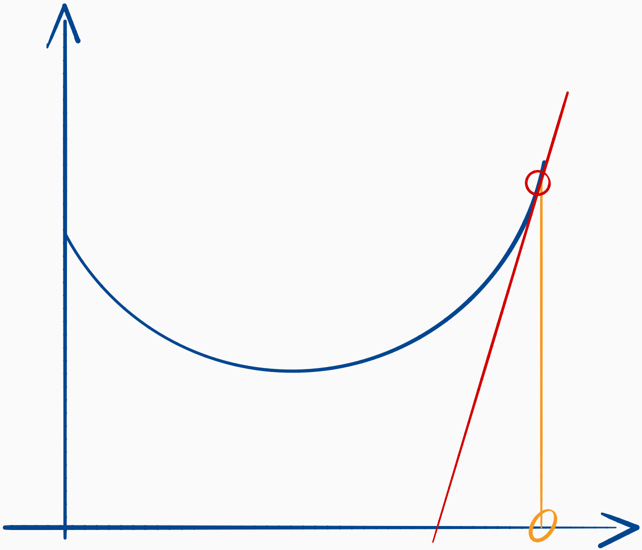
Proposition

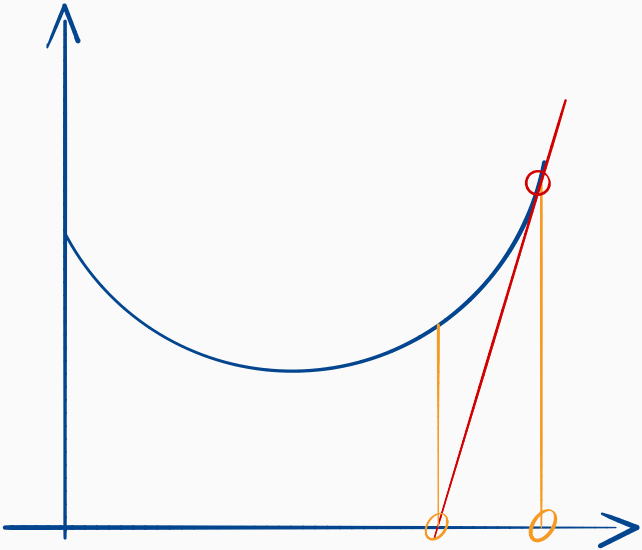
For every $r \in \mathbb{N}$ and every upper-level decision $x \in \Omega_{\mathrm{u}}$, it holds

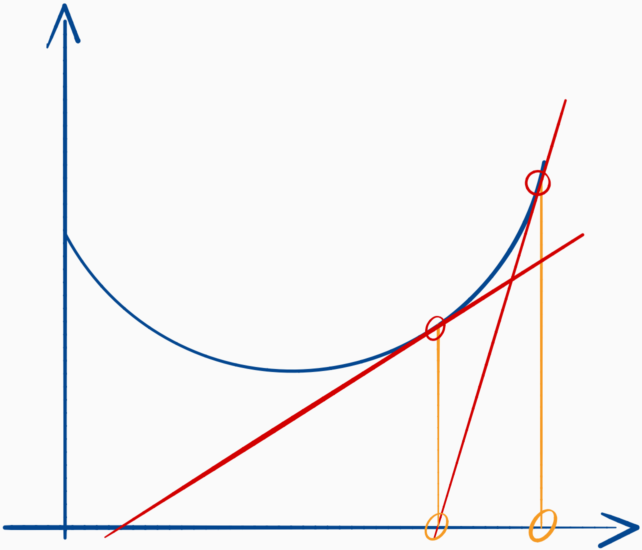
$$\underline{\varphi}^r(x) \leq \underline{\varphi}^{r+1}(x) \leq \varphi(x)$$

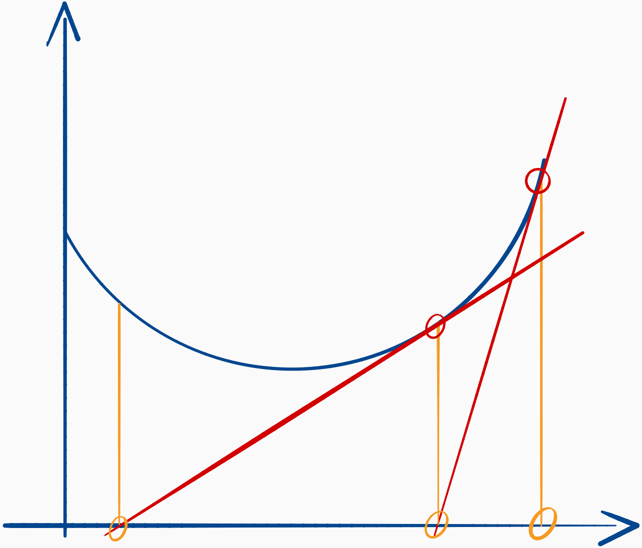


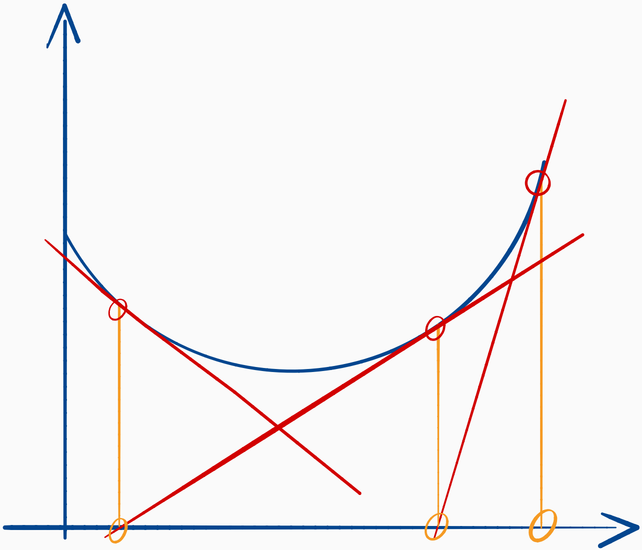


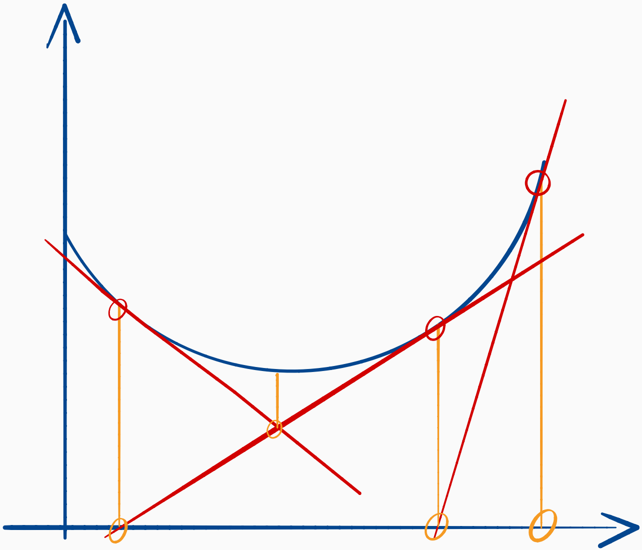


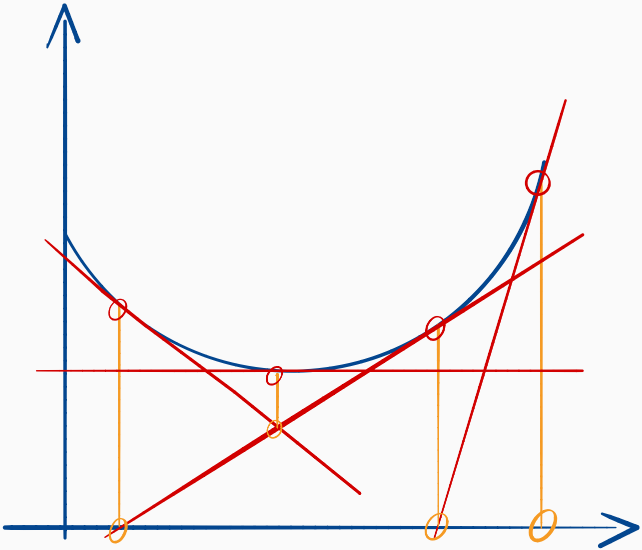


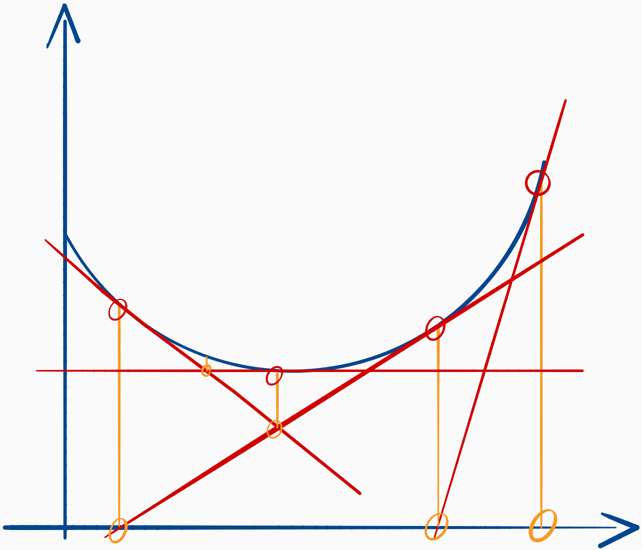


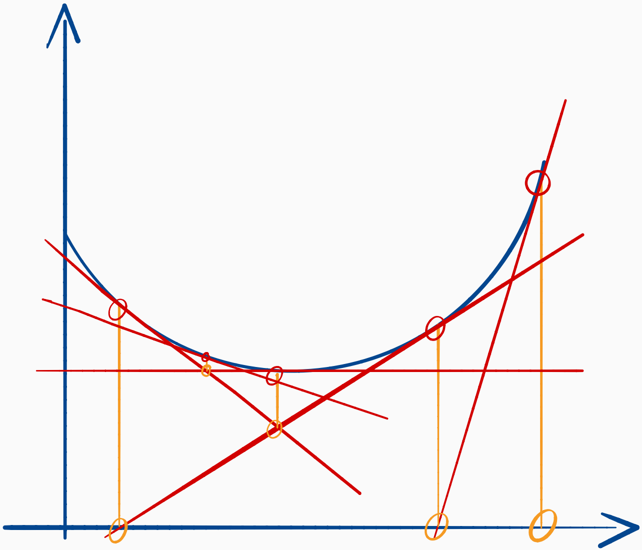


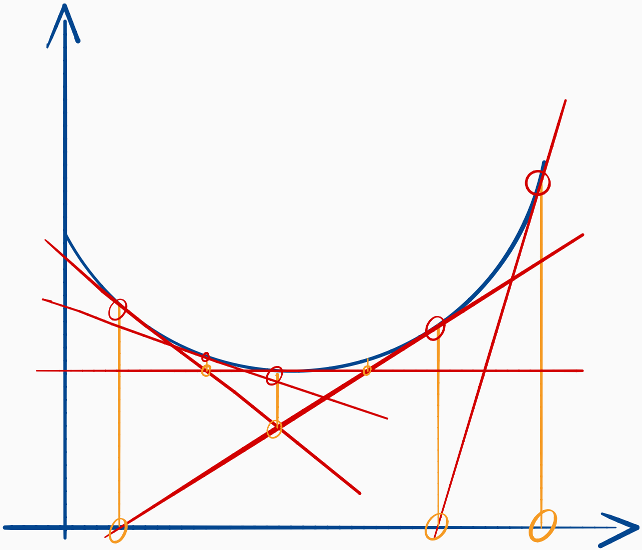


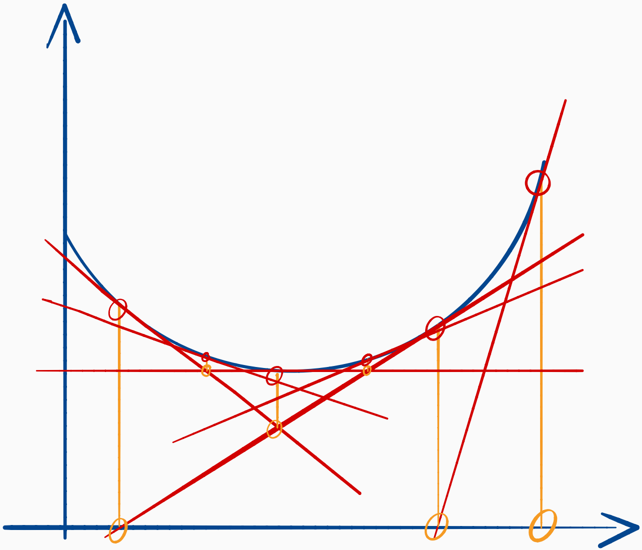


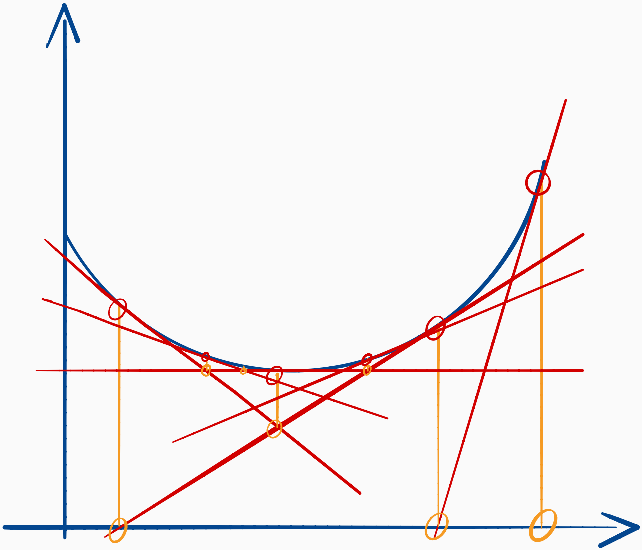


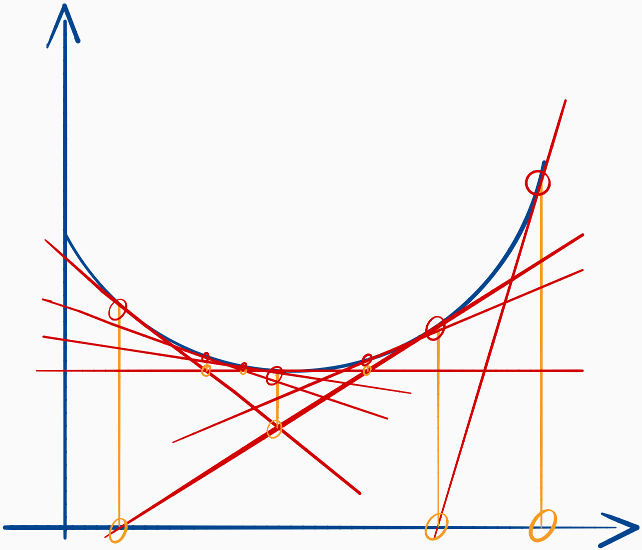












Modified variant of the single-level reformulation

$$
\begin{aligned}
\min_{x,y} \quad & F(x,y) \\
\text{s.t.} \quad & G(x,y) \leq 0, \quad g(x,y) \leq 0 \\
& E^r y \leq e^r \\
& f(x,y) \leq \underline{\varphi}^r(x) \\
& x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y}
\end{aligned}
$$

A "First-Relax-Then-Reformulate" Approach

Modified variant of the single-level reformulation

$$
\begin{aligned}
\min_{x,y} \quad & F(x,y) \\
\text{s.t.} \quad & G(x,y) \leq 0, \quad g(x,y) \leq 0 \\
& E^r y \leq e^r \\
& f(x,y) \leq \underline{\varphi}^r(x) \\
& x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y}
\end{aligned}
$$

Feasibility problem

$$
\begin{aligned}
\min_{x,y,s} \quad & s \\
\text{s.t.} \quad & G(x,y) \leq 0, \quad g(x,y) \leq 0 \\
& E^r y \leq e^r \\
& f(x,y) \leq \underline{\varphi}^r(x) + s \\
& x \in \mathbb{R}^{n_x}, \quad y \in \mathbb{R}^{n_y}
\end{aligned}
$$

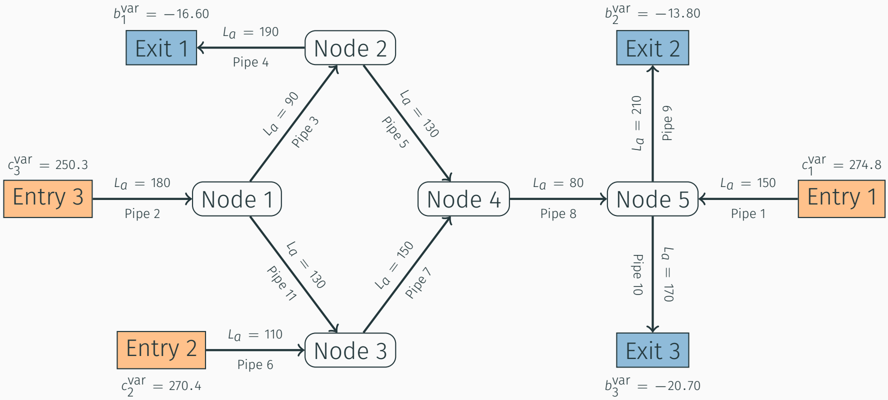
Algorithm "First-Relax-Then-Reformulate".

Choose $\delta_b > 0$, set $r = 0$, $s = 0$, $\chi = \infty$, $E^0 = [0 \dots 0] \in \mathbb{R}^{1\times n_y}$, $e^0 = 0 \in \mathbb{R}$.
while $\chi > \delta_b$ or $s > 0$ **do**
  Construct E^{r+1} and e^{r+1}
  if the modified variant of the single-level reformulation is feasible **then**
    Solve this problem to obtain (x^{r+1}, y^{r+1}) and set $s = 0$.
  else if the feasibility problem is feasible **then**
    Solve this problem to obtain (x^{r+1}, y^{r+1}, s).
  else
    Return "The original problem is infeasible."
  end if
  Set $r \leftarrow r + 1$ and $\chi = b(y^r)$.
end while
Return $(\bar{x}, \bar{y}) = (x^r, y^r)$.

Algorithm "First-Relax-Then-Reformulate".

Choose $\delta_b > 0$, set $r = 0$, $s = 0$, $\chi = \infty$, $E^0 = [0 \dots 0] \in \mathbb{R}^{1 \times n_y}$, $e^0 = 0 \in \mathbb{R}$.
while $\chi > \delta_b$ or $s > 0$ **do**
    Construct E^{r+1} and e^{r+1}
    if the modified variant of the single-level reformulation is feasible **then**
        Solve this problem to obtain (x^{r+1}, y^{r+1}) and set $s = 0$.
    else if the feasibility problem is feasible **then**
        Solve this problem to obtain (x^{r+1}, y^{r+1}, s).
    else
        Return "The original problem is infeasible."
    end if
    Set $r \leftarrow r + 1$ and $\chi = b(y^r)$.
end while
Return $(\bar{x}, \bar{y}) = (x^r, y^r)$.

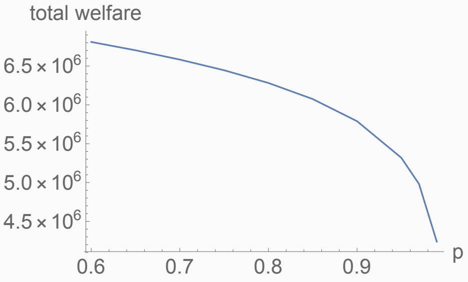
Theorem: If the algorithm terminates, then $(\bar{x}, \bar{y})$ is $(0, 0, \delta_b, 0)$-feasible for original bilevel problem.

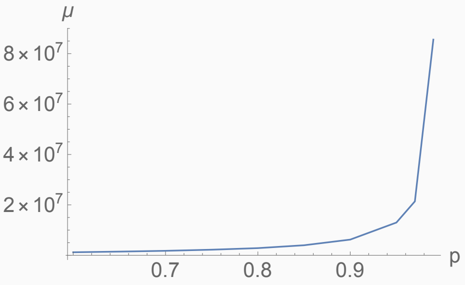

$b_1^{var} = -16.60$
Exit 1
$L_a = 190$
Pipe 4
Node 2
$b_2^{var} = -13.80$
Exit 2
$L_a = 90$
Pipe 3
$L_a = 130$
Pipe 5
$L_a = 210$
Pipe 9
$c_3^{var} = 250.3$
Entry 3
$L_a = 180$
Pipe 2
Node 1
Node 4
$L_a = 80$
Pipe 8
Node 5
$L_a = 150$
Pipe 1
$c_1^{var} = 274.8$
Entry 1
$L_a = 130$
Pipe 11
$L_a = 150$
Pipe 7
Pipe 10
$L_a = 170$
Entry 2
$L_a = 110$
Pipe 6
Node 3
Exit 3
$c_2^{var} = 270.4$
$b_3^{var} = -20.70$

Numerical Results

	Bisection	Bounding		δ-Feasibility		Total		
p	Runtime	#Iter.	Runtime	#Iter.	Runtime	#Iter.	Runtime	Gap
0.60	12.13	32	36.80	10	28.97	42	77.9	0.001
0.65	14.15	28	32.00	16	40.71	44	86.86	0.001
0.70	11.13	26	29.70	13	39.70	39	80.53	0.001
0.75	9.04	25	28.55	6	14.19	31	51.78	0.002
0.80	7.98	25	29.06	4	6.26	29	43.3	0.005
0.85	11.08	21	24.01	3	7.41	24	42.5	0.006
0.90	11.05	23	26.34	8	27.52	31	64.91	0.017
0.95	5.96	24	27.99	6	14.14	30	48.09	0.010
0.96	7.56	22	24.56	3	4.17	25	36.29	0.011
0.97	6.94	21	23.96	4	9.20	25	40.10	0.015
0.98	4.63	25	93.68	9	106.31	34	204.62	0.032
0.99	6.96	26	29.76	10	1250.65	36	1287.37	0.187

Total Welfare and Price of Load Coverage





Open Access | Published: 13 May 2022

On convex lower-level black-box constraints in bilevel optimization with an application to gas market models with chance constraints

Holger Heitsch, René Henrion, Thomas Kleinert & Martin Schmidt

Journal of Global Optimization (2022) | Cite this article

183 Accesses | **1** Altmetric | Metrics

Abstract

Bilevel optimization is an increasingly important tool to model hierarchical decision making. However, the ability of modeling such settings makes bilevel problems hard to solve in theory and practice. In this paper, we add on the general difficulty of this class of problems by further incorporating convex black-box constraints in the lower level. For this setup, we develop a cutting-plane algorithm that computes approximate bilevel-feasible points. We apply this method to a bilevel model of the European gas market in which we use a joint chance constraint to model uncertain loads. Since the chance constraint is not available in closed form, this fits into the black-box setting studied before. For the applied model, we use further problem-specific insights to derive bounds on the objective value of the bilevel problem. By doing so, we are able to show that we solve the application problem to approximate global optimality. In our numerical case study we are thus able to evaluate the welfare sensitivity in dependence of the achieved safety level of uncertain load coverage.

An Open Problem:
Continuous & Nonconvex Lower Levels

Nonconvexities in the Lower Level

Upper-level problem

$$\begin{aligned} \text{``}\min_x\text{''} \quad & F(x, y) \\ \text{s.t.} \quad & G(x, y) \geq 0, \quad y \in \mathcal{S}(x) \end{aligned}$$

Lower-level problem

$$\begin{aligned} \min_y \quad & f(x, y) \\ \text{s.t.} \quad & g(x, y) \geq 0 \end{aligned}$$

Who can solve this problem?

Upper level

$$\begin{aligned}
\max_{x \in \mathbb{R}^2} \quad & F(x, y) = x_1 - 2y_{n+1} + y_{n+2} \\
\text{s.t.} \quad & (x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2] \\
& y \in S(x)
\end{aligned}$$

- $\underline{x}$, $\bar{x} \in \mathbb{R}^2$ with $1 \le \underline{x}_i < \bar{x}_i$, $i \in \{1, 2\}$
- Upper level is an LP
  with simple bound constraints
- Upper level has no coupling constraints

Who can solve this problem?

Upper level

$$\max_{x \in \mathbb{R}^2} \quad F(x,y) = x_1 - 2y_{n+1} + y_{n+2}$$
$$\text{s.t.} \quad (x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$$
$$y \in S(x)$$

Lower level

$$\max_{y \in \mathbb{R}^{n+2}} \quad f(x,y) = y_1 - y_n \left(x_1 + x_2 - y_{n+1} - y_{n+2}\right)$$
$$\text{s.t.} \quad y_1 + y_n = \frac{1}{2}$$
$$y_i^2 \le y_{i+1}, \quad i \in \{1, \dots, n-1\}$$
$$y_i \ge 0, \quad i \in \{1, \dots, n\}$$
$$y_{n+1} \in [0, x_1]$$
$$y_{n+2} \in [-x_2, x_2]$$

- $\underline{x}$, $\bar{x} \in \mathbb{R}^2$ with $1 \le \underline{x}_i < \bar{x}_i$, $i \in \{1, 2\}$
- Upper level is an LP with simple bound constraints
- Upper level has no coupling constraints
- Feasible set of lower level is non-empty and compact for all feasible leader decisions
- Slater's CQ is also satisfied for all feasible leader decisions
- All constraints are linear except for some convex-quadratic inequality constraints
- The coefficients/right-hand sides are either 1 or 1/2
- Bilinear objective function

Exact feasibility

$$\max_{y \in \mathbb{R}^{n+2}} \quad f(x,y) = y_1 - y_n\left(x_1 + x_2 - y_{n+1} - y_{n+2}\right)$$

$$\text{s.t.} \quad y_1 + y_n = \frac{1}{2}$$

$$y_i^2 \leq y_{i+1}, \quad i \in \{1, \dots, n-1\}$$

$$y_i \geq 0, \quad i \in \{1, \dots, n\}$$

$$y_{n+1} \in [0, x_1]$$

$$y_{n+2} \in [-x_2, x_2]$$

Result #1

For every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$, a feasible follower's decision y satisfies $y_n > 0$.

Exact feasibility

$$\begin{aligned} \max_{y \in \mathbb{R}^{n+2}} \quad & f(x,y) = y_1 - y_n\,(x_1 + x_2 - y_{n+1} - y_{n+2}) \\ \text{s.t.} \quad & y_1 + y_n = \frac{1}{2} \\ & y_i^2 \leq y_{i+1}, \quad i \in \{1, \dots, n-1\} \\ & y_i \geq 0, \quad i \in \{1, \dots, n\} \\ & y_{n+1} \in [0, x_1] \\ & y_{n+2} \in [-x_2, x_2] \end{aligned}$$

Result #1

For every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$, a feasible follower's decision y satisfies $y_n > 0$.

Result #2

For every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$, the set of optimal solutions of the lower-level problem is a singleton.

Exact feasibility

$$
\begin{aligned}
\max_{y \in \mathbb{R}^{n+2}} \quad & f(x,y) = y_1 - y_n\left(x_1 + x_2 - y_{n+1} - y_{n+2}\right) \\
\text{s.t.} \quad & y_1 + y_n = \frac{1}{2} \\
& y_i^2 \le y_{i+1}, \quad i \in \{1, \dots, n-1\} \\
& y_i \ge 0, \quad i \in \{1, \dots, n\} \\
& y_{n+1} \in [0, x_1] \\
& y_{n+2} \in [-x_2, x_2]
\end{aligned}
$$

Result #1

For every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$, a feasible follower's decision y satisfies $y_n > 0$.

Result #2

For every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$, the set of optimal solutions of the lower-level problem is a singleton.

Result #3

The bilevel problem has a unique solution given by $x^* = (\underline{x}_1, \bar{x}_2)$ with an optimal objective function value of $F^* = \underline{x}_1 + \bar{x}_2$.

Definition

Let $0 < \varepsilon \in \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$, and $g : \mathbb{R}^n \to \mathbb{R}^m$ be given. A point $x \in \mathbb{R}^n$ is called ε-feasible for the problem $\max_{x \in \mathbb{R}^n} \{f(x) : g(x) \leq 0\}$ if $g_i(x) \leq 0$ holds for all $i \in \{1, \ldots, m\} \setminus N$ and if $\max_{i \in N} g_i(x) \leq \varepsilon$ holds, where $N \subseteq \{1, \ldots, m\}$ denotes the index set of all nonlinear constraints.

Definition

Let $0 < \varepsilon \in \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$, and $g : \mathbb{R}^n \to \mathbb{R}^m$ be given. A point $x \in \mathbb{R}^n$ is called ε-feasible for the problem $\max_{x \in \mathbb{R}^n} \{f(x) \colon g(x) \leq 0\}$ if $g_i(x) \leq 0$ holds for all $i \in \{1, \ldots, m\} \setminus N$ and if $\max_{i \in N} g_i(x) \leq \varepsilon$ holds, where $N \subseteq \{1, \ldots, m\}$ denotes the index set of all nonlinear constraints.

Result #4

Unless $\varepsilon < 2^{-2^{n-1}}$, there is an ε-feasible follower's decision y with $y_n = 0$ for every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$.

Definition

Let $0 < \varepsilon \in \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$, and $g : \mathbb{R}^n \to \mathbb{R}^m$ be given. A point $x \in \mathbb{R}^n$ is called ε-feasible for the problem $\max_{x \in \mathbb{R}^n} \{f(x) \colon g(x) \leq 0\}$ if $g_i(x) \leq 0$ holds for all $i \in \{1, \ldots, m\} \setminus N$ and if $\max_{i \in N} g_i(x) \leq \varepsilon$ holds, where $N \subseteq \{1, \ldots, m\}$ denotes the index set of all nonlinear constraints.

Result #4

Unless $\varepsilon < 2^{-2^{n-1}}$, there is an ε-feasible follower's decision y with $y_n = 0$ for every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$.

Result #5

Unless $\varepsilon < 2^{-2^{n-1}}$, the set of ε-feasible follower's solutions is not a singleton for every feasible leader's decision $(x_1, x_2) \in [\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$.

Result #3 (revisited)

The bilevel problem has a unique solution given by $x^* = (\underline{x}_1, \bar{x}_2)$ with an optimal objective function value of $F^* = \underline{x}_1 + \bar{x}_2$.

Result #3 (revisited)

The bilevel problem has a unique solution given by $x^* = (\underline{x}_1, \bar{x}_2)$ with an optimal objective function value of $F^* = \underline{x}_1 + \bar{x}_2$.

Result #6

Let $\varepsilon \geq 2^{-2^{n-1}}$ and suppose that we allow for ε-feasible follower's solutions.

Then, the optimistic optimal solution of the bilevel problem is given by $x^*_{\text{o}} = (\bar{x}_1, \bar{x}_2)$ with an optimal objective function value of $F^*_{\text{o}} = \bar{x}_1 + \bar{x}_2$.

The pessimistic optimal solution is given by $x^*_{\text{p}} = (\underline{x}_1, \underline{x}_2)$ with an optimal objective function value of $F^*_{\text{p}} = -\underline{x}_1 - \underline{x}_2$.

Result #3 (revisited)

The bilevel problem has a unique solution given by $x^* = (\underline{x}_1, \bar{x}_2)$ with an optimal objective function value of $F^* = \underline{x}_1 + \bar{x}_2$.

Result #6

Let $\varepsilon \geq 2^{-2^{n-1}}$ and suppose that we allow for ε-feasible follower's solutions.

Then, the optimistic optimal solution of the bilevel problem is given by $x_o^* = (\bar{x}_1, \bar{x}_2)$ with an optimal objective function value of $F_o^* = \bar{x}_1 + \bar{x}_2$.

The pessimistic optimal solution is given by $x_p^* = (\underline{x}_1, \underline{x}_2)$ with an optimal objective function value of $F_p^* = -\underline{x}_1 - \underline{x}_2$.

- By the way: $n \geq \log_2(\log_2(1/\varepsilon^2))$
- For $\varepsilon = 10^{-8}$, the problem gets unsolvable for $n = 6$

Well ... and now?

Is this an impossibility result
for computationally solving bilevel problems
with continuous and nonconvex lower-level problems?

Some Related Literature

- A Survey on Mixed-Integer Programming Techniques in Bilevel Optimization
  Thomas Kleinert, Martine Labbé, Ivana Ljubic, Martin Schmidt
  EURO Journal on Computational Optimization
- On Convex Lower-Level Black-Box Constraints in Bilevel Optimization with an Application to Gas Market Models with Chance Constraints
  Holger Heitsch, René Henrion, Thomas Kleinert, Martin Schmidt
  Journal of Global Optimization
- On a Computationally Ill-Behaved Bilevel Problem with a Continuous and Nonconvex Lower Level
  Yasmine Beck, Daniel Bienstock, Johannes Thürauf, Martin Schmidt