A max-min reformulation approach to nonconvex bilevel optimization

Rongzhu Ke[*] Christopher Thomas Ryan[†] Jin Zhang[‡]

November 17, 2023

Abstract

We use a max-min reformulation approach to derive necessary optimality conditions for nonconvex bilevel programs. The existing literature has concentrated on hybrids of two basic approaches: (a) the KKT approach and (b) the value function approach that posits calmness conditions and employs nonsmooth analysis. Both methods reformulate the problem into a single-level problem, possibly with complementarity or nonsmooth constraints. We explore an alternative approach based on a max-min reformulation of the problem. This produces a parsimonious optimality condition that involves a single alternative best response of the follower rather than an enumeration of best responses, which is common in other methods. We provide examples where our optimality conditions hold but fail the constraint qualifications of other approaches.

Key Words: Bilevel optimization, optimality conditions, penalty function

1 Introduction

We explore bilevel programming problems of the form

$$\begin{aligned} \max_{x\in X,y} \quad & F(x,y) \\ \text{subject to} \quad & y \in S(x) \\ & G(x,y) \geq 0 \end{aligned} \tag{BLPP}$$

where X is a compact subset of $\mathbb{R}^n$ and $S(x)$ denotes the set of optimal solutions to the lower-level problem (LLP)

$$\begin{aligned} \max_{y\in Y} \quad & f(x,y) \\ \text{subject to} \quad & g(x,y) \geq 0 \end{aligned} \tag{LPP}$$

with $F, f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$, and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ where n, m, q and p are nonnegative integers. All functions are continuously differentiable, and f is twice continuously

[*]Department of Economics, Zhejiang University, Hangzhou, Zhejiang, China. E-mail: rzke455@zju.edu.cn

[†]Sauder School of Business, University of British Columbia, Vancouver, BC, Canada. E-mail: chris.ryan@sauder.ubc.ca

[‡]Corresponding author: Department of Mathematics and SUSTech International Center for Mathematics, Southern University of Science and Technology, and National Center for Applied Mathematics Shenzhen, Shenzhen, Guangdong, China. E-mail: zhangj9@sustech.edu.cn.

differentiable. There are no convexity requirements and, in particular, the lower-level problem need not be a convex optimization problem. We assume that Y is a compact set and that (BLPP) is feasible and possesses an optimal solution (x^*, y^*) that is interior to $X \times Y$. Moreover, we will assume that $S(x^*)$ lies in the interior of Y. The over-arching goal of the paper is to provide optimality conditions that are satisfied by (x^*, y^*).[1]

Bilevel programs are challenging. Their study inherently involves parametric, nonconvex and nonsmooth optimization at its core. These are natural implications of the properties of the set $S(x)$ as a mapping of x. However, understanding bilevel optimization is worthwhile because of its numerous and far-reaching applications (for surveys of applications, see Colson [2] and Dempe [3]).

Considerable effort has been spent in deriving optimality conditions for (BLPP). Naive applications of standard single-level theory are well-documented to fail in general (see details in Dempe and Zemkoho [5] and Ye and Zhu [13]). There are two broad approaches to deriving successful conditions. The first is classical and assumes that the lower-level problem is a convex optimization problem and hence replaceable by Karush-Kuhn-Tucker (KKT) conditions as constraints to the leader's problem (for a modern adaptation of the approach, see, for instance, Dempe et. al. [4]). However, as demonstrated in [12], when the lower-level problem is not convex the presence of nonoptimal stationary solutions to the lower-level problem implies that the resulting necessary optimality conditions do not characterize optimal solutions.

In response to these limitations, Ye and Zhu proposed an alternative approach that applied to nonconvex bilevel programs in Ye and Zhu [13]. Their approach involves using the value function of the lower-level problem to create a single-level optimization problem. They give constraint qualifications for when Fritz-John (FJ) and KKT-type necessary optimality conditions hold. One notable condition is *partial calmness* which allows the value function to be handled in the objective of the resulting single-level problem rather than in the constraints. This yields clean optimality conditions that apply to a variety of cases. Later in [14], Ye and Zhu leverage both the KKT and value function approaches to yield new constraint qualifications and optimality conditions that apply even more broadly. Of note are *weak calmness* constraint qualifications based on the linearization cone of the bilevel problem with an additional constraint corresponding to the Lagrangian of the lower-level problem. Sufficient conditions to establish calmness are also provided. After these two seminal papers by Ye and Zhu, other constraint qualifications built on both the KKT and value function approaches were discovered. Of note in this direction are papers by Dempe and co-authors (for instance, Dempe et. al. [4], Dempe and Zemkoho [5] and Dempe and Zemkoho [6]).

The KKT and value function approaches convert (BLPP) to a single-level optimization problem. The resulting single-level optimization problems have additional complexity beyond a standard nonconvex optimization problem. In the KKT approach, complementarity constraints are included. In the value function approach, a nonsmooth function enters the constraints. Known results on complementarity and nonsmooth optimization problems are adapted to the bilevel setting to derive optimality conditions.

Our point of departure is that we analyze a new single-level reformulation of (BLPP). To do so, we draw inspiration from the classic paper of Mirrlees [12]. In this technique, the follower's optimization problem is replaced with infinitely many *incentive compatibility* constraints of the form $f(x, y) - f(x, z) \geq 0$ for all $z \in Y$ to encode that (x, y) must be chosen with y as a best response to x. The derivation of an inner minimization in our max-min formulation derives from the fact that the optimization problem $\min_{z \in Y} f(x, y) - f(x, z)$ must have an optimal value greater than or equal to zero. We then employ an exact penalization method to the resulting *min-max*

[1] For clarity of exposition, we do not take up the issue of corner solutions. The theory presented here can be extended in a relatively straightforward manner to this setting using familiar techniques. Problems with equality constraints are easily adapted to our analysis, but we assume none are present for brevity.

problem that arises from relaxing the max-min problem to derive necessary optimality conditions for the equivalent bilevel problem.

Penalty functions are an established tool to study bilevel problems. Most papers use them to design algorithms to solve bilevel programs numerically and focus on issues of exactness and convergence (see, for example, Ishizuka and Aiyoshi [8] and Meng et. al. [11]). Liu et. al. [9] derive optimality conditions for convex bilevel programs using penalty functions. Marcotte and Zhu [10] study generalized bilevel programs using penalty functions and derive optimality conditions under certain convexity conditions. By contrast, our focus is nonconvex problems. Moreover, Liu et. al. [9] penalize the KKT reformulation where we penalize our max-min reformulation.

1.1 Motivating example

The following example provides an instance where the usual constraint qualifications in existing literature (namely the work of Ye and Zhu in [14]) fail, but we will show later that it is amenable to the approach we develop in this paper.

Example 1. Consider the following bilevel program:

$$\begin{aligned} \max_{x,y} \quad & -(x_1-\tfrac{6}{5})^2+\tfrac{1}{2}(x_2-\tfrac{y}{2})^2-(y-\tfrac{1}{4})^2-\tfrac{1}{2}x_2 \\ \text{subject to} \quad & \tfrac{\sqrt{2}-1}{2}x_1-y+1/2\geq 0 \\ & y\in \arg\max_{y\in Y} -(x_1-1)(x_1-y)y-(y^2-1)^2(y^2-\tfrac{1}{2})^2. \end{aligned}$$

where $Y=[-1.5,1.5]$.

Since the solution $(x^*,y^*)=(1,\frac{\sqrt{2}}{4}+\frac{1}{2},\frac{1}{\sqrt{2}})$ satisfies the upper-level constraint $G(x,y)=\frac{\sqrt{2}-1}{2}x_1-y+1/2\geq 0$ one can show that it is a bilevel optimal solution. Note that x_2 is chosen in reaction to x_1 and y since x_2 does not show up in the constraints. This gives the first-order condition $(x_2-\frac{y}{2})=\frac{1}{2}$.

To be consistent with Ye and Zhu's notation (which solves minimization problems instead of maximization ones), we modify the problem into minimization form with

$$f(x,y)=(x_1-1)(x_1-y)y+(y^2-1)^2(y^2-\tfrac{1}{2})^2.$$

Therefore, the problem becomes

$$\begin{aligned} \min_{x_1,y} \quad & (x_1-\tfrac{6}{5})^2+(y-\tfrac{1}{4})^2-\tfrac{1}{8}+\tfrac{1}{2}(\tfrac{y}{2}+\tfrac{1}{2}) \\ \text{subject to} \quad & -\tfrac{\sqrt{2}-1}{2}x_1+y-\tfrac{1}{2}\leq 0 \\ & y\in \arg\min(x_1-1)(x_1-y)y+(y^2-1)^2(y^2-\tfrac{1}{2})^2. \end{aligned}$$

Hence$\nabla_{x_1}f(x^*,y^*)=\frac{\sqrt{2}-1}{2}$, $\nabla_{x_1}f(x^*,z^2)=-\frac{1+\sqrt{2}}{2}$, and $\nabla_{x_1}f(x^*,z^3)=-2$, where $(z^2,z^3)=(-\frac{\sqrt{2}}{2},-1)$. This means we can obtain that $W(x^*)$ is the set of directional derivatives of the lower-level objective function $f(x,y)$ at x^* in the direction d_x, i,e,

$$W(x^*)=\{\nabla_x f(x^*,y'):y'\in S(x^*)\}=\{(-2,0)^\top,(-\tfrac{1+\sqrt{2}}{2},0)^\top,(\tfrac{\sqrt{2}-1}{2},0)^\top\}$$

where $W(x^*)$ is as defined in equation (1.4) of Ye and Zhu [14]. Therefore, the MPEC linearization cone (Definition 3.4 in Ye and Zhu [14]) is

$$\begin{aligned} \mathcal{L}^{\mathrm{MPEC}}(x^*,y^*) &= \{d\in\mathbb{R}^3:\nabla(\nabla_y f(x^*,y^*))d=0,\nabla G(x^*,y^*)d\leq 0\} \\ &= \{d\in\mathbb{R}^3:-(\sqrt{2}-1)d_{x_1}+d_y=0,-\tfrac{\sqrt{2}-1}{2}d_{x_1}+d_y\leq 0\}. \end{aligned}$$

Note that $d \in \mathcal{L}^{MPEC}(x^*, y^*)$ implies that only $d_{x_1} \leq 0$ is feasible, by $(\sqrt{2}-1)d_{x_1} = d_y \leq \frac{\sqrt{2}-1}{2} d_{x_1}$. MPEC-weak calmness at (x^*, y^*) with modulus $\mu > 0$ requires

$$[\nabla F(x^*, y^*) + \mu \nabla f(x^*, y^*)]^\top d - \mu \min_{\xi \in W(x^*)} \xi^\top d_x \geq 0 \text{ for all } d \in \mathcal{L}^{\text{MPEC}}(x^*, y^*).$$

By $d_y = (\sqrt{2}-1)d_{x_1}$, the above inequality is equivalent to

$$[\nabla F(x^*, y^*)^\top d + \mu \tfrac{\sqrt{2}-1}{2} d_{x_1} - \mu \tfrac{\sqrt{2}-1}{2} d_{x_1} \geq 0 \text{ for any } d \in \mathcal{L}^{\text{MPEC}}(x^*, y^*),$$

which amounts to $\nabla_{x_1} F(x^*, y^*) d_{x_1} + \nabla_y F(x^*, y^*) d_y + \nabla_{x_2} F(x^*, y^*) d_{x_2} \geq 0$ since $d_{x_1} \leq 0$ when $d \in \mathcal{L}^{\text{MPEC}}(x^*, y^*)$ and $\xi = (\frac{\sqrt{2}-1}{2}, 0)^\top$ is chosen from $W(x^*)$. However, for a given $d_{x_1} < 0$,

$$\begin{aligned}
\nabla F(x^*, y^*)^\top d &= \nabla_{x_1} F(x^*, y^*) d_{x_1} + \nabla_y F(x^*, y^*) d_y + \nabla_{x_2} F(x^*, y^*) d_{x_2} \\
&= [\nabla_x F(x^*, y^*) + \nabla_y F(x^*, y^*)(\sqrt{2}-1)] d_{x_1} + \nabla_{x_2} F(x^*, y^*) d_{x_2} \\
&= [-\tfrac{2}{5} + (\sqrt{2} - \tfrac{1}{4})(\sqrt{2}-1)] d_{x_1} - 0 \cdot d_{x_2} < 0.
\end{aligned}$$

Therefore, the MPEC's weak calmness is not satisfied. Since there is no lower-level constraint, MPEC-weakly calm is the same as weakly calm, and so the constraint qualification in Theorem 3.2 in Ye and Zhu [14] does not apply to this instance.

Motivated by the above observations, we employ the max-min reformulation of the bilevel problem to derive necessary optimality conditions that can apply to this instance. We return to Example 1 in Example 4.

1.2 Contributions

Existing optimal conditions require more detailed knowledge of the set $S(x)$ than our approach. For example, consider the value function approach. When this approach is used, the limiting subdifferential of value function $\max_z \{f(x,z) : g(x,z) \geq 0\}$ is a fundamental object, but this involves the union of all solutions from the set $S(x)$. Moreover, the action object of the subdifferential operation is $-\max_z \{f(x,z) : g(x,z) \geq 0\}$. For the required condition

$$\partial(-\max_z \{f(x,z) : g(x,z) \geq 0\}) = -\partial(\max_z \{f(x,z) : g(x,z) \geq 0\})$$

to hold, we need to use a Clarke subdifferential, which involves the convex hull of the union of all solutions from $S(x)$ (see details in Guo et al. [7]). However, obtaining all the information about $S(x)$ is difficult. Therefore, our first contribution is summarized as follows.

Contribution 1. Based on the technique of reformulating (BLPP) as a max-min problem and using a penalty function approach, the optimality conditions we derived involve a *single* alternate best response from the set $S(x)$, rather than considering many alternate best responses.

Optimality conditions based on calmness-like conditions typically have constraint qualifications with the following flavor: (i) they involve the leader's objective function, and (ii) the qualification must be valid for a collection of direction vectors. For instance, Theorem 3.2 in Ye and Zhu [14] requires the inner product of the leader and follower's gradients with a family of directions d from a cone to be nonnegative. Therefore, our second contribution is summarized as follows.

Contribution 2. The constraint qualifications we derived are more "classical" in form—they involve only the constraints of the leader's problem, which includes the follower's objective function and the existence of a single direction d.

The rest of the paper is organized as follows. Section 2, we reformulate (BLPP) into a max-min problem and relax this max-min problem into a local minimax problem. Section 3 is the main section of the paper, where we set up our penalty function and derive Fritz-John optimality conditions for (BLPP). This is quite involved since we need to show that our reformulation can deliver the first-order condition by ensuring there exists an optimal interior point solution to the minimax problem. Moreover, careful attention is paid to establishing differentiability properties of the value function of the inner minimization problem for sufficiently large penalization levels. Section 4 presents our takeaway results. Here, we give KKT-like necessary optimality conditions for (BLPP) under a variety of constraint qualifications. Section 5 provides examples that show how our conditions apply in cases not covered by other optimality conditions in the literature. In particular, we return to Example 1 and show how to derive optimal conditions using our methodology.

2 The max-min reformulation

In this section, we reformulate (BLPP) into a max-min problem and use minimax inequality to relax the resulting max-min problem.

2.1 Reformulation into a max-min problem

We begin by transforming the bilevel program to a "single-level" equivalent problem, albeit one with a "max-min" structure. This involves introducing an auxiliary decision variable z. First, observe that (BLPP) is equivalent to:

$$
\begin{aligned}
\max_{x,y} \quad & F(x,y) \\
\text{subject to} \quad & \min_{z:g(x,z)\geq 0} \{f(x,y) - f(x,z)\} \geq 0 && (1)\\
& G(x,y) \geq 0 && (2)\\
& g(x,y) \geq 0. && (3)
\end{aligned}
$$

The auxiliary variable z plays the role of an alternative choice for the follower. The first step is to pull the minimization operator out from the constraint (1) and behind the objective function to form a single-level max-min problem. This requires handling the possibility that a choice of x does not implement y, in which case (1) is violated. To deal with this, we define the extended-real valued function from $X \times Y \times Y \to \mathbb{R} \cup \{-\infty\}$,

$$
F^I(x,y,z) := \begin{cases} F(x,y) & \text{if } f^I(x,y) - f^I(x,z) \geq 0 \\ -\infty & \text{otherwise,} \end{cases} \tag{4}
$$

where

$$
f^I(x,y) := \begin{cases} f(x,y) & \text{if } g(x,y) \geq 0 \\ -\infty & \text{otherwise} \end{cases} \tag{5}
$$

for all $x \in X$ and $y, z \in Y$. The reformulation is:

$$
\begin{aligned}
\max_{x\in X, y\in Y} \min_{z\in Y} \quad & F^I(x,y,z) \\
\text{subject to} \quad & G(x,y) \geq 0 \\
& g(x,y) \geq 0.
\end{aligned} \tag{Max-Min}
$$

The explicit constraints after the "subject to" only involve the primary variables x and y. The objective $F^I(x,y,z)$ is well-defined, i.e., by the constraint $g(x,y) \geq 0$, we avoid $f^I(x,y) - f^I(x,z) = -\infty + \infty$.

We characterize the relationship between the optimal solution to (Max-Min) and (BLPP). First, we present a useful lemma. We say a solution $(\hat{x}, \hat{y})$ to (BLPP) is *bilevel feasible* is $\hat{x} \in X$, $\hat{y} \in Y$, $\hat{y} \in S(\hat{x})$, and $G(\hat{x}, \hat{y}) \geq 0$.

Lemma 1. Let $(\hat{x}, \hat{y})$ be such that $G(\hat{x}, \hat{y}) \geq 0$ and $g(\hat{x}, \hat{y}) \geq 0$. Then $(\hat{x}, \hat{y})$ is bilevel feasible if and only if $\min_z F^I(\hat{x}, \hat{y}, z) > -\infty$. Moreover, when $(\hat{x}, \hat{y})$ is bilevel feasible, $\min_z F^I(\hat{x}, \hat{y}, z) = F(\hat{x}, \hat{y})$.

Proof. Suppose $(\hat{x}, \hat{y})$ is not bilevel feasible. Then there exists a $\hat{z}$ with $g(\hat{x}, \hat{z}) \geq 0$ such that $f(\hat{x}, \hat{y}) < f(\hat{x}, \hat{z})$ and hence $F^I(\hat{x}, \hat{y}, \hat{z}) = -\infty$. This implies $\min_z F^I(\hat{x}, \hat{y}, z) = -\infty$.

Conversely, suppose $(\hat{x}, \hat{y})$ is bilevel feasible. For any z such that $g(\hat{x}, z) \ngeq 0$ we have $f^I(\hat{x}, z) = -\infty$ and so $f^I(\hat{x}, \hat{y}) - f^I(\hat{x}, z) \geq 0$ is sure to hold, implying $F^I(\hat{x}, \hat{y}, z) = F(\hat{x}, \hat{y}) > -\infty$. On the other hand, if z satisfies $g(\hat{x}, z) \geq 0$ then $f^I(\hat{x}, \hat{y}) \geq f^I(\hat{x}, z)$ since $(\hat{x}, \hat{y})$ is bilevel feasible. This implies $F^I(\hat{x}, \hat{y}, z) = F(\hat{x}, \hat{y}) > -\infty$. □

Now, based on Lemma 1, we have the following theorem.

Theorem 1. (Max-Min) and (BLPP) are equivalent problems in the sense that (a) if (x^*, y^*, z^*) is an optimal solution to (Max-Min) then (x^*, y^*) is an optimal solution to (BLPP), and (b) if (x^*, y^*) is an optimal solution to (BLPP) then (x^*, y^*, z) is an optimal solution to (Max-Min) for any $z \in Y$. In either case, (Max-Min) and (BLPP) have the same optimal objective value.

Proof. Let val(Max-Min) be the optimal objective value of problem (Max-Min). To establish part (a), first note that val(Max-Min) $> -\infty$. This follows since a bilevel optimal solution $(\hat{x}, \hat{y})$ exists and so by Lemma 1, val(Max-Min) $\geq \min_z F^I(\hat{x}, \hat{y}, z) = F(\hat{x}, \hat{y}) > -\infty$. Also by Lemma 1, any choice in $\arg\max_{x,y} \min_z \left\{F^I(x, y, z) : G(x, y) \geq 0, g(x, y) \geq 0\right\}$ is bilevel feasible since all other choices have an objective value of $-\infty$. In particular, (x^*, y^*) is bilevel feasible. Since, again by Lemma 1, $F^I(x, y, z) = F(x, y)$ for any (x, y) that are bilevel feasible, the optimality of (x^*, y^*, z^*) implies, for any bilevel feasible (x, y), we have

$$F(x^*, y^*) = F^I(x^*, y^*, z^*) \geq F^I(x, y, z^*) = F(x, y)).$$

This means (x^*, y^*) is an optimal solution to (BLPP). The fact that val(Max-Min) = val(BLPP) then immediately follows from $F(x^*, y^*) = F^I(x^*, y^*, z^*)$.

Now for part (b). Since (x^*, y^*) is an optimal solution to (BLPP), we know $G(x^*, y^*) \geq 0$, $g(x^*, y^*) \geq 0$ and (by Lemma 1) $\min_z F^I(x^*, y^*, z) = F(x^*, y^*) > -\infty$. Hence, any choice of $(x, y) \in \arg\max_{x,y} \min_z \left\{F^I(x, y, z) : G(x, y) \geq 0, g(x, y) \geq 0\right\}$ must be bilevel feasible and has a (Max-Min) objective value $F(x, y)$. Since $F(x^*, y^*) \geq F(x, y)$ by bilevel optimality, this implies

$$(x^*, y^*) \in \arg\max_{x,y} \min_{z} \left\{F^I(x, y, z) : G(x, y) \geq 0, g(x, y) \geq 0\right\}.$$

Note that for any choice of z, $F^I(x^*, y^*, z) = F(x^*, y^*)$. We conclude (x^*, y^*, z) is an optimal solution to (Max-Min) for any choice of z. □

2.2 A local min-max relaxation

Proceeding from problem (Max-Min), we propose a minimax relaxation based on the minimax inequality. By this relaxation, we will select a proper z^* that can be used to characterize a necessary optimality condition. Our penalty function crucially depends on the relaxation problem and the local property of its solution.

Let (x^*, y^*) be a solution to the original bilevel problem that we want to characterize. For a sufficiently small $\epsilon > 0$, the norm-ball neighborhood $\mathcal{N}_\epsilon(x^*, y^*)$ surrounding (x^*, y^*) has the following property

$$\begin{aligned}\text{val(Max-Min)} &= \max_{(x,y)\in\mathcal{N}_\varepsilon(x^*,y^*)} \min_{z\in Y}\{F^I(x,y,z) : G(x,y)\geq 0, \nabla_y f(x,y) - \nabla_y f(x^*,y^*) = 0, g(x,y)\geq 0\} \\ &\leq \min_{z\in Y} \max_{(x,y)\in\mathcal{N}_\varepsilon(x^*,y^*)}\{F^I(x,y,z) : G(x,y)\geq 0, \nabla_y f(x,y) - \nabla_y f(x^*,y^*) = 0, g(x,y)\geq 0\}.\end{aligned}$$

Since for any $z \in Y$, there must exist some $(x, y) \in \mathcal{N}_\epsilon(x^*, y^*)$ to make $F^I(x, y, z) > -\infty$, then we can formally introduce the following minimax problem, which is equivalent to the last line of the above presentation:

$$\min_{z\in Y} \max_{(x,y)\in\mathcal{N}_\epsilon(x^*,y^*)}\{F(x,y) : f^I(x,y) - f^I(x,z)\geq 0, \nabla_y f(x,y) - \nabla_y f(x^*,y^*) = 0, G(x,y)\geq 0, g(x,y)\geq 0\}. \tag{Min-Max|ϵ}$$

For any $\epsilon > 0$, the inner maximization in (Min-Max|ϵ) plays its own important role in the development of this paper. Given an $\epsilon > 0$ and any z, define the function:

$$\Theta(z,\epsilon) := \max_{(x,y)\in\mathcal{N}_\epsilon(x^*,y^*)}\{F(x,y) : f^I(x,y) - f^I(x,z)\geq 0, \nabla_y f(x,y) - \nabla_y f(x^*,y^*) = 0, G(x,y)\geq 0, g(x,y)\geq 0\}. \tag{Inner|z,ϵ}$$

We let $(x^\#(z, \epsilon), y^\#(z, \epsilon))$ denote an arbitrary optimal solution of (Inner|z,ϵ).

For (Inner|z,ϵ) problem, we know that given z and any $\epsilon > 0$, the inner maximization must possess a solution by the upper semicontinuity of the objective function and constraints. It is not immediately clear whether the outer minimization problem has a solution. Let's postpone the discussion of this existence issue for now and suppose the outer minimization in the above minimax problem has a solution for any sufficiently small $\epsilon > 0$. We will provide a sufficient condition for the existence in Appendix A.

Assumption 1. There exists a $\underline{\epsilon} > 0$ such that for all $0 < \epsilon < \underline{\epsilon}$, problem $\min_{z\in Y} \Theta(z, \epsilon)$ has an optimal solution.

Lemma 2. Suppose Assumption 1 holds. As $\epsilon \to 0$ there exists a continuous branch $z(\epsilon) \in \arg\min_{z\in Y} \Theta(z, \epsilon)$ such that $z(\epsilon) \to z^*$.

Proof. Under Assumption 1, we have that $\arg\min_{z\in Y} \Theta(z, \epsilon)$ is lower hemi-continuous in ϵ. Therefore, as $\epsilon \to 0$ from the right, there must be a continuous path connecting $z(\epsilon)$ and its limit z^*. □

We also introduce the following notation. Setting $z(\epsilon)$ equal to the selection of optimal solution to $\min_{z\in Y} \Theta(z, \epsilon)$ as $\epsilon \to 0$ in the previous lemma, we may define

$$x^\#(\epsilon) := x^\#(z(\epsilon), \epsilon). \tag{6}$$

and similarly define $y^\#(\epsilon)$.

For convenience, we also propose the following regularity assumption:

Assumption 2. The matrix $\nabla^2_{yy} f(x^*, y^*)$ is invertible.

This regularity condition tells us that there exists an $\hat{\epsilon}$ sufficiently small such that $\nabla_y f(x, y) = \nabla_y f(x^*, y^*)$ determines a function $y(x)$ for $(x, y) \in \mathcal{N}_{\hat{\epsilon}}(x^*, y^*)$ where $\nabla_y f(x, y(x)) = \nabla_y f(x^*, y^*)$ for all $(x, y(x)) \in \mathcal{N}_{\hat{\epsilon}}(x^*, y^*)$. With this relationship, we will have occasion to study the ball $\mathcal{N}_\epsilon(x^*)$ around x^* only and leverage this to say things about the ball $N_{\epsilon''}(x^*, y^*)$ for possibly a different ϵ''. The idea is that when $x \in \mathcal{N}_\epsilon(x^*)$ then $y(x) \in \mathcal{N}_{\epsilon'}(y^*)$ for some ϵ' by continuity. This, $(x, y(x)) \in \mathcal{N}_{\epsilon''}(x^*, y^*)$ for $\epsilon'' = \max\{\epsilon, \epsilon'\}$.

Although Assumption 2 is a bit strong, it can be regarded as a linearly independent constraint qualification for the inner problem (Inner$|z$,ϵ) with equality constraint $\nabla_y f(x, y(x)) = \nabla_y f(x^*, y^*)$. For our theory to work, this condition may be over-sufficient, but it makes the presentation a lot smoother.

Given $z(\epsilon) \in \arg\min_{z\in Y} \Theta(z, \epsilon)$ by Assumption 1, next we want to show that for sufficiently small ϵ, the inner problem (Inner$|z$,ϵ) possesses an interior solution, where "interior" means the interior to the neighborhood $\mathcal{N}_\epsilon(x^*, y^*)$. This interiority yields the first-order condition, which we later use to characterize a necessary condition using a penalty function.

3 Fritz-John optimality conditions for bilevel programs

In this section, we show the following theorem, which gives a Fritz-John type necessary condition for characterizing (x^*, y^*) using just one alternative optimizer $z^* \in S(x^*)$.

Theorem 2 (Fritz-John optimality conditions)**.** Let (x^*, y^*, z^*) be a given optimal solution to (Max-Min) and z^* in $S(x^*)$ and suppose Assumptions 1 and 2 hold. Then there exist nonnegative multipliers $\lambda_{G,j}$ for $j = 1, \ldots, q$, $\lambda_{g,\ell}$, and κ_ℓ, for $\ell = 1, \ldots, p$, $\mu_0, \beta \geq 0$ and unsigned θ_i for $i = 1, \ldots, m$, not all zero, such that:

$$\begin{aligned} 0_n = {} & \mu_0 \nabla_x F(x^*, y^*) + \sum_{j=1}^{q} \lambda_{G,j} \nabla_x G_j(x^*, y^*) + \sum_{\ell=1}^{p} \lambda_{g,\ell} \nabla_x g_\ell(x^*, y^*) && \text{(7a)} \\ & + \sum_{\ell \in A_g(x^*,z^*)} \kappa_\ell \nabla_x g_\ell(x^*, z^*) + \sum_{i=1}^{m} \theta_i \nabla_x \tfrac{\partial f(x^*,y^*)}{\partial y_i} + \beta(\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*)) \\ 0_m = {} & \mu_0 \nabla_y F(x^*, y^*) + \sum_{j=1}^{q} \lambda_{G,j} \nabla_y G_j(x^*, y^*) + \sum_{\ell=1}^{p} \lambda_{g,\ell} \nabla_y g_\ell(x^*, y^*) && \text{(7b)} \\ & + \sum_{i=1}^{m} \theta_i \nabla_y \tfrac{\partial f(x^*,y^*)}{\partial y_i} + \beta \nabla_y f(x^*, y^*) \end{aligned}$$

and the following complementary slackness conditions hold

$$\begin{aligned} \lambda_{G,j} G_j(x^*, y^*) &= 0 \quad \text{for } j = 1, \ldots, q, && \text{(8a)} \\ \lambda_{g,\ell} g_\ell(x^*, y^*) &= 0 \quad \text{for } \ell = 1, \ldots, p, && \text{(8b)} \\ \beta(f(x^*, y^*) - f(x^*, z^*)) &= 0. && \text{(8c)} \end{aligned}$$

3.1 Discussion on single-dimensional case

For ease of exposition, we prove Theorem 2 first in the single dimensional case where $X \subset \mathbb{R}$. This case has its own important value in applications (for example, in economics and finance). By Assumption 2, for $\epsilon > 0$ the inner optimization problem (Inner$|z$,ϵ) becomes

$$\max_{x\in[x^*-\epsilon, x^*+\epsilon]} \{F(x, y(x)) : f^I(x, y) - f^I(x, z(\epsilon)) \geq 0, , G(x, y(x)) \geq 0, g(x, y(x)) \geq 0\}, \tag{9}$$

where $y(x)$ solves $\nabla_y f(x, y(x)) - \nabla_y f(x^*, y^*) = 0$ for y when $x \in [x^* - \epsilon, x^* + \epsilon]$.

The idea of proof is divided into two major steps. Step 1 shows that a solution x (and y) of the inner problem must be in the interior of the neighborhood $[x^* - \epsilon, x^* + \epsilon]$. Step 2 shows the desired necessary condition using a penalty function. The key result in the first step is the following.

Proposition 1. Suppose Assumptions 1 and 2 hold and take a sufficiently small $\epsilon > 0$. For the single-dimensional case $X \subset \mathbb{R}$, either (i) x^* is a stationary point of $F(x, y(x))$; or (ii) every optimal solution $(x^{\#}(\epsilon), y^{\#}(\epsilon))$ to (Inner$|z,\epsilon$)—where $x^{\#}(\epsilon)$ is as defined in (6)—is in the interior of $\mathcal{N}_\epsilon(x^*, y^*)$.

Proof. (i) If x^* is a stationary point of $F(x, y(x))$ for $x \in [x^* - \epsilon, x^* + \epsilon]$, then we have the first-order condition

$$\tfrac{d}{dx} F(x^*, y(x^*)) = \nabla_x F(x^*, y(x^*)) + \nabla_y F(x^*, y(x^*)) \nabla_x y(x^*) = 0. \tag{10}$$

Taking the total derivative of $\nabla_y f(x, y(x)) = \nabla_y f(x^*, y^*)$ with respect to x at $(x, y) = (x^*, y^*)$ yields

$$\nabla_x y(x^*) = -(\nabla^2_{yy} f(x^*, y^*))^{-1} \nabla_x \nabla_y f(x^*, y^*).$$

This allows us to rewrite (10) as

$$\nabla_x F(x^*, y^*) + \theta \nabla_x \nabla_y f(x^*, y^*) = 0,$$

where

$$\theta = -\nabla_y F(x^*, y(x^*))(\nabla^2_{yy} f(x^*, y^*))^{-1}.$$

Moreover, the definition of θ implies that

$$\nabla_y F(x^*, y(x^*)) + \theta(\nabla^2_{yy} f(x^*, y^*)) = 0$$

holds. Therefore, we obtain the desired necessary conditions

$$\begin{aligned} \nabla_x F(x^*, y^*) + \theta \nabla^2_{xy} f(x^*, y^*) = 0 \\ \nabla_y F(x^*, y^*) + \theta \nabla^2_{yy} f(x^*, y^*) = 0, \end{aligned}$$

for some nonzero vector θ with $\beta = 0$ and $\mu_0 = 1$.

(ii) It remains to consider case where $\frac{d}{dx} F(x^*, y(x^*)) \neq 0$. Then for sufficiently small $\epsilon > 0$, $F(x, y(x))$ must be either increasing or decreasing in $x \in [x^* - \epsilon, x^* + \epsilon]$. We only discuss the case where $F(x, y(x))$ is increasing, the decreasing case is symmetric, so we do not repeat the analysis.

By Assumption 2, $\nabla_y f(x, y(x)) = \nabla_y f(x^*, y^*)$ gives a continuous function $y(x)$ in the interval $x \in [x^* - \epsilon, x^* + \epsilon]$. Therefore, $y(x) \in [y(x^* - \epsilon), y(x^* + \epsilon)]$ for $x \in [x^* - \epsilon, x^* + \epsilon]$.

Recall the definition of $x^{\#}(\epsilon)$ in (6) and suppose that $x^{\#}(\epsilon)$ is not in the interior of $[x^* - \epsilon, x^* + \epsilon]$. Then we have $x^{\#}(\epsilon) = x^* + \epsilon$. It can not be $x^{\#}(\epsilon) = x^* - \epsilon$ because that $F(x^* - \epsilon, y(x^* - \epsilon)) < F(x^*, y^*)$ contradicts $F(x^* - \epsilon, y(x^* - \epsilon)) = \min_z \Theta(z, \epsilon) \geq F(x^*, y^*)$ by definition. Consider

$$\hat{z} \in \arg\max_{z \in Y} f^I(x^* + \epsilon, z),$$

and we have

$$\begin{aligned} & F(x^* + \epsilon, y(x^* + \epsilon)) \\ & = \max_{(x,y) \in \mathcal{N}_\epsilon(x^*, y^*)} \{F(x, y) : f^I(x, y) - f^I(x, z(\epsilon)) \geq 0, \nabla_y f(x, y) = \nabla_y f(x^*, y^*), G(x, y) \geq 0, g(x, y) \geq 0\} \\ & \leq \max_{(x,y) \in \mathcal{N}_\epsilon(x^*, y^*)} \{F(x, y) : f^I(x, y) - f^I(x, \hat{z}) \geq 0, \nabla_y f(x, y) = \nabla_y f(x^*, y^*), G(x, y) \geq 0, g(x, y) \geq 0\} \\ & \leq \max_{x \in \mathcal{N}_\epsilon(x^*)} \{F(x, y(x)) : f^I(x, y(x)) - f^I(x, \hat{z}) \geq 0, G(x, y(x)) \geq 0, g(x, y(x)) \geq 0\} \\ & \leq F(x^* + \epsilon, y(x^* + \epsilon)), \end{aligned}$$

where the first inequality is by the definition of $z(\epsilon)$ as a minimizer, the second inequality is by the definition of $y(x)$ that solves $\nabla_y f(x,y) = \nabla_y f(x^*,y^*)$, and the last inequality is by the fact that $F(x, y(x))$ is increasing.

Therefore, all inequalities become equalities. By the definition of $\hat{z}$, $(x^* + \epsilon, y(x^* + \epsilon))$ is thus bilevel feasible, which yields a contradiction

$$F(x^* + \epsilon, y(x^* + \epsilon)) > F(x^*, y^*) \geq F(x^* + \epsilon, y(x^* + \epsilon))$$

as (x^*, y^*) is bilevel optimal. As a result, $(x^{\#}(\epsilon), y^{\#}(\epsilon))$ must be interior. □

From Lemma 2, we can construct a subsequence

$$\epsilon_n \to 0 \text{ such that } z(\epsilon_n) \to z^*. \tag{11}$$

For convenience, we denote

$$(x^{\#}(\epsilon_n), y^{\#}(\epsilon_n), z(\epsilon_n)) = (x^n, y^n, z^n). \tag{12}$$

Observe that for each n, (x^n, y^n, z^n) is an optimal solution to (Min-Max$|\epsilon_n$).

We now introduce the following penalty function:

$$\begin{aligned} F^{k,n}(x,y) := F(x,y) &- \underbrace{\tfrac{k}{2}\sum_{j=1}^{q}(G_j^-(x,y))^2}_{(i)} - \underbrace{\tfrac{k}{2}\sum_{\ell=1}^{p}(g_\ell^-(x,y))^2}_{(ii)} - \underbrace{\tfrac{k}{2}\sum_{\ell=1}^{p}(g_\ell^-(x,z^n))^2}_{(iii)} - \\ &\underbrace{\tfrac{k}{2}(\min\{0, f^k(x,y) - f(x,z^n)\})^2}_{(iv)} - \underbrace{\tfrac{k}{2}\|\nabla_y f(x,y) - \nabla_y f(x^n,y^n)\|^2}_{(v)} - \underbrace{\tfrac{\alpha_2}{2}\|y - y^n\|^2}_{(vi)} - \underbrace{\tfrac{\alpha_1}{2}\|x - x^n\|^2}_{(vii)}, \end{aligned} \tag{13}$$

where α_1 and α_2 are positive constants, $g_\ell^-(x,y) = \min\{0, g_\ell(x,y)\}$, and similar to G_j^-, and

$$f^k(x,y) = f(x,y) - \frac{k}{2}\sum_{\ell=1}^{p}(g_\ell^-(x,y))^2.$$

The above penalty function is very similar to the single-level problem (as studied in Chapter 3 of [1]). The only difference is term (iv), which is designed to represent the inequality constraint $f^I(x,y) - f^I(x,z^n) \geq 0$. The function $f^k(x,y)$ is thus naturally an approximation of $f^I(x,y)$. This design captures the idea that when terms (ii) and (iii) converge to zero, i.e., constraint $g(x,y) \geq 0$ and $g(x,z^n) \geq 0$ are satisfied, then the difference $f^k(x,y) - f(x,z^n)$ is likely to be dominated by the difference $f(x,y) - f(x,z^n)$.

Let

$$(x^{k,n}, y^{k,n}) \in \arg\max_{(x,y)\in\mathcal{N}_{\epsilon_n}(x^*,y^*)} F^{k,n}(x,y) \tag{14}$$

be a solution to maximizing the penalty function for each k and n. We want to establish that the penalty function is exact given any n, and $(x^{k,n}, y^{k,n}) \to (x^n, y^n)$ as $k \to \infty$.

Lemma 3 (Exactness Lemma)**.** Suppose Assumptions 1 and 2 hold. Let (x^n, y^n, z^n) denote the sequence of optimal solutions to (Min-Max$|\epsilon_n$) defined in (12) and $(x^{k,n}, y^{k,n})$ denote the sequence defined in (14) of maximizers of the penalty function (13). Then, $\lim_{k\to\infty} F^{k,n}(x^{k,n}, y^{k,n}) = F(x^n, y^n)$ and $(x^{k,n}, y^{k,n}) \to (x^n, y^n)$ as $k \to \infty$.

Proof. We first show the exactness $\lim_{k\to\infty} F^{k,n}(x^{k,n}, y^{k,n}) = F(x^n, y^n)$. Direction "$\geq$" is relatively easy. As $(x^n, y^n) \in \mathcal{N}_{\epsilon_n}(x^*, y^*)$ is feasible to the unconstrained maximation problem $\max_{(x,y)\in\mathcal{N}_{\epsilon_n}(x^*,y^*)} F^{k,n}(x, y)$ and $g(x^n, y^n) \geq 0$, $G(x^n, y^n) \geq 0$ and $g(x^n, z^n) \geq 0$ are satisfied, then

$$f^k(x^n, y^n) - f(x^n, z^n) = f(x^n, y^n) - f(x^n, z^n) \geq 0.$$

Therefore,

$$F^{k,n}(x^{k,n}, y^{k,n}) \geq F^{k,n}(x^n, y^n) = F(x^n, y^n)$$

and the inequality pass the limit as $k \to \infty$.

It remains to show direction "$\leq$". Denote $(x^{\infty,n}, y^{\infty,n})$ as the limit of the $(x^{k,n}, y^{k,n})$ sequence (taking a subsequence if necessary). Therefore $g(x^{\infty,n}, y^{\infty,n}) \geq 0$, $G(x^{\infty,n}, y^{\infty,n}) \geq 0$ and $\nabla_y f(x^{\infty,n}, y^{\infty,n}) = \nabla_y f(x^n, y^n)$ should be satisfied, otherwise, $F^{k,n}(x^{k,n}, y^{k,n}) \to -\infty$, contradicting direction "$\geq$" as we have shown. To show that $(x^{\infty,n}, y^{\infty,n})$ is feasible for inner maximization problem, it remains to show

$$f^I(x^{\infty,n}, y^{\infty,n}) \geq f^I(x^{\infty,n}, z^n).$$

We show the above inequality by contradiction. Suppose not true, that is, $f^I(x^{\infty,n}, y^{\infty,n}) < f^I(x^{\infty,n}, z^n)$. Given $g(x^{\infty,n}, y^{\infty,n}) \geq 0$, we have

$$f^I(x^{\infty,n}, y^{\infty,n}) = f(x^{\infty,n}, y^{\infty,n}) < f^I(x^{\infty,n}, z^n),$$

which implies that $f^I(x^{\infty,n}, z^n) > -\infty$, so it must be $f^I(x^{\infty,n}, z^n) = f(x^{\infty,n}, z^n)$. We obtain the following contradiction,

$$0 \leq \lim_{k\to\infty} [f^k(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)] \leq f(x^{\infty,n}, y^{\infty,n}) - f(x^{\infty,n}, z^n) < 0,$$

where the first inequality $0 \leq \lim_{k\to\infty}[f^k(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)]$ is by the property of penalty (otherwise, it contradicts direction "$\geq$"). Therefore, it must be $f^I(x^{\infty,n}, y^{\infty,n}) \geq f^I(x^{\infty,n}, z^n)$ and $(x^{\infty,n}, y^{\infty,n})$ is feasible for the inner maximization problem (Inner$|z,\epsilon$). Therefore,

$$\lim_{k\to\infty} \max_{x,y\in\mathcal{N}_{\epsilon_n}(x^*,y^*)} F^{k,n}(x, y) \leq F(x^{\infty,n}, y^{\infty,n}) \leq F(x^n, y^n),$$

which shows the desired direction.

Finally, all negative terms in the $F^{k,n}(x^{k,n}, y^{k,n})$ converge to zero. In particular, this means $||x^{\infty,n} - x^n|| = 0$ and $||y^{\infty,n} - y^n|| = 0$ and so $(x^{k,n}, y^{k,n}) \to (x^{\infty,n}, y^{\infty,n}) = (x^n, y^n)$ as $k \to \infty$. □

Corollary 1. Suppose Assumptions 1 and 2 hold. Let ϵ_n denote the sequence defined in (11) and $(x^{k,n}, y^{k,n})$ denote the sequence defined in (14). For sufficiently large k, $(x^{k,n}, y^{k,n})$ is an interior point of $\mathcal{N}_{\epsilon_n}(x^*, y^*)$.

Proof. It is straightforward from $(x^{k,n}, y^{k,n}) \to (x^n, y^n)$, since (x^n, y^n) is interior of $\mathcal{N}_{\epsilon_n}(x^*, y^*)$. □

By Corollary 1, we obtain the following first-order condition:

$$\lim_{k\to\infty} \nabla_{x,y} F^{k,n}(x^{k,n}, y^{k,n}) = 0, \text{ for any } n.$$

Sending $n \to \infty$ in this first-order condition, it is not difficult to obtain the desired conclusion for the single-dimensional case $X \subset \mathbb{R}$. We omit the proof here and provide the more general proof when x is multi-dimensional (see Section 3.3).

3.2 Discussion on multi-dimensional case

We now extend the idea to multi-dimensional $X \subset \mathbb{R}^n$. If x^* is a stationary point of $F(x, y(x))$, we are done. However, if x^* is not a stationary point of $F(x, y(x))$, in the multi-dimensional case, we introduce two sets to obtain the result.

Recall that under Assumption 2, the inner maximization problem (Inner$|z,\epsilon$) for a given $\epsilon > 0$ can be expressed as:

$$\max_{x \in \mathcal{N}_\epsilon(x^*)} \{F(x, y(x)) : f^I(x, y(x)) - f^I(x, z(\epsilon)) \geq 0, , G(x, y(x)) \geq 0, g(x, y(x)) \geq 0\}, \tag{15}$$

where $y(x)$ solves $\nabla_y f(x, y(x)) - \nabla_y f(x^*, y^*) = 0$ for y when $x \in \mathcal{N}_\epsilon(x^*)$.

The feasible set of (15), given z, is:

$$\{x \in \mathcal{N}_\epsilon(x^*) : \phi(x, z) \geq 0\}$$

where $\phi(x, z) = \min\{f^I(x, y(x)) - f^I(x, z), G(x, y(x)), g(x, y(x))\}$. The lower-level set, given value $F(x^{\#}(\epsilon), y^{\#}(\epsilon))$, is given by

$$\mathcal{F}_\epsilon = \{x \in \mathcal{N}_\epsilon(x^*) : F(x, y(x)) \leq F(x^{\#}(\epsilon), y^{\#}(\epsilon))\},$$

where $(x^{\#}(\epsilon), y^{\#}(\epsilon))$ is as defined in (6).

Lemma 4. Suppose Assumptions 1 and 2 hold and let ϵ_n be the sequence defined in (11) and (x^n, y^n, z^n) be defined as in (12). Then, there exists a subsequence of the ϵ_n such that (x^n, y^n) (abusing notation to let n index this subsequence) is in the interior of $\mathcal{N}_{\epsilon_n}(x^*, y^*)$.

Proof. Suppose that there is no such subsequence of the ϵ_n, i.e., for any sufficiently small $\epsilon > 0$, there exists no interior solution within ball $\mathcal{N}_\epsilon(x^*)$. We want to derive a contradiction.

Note that the inner problem solution $x^{\#}(\epsilon)$ is a corner solution, i.e., there exists a direction d_x such that $x^* + \epsilon d_x = x^{\#}(\epsilon)$. Therefore, for given $z^{\#}(\epsilon)$, we have

$$\{x \in \mathcal{N}_\epsilon(x^*) : \phi(x, z^{\#}(\epsilon)) \geq 0\} \subset \mathcal{F}_\epsilon \tag{16}$$

and

$$x^{\#}(\epsilon) \in \{x \in \mathcal{N}_\epsilon(x^*) : \phi(x, z^{\#}(\epsilon)) \geq 0\} \cap \mathcal{F}_\epsilon. \tag{17}$$

The first relationship is easy to understand because otherwise

$$\max_{x \in \{x \in N_\varepsilon(x^*) : \phi(x, z^{\#}(\varepsilon) \geq 0\}} F(x, y(x)) > F(x^{\#}(\varepsilon), y^{\#}(\varepsilon)),$$

and the second relationship is by the definition of corner solution $x^{\#}(\varepsilon)$.

Next, we indicate the following equivalence:

$$\begin{aligned} &F(x^{\#}(\epsilon), y^{\#}(\epsilon)) \\ =&\min_{z \in Y} \max_{x \in \mathcal{F}_\epsilon} \max_{y \in y(\mathcal{F}_\epsilon)} \{F(x, y) : f^I(x, y) - f^I(x, z) \geq 0, \nabla_y f(x, y) = \nabla_y f(x^*, y^*), G(x, y) \geq 0, g(x, y) \geq 0\} \end{aligned}$$

where $y(\mathcal{F}_\epsilon)$ denote the image of $x \in \mathcal{F}_\epsilon$ through the mapping $y(x)$.

To see the above equivalence, first of all, "$\geq$" direction is easy as $\mathcal{F}_\epsilon \subset \mathcal{N}_\epsilon(x^*)$. It remains to show "$\leq$" direction. Denote

$$\hat{\Theta}(z, \epsilon) := \max_{x \in \mathcal{F}_\epsilon} \max_{y \in y(\mathcal{F}_\epsilon)} \{F(x, y) : f^I(x, y) - f^I(x, z) \geq 0, \nabla_y f(x, y) = \nabla_y f(x^*, y^*), G(x, y) \geq 0, g(x, y) \geq 0\}.$$

Let $z'(\epsilon) \in \arg\min_{z\in Y} \hat{\Theta}(z,\epsilon)$. Then the feasible set

$$\{x \in \mathcal{N}_\epsilon(x^*) : \phi(x, z'(\epsilon)) \geq 0\} \subset \mathcal{F}_\epsilon,$$

otherwise, it violates the direction "$\geq$". Then we have

$$\begin{aligned}
& F(x^\#(\epsilon), y^\#(\epsilon)) \\
\leq\; & \max_{x\in\mathcal{N}_\epsilon(x^*)} \max_{y\in y(\mathcal{N}_\epsilon(x^*))} \{F(x,y) : f^I(x,y) - f^I(x,z'(\epsilon)) \geq 0, \nabla_y f(x,y) = \nabla_y f(x^*,y^*), G(x,y) \geq 0, g(x,y) \geq 0\} \\
=\; & \max_{x\in\mathcal{F}_\epsilon} \max_{y\in y(\mathcal{F}_\epsilon)} \{F(x,y) : f^I(x,y) - f^I(x,z'(\epsilon)) \geq 0, \nabla_y f(x,y) = \nabla_y f(x^*,y^*), G(x,y) \geq 0, g(x,y) \geq 0\},
\end{aligned}$$

which shows the desired direction.

Now, we show a contradiction. The idea is similar to what we did for the single-dimensional case. Let $\hat{z} \in \arg\max_{z\in Y} f^I(x,z)$. Then

$$\begin{aligned}
& F(x^\#(\epsilon), y^\#(\epsilon)) \\
=\; & \min_{z\in Y} \max_{x\in\mathcal{F}_\epsilon} \max_{y\in y(\mathcal{F}_\epsilon)} \{F(x,y) : f^I(x,y) - f^I(x,z) \geq 0, \nabla_y f(x,y) = \nabla_y f(x^*,y^*), G(x,y) \geq 0, g(x,y) \geq 0\} \\
\leq\; & \max_{x\in\mathcal{F}_\epsilon} \max_{y\in y(\mathcal{F}_\epsilon)} \{F(x,y) : f^I(x,y) - f^I(x,\hat{z}) \geq 0, \nabla_y f(x,y) = \nabla_y f(x^*,y^*), G(x,y) \geq 0, g(x,y) \geq 0\} \\
\leq\; & F(x^\#(\epsilon), y^\#(\epsilon)),
\end{aligned}$$

therefore, all inequalities become equalities, and thus $(x^\#(\epsilon), y^\#(\epsilon))$ is bilevel feasible, which yields a contradiction

$$F(x^*, y^*) < F(x^\#(\varepsilon), y^\#(\varepsilon)) \leq F(x^*, y^*)$$

where (x^*, y^*) is bilevel optimal. Therefore, for sufficiently small $\epsilon > 0$ such that $\mathcal{F}_\epsilon$ can be well-defined, then the solution to the inner maximization problem must be an interior point of $\mathcal{F}_\epsilon$ and so it must be an interior point of $\mathcal{N}_\epsilon(x^*)$. □

The above lemma yields the following.

Lemma 5. Suppose Assumptions 1 and 2 hold and ϵ_n be the sequence defined in (11). We have the following two first-order conditions:

$$\begin{array}{c} \lim_{k\to\infty} \nabla_x F^{k,n}(x^{k,n}, y^{k,n}) = 0_n \\ \lim_{k\to\infty} \nabla_y F^{k,n}(x^{k,n}, y^{k,n}) = 0_m \end{array} \tag{18}$$

Proof. By Lemma 4, we have that $(x^{k,n}, y^{k,n})$ is interior of $\mathcal{N}_{\epsilon_n}(x^*, y^*)$. Therefore, similar to the single-dimensional case, first-order conditions can be directly obtained. □

3.3 Fritz-John optimality conditions

Based on the above results, we begin to prove Theorem 2. The proof builds on ideas established in the more straightforward setting of Proposition 3.3.5 in [1].

Proof of Theorem 2. We focus on establishing the condition with respect to x (equation (7a)). The condition for y follows by analogous reasoning and is omitted for brevity.

In the previous two subsections, we furnished the arguments to establish that (18) provides optimality conditions for (BLPP). Hence, we study the optimality conditions of the kth iterate of the penalized problem:

$$\nabla_x F^{k,n}(x^{k,n}, y^{k,n}) \to 0_n$$

using Lemma 5. Then we get:

$$0_n \leftarrow \nabla_x F(x^{k,n}, y^{k,n}) - k \sum_{j=1}^{q} G_j^-(x^{k,n}, y^{k,n}) \nabla_x G_j(x^{k,n}, y^{k,n}) - k \sum_{\ell=1}^{p} g_\ell^-(x^{k,n}, y^{k,n}) \nabla_x g_\ell(x^{k,n}, y^{k,n})$$

$$- k \sum_{\ell=1}^{p} g_\ell^-(x^{k,n}, z^n) \nabla_x g_\ell(x^{k,n}, z^n) - \alpha_1(x^{k,n} - x^n) - k \nabla^2_{yx} f(x^{k,n}, y^{k,n})(\nabla_y f(x^{k,n}, y^{k,n}) - \nabla_y f(x^n, y^n))$$

$$- k\Bigg(\min \Big\{0, f(x^{k,n}, y^{k,n}) - \tfrac{k}{2} \sum_{\ell=1}^{p} (g_\ell^-(x^{k,n}, y^{k,n}))^2 - f(x^{k,n}, z^n)] \Big\} \Bigg) \times$$

$$\Bigg(\nabla_x f(x^{k,n}, y^{k,n}) - k \sum_{\ell=1}^{p} g_\ell^-(x^{k,n}, y^{k,n}) \nabla_x g_\ell(x^{k,n}, y^{k,n}) - \nabla_x f(x^{k,n}, z^n) \Bigg).$$

We may re-express the above optimality condition as:

$$0_n \leftarrow \nabla_x F(x^{k,n}, y^{k,n}) + \sum_{j=1}^{q} \eta_{G,j}^{k,n} \nabla_x G_j(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} (\eta_{g,\ell}^{k,n} + \omega_\ell^{k,n}) \nabla_x g_\ell(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} \xi_\ell^{k,n} \nabla_x g_\ell(x^{k,n}, z^n)$$

$$+ \sum_{i=1}^{m} \vartheta_i^{k,n} \nabla_x \tfrac{\partial f(x^{k,n}, y^{k,n})}{\partial y_i} + \gamma^{k,n}(\nabla_x f(x^{k,n}, y^{k,n}) - \nabla_x f(x^{k,n}, z^n)) - \alpha_1(x^{k,n} - x^n), \tag{20}$$

where

$$\eta_{G,j}^{k,n} := -k G_j^-(x^{k,n}, y^{k,n}) \qquad \text{for } j = 1, \ldots, q, \tag{21a}$$

$$\eta_{g,\ell}^{k,n} := -k g_\ell^-(x^{k,n}, y^{k,n}) \qquad \text{for } \ell = 1, \ldots, p, \tag{21b}$$

$$\xi_\ell^{k,n} := -k g_\ell^-(x^{k,n}, z^n) \qquad \text{for } \ell = 1, \ldots, p, \tag{21c}$$

$$\vartheta_i^{k,n} := -k(\tfrac{\partial}{\partial y_i} f(x^{k,n}, y^{k,n}) - \tfrac{\partial}{\partial y_i} f(x^n, y^n)) \qquad \text{for } i = 1, \ldots, m, \tag{21d}$$

$$\gamma^{k,n} := -k(\min\{0, f^k(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)\}), \text{ and} \tag{21e}$$

$$\omega_\ell^{k,n} := \gamma^{k,n} \eta_{g,\ell}^{k,n} \qquad \text{for } \ell = 1, \ldots, p, \tag{21f}$$

Denote

$$\delta^{k,n} = \sqrt{1 + \sum_{j=1}^{q} (\eta_{G,j}^{k,n})^2 + \sum_{\ell=1}^{p} (\eta_{g,\ell}^{k,n})^2 + \sum_{\ell=1}^{p} (\xi_\ell^{k,n})^2 + \sum_{i=1}^{m} (\vartheta_i^{k,n})^2 + (\gamma^{k,n})^2 + \sum_{\ell=1}^{p} (\omega_\ell^{k,n})^2} \tag{22}$$

and set $\mu_0^{k,n} := \frac{1}{\delta^{k,n}}$, $\lambda_{G,j}^{k,n} := \frac{\eta_{G,j}^{k,n}}{\delta^{k,n}}$ for $j = 1, \ldots, q$, $\lambda_{g,\ell}^{k,n} := \frac{\eta_{g,\ell}^{k,n}}{\delta^{k,n}}$, $\kappa_\ell^{k,n} := \frac{\xi_\ell^{k,n}}{\delta^{k,n}}$, and $\rho_\ell^{k,n} := \frac{\omega_\ell^{k,n}}{\delta^{k,n}}$ for $\ell = 1, \ldots, p$, $\theta_i^{k,n} := \frac{\vartheta_i^{k,n}}{\delta^{k,n}}$ for $i = 1, \ldots, m$, and $\beta^{k,n} := \frac{\gamma^{k,n}}{\delta^{k,n}}$. Dividing (20) by $\delta^{k,n}$ yields

$$0_n \leftarrow \mu_0^{k,n} \nabla_x F(x^{k,n}, y^{k,n}) + \sum_{j=1}^{q} \lambda_{G,j}^{k,n} \nabla_x G_j(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} (\lambda_{g,\ell}^{k,n} + \rho_\ell^{k,n}) \nabla_x g_\ell(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} \kappa_\ell^{k,n} \nabla_x g_\ell(x^{k,n}, z^n)$$

$$+ \sum_{i=1}^{m} \theta_i^{k,n} \nabla_x \tfrac{\partial f(x^{k,n}, y^{k,n})}{\partial y_i} + \beta^{k,n}(\nabla_x f(x^{k,n}, y^{k,n}) - \nabla_x f(x^{k,n}, z^n)) - \tfrac{\alpha_1}{\delta^{k,n}}(x^{k,n} - x^n). \tag{23}$$

From (22) we have

$$(\mu_0^{k,n})^2 + \sum_{j=1}^{q}(\lambda_{G,j}^{k,n})^2 + \sum_{\ell=1}^{p}(\lambda_{g,\ell}^{k,n})^2 + \sum_{\ell=1}^{p}(\kappa_\ell^{k,n})^2 + \sum_{i=1}^{m}(\theta_i^{k,n})^2 + (\beta^{k,n})^2 + \sum_{\ell=1}^{p}(\rho_\ell^{k,n})^2 = 1, \tag{24}$$

implies that the sequence $(\mu_0^{k,n}, \lambda_G^{k,n}, \lambda_g^k, \kappa^{k,n}, \theta^{k,n}, \beta^{k,n}, \rho^{k,n})_{k=1}^{\infty}$ is bounded and so (since a sequence in Euclidean space) contains a convergent subsequence to some limit $(\mu_0, \lambda_G, \lambda_g, \kappa, \theta, \beta, \rho)$, as $k \to \infty$ and $n \to \infty$. From (24) we know the limit vector $(\mu_0, \lambda_G, \lambda_g, \kappa, \theta, \beta, \rho)$ is not the zero vector.

Now, redefine the sequence k, n to index this convergent subsequence and take $k \to \infty$ and $n \to \infty$ on both sides of (23). The continuous differentiability of F, G, f and g and the fact that $(x^{k,n}, y^{k,n}) \to (x^*, y^*)$ and $z^n \to z^*$ yields (7a) and noting the fact $\frac{\alpha_1}{\delta^{k,n}}(x^{k,n} - x^n) \to 0$. The nonnegativity of $(\mu_0, \lambda_G, \lambda_g, \kappa, \beta, \rho)$ follows from (21a)–(21e). The complementarity conditions follow by standard arguments that we will not repeat here (see Proposition 3.3.5 in Bertsekas [1]). □

4 KKT optimality conditions for bilevel programs

This section proposes some constraint qualifications for when the Fritz-John multiplier $\mu_0 = 1$.

To state our conditions, we use the following notation: $A_G(x^*, y^*) := \{j : G_j(x^*, y^*) = 0\}$, $A_g(x^*, y^*) := \{\ell : g_\ell(x^*, y^*) = 0\}$, and $A_g(x^*, z^*) := \{\ell : g_\ell(x^*, z^*) = 0\}$. The first set of conditions assumes there exists a $z^* \in S(x^*)$ such that there exists a nonzero vector $d = (d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^m$ such that:

(Q1) $\nabla G_j(x^*, y^*)^\top d > 0$ for all $j \in A_G(x^*, y^*)$,

(Q2) $\nabla g_\ell(x^*, y^*)^\top d > 0$ for all $\ell \in A_g(x^*, y^*)$,

(Q3) $\nabla_x g_\ell(x^*, z^*)^\top d_x > 0$ for all $\ell \in A_g(x^*, z^*)$,

(Q4) $\nabla \frac{\partial f(x^*, y^*)}{\partial y_i}^\top d = 0$ for all $i = 1, \ldots, m$,

(Q5) the columns of $\nabla^2_{yy} f(x^*, y^*)$ are linearly independent,

(Q6) $\nabla f(x^*, y^*)^\top d \geq \nabla f(x^*, z^*)^\top d$,

(Q7) $||\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*)|| \neq 0$,

Theorem 3 (KKT optimality conditions)**.** Let (x^*, y^*) be an optimal solution to (BLPP) and $z^* \in S(x^*)$. Suppose Assumptions 1 and 2 and the constraint qualifications (Q1)–(Q7) hold. Then, there exist nonnegative multipliers $\lambda_{G,j}$ for $j = 1, \ldots, q$, $\lambda_{g,\ell}$, ρ_ℓ and κ_ℓ for $\ell = 1, \ldots, p$, $\beta \geq 0$ and unsigned θ_i for $i = 1, \dots, m$ such that:

$$0_n = \nabla_x F(x^*, y^*) + \sum_{j=1}^{q} \lambda_{G,j} \nabla_x G_j(x^*, y^*) + \sum_{\ell=1}^{p} (\lambda_{g,\ell} + \rho_\ell) \nabla_x g_\ell(x^*, y^*) \tag{25a}$$
$$+ \sum_{\ell \in A_g(x^*, z^*)} \kappa_\ell \nabla_x g_\ell(x^*, z^*) + \sum_{i=1}^{m} \theta_i \nabla_x \tfrac{\partial f(x^*, y^*)}{\partial y_i} + \beta(\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*))$$
$$0_m = \nabla_y F(x^*, y^*) + \sum_{j=1}^{q} \lambda_{G,j} \nabla_y G_j(x^*, y^*) + \sum_{\ell=1}^{p} (\lambda_{g,\ell} + \rho_\ell) \nabla_y g_\ell(x^*, y^*) \tag{25b}$$
$$+ \sum_{i=1}^{m} \theta_i \nabla_y \tfrac{\partial f(x^*, y^*)}{\partial y_i} + \beta \nabla_y f(x^*, y^*)$$

and the complementary slackness conditions (8) hold.

Proof. It suffices to show $\mu_0 \neq 0$ (and hence $\mu_0 > 0$ since $\mu_0 \geq 0$) in (7a) and (7b). Indeed, if $\mu_0 \neq 0$ then one can divide (7a) and (7b) through by μ_0 to yield (25a)–(25b). By way of contradiction, suppose that $\mu_0 = 0$. We first claim that at least one of $\lambda_{G,j}$, $\lambda_{g,\ell}$, ρ_ℓ, κ_ℓ, and β are positive (recall all these coefficients are nonnegative). Indeed, otherwise we have $0_m = \sum_{i=1}^{m} \theta_i \nabla_y \frac{\partial f(x^*, y^*)}{\partial y_i}$, which violates (Q5). We next claim that $\lambda_{G,j}, \lambda_{g,\ell}$, ρ_ℓ, and κ_ℓ are all zero. If not, then multiplying (7a) and (7b) through by the vector d given in the statement of the theorem yields

$$0 = \sum_{j \in A_G(x^*, y^*)} \lambda_{G,j} \nabla G_j(x^*, y^*)^\top d + \sum_{\ell \in A_g(x^*, y^*)} (\lambda_{g,\ell} + \rho_\ell) \lambda_{g,\ell} \nabla g_\ell(x^*, y^*)^\top d$$
$$+ \sum_{\ell \in A_g(x^*, z^*)} \kappa_\ell \nabla_x g_\ell(x^*, z^*)^\top d_x + \sum_{i=1}^{m} \theta_i \nabla \tfrac{\partial f(x^*, y^*)}{\partial y_i}^\top d$$
$$+ \beta(\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*))^\top d_x + \beta \nabla_y f(x^*, y^*)^\top d_y.$$

Then from (Q1)–(Q6) and the fact that not all of $\lambda_{G,j}, \lambda_{g,\ell}$, ρ_ℓ and κ_ℓ are zero, gives a contradiction of $0 > 0$ in the above expression. Note also that since (Q6) holds strictly, we may conclude that $\beta = 0$. Since at least one of λ_{G_j},$\lambda_{g,\ell}$, κ_ℓ, ρ_ℓ and β must be positive, we have already eliminated this possibility. Hence, to avoid contradiction, we assume (Q6) holds with equality and $\beta > 0$.

In fact, we show something stronger, that $\eta_{G,j}^{k,n}$, $\eta_{g,\ell}^{k,n}$, $\omega_\ell^{k,n}$, and $\xi_\ell^{k,n}$ (as defined in (21)) are all bounded. Indeed, suppose otherwise, then dividing (20) by the norm of $(\eta_G^{k,n}, \eta_g^{k,n}, \omega^{k,n}, \xi^{k,n})$ we get a similar contradiction as above when multiplying through by vector d. Given that $\eta_{G,j}^{k,n}$, $\eta_{g,\ell}^{k,n}$, $\omega_\ell^{k,n}$, and $\xi_\ell^{k,n}$ are all bounded, for $\mu_0 = 0$ to happen, the only possibility is either $\gamma^{k,n} \to \infty$ or $||\vartheta^{k,n}|| \to \infty$ or both. There are remaining cases to consider

Case 1: $\gamma^{k,n} = o(||\vartheta^{k,n}||)$

In this case, as $\gamma^{k,n} \to \infty$, we divide the first-order condition $\nabla_y F^{k,n}(x^{k,n}, y^{k,n}) = 0$ by $||\vartheta^{k,n}||$, it remains

$$0 \leftarrow \tfrac{\gamma^{k,n}}{||\vartheta^{k,n}||} [\nabla_y f(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} \eta_{g,\ell}^{k,n} \nabla_y g_\ell(x^{k,n}, y^{k,n})] + \sum_{i=1}^{m} \tfrac{\vartheta_i^{k,n}}{||\vartheta^{k,n}||} \nabla_y \tfrac{\partial}{\partial y_i} f(x^{k,n}, y^{k,n}),$$

where all other terms such as $\frac{\eta_{G,j}^{k,n}}{||\vartheta^{k,n}||}$ or $\frac{\eta_{g,\ell}^{k,n}}{||\vartheta^{k,n}||}$ goes to zero as we have argued that $(\eta_G^{k,n}, \eta_g^{k,n}, \omega^{k,n}, \xi^{k,n})$

are bounded. Therefore, we obtain

$$0 \leftarrow \sum_{i=1}^{m} \frac{\vartheta_i^{k,n}}{||\vartheta^{k,n}||} \nabla_y \frac{\partial}{\partial y_i} f,$$

and therefore contradicts the full rank qualification (Q5).
Case 2: $\gamma^{k,n} = \Omega(||\vartheta^{k,n}||)$

In this case $\gamma^{k,n} \to \infty$ and we divide the first-order condition by $\gamma^{k,n} \to \infty$, it remains

$$0 \leftarrow \nabla f(x^{k,n}, y^{k,n}) - \nabla f(x^{k,n}, z^n) + \sum_{\ell=1}^{p} \eta_{g,\ell}^{k,n} \nabla g_\ell(x^{k.n}, y^{k,n}) + \sum_{i=1}^{m} \frac{\vartheta_i^{k,n}}{\gamma^{k,n}} \nabla \frac{\partial}{\partial y_i} f(x^{k,n}, y^{k,n}),$$

while all other coefficients in $\nabla F^{k,n}(x^{k,n}, y^{k,n})$ are bounded.

First, it must be $\eta_{g,\ell}^{k,n} \to 0$, otherwise, we multiply by vector d from (Q1)–(Q6), there will be a contradiction. Second, it must be $\gamma^{k,n} = \Omega(||\vartheta^{k,n}||)$, otherwise, we have $\nabla f(x^{k,n}, y^{k,n}) - \nabla f(x^{k,n}, z^n) \to 0$, which violates (Q7). Third, we claim that

$$\nabla_y f(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} \eta_{g,\ell}^{k,n} \nabla_y g_\ell(x^{k.n}, y^{k,n}) \to 0. \tag{26}$$

The argument follows from the fact that $y^{k,n} \to y^*$ (as $k \to \infty, n \to \infty$) is an asymptotic maximizer of $f^k(x^{k,n}, y)$ over $y \in Y$. To see this, it suffices to show

$$\lim_{k\to\infty, n\to\infty} \max_{y \in Y} f^k(x^{k,n}, y) \le \lim_{k\to\infty, n\to\infty} f^k(x^{k,n}, y^{k,n}) = f(x^*, y^*).$$

Let $\hat{y}^{k,n} \in \arg\max_{y\in Y} f^k(x^{k,n}, y)$, and denote its subsequential limit as $\hat{y}^\infty$, then $\hat{y}^\infty$ satisfies $g_\ell(x^*, \hat{y}^\infty) \ge 0$ and therefore,

$$\lim_{k\to\infty, n\to\infty} \max_{y \in Y} f^k(x^{k,n}, y) \le \lim_{k\to\infty, n\to\infty} f^k(x^{k,n}, \hat{y}^{k,n}) = f(x^*, \hat{y}^\infty) \le f(x^*, y^*).$$

Therefore, as $\eta_{g,\ell}^{k,n} \to 0$ is bounded as we have argued, we obtain

$$0 \leftarrow \nabla_y f^k(x^{k,n}, y^{k,n}) = \nabla_y f(x^{k,n}, y^{k,n}) + \sum_{\ell=1}^{p} \eta_{g,\ell}^{k,n} \nabla_y g_\ell(x^{k.n}, y^{k,n}).$$

Next from the above result and $\eta_{g,\ell}^{k,n} \to 0$, we also have $\nabla_y f^k(x^{k,n}, y^{k,n}) \to 0$, which implies $\nabla_y f(x^*, y^*) = 0$ and thus $\nabla_y f(x^n, y^n) = \nabla_y f(x^*, y^*) = 0$. Finally, we obtain a contradiction by showing $\gamma^{k,n} \nabla_y f(x^{k,n}, y^{k,n}) \to 0$. To see this, note that from exactness, we have

$$k||\nabla_y f(x^{k,n}, y^{k,n}) - \nabla_y f(x^n, y^n)||^2 = k||\nabla_y f(x^{k,n}, y^{k,n})||^2 \to 0,$$

which implies $||\nabla_y f(x^{k,n}, y^{k,n})|| = o(k^{-1/2})$. By the exactness again, we know that $\gamma^{k,n} = o(k^{1/2})$, therefore,

$$\gamma^{k,n} \nabla_y f(x^{k,n}, y^{k,n}) \to 0.$$

Therefore, returning to the first-order condition $\nabla_y F^{k,n}(x^{k,n}, y^{k,n}) = 0$. Since $||\vartheta^{k,n}|| \to \infty$, we divide the above $\nabla_y F^{k,n}(x^{k,n}, y^{k,n})$ by $||\vartheta^{k,n}||$ and ignore those small terms, it yields

$$\frac{1}{||\vartheta^{k,n}||} \sum_{\ell=1}^{p} \omega_{g,\ell}^{k,n} \nabla_y g_\ell(x^{k.n}, y^{k,n}) + \sum_{i=1}^{m} \frac{\vartheta_i^{k,n}}{||\vartheta^{k,n}||} \nabla_y \frac{\partial f(x^{k,n}, y^{k,n})}{\partial y_i} \to \sum_{i=1}^{m} \frac{\vartheta_i^{k,n}}{||\vartheta^{k,n}||} \nabla_y \frac{\partial f(x^{k,n}, y^{k,n})}{\partial y_i} \to 0,$$

which contradicts (Q5). $\square$

We now state an alternate optimality condition for when (Q7) fails to hold.

Theorem 4. Let (x^*, y^*) be an optimal solution to (BLPP) and $z^* \in S(x^*)$. Suppose Assumptions 1 and 2 and (Q1)–(Q6) hold, but (Q7) does not hold. Then, the negation of (Q7) and (25b) characterize (x^*, y^*). Moreover, if there does not exist a nonzero vector $\hat{d}$ (with $||\hat{d}|| = 1$) such that

$$\hat{d}^\top \nabla^2 G_j(x^*, y^*)^\top \hat{d} = \nabla G_j(x^*, y^*)^\top \hat{d} = 0 \text{ for every } j \in A_G(x^*, y^*), \tag{27a}$$

$$\hat{d}_x^\top \nabla^2_{xx} g_\ell(x^*, z^*)^\top \hat{d}_x = \nabla_x g_\ell(x^*, z^*)^\top \hat{d}_x = 0 \text{ for every } \ell \in A_g(x^*, z^*), \tag{27b}$$

$$\left(\nabla \frac{\partial f(x^*, y^*)}{\partial y_i}\right)^\top \hat{d} = 0 \text{ for every } i = 1, ..., m, \text{ and} \tag{27c}$$

$$\hat{d}^\top \nabla^2 g_\ell(x^*, y^*)^\top \hat{d} = \nabla g_\ell(x^*, y^*)^\top \hat{d} = 0 \text{ for every } \ell \in A_g(x^*, y^*), \tag{27d}$$

then there exist nonnegative multipliers $\lambda_{G,j}$ for $j = 1, \ldots, q$, $\lambda_{g,\ell}$, and κ_ℓ for $\ell = 1, \ldots, p$, $\beta \geq 0$ and unsigned θ_i for $i = 1, \ldots, m$ such that optimality conditions (25) hold, as well as complementary slackness conditions (8) hold.

Proof. In the proof of Theorem 3, note that (Q1-Q6) is sufficient to establish the optimality condition for y by (25b).

To establish the "moreover", we assume, by way of contradiction, that there exist no such optimality condition of the form (25a)-(25b) with $\mu_0 \neq 0$. In particular, we may assume that μ_0 as defined in Theorem 2 is zero.

The argument here is identical to that of Theorem 3 up until the end of Case 1. Thus we are in the situation of case 2: $\gamma^{k,n} = \Omega(||\vartheta^{k,n}||)$. This implies that $\gamma^{k,n} \to \infty$ and the other multipliers $(\eta_G^{k,n}, \eta_g^{k,n}, \omega^{k,n}, \xi^{k,n})$ are bounded. This further implies $\eta_{g,\ell}^{k,n} \to 0$. As we have argued in the Case 2 of the Proof of Theorem 3, the asymptotic optimality condition (26) still holds, and it further implies $\nabla_y f(x^{k,n}, y^{k,n}) \to 0$ and thus $\nabla_y f(x^n, y^n) = 0$.

Now given $\nabla_y f(x^n, y^n) = 0$ and $||\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*)|| = 0$, by the definition

$$\begin{aligned}
\gamma^{k,n} &= -k(\min\{0, f^k(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)\}) \\
&= -k[f(x^{k,n}, y^{k,n}) - \tfrac{k}{2}\sum_{\ell=1}^{p} g_\ell^-(x^{k,n}, y^{k,n})^2 - f(x^{k,n}, z^n)] \\
&= -k[f(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)] + o(1) \\
&= -k(\tfrac{1}{2}(d^{k,n})^\top \nabla^2 (f(x^n, y^n) - f(x^n, z^n))^\top d^{k,n}) \\
&= O(k||d^{k,n}||^2),
\end{aligned}$$

where $d^{k,n} = (x^{k,n} - x^n, y^{k,n} - y^n)$, and the second equality is by the fact $f^k(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)$ must approach zero from below, the third one is by the fact $\eta_{g,\ell}^{k,n} = k g_\ell^-(x^{k,n}, y^{k,n}) \to 0$, and the second last one is by Taylor expension of $f(x^{k,n}, y^{k,n}) - f(x^{k,n}, z^n)$ around (x^n, y^n) and using the fact $f(x^n, y^n) - f(x^n, z^n) = 0$, $\nabla_y f(x^n, y^n) = 0$ and $||\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*)|| = 0$. Therefore, $\gamma^{k,n} \to \infty$ must imply

$$k||d^{k,n}||^2 \to \infty.$$

We thus use the above divergence to show the result. Consider, for example, the coefficient $\eta_{G,j}^{k,n}$ for $j \in A_G(x^*, y^*)$. Since $\eta_{G,j}^{k,n}$ is bounded as k goes to infinity, we have that $\eta_{G,j}^{k,n} = -kG_j(x^{k,n}, y^{k,n})$ for sufficiently large k. By taking the Taylor series expansion of G_j around (x^n, y^n), we have

$$-\frac{\eta_{G,j}^{k,n}}{k||d^{k,n}||} = \frac{\nabla G_j(x^n, y^n)^\top d^{k,n} + o(d^{k,n})}{||d^{k,n}||} \to \nabla G_j(x^*, y^*)^\top \hat{d} = 0,$$

where we let $\hat{d} = \lim_{k\to\infty, n\to\infty} \frac{d^{k,n}}{\|d^{k,n}\|}$ and using the fact that $G_j(x^*, y^*) = 0$ since $j \in A_G(x^*, y^*)$. Similarly, we have

$$\nabla g_\ell(x^*, y^*)^\top \hat{d} = 0 \text{ for } \ell \in A_g(x^*, y^*)$$
$$\nabla_x g_\ell(x^*, z^*)^\top \hat{d}_x = 0 \text{ for } \ell \in A_g(x^*, z^*).$$

It remains to establish $\left(\nabla \frac{\partial f(x^*,y^*)}{\partial y_i}\right)^\top \hat{d} = 0$. Using the fact that $\gamma^{k,n} = \Omega(||\vartheta^{k,n}||)$ and thus $\frac{||\vartheta^{k,n}||}{k\|d^{k,n}\|} \to 0$, we follow the same Taylor expansion to obtain the desired result.

For the second-order expression, again, we consider

$$\eta_{G,j}^{k,n} = -kG_j(x^{k,n}, y^{k,n}) \quad \text{for} j \in A_G(x^*, y^*).$$

Given the $\nabla G_j(x^n, y^n)^\top \frac{d^{k,n}}{\|d^{k,n}\|} \to 0$, we uses the second-order Taylor expansion

$$0 \leftarrow -\frac{\eta_{G,j}^{k,n}}{k\|d^{k,n}\|^2} = \frac{\min\{0, G_j(x^{k,n}, y^{k,n})\}}{\|d^{k,n}\|^2} = \frac{(d^{k,n})^\top \nabla^2 G_j(x^*, y^*)(d^{k,n})}{\|d^{k,n}\|^2} \to \hat{d}^\top \nabla^2 G_j(x^*, y^*)^\top \hat{d}.$$

The other second-order case follows the same reasoning. □

A useful sufficient condition for the above theorem to hold is that the gradient of a binding constraints set $\nabla G(x^*, y^*)$, $\nabla g(x^*, y^*)$, $\nabla_x g(x^*, z^*)$, or the first-order condition $\nabla \frac{\partial f(x^*,y^*)}{\partial y_i}$ has a linearly independent column. Then, we cannot find any nonzero vector to make $\nabla G(x^*, y^*)^\top \hat{d} = 0$ or so. Then, we obtain the desired necessary optimality conditions. Another sufficient condition is that the matrix $(\nabla G_j(x^*, y^*), \nabla g_\ell(x^*, y^*), \nabla_x g_\ell(x^*, z^*), \nabla\nabla_y f(x^*, y^*))$ is of full rank, then there does not exist any nonzero vector $\hat{d}$ to make

$$(\nabla G_j(x^*, y^*), \nabla g_\ell(x^*, y^*), \nabla_x g_\ell(x^*, z^*), \nabla\nabla_y f(x^*, y^*))^\top \hat{d} = 0.$$

Therefore, it ensures that the non-existence of any nonzero vector $\hat{d}$ to make

$$\nabla G_j(x^*, y^*)^\top \hat{d} = \nabla g_\ell(x^*, y^*)^\top \hat{d} = \nabla_x g_\ell(x^*, z^*)^\top \hat{d}_x = \nabla\nabla_y f(x^*, y^*)^\top \hat{d} = 0.$$

Of course, the linear independent condition will imply (Q1-Q6), which can be regarded as a strength of (Q1-Q6). This is an analog to the strengthening from MFCQ to LICQ in the single-level problem. We summarize these results as the following corollary.

Corollary 2. Let (x^*, y^*) be an optimal solution to (BLPP) and $z^* \in S(x^*)$ and suppose Assumptions 1 and 2 hold. If (i) the gradient $(\nabla G_j(x^*, y^*), \nabla g_\ell(x^*, y^*), \nabla_x g_\ell(x^*, z^*), \nabla\nabla_y f(x^*, y^*))$ is of full rank, or (ii) (Q1-Q6) hold and one of binding constraint set $\nabla G_j(x^*, y^*), \nabla g_\ell(x^*, y^*), \nabla_x g_\ell(x^*, z^*)$, or equality constraint $\nabla\nabla_y f(x^*, y^*)$ contains a linear independent column, then there exist non-negative multipliers $\lambda_{G,j}$ for $j = 1, \dots, q$, $\lambda_{g,\ell}$, and κ_ℓ for $\ell = 1, \dots, p$, $\beta \geq 0$ and unsigned θ_i for $i = 1, \dots, m$ such that optimality conditions (25) hold, as well as complementary slackness conditions (8) hold.

Proof. Part (i) is straightforward, by using the contradiction $\mu_0 = 0$. For part (ii), it follows by Theorem 4. □

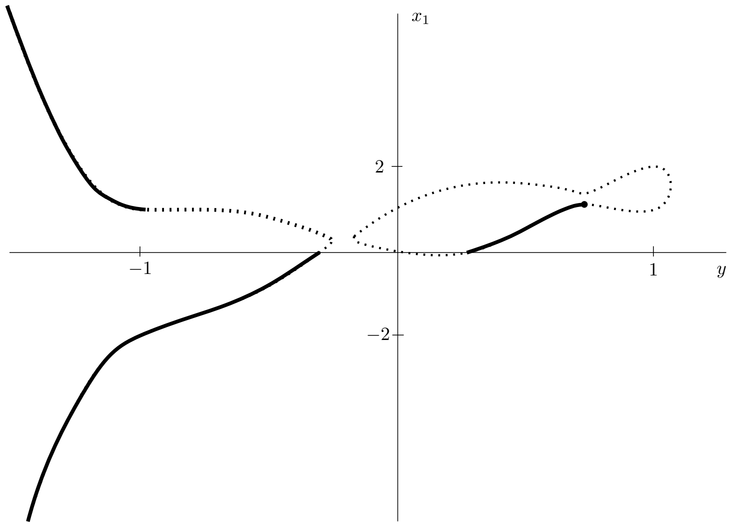

Figure 1: The best response curve (thick line) for Example 2 in terms of x_1 and y. The dashed curve is the locus of stationary points for the lower-level problem. The horizontal axis corresponds to the follower's decision variable y, and the vertical axis corresponds to the leader's first decision variable x_1.

5 Illustrative examples

This section provides three examples of nonconvex bilevel programming problems that illustrate our constraint qualification and necessary optimality condition. These examples are chosen because they fail the constraint qualifications of other optimality conditions found in the literature. Example 2 and Example 1 illustrate how our constraint qualification and optimality conditions apply where the "calmness" conditions of Ye and Zhu [14] do not. The Example 5 is adapted from Ye and Zhu [14] and shows our optimality condition applies when the traditional KKT and value function approaches are invalid.

5.1 Comparison with Theorem 2.1 of Ye and Zhu [14]

Example 2. Consider the following bilevel program:

$$\begin{aligned} \max_{x,y} \quad & -(x_1 - \tfrac{6}{5})^2 + \tfrac{1}{2}(x_2 - \tfrac{y}{2})^2 - (y - \tfrac{1}{4})^2 - \tfrac{1}{2}x_2 \\ \text{subject to} \quad & y \in \arg\max_{y \in Y} -(x_1 - 1)(x_1 - y)y - (y^2 - 1)^2(y^2 - \tfrac{1}{2})^2 \end{aligned}$$

and $Y = [-1.5, 1.5]$.

Observe that the lower-level problem is monotone in x_2, which does not interact with y. The key step is to choose x_1 to implement y. Figure 1 illustrates the best response curve; that is, the graph of $S(x_1)$.

One can verify that the bilevel optimal solution is $(x^*, y^*) = (1, \frac{\sqrt{2}}{4} + \frac{1}{2}, \frac{1}{\sqrt{2}})$. First note that x_2 is chosen on the top of x_1 and y first (unconstrained). This gives the first-order condition

$(x_2 - \frac{y}{2}) = \frac{1}{2}$. Therefore, the problem becomes

$$\begin{aligned}
\min_{x_1,y} \quad & (x_1 - \tfrac{6}{5})^2 + (y - \tfrac{1}{4})^2 - \tfrac{1}{8} + \tfrac{1}{2}(\tfrac{y}{2} + \tfrac{1}{2}) \\
\text{subject to} \quad & y \in \arg\min (x_1 - 1)(x_1 - y)y + (y^2 - 1)^2(y^2 - \tfrac{1}{2})^2.
\end{aligned}$$

The first-order condition curve (they have two branches, upper and lower) are:

$$\begin{aligned}
x_1^u(y) &= y + \tfrac{1}{2} + \tfrac{1}{2}\sqrt{-32y^7 + 72y^5 - 52y^3 + 4y^2 + 8y + 1} \\
x_1^l(y) &= y + \tfrac{1}{2} - \tfrac{1}{2}\sqrt{-32y^7 + 72y^5 - 52y^3 + 4y^2 + 8y + 1}.
\end{aligned}$$

However, only part of the first-order condition curve belongs to the best response. Clearly, the closest branch of the best response curve to the center $(\frac{6}{5}, \frac{3}{8})$ is on the lower branch of the first-order condition curve $\{x_1^l(y) : 0 \leq y \leq \frac{\sqrt{2}}{2}\}$. We can verify that the objective function is minimized along the lower branch at the point $y^* = \frac{\sqrt{2}}{2}$ and thus accordingly $x_1^* = x_1^l(y^*) = 1$. Thus, $x_2^* = \frac{\sqrt{2}}{4} + \frac{1}{2}$.

We will use alternate best response $z^* = -\frac{1}{\sqrt{2}}$. We show that, under these values, our constraint qualification in Theorem 3 holds. Observe that (Q1)–(Q3) are vacuous since there are neither upper nor lower-level constraints. It remains to check (Q4)–(Q7).

Observe that (Q4) amounts to $(\sqrt{2} - 1)d_{x_1} = d_y$. Condition (Q5) amounts to $d_{x_1} \leq 0$. Observe that $\nabla_y f(x^*, y^*) = 0$, as the lower-level problem is unconstrained. We thus have tremendous freedom to choose the vector d. In fact, taking $d_{x_1} = -1$, $d_{x_2} = 1$ and $d_y = 1 - \sqrt{2}$ suffices. Finally, (Q6) is automatically satisfied since there is a single column, and it is nonzero ($\nabla^2_{yy} f(x^*, y^*) = -1$). Also, observe for (Q7) that

$$||\nabla_x f(x^*, y^*) - \nabla_x f(x^*, z^*)|| = ||(-\sqrt{2}, 0)^\top|| \neq 0.$$

Hence, the optimality condition in Theorem 3 applies to this problem and (25) holds. The first-order conditions have

$$\begin{aligned}
\tfrac{2}{5} - \beta\sqrt{2} + \theta(\sqrt{2} - 1) &= 0 \\
-2(\tfrac{\sqrt{2}}{2} - \tfrac{1}{4}) + \tfrac{1}{4} + \beta \cdot 0 - \theta &= 0 \\
(\tfrac{\sqrt{2}}{4} + \tfrac{1}{2} - \tfrac{\sqrt{2}}{4}) - \tfrac{1}{2} + \beta \cdot 0 + \theta \cdot 0 &= 0,
\end{aligned}$$

which gives unique Lagrange multipliers $\theta = \frac{3}{4} - \sqrt{2}$ and $\beta = \frac{1}{\sqrt{2}}(\frac{2}{5} - (\sqrt{2} - 1)(\sqrt{2} - \frac{3}{4})) > 0$.

We now show that this example does not satisfy the constraint qualification set in Theorem 2.1 of Ye and Zhu [14]. To be consistent with Ye and Zhu's notation, we modify the problem into minimization form with

$$F(x, y) = (x_1 - \tfrac{6}{5})^2 + (y - \tfrac{1}{4})^2 - \tfrac{1}{2}(x_2 - \tfrac{y}{2})^2 + \tfrac{1}{2}x_2$$

and

$$f(x, y) = (x_1 - 1)(x_1 - y)y + (y^2 - 1)^2(y^2 - \tfrac{1}{2})^2.$$

The constraint qualification in Theorem 2.1 of Ye and Zhu [14] is that there exists a nonnegative real multiplier μ such that $\nabla F_\mu(x^*, y^*)^\top d \geq 0$ holds for all $d = (d_x, d_y)$ satisfying $\nabla^2_{yx} f(x^*, y^*)d_x + \nabla^2_{yy} f(x^*, y^*)d_y = 0$ where $F_\mu(x, y) = F(x, y) + \mu(f(x, y) - V(x))$ and $V(x) = \min\{f(x, y) : y \in Y\}$.

The optimal solution is $x^* = (1, \frac{\sqrt{2}}{4} + \frac{1}{2})$ and $y^* = \frac{\sqrt{2}}{2}$, $\nabla_{x_1} f(x^*, y^*) = \frac{\sqrt{2}-1}{2}$, $\nabla_{x_1} f(x^*, z^1) = 0$, $\nabla_{x_1} f(x^*, z^2) = -\frac{\sqrt{2}+1}{2}$, and $\nabla_{x_1} f(x^*, z^3) = -2$, where $(z^1, z^2, z^3) = (1, -\frac{\sqrt{2}}{2}, -1)$. Using Danskin's theorem (Proposition 2.1 in Ye and Zhu [14]), we have

$$V'(x^*, d_x) = \min\{\nabla_x f(x^*, y^*)^\top d_x, \nabla_x f(x^*, z^i)^\top d_x\} = \begin{cases} -2d_{x_1} + 0 \cdot d_{x_2} & \text{if } d_{x_1} > 0 \\ \frac{\sqrt{2}-1}{2} d_{x_1} + 0 \cdot dx_2 & \text{if } d_{x_1} \le 0, \end{cases}$$

where $\nabla_{x_2} f(x^*, y^*) = \frac{1}{2}$ is constant. Note that the tangent plane of the first-order condition is

$$\{d \in \mathbb{R}^3 : \nabla^2_{x_1 y} f(x^*, y^*) d_{x_1} + \nabla^2_{x_2 y} f(x^*, y^*) d_{x_2} + \nabla^2_{yy} f(x^*, y^*) d_y = 0\}$$

and amounts to $-(\sqrt{2}-1)d_{x_1} + d_y = 0$, or equivalently $d_y = (\sqrt{2}-1)d_{x_1}$ and d_{x_2} is totally free.

Therefore, we have

$$\begin{aligned} F'_\mu((x^*, y^*), d) &= \nabla F(x^*, y^*)^\top d + \mu[\nabla_x f(x^*, y^*) - V'(x^*, d_x)] d_x \\ &= \begin{cases} \nabla F(x^*, y^*)^\top d + \mu\left((\frac{\sqrt{2}-1}{2} + 2) d_{x_1} + 0 \cdot d_{x_2} \right) & \text{if } d_{x_1} > 0 \\ \nabla F(x^*, y^*)^\top d + \mu\left((\frac{\sqrt{2}-1}{2} - \frac{\sqrt{2}-1}{2}) d_{x_1} + 0 \cdot d_{x_2} \right) & \text{if } d_{x_1} \le 0. \end{cases} \end{aligned}$$

Ye's condition requires that $F'_\mu((x^*, y^*), d) \ge 0$ for any $d_y = (\sqrt{2}-1)d_{x_1}$ and d_{x_2}. Consider the direction $d_{x_1} < 0$, however we have

$$\begin{aligned} \nabla F(x^*, y^*)^\top d &= \nabla_{x_1} F(x^*, y^*) d_{x_1} + \nabla_y F(x^*, y^*) d_y + \nabla_{x_2} F(x^*, y^*) d_{x_2} \\ &= [\nabla_x F(x^*, y^*) + \nabla_y F(x^*, y^*)(\sqrt{2}-1)] d_{x_1} + \nabla_{x_2} F(x^*, y^*) d_{x_2} \\ &= [-\tfrac{2}{5} + (\sqrt{2} - \tfrac{1}{4})(\sqrt{2}-1)] d_{x_1} - 0 \cdot d_{x_2} < 0. \end{aligned}$$

Hence, the constraint qualification of Theorem 2.1 in Ye and Zhu [14] fails.

We now modify Example 2 to satisfy both Ye and Zhu's constraint qualification and ours. The problem is

Example 3. Consider the following bilevel problem:

$$\begin{aligned} \max_{x,y} \quad & -(x-2)^2 - (y - \tfrac{1}{3})^2 \\ \text{subject to} \quad & y \in \arg\max_{y \in Y} -(x-1)(x-y)y - (y^2-1)^2(y^2 - \tfrac{1}{2})^2. \end{aligned}$$

and $Y = [-1.5, 1.5]$.

In this example, the optimal solution is $x^* = 1$ and $y^* = \frac{\sqrt{2}}{2}$. As shown before, there are three alternative best responses given $x^* = 1$. We choose $z^* = -\frac{1}{\sqrt{2}}$. Then

$$\begin{aligned} F'_\mu((x^*, y^*), d) &= \nabla F(x^*, y^*)^\top d + \mu[\nabla_x f(x^*, y^*) - V'(x^*, d_x)] d_x \\ &= \begin{cases} \nabla F(x^*, y^*)^\top d + \mu\left((\frac{\sqrt{2}-1}{2} + 2) d_x \right) & \text{if } d_x > 0 \\ \nabla F(x^*, y^*)^\top d + \mu\left((\frac{\sqrt{2}-1}{2} - \frac{\sqrt{2}-1}{2}) d_x \right) & \text{if } d_x \le 0. \end{cases} \end{aligned}$$

Note that for the tangent plane $-(\sqrt{2}-1)d_x + d_y = 0$, for any direction $d_x < 0$, we have

$$\begin{aligned} \nabla F(x^*, y^*)^\top d &= 2(1-2)d_x + 2(\tfrac{\sqrt{2}}{2} - \tfrac{1}{3}) d_y \\ &= d_x(-2 + 2(\tfrac{\sqrt{2}}{2} - \tfrac{1}{3})(\sqrt{2}-1)) \\ &> 0. \end{aligned}$$

Therefore, Ye and Zhu's constraint qualification is satisfied since there exists some $\mu \geq 0$ such that

$$F'_\mu((x^*, y^*), d) \geq 0$$

is satisfied for all directions d on the tangent plane. Therefore, one can use all alternate best response to characterize (x^*, y^*). However, according to our theory, we only need to use one alternative best response z^*.

The first-order conditions are

$$\begin{aligned} 2 - \beta\sqrt{2} + \theta(\sqrt{2} - 1) &= 0 \\ -2(\tfrac{\sqrt{2}}{2} - \tfrac{1}{3}) + \beta \cdot 0 - \theta &= 0, \end{aligned}$$

which yields

$$\theta = 2(\tfrac{\sqrt{2}}{2} - \tfrac{1}{3})$$

and

$$\beta = \frac{2-(\sqrt{2}-1)(\sqrt{2}-\frac{2}{3})}{\sqrt{2}} > 0.$$

5.2 Comparison with Theorem 3.2 of Ye and Zhu [14]

The following example, first mentioned in Section 1.1, is a constrained version of Example 2 with a single upper-level constraint.

Example 4. Consider the following bilevel program:

$$\begin{aligned} \max_{x,y} \quad & -(x_1 - \tfrac{6}{5})^2 + \tfrac{1}{2}(x_2 - \tfrac{y}{2})^2 - (y - \tfrac{1}{4})^2 - \tfrac{1}{2}x_2 \\ \text{subject to} \quad & \tfrac{\sqrt{2}-1}{2}x_1 - y + 1/2 \geq 0 \\ & y \in \arg\max_{y \in Y} -(x_1 - 1)(x_1 - y)y - (y^2 - 1)^2(y^2 - \tfrac{1}{2})^2. \end{aligned}$$

We have shown that the constraint qualifications in Theorem 3.2 of Ye and Zhu [14] fail, but our constraint qualifications hold in this example.

Since the solution $(x^*, y^*) = (1, \frac{\sqrt{2}}{4} + \frac{1}{2}, \frac{1}{\sqrt{2}})$ satisfies the upper-level constraint $G(x, y) = \frac{\sqrt{2}-1}{2}x_1 - y + 1/2 = 0$ it remains a bilevel optimal solution. Constraint qualification (Q1) requires that $\frac{\sqrt{2}-1}{2}d_{x_1} > d_y$. Carrying over from Example 2 we also require that $(\sqrt{2} - 1)d_{x_1} = d_y$ and $d_{x_1} \leq 0$. Taking $d_{x_1} = -1$, $d_{x_2} = 0$ and $d_y = 1 - \sqrt{2}$ suffices.

5.3 Comparison with classical conditions

Our final example is adapted from Ye and Zhu [14] (Example 4.3) and converted into our format. Ye and Zhu used this example to illustrate how the classic KKT approach and value functions approaches may fail, but the weak calmness condition of Ye and Zhu [14] nonetheless holds. We show that this problem also satisfies our constraint qualification and provides an alternate characterization of the optimal solution. This example demonstrates how our constraint qualification can apply when more classical approaches fail.

Example 5. Consider the following bilevel program:

$$\begin{aligned} \max_{x,y} \quad & -(x - 0.5)^2 - (y - 2)^2 \\ \text{subject to} \quad & y \in S(x) := \arg\max_{y \in [-4,4]} \{3y - y^3 : y \geq x - 3\}. \end{aligned}$$

In terms of our notation, we have $F(x,y) = -(x-0.5)^2 - (y-2)^2$, $f(x,y) = 3y - y^3$, $g(x,y) = -x+y+3$ and $Y = [-4,4]$. Note that G is not defined. In [14], Ye and Zhu argue that $x^* = y^* = 1$ is a bilevel optimal solution. There is a unique alternate best response $z^* = -2$. Observe that both y^* and z^* are in the interior of Y.

We now examine the constraint qualifications of Theorems 3 and 4. First, observe that (Q7) does not hold, and so Theorem 4 applies. It remains to check (Q1) to (Q6). Condition (Q1) holds vacuously. Since $g_\ell(x^*, y^*) \neq 0$, condition (Q2) also holds vacuously.

The remaining conditions concern simultaneously satisfying some inequalities involving d_x and d_y. Condition (Q3) restricts $d_x < 0$ since $g_x(x^*, z^*) = -1$. Condition (Q4) amounts to $-6d_y = 0$ (since $\nabla^2_{xy} f(x,y) = 0$ and $\nabla^2_{yy} f(x^*, y^*) = -6$) and so $d_y = 0$. Condition (Q5) holds trivially since $\nabla f(x^*, y^*) = (0,0)^\top$. Condition (Q6) holds vacuously since $f_y(x^*, y^*) = 0$. Together this implies that d satisfying $d_x < 0$ and $d_y = 0$ works for (Q1) to (Q6).

Finally, we show there does not exist a $\hat{d}$ such that (27) holds. Indeed, from (27b) we have $\nabla_x g(x^*, z^*)^\top \hat{d}_x = 0$, which implies that $-\hat{d}_x = 0$ or $\hat{d}_x = 0$. Moreover, from (27c) we know $\left(\nabla \frac{\partial f(x^*, y^*)}{\partial y_i}\right)^\top \hat{d} = 0$, which implies $\nabla^2_{yy} f(x^*, y^*) \hat{d}_y = -6\hat{d}_y = 0$ or $\hat{d}_y = 0$. This implies $\hat{d}_x = \hat{d}_y = 0$ and so there does not exist a nonzero $\hat{d}$ such that (27) holds.

By Theorem 4, this implies that optimality conditions (25) and complementary slackness conditions (8) hold. Moreover, from (25a) we have: $-1 = -\kappa + \lambda_g + \rho_\ell$ and and from (25b) $-2 = \lambda_g + \rho_\ell - 6\theta$. Any multipliers satisfying these inequalities provide a characterization of optimality.

6 Conclusion

We have presented a novel optimality condition for nonconvex bilevel programs. Our approach uses a penalty function that penalizes a max-min reformulation of the bilevel problem to deliver a "classical" constraint qualification involving only the gradients of the constraint functions (including the follower's objective function f) evaluated at the optimal solution and a single alternate best response, along with a constraint qualification that posits the existence of a single direction vector d appropriately positioned with respect to the gradients of the constraints. This bears a strong resemblance to classical optimality conditions and constraint qualifications in the single-level setting. By contrast, previous approaches to developing constraint qualifications for the bilevel case that involve the follower's objective, many alternate best responses, and whose constraint qualifications typically involve a set of directions perpendicular to some tangent surface. Our examples illustrate how our constraint qualification is milder, and optimality conditions tighter than some of the existing results in the literature.

Acknowledgements

Rongzhu Ke is supported by the Natural Science Fundation of China (No. 72192803, 72192800). Christopher Thomas Ryan is supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant RGPIN-2020-06488. Jin Zhang is supported by the National Science Foundation of China (No. 12222106), Guangdong Basic and Applied Basic Research Foundation (No. 2022B1515020082), and Shenzhen Science and Technology Program (No. RCYX20200714114700072). We also thank Yao Wei Yang Peixuan for their able assistance.

References

[1] D.P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 1999.

[2] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–56, 2007.

[3] S. Dempe. *Foundations of Bilevel Programming.* Springer, 2002.

[4] S. Dempe, J. Dutta, and B.S. Mordukhovich. New necessary optimality conditions in optimistic bilevel programming. *Optimization*, 56(5-6):577–604, 2007.

[5] S. Dempe and A.B. Zemkoho. The generalized Mangasarian-Fromowitz constraint qualification and optimality conditions for bilevel programs. *Journal of Optimization Theory and Applications*, 148(1):46–68, 2011.

[6] S. Dempe and A.B. Zemkoho. The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. *Mathematical Programming*, pages 1–27, 2013.

[7] Lei Guo, Jane J Ye, and Jin Zhang. Sensitivity analysis of the maximal value function with applications in nonconvex minimax programs. *Mathematics of Operations Research*, 2023.

[8] Y. Ishizuka and E. Aiyoshi. Double penalty method for bilevel optimization problems. *Annals of Operations Research*, 34(1):73–88, 1992.

[9] G.S. Liu, J.Y. Han, and J.Z. Zhang. Exact penalty functions for convex bilevel programming problems. *Journal of Optimization Theory and Applications*, 110(3):621–643, 2001.

[10] P. Marcotte and D.L. Zhu. Exact and inexact penalty methods for the generalized bilevel programming problem. *Mathematical Programming*, 74(2):141–157, 1996.

[11] Z. Meng, C. Dang, R. Shen, and M. Jiang. An objective penalty function of bilevel programming. *Journal of Optimization Theory and Applications*, 153(2):377–387, 2012.

[12] J.A. Mirrlees. The theory of moral hazard and unobservable behaviour: Part I. *The Review of Economic Studies*, 66(1):3–21, 1999.

[13] J.J. Ye and D.L. Zhu. Optimality conditions for bilevel programming problems. *Optimization*, 33(1):9–27, 1995.

[14] J.J. Ye and D.L. Zhu. New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches. *SIAM Journal on Optimization*, 20(4):1885, 2010.

A Justification of Assumption 1

We now justify Assumption 1 in the single-dimensional case $x \in \mathbb{R}$ first. Please note that this proof requires quite a few ideas from the development in Section 3, so it is best understood after reading this section.

First, note that if x^* is a stationary point of $F(x, y(x))$ (where $y(x)$ is as defined after (9)) then we can directly derive the Fritz-John and KKT necessary optimality conditions for (BLPP) without the local min-max relaxation technique. See the proof of Proposition 1(i) for details of this.

Thus, we may assume that x^* is not a stationary point of $F(x, y(x))$, we can choose ϵ small enough so that for $(x, y(x)) \in \mathcal{N}_\epsilon(x^*, y^*)$ so that for $x \in \mathcal{N}_\epsilon(x^*)$, $F(x, y(x))$ does not have any stationary point (i.e., $F(x, y(x))$ is strictly monotone if x is single-dimensional). We now derive a contradiction.

We start by defining the following useful notion.

Definition 1 (COI)**.** A function $\Phi : X \to \mathbb{R}$ is in *conflict of interest (COI)* with $\Psi : X \to \mathbb{R}$ at $\tilde{x} \in X$, if there exists an $\epsilon > 0$, $\Phi(x)$ is strictly increasing (decreasing) and $\Psi(x)$ is strictly decreasing (increasing) on $x \in \mathcal{N}_\epsilon(\tilde{x})$.

Let $\epsilon > 0$ and require without loss that our sequence ϵ^n defined in (11) has $\epsilon_n < \epsilon$. We define (x^n, y^n, z^n) as before in (12), i.e., z^n is an infimizing sequence, and (x^n, y^n) is a corresponding inner solution sequence of the inner maximization problem given z^n. Let $(x^\infty, y^\infty, z^\infty)$ be the limit of the subsequence of the (x^n, y^n, z^n) (possibly restricting to a subsequence).

Now given z^n, $f(x, y(x)) - f(x, z^n)$ must be in COI with $F(x, y(x))$ for $x \in \mathcal{N}_\epsilon(x^n)$. If they are not in COI around x^* then we have the necessary existence of Assumption 1 at the boundary. Thus, we may assume $\frac{d}{dx}[f(x, y(x)) - f(x, z^n)] > 0$ if $\frac{d}{dx}F(x, y(x)) < 0$, and vise versa. We assume $F(x, y(x))$ is increasing without loss.

We first claim $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ and provide proof in the following lemma. The statement of the proof requires the following definition. We say a continuously differentiable function $f : X \to \mathbb{R}$ *oscillates infinitely* at x^∞ if x^∞ is the nonisolated zero of the function $f(\cdot) - f(x^\infty)$.

Lemma 6. Suppose that for every choice of $\underline{\epsilon}$ there exists an $0 < \epsilon < \underline{\epsilon}$ such that the problem $\min_{z \in Y} \Theta(z, \epsilon)$ does not have an optimal solution. If $f(x, y(x)) - f(x, z^\infty)$ does not oscillate infinitely at x^∞, we have $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ for all $x \in [x^\infty, x^\infty + \epsilon']$ for some $\epsilon' > 0$, and if $f(x, y(x)) - f(x, z^\infty)$ oscillates infinitely at x^∞, then $f(x, y(x)) - f(x, z^\infty) \leq 0$ for all $x \in [x^\infty, x^\infty + \epsilon']$ and there is a sequence $\bar{x}_n \to x^\infty$, such that $f(x, y(x)) - f(x, z^\infty) = 0$ and $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ at $x = \bar{x}_n$ for all n.

Proof. We consider the case where $F(x, y(x))$ is strictly increasing in $x \in \mathcal{N}_\epsilon(x^*)$, and the other case is similar. Since $\frac{d}{dx}F(x, y(x)) > 0$, the COI condition implies $\frac{d}{dx}[f(x, y(x)) - f(x, z^n)] < 0$ for $x \in \mathcal{N}_\epsilon(x^n)$. Indeed, in this case

$$x^n = \max\{x \in \mathcal{N}_\epsilon(x^*) : f(x, y(x)) - f(x, z^n) \geq 0\}.$$

If at the limit $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] < 0$ for $x \in \mathcal{N}_\epsilon(x^\infty)$, for the non-existence to happen, it must be that $f(x, y(x)) - f(x, z^\infty)$ will increase and cross x-axis again at some point $x' \in (x^\infty + \epsilon', x^* + \epsilon)$, where $x^* - x^\infty + \epsilon \geq \epsilon' > 0$. We consider the problem

$$\Theta(z, \epsilon') = \max_{(x,y) \in \mathcal{N}_{(x^\infty + \epsilon' - x^*)}(x^*, y^*)} \{F(x, y) : f(x, y) - f(x, z) \geq 0, \nabla_y f(x, y) = 0\}.$$

It is clearly that

$$\inf_z \Theta(z, (x^\infty + \epsilon' - x^*)) = \inf_z \Theta(z, \epsilon) = \Theta^*,$$

since $(x^n, y^n) \in \mathcal{N}_{(x^\infty + \epsilon' - x^*)}(x^*, y^*)$ is feasible for the infimizing sequence z^n. If there is a better sequence for $\inf_z \Theta(z, (x^\infty + \epsilon' - x^*)) < \Theta^*$, it should be also a sharper infimizing sequence for $\inf_z \Theta(z, \epsilon)$, a contradiction. Therefore, for the new minimax problem, as $z^n \to z^\infty$ and

$$x^n \to x^\infty = \max\{x \in \mathcal{N}_{(x^\infty + \epsilon' - x^*)}(x^*) : f(x, y(x)) - f(x, z^\infty) \geq 0\}$$

is attainable. We obtain $\Theta^* = \Theta(z^\infty, \epsilon)$, a contradiction.

Similarly, we can show the case $\frac{d}{dx}F(x, y(x)) < 0$ for $(x, y(x)) \in \mathcal{N}_\epsilon(x^*, y^*)$.

Now, we discuss two cases regarding constraint function.

Case 1. If $f(x, y(x)) - f(x, z^\infty)$ does not oscillate infinitely over $x \in [x^\infty, x^\infty + \epsilon']$ for some $\epsilon' > 0$.

In this case, $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)]$ cannot change the sign for $x \in (x^\infty, x^\infty + \epsilon')$ for arbitrarily small $\epsilon' > 0$. Then $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ at $x = x^\infty$. However, if $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ at $x = x^\infty$ and $f(x, y(x)) - f(x, z^\infty) = 0$ also has a local unique solution at $x = x^\infty$, we also have the continuity of

$$x(z^n) := \max\{x \in \mathcal{N}_\epsilon(x^*) : f(x, y(x)) - f(x, z^n) \geq 0\},$$ [2]

which implies the attainability of minimum.

Case 2. If $f(x, y(x)) - f(x, z^\infty)$ oscillates infinitely over $x \in [x^\infty, x^\infty + \epsilon']$.

In this case, $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)]$ changes the sign for $x \in (x^\infty, x^\infty + \epsilon')$ for arbitrarily small $\epsilon' > 0$. But the amplitude of $f(x, y(x)) - f(x, z^\infty)$ is getting smaller and smaller and $f(x, y(x)) - f(x, z^\infty) \leq 0$ over $x \in [x^\infty, x^\infty + \epsilon']$, then we can find a sequence $\bar{x}_n \to x^\infty$ such that $\frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$ and $f(x, y(x)) - f(x, z^\infty) = 0$ at $x = \bar{x}_n$.

Therefore, if $f(x, y(x)) - f(x, z^\infty)$ does not oscillate infinitely, the case for the non-existence to happen is

$$\tfrac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$$

for all $x \in [x^\infty, x^\infty + \epsilon']$ for arbitrarily small $\epsilon' > 0$. If $f(x, y(x)) - f(x, z^\infty)$ oscillates infinitely, we can find the sequence $\bar{x}_n \to x^\infty$, then the case for the non-existence to happen is $f(x, y(x)) - f(x, z^\infty) \leq 0$ for all $x \in [x^\infty, x^\infty + \epsilon']$ and

$$f(x, y(x)) - f(x, z^\infty) = 0 \quad \text{and} \quad \frac{d}{dx}[f(x, y(x)) - f(x, z^\infty)] = 0$$

at $x = \bar{x}_n$. □

Based on Lemma 6, in either case, we derive a contradiction. Let $z^\# \in \arg\max_{z \in Y} f(x^\infty, z)$. Without loss of generality, we assume that $f(x, y(x)) - f(x, z^\#)$ does not cross the x-axis from below at any $x \in [x^\infty, x^\infty + \epsilon']$. Note that $f(x^\infty, y(x^\infty)) - f(x^\infty, z^\#) < 0$ and $f(x^*, y(x^*)) - f(x^*, z^\#) \geq 0$. Therefore,

$$f(x, y(x)) - f(x, z^\#) < 0, \qquad \text{for any} \quad x^\infty + \epsilon' \geq x \geq x^\infty.$$

Therefore, when we choose $z^\#$ instead of $z^n \to z^\infty$, the feasible choice of $x < x^\infty$. Therefore, $x(z^\#) < x^\infty \leq x^n$ and

$$F(x(z^\#), y(x(z^\#))) < F(x^\infty, y(x^\infty)) = \Theta^* = \inf_z \Theta(z, \epsilon) = \inf_z \Theta(z, (x^\infty + \epsilon' - x^*)),$$

which contradicts the definition of Θ^*. Therefore, we rule out the non-existence.

Next, we generalize the existing result to a more general setting. If there are constraints, we may consider the following problem, which is equivalent to (Min-Max|ϵ).

$$\begin{aligned}
\min_{z \in Y} \max_{(x,y) \in \mathcal{N}_\epsilon(x^*, y^*)} \quad & F(x, y) \\
\text{subject to} \quad & f^I(x, y) - f^I(x, z) \geq 0 \\
& \nabla_y f(x, y) = \nabla_y f(x^*, y^*) \\
& G(x, y) \geq 0,
\end{aligned}$$

[2] Such a maximum exists before x is single-dimensional and the underlying set is closed. Excuse also the slight abuse of the notation $x(\cdot)$.

where

$$f^I(x,y) = f(x,y) \text{ if } g(x,y) \geq 0 \text{ and } -\infty \text{ otherwise}$$

is the characteristic function of the lower problem.

If (x^*, y^*) is a stationary point of the $F(x, y(x))$ we are done. Suppose that (x^*, y^*) is not a stationary point of $F(x, y(x))$ for small ball $x \in \mathcal{N}_\epsilon(x^*)$. We now derive a contradiction of the fact that the minimax problem does not have a solution. We have the definition of $(x^\infty, y^\infty, z^\infty)$ as before, then we define

$$\eta^n \in \arg \max_{\eta\in(-\epsilon,\epsilon)} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : f^I(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) - f^I(x^\infty + \eta d_x, z^n) \geq 0,$$
$$g(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) \geq 0, G(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) \geq 0\}.$$

If the non-existence occurs, then there at least exists a direction d_x such that the following problem

$$\min_{z\in Y} \max_{\eta\in[-\epsilon,\epsilon]} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : f^I(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) - f^I(x^\infty + \eta d_x, z^n) \geq 0,$$
$$g(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) \geq 0, G(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) \geq 0\}.$$

does not exist a solution, i,e, there must exist a direction d_x such that constraint functions are flat along that direction. To understand that, we define

$$\begin{aligned} \phi(\eta, z) =: \quad & \min\{f^I(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) - f^I(x^\infty + \eta d_x, z), \\ & G(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)), g(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))\}. \end{aligned}$$

Note that by similar reasoning as in the single-dimensional case, it is impossible for $\phi(\eta, z^\infty)$ to cross η-axis from below at some point $\eta' \in (0, \epsilon)$. (otherwise, $\eta = 0$ as the minimizer is attainable). Hence, for the non-existence to occur, it must be $\phi(\eta, z^\infty) = 0$ for a positive measure $\eta \in [0, \epsilon']$.

Therefore, it is equivalent to have

$$\eta^n \in \arg \max_{\eta\in(-\epsilon,\epsilon)} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : \phi(\eta, z^n) \geq 0\},$$

and $\eta^n \to 0$ by the definition of maximum.

Meanwhile, the non-existence implies that $\eta = 0$ is not a local maximum of $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ given direction d_x. We discuss two cases.

Case 1. If $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ does not oscillate infinitely over $\eta \in [-\epsilon, \epsilon]$.

In this case, $\eta = 0$ is a local minimum of $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ for given direction d_x or $F(x, y(x))$ is strictly monotone along that direction d_x.

Then if $F(x, y(x))$ is strictly increasing or decreasing along that direction, that is,

$$\tfrac{d}{d\eta} F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) > 0 \text{ or } < 0$$

for $\eta \in (-\epsilon, \epsilon)$ where $\epsilon > 0$ is small. Without loss of generality, we consider the "> 0" case. The case "< 0" is similar.

Next, we consider the single-dimension problem given d_x and derive a contradiction. First, for the sequence $z^n \to z^\infty$, the non-existence implies that $\phi(\eta, z^n)$ is decreasing in $\eta \in (-\epsilon, \epsilon)$, otherwise, the problem

$$\max_{\eta\in[-\epsilon,\epsilon]} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : \phi(\eta, z^n) \geq 0\}$$

will have a solution $\eta^n = \epsilon$, which never converge to 0, a contradiction. Next, we choose $z^\# \in \arg\max_{z\in Y} f^I(x^\infty, z)$, then

$$\phi(\eta, z^\#) < 0 = \phi(\eta, z^\infty)$$

for any $\eta \in [0, \epsilon']$ and small $\epsilon > \epsilon' > 0$. And $\phi(\eta, z^\#)$ is decreasing in $\eta \in (-\epsilon', \epsilon')$. Therefore, when choosing $z^\#$, we obtain a smaller solution

$$0 < \eta^\# \in \arg\max_{\eta\in[-\epsilon',\epsilon']} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : \phi(\eta, z^\#) \geq 0\}.$$

As a result, $F(x^\infty + \eta^\# d_x, y^*(x^\infty + \eta^\# d_x)) < F(x^\infty, y^*(x^\infty))$, which is a contradiction with the definition of infimum $F(x^\infty, y^*(x^\infty))$ as $\eta^n \to 0$.

If $\eta = 0$ is a local minimum of $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ for given direction d_x. We consider $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ is increasing in $\eta \in [0, \epsilon)$ and decreasing in $\eta \in (-\epsilon, 0]$. Then $\phi(\eta, z)$ is decreasing in $\eta \in [0, \epsilon)$ and increasing in $\eta \in (-\epsilon, 0]$, that is, $\eta = 0$ is a local maximum of constraint function $\phi(\eta, z)$.

As the sequence $z^n \to z^\infty$, then η^n converge to 0. If the non-existence occurs, it must be $\phi(\eta, z^\infty) = 0$ for a positive measure $\eta \in [0, \epsilon']$, then the problem

$$\max_{\eta\in[0,\epsilon']} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : \phi(\eta, z^\infty) = 0\}$$

will have a solution $\eta = \epsilon'$, that is,

$$F(x^\infty + \epsilon' d_x, y^*(x^\infty + \epsilon' d_x)) > F(x^\infty, y^*(x^\infty)),$$

which is contradiction with $(x^\infty, y^*(x^\infty))$ is a corresponding inner solution of inner maximization problem given z^∞.

Case 2. If $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ oscillates infinitely over η for any sufficiently small interval $[-\epsilon, \epsilon]$.

In this case, firstly, although it is difficult to hold the sign of $\frac{d}{d\eta} F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ unchanged. But we can find a sequence of $\bar{\eta}_n \to 0^+$ such that $F(x^\infty + \bar{\eta}_n d_x, y^*(x^\infty + \bar{\eta}_n d_x)) = \max_{\eta\in[0,\bar{\eta}_n]} F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ is a monotone sequence (increasing or decreasing) and the sequence $\phi(\bar{\eta}_n, z^\#)$ is also monotone but goes the opposite direction as $F(x^\infty + \bar{\eta}_n d_x, y^*(x^\infty + \bar{\eta}_n d_x))$. We only consider the case where $F(x^\infty + \bar{\eta}_n d_x, y^*(x^\infty + \bar{\eta}_n d_x))$ is an increasing sequence, and the logic for the decreasing sequence is similar. Then $\phi(\bar{\eta}_n, z^\infty) \leq 0$ and we can argue that $\phi(\bar{\eta}_n, z^\infty)$ should be flat over the sequence $\bar{\eta}_n$ for the non-existence to occur, i.e., $\phi(\bar{\eta}_n, z^\infty) = 0$ for any sufficiently large $n \geq n_0$. If so, we choose $z^\# \in \arg\max_{z\in Y} f^I(x^\infty, z)$, then

$$\phi(\eta, z^\#) < 0 = \phi(\bar{\eta}_n, z^\infty)$$

for any $\eta \in [0, \epsilon']$ and small $\epsilon > \epsilon' > 0$. And $\phi(\eta, z^\#)$ is decreasing in $\eta \in (-\epsilon', \epsilon')$. Therefore, when choosing $z^\#$, we obtain a smaller solution

$$0 < \eta^\# \in \arg\max_{\eta\in(-\epsilon',\epsilon')} \{F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x)) : \phi(\eta, z^\#) \geq 0\}.$$

As a result, $F(x^\infty + \eta^\# d_x, y^*(x^\infty + \eta^\# d_x)) < F(x^\infty, y^*(x^\infty))$. Note that a contradiction with the definition of infimum $F(x^\infty, y^*(x^\infty))$ as $\eta^n \to 0$.

Next, if $\eta = 0$ is a local minimum of $F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$, we can also find a sequence of $\bar{\eta}_n \to 0^+$ such that $F(x^\infty + \bar{\eta}_n d_x, y^*(x^\infty + \bar{\eta}_n d_x)) = \max_{\eta\in[0,\bar{\eta}_n]} F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ is a increasing sequence and $F(x^\infty - \bar{\eta}_n d_x, y^*(x^\infty - \bar{\eta}_n d_x)) = \max_{\eta\in[-\bar{\eta}_n,0]} F(x^\infty + \eta d_x, y^*(x^\infty + \eta d_x))$ is a decreasing seqeunce. Then we can also get $\eta = 0$ is a local maximum of $\phi(\eta, z)$.

Similar to the Case 1, if the non-existence occurs, we can derive a contradiction with $(x^\infty, y^*(x^\infty))$ being the inner solution of given z^∞.

Based on the above verification, we provide a sufficient condition for Assumption 1 via the following proposition.

Proposition 2. Suppose x^* is not a stationary point of $F(x, y(x))$ and Assumption 2 holds. Then, for any sufficiently small ϵ problem $\min_{z\in Y} \Theta(z, \epsilon)$ possesses an optimal solution.

As we mentioned earlier, this suffices for taking Assumption 1 since if x^* is a stationary point of $F(x, y(x))$, then we have direct methods to derive necessary optimality conditions.