Bilevel programs: directional constraint qualifications and necessary optimality conditions

Jane Ye

University of Victoria, Canada

Lecture 3 at the Forum on Developments and Origins of Operations Research
November 26, 2021
Organizers: The Mathematical Programming Branch of OR Society of China
& Southern University of Science and Technology

Outline

1. Directional optimality conditions

2. Application to bilevel programming

Bilevel program

$$\text{(BP)} \quad \min_{x,y} \quad F(x,y)$$
$$s.t. \quad y \in S(x),$$

where $S(x)$ denotes the solution set of the lower level program

$$(P_x) \quad \min_y f(x,y) \quad \text{s.t. } g(x,y) \leq 0,$$

where $F, f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are continuously differentiable.

Recall the following reformulations using the value function $V(x) := \inf_y \{f(x, y) : g(x, y) \leq 0\}$:

$$(VP) \qquad \min_{x,y} \quad F(x, y)$$
$$s.t. \quad f(x, y) - V(x) \leq 0,$$
$$g(x, y) \leq 0,$$

$$(CP) \qquad \min_{x,y,u} F(x, y)$$
$$\text{s.t.}\ f(x, y) - V(x) \leq 0,$$
$$\nabla_y f(x, y) + \nabla_y g(x, y)^T u = 0,$$
$$g(x, y) \leq 0, \ u \geq 0, \ u^T g(x, y) = 0.$$

If the value function is Lipschitz continuous, then (VP) and (CP) can be written as the nonsmooth optimization problem:

$$(P) \qquad \min \quad f(x) \qquad s.t. \quad g(x) \leq 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is smooth and $g : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous.

Regular and Limiting Subdifferentials

- For a function $\varphi : \mathbb{R}^n \to [-\infty, \infty]$ and a point $\bar{x} \in \mathbb{R}^n$ where $\varphi(\bar{x})$ is finite, the regular (or Fréchet) subdifferential of φ at $\bar{x}$:

$$\widehat{\partial}\varphi(\bar{x}) := \left\{\xi : \varphi(x) \geq \varphi(\bar{x}) + \xi^T(x - \bar{x}) + o(\|x - \bar{x}\|) \ \forall x\right\}.$$

- The regular subdifferential may be empty. The limiting/Mordukhovich subdifferential of φ at $\bar{x}$ defined by

$$\partial\varphi(\bar{x}) := \left\{\lim_k \xi^k | x^k \to \bar{x}, \varphi(x^k) \to \varphi(\bar{x}),\ \xi^k \in \widehat{\partial}\varphi(x^k)\right\}$$

is more robust.

- If φ is Lipschitz continuous around $\bar{x}$, the Clarke subdifferential

$$\partial^c\varphi(\bar{x}) = \mathrm{co}\varphi(\bar{x}).$$

Theorem (Nonsmooth necessary optimality condition)

If $\bar{x}$ is a local minimum for (P), then FJ condition holds:
$\exists(\lambda_0, \lambda) \neq 0$ *such that*

$$0 \in \lambda_0 \nabla f(\bar{x}) + \partial g(\bar{x})^T \lambda,$$
$$\lambda \perp g(\bar{x}), \quad \lambda_0, \lambda \geq 0.$$

If there is no nonzero abnormal multipliers (nonsmooth MFCQ):

$$\begin{cases} 0 \in \partial g(\bar{x})^T \lambda, \\ 0 \leq \lambda \perp g(\bar{x}), \end{cases} \implies \lambda = 0$$

then KKT condition holds: there exists $\lambda \geq 0$ such that

$$0 \in \nabla f(\bar{x}) + \partial g(\bar{x})^T \lambda,$$
$$\lambda \perp g(\bar{x}).$$

But the nonsmooth MFCQ/NNAMCQ never hold for problem (VP) and (CP)! So we need weaker constraint qualificaitons.

Clarke calmness condition

Definition (Clarke)

Suppose that $\bar{x}$ be a local optimal solution of (P). We say (P) is Clarke calm at $\bar{x}$ if $\bar{x}$ also solves the penalized problem

$$\min \; f(x) + \rho \|g_+(x)\|$$
$$s.t. \; x \in \bar{x} + \mathbb{B}_\epsilon$$

for some $\rho, \epsilon \geq 0$.

- Under the Clarke calmness condition, a local optimal solution must satisfy the KKT condition.

Definition (Metric subregularity/local error bound/calmness)

The set-valued map $G(x) := -g(x) + \mathbb{R}^m_-$ *is metrically subregular at* $(\bar{x}, 0) \in gphG$ *if there are positive reals* $\epsilon > 0$ *and* $\kappa > 0$ *such that*

$$\mathrm{dist}(x, G^{-1}(0)) \leq \kappa \|g_+(x)\| \quad \forall x \in \bar{x} + \mathbb{B}_\epsilon,$$

where $G^{-1}(0) := \{x | g(x) \leq 0\}$.

- The set-valued map $G(x) := -g(x) + \mathbb{R}^m_-$ is metrically subregular at $(\bar{x}, 0)$ if and only if the perturbed feasible map $\mathcal{F}(v) := \{x | g(x) \leq v\}$ is calm at $(0, \bar{x}) \in \mathrm{gph}\mathcal{F}$.

Sufficient conditions for metric subregularity

Definition (Quasi-normality; Guo, JY and Zhang, 2013)

Let $g(\bar{x}) \leq 0$. g is Lipschitz continuous. We say the quasi-normality holds at $\bar{x}$ if

$$\begin{array}{l} 0 \in \partial g(\bar{x})^T \lambda, \quad 0 \leq \lambda \perp g(\bar{x}), \\ \exists x^k \to \bar{x} \text{ s.t. } g_i(x^k) > 0 \text{ if } \lambda_i > 0 \end{array} \Longrightarrow \lambda = 0$$

Constraint qualifications and their relationships

Nonsmooth MFCQ/NNAMCQ

$\Longrightarrow$ quasi-normality

$\Longrightarrow$ metric subregularity/calmness

$\Longrightarrow$ Clarke calmness

$\Longrightarrow$ KKT condition for a local minimum.

Constraint qualifications for (VP) or (CP)

Since Nonsmooth MFCQ/NNAMCQ does not hold for (VP) and (CP), we have

quasi-normality

$\Longrightarrow$ **metric subregularity/calmness**

$\Longrightarrow$ **Clarke calmness**

$\Longrightarrow$ **KKT condition for a local minimum.**

KKT condition for (VP)

Theorem (JY and Zhu 1995)

Let $(\bar{x}, \bar{y})$ *be a local minimizer of* (BP). *Suppose a constraint qualification for problem (VP) holds at* $(\bar{x}, \bar{y})$, *MFCQ holds at each* $y \in S(\bar{x})$ *and the feasible region for lower level problem is uniformly bounded. Then there exist* $\lambda^i \geq 0, \sum_{i=1}^{n+1} \lambda^i = 1, y^i \in S(\bar{x})$, $(\lambda, \mu, \mu_g^i) \in \mathbb{R}_+^{1+p+p}$ *such that*

$$0 = \nabla_x F(\bar{x}, \bar{y}) + \lambda \left(\nabla_x f(\bar{x}, \bar{y}) - \sum_{i=1}^{n+1} \lambda^i (\nabla_x f(\bar{x}, y^i) + \nabla_x g(\bar{x}, y^i)^T \mu_g^i) \right)$$

$$+ \nabla_x g(\bar{x}, \bar{y})^T \mu,$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) + \lambda \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^T \mu,$$

$$0 = \nabla_y f(\bar{x}, y^i) + \nabla_y g(\bar{x}, y^i)^T \mu_g^i, \quad \mu_g^i \perp g(\bar{x}, y^i),$$

$$\mu \perp g(\bar{x}, \bar{y}).$$

Proof

There exist $\lambda \geq 0$ and $\mu \geq 0$ such that

$$\begin{aligned} 0 \;\; &\in \;\; \nabla F(\bar{x},\bar{y}) + \lambda(\nabla f(\bar{x},\bar{y}) - \partial^c V(\bar{x}) \times \{0\}) + \nabla g(\bar{x},\bar{y})^T \mu, \\ & \;\;\;\; \mu \perp g(\bar{x},\bar{y}). \end{aligned}$$

Under the assumptions, V is Lipschitz continuous and

$$\partial^c V(\bar{x}) \subseteq co\{\nabla_x f(\bar{x},y) + \nabla_x g(\bar{x},y)^T \lambda : y \in S(\bar{x}), \lambda \in \mathrm{KT}(\bar{x},y)\}.$$

Hence there exist $\lambda^i \geq 0, \sum_{i=1}^{n+1} \lambda^i = 1,\ y^i \in S(\bar{x}), \mu_g^i \in KT(\bar{x},y^i)$ such that

$$\xi \in \partial^c V(\bar{x}) \Rightarrow \xi = \sum_{i=1}^{n+1} \lambda^i(\nabla_x f(\bar{x},y^i) + \nabla_x g(\bar{x},y^i)^T \mu_g^i).$$

Definition (Inner Semicontinuity)

Given $\bar{y} \in S(\bar{x})$, we say that the set-valued map $S(x) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is inner semicontinuous at $(\bar{x}, \bar{y})$ if for any sequences $x^k \to \bar{x}$, there exists a sequence $y^k \in S(x^k)$ converging to $\bar{y}$.

If $S(x)$ is single valued, then the concept of inner semicontinuity reduced to the continuity of the function.

KKT condition for (VP), inner semicontinuous case

Theorem (cf. Theorem 6.21 of Mordukhovich 2018)

Let $(\bar{x}, \bar{y})$ be a local minimizer of (BP). Suppose that $S(x)$ is inner semicontinuous at $(\bar{x}, \bar{y})$, a constraint qualification for (VP) holds at $(\bar{x}, \bar{y})$, and MFCQ for problem $(P_{\bar{x}})$ holds at $\bar{y}$. Then there exists a vector $(\lambda, \mu, \mu_g) \in \mathbb{R}_+^{1+p+p}$ satisfying

$$
\begin{aligned}
0 &= \nabla_x F(\bar{x}, \bar{y}) + \nabla_x g(\bar{x}, \bar{y})^T(\mu - \lambda \mu_g),\\
0 &= \nabla_y F(\bar{x}, \bar{y}) + \lambda \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^T \mu, \quad \mu \perp g(\bar{x}, \bar{y}),\\
0 &= \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^T \mu_g, \mu_g \perp g(\bar{x}, \bar{y}),
\end{aligned}
$$

Proof. There exist $\lambda \geq 0, \mu \geq 0$ such that KKT condition holds:

$$
\begin{aligned}
0 \quad \in \quad & \nabla F(\bar{x}, \bar{y}) + \lambda(\nabla f(\bar{x}, \bar{y}) - \partial^c V(\bar{x}) \times \{0\}) + \nabla g(\bar{x}, \bar{y})^T \mu,\\
& \mu \perp g(\bar{x}, \bar{y}).
\end{aligned}
$$

$$
\partial^c V(\bar{x}) \subseteq \left\{ \nabla_x f(\bar{x}, \bar{y}) + \nabla_x g(\bar{x}, \bar{y})^T \mu_g : \mu_g \in \mathrm{KT}(\bar{x}, \bar{y}) \right\}.
$$

Reall that if there is no nonzero abnormal multipliers:

$$\begin{cases} 0 \in \partial g(\bar{x})^T \lambda, \\ 0 \le \lambda \perp g(\bar{x}), \end{cases} \implies \lambda = 0$$

then there exists $\lambda \ge 0$ such that

$$0 \in \nabla f(\bar{x}) + \partial g(\bar{x})^T \lambda,$$
$$\lambda \perp g(\bar{x}).$$

If $\partial g(\bar{x})$ can be replaced by a smaller set, then the NNAMCQ becomes weaker and KKT condition becomes sharper. This motivates us to consider directional constraint qualification and directional KKT condition.

- Consider problem (P) : $\min f(x)$ s.t. $g(x) \leq 0$, where f is smooth, $g = (g_1, \ldots, g_m)$ and g_i is directionally differentiable and Lipschitz continuous at $\bar{x}$.
- Let $\bar{I}_g := \{i | g_i(\bar{x}) = 0\}$.
- Define $g' := (g'_1, \ldots, g'_m)$ and $g'_i(\bar{x}; d) = 0$ if $i \notin \bar{I}_g$.

The linearized cone:

$$L(\bar{x}) := \{d | g'(\bar{x}; d) \leq 0\}.$$

The critical cone:

$$C(\bar{x}) = \{d \in L(\bar{x}) | \nabla f(\bar{x})^T d \leq 0\}.$$

Note that if $C(\bar{x}) = \{0\}$, then $\bar{x}$ is already a local minimum. Otherwise if $C(\bar{x}) \neq \{0\}$, then one can use a nonzero critical direction to obtain a sharper optimality condition.

Directional Clarke subdifferentials

- $x^k \xrightarrow{d} \bar{x} \iff$ there exist $t_k \downarrow 0, d^k \to d$ such that $x^k = \bar{x} + t_k d^k$.

Definition (Ginchev and Mordukhovich 2011)

Let $\varphi : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\varphi(\bar{x})$ is finite. The limiting subdifferential of φ at $\bar{x}$ in direction $d \in \mathbb{R}^n$ is defined as

$$\partial\varphi(\bar{x}; d) := \left\{ \lim_k \xi^k | \exists x^k \xrightarrow{d} \bar{x}, \varphi(x^k) \to \varphi(\bar{x}),\ \xi^k \in \widehat{\partial}\varphi(x^k) \right\}.$$

- Let $\varphi(x) = |x|$. The directional limiting subdifferential at 0 is

$$\partial\varphi(0; 1) = \{1\}, \quad \partial\varphi(0; -1) = \{-1\}.$$

- But the limiting subdifferential $\partial\varphi(0) = [-1, 1]$.

When φ is Lipschitz continuous, we define the directional Clarke subdifferential as $\partial^c\varphi(\bar{x}; d) := \mathrm{co}\partial\varphi(\bar{x}; d)$.

Directional optimality conditions, Gfrerer 2013

- Suppose $\bar{x}$ is a local optimal solution of (P). Let $d \in C(\bar{x})$. Then there exists nonnegative numbers λ_0, λ not all equal to zero such that Fritz John type directional optimality condition holds at $\bar{x}$ in direction d:

$$\begin{aligned} &0 \in \lambda_0 \nabla f(\bar{x}) + \partial g(\bar{x}; d)^T \lambda, \\ &0 \leq \lambda \perp g(\bar{x}), \quad \lambda \perp g'(\bar{x}; d). \end{aligned}$$

- If

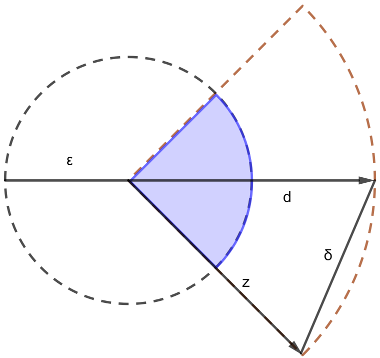
$$\left\{ \begin{array}{l} 0 \in \partial g(\bar{x}; d)^T \lambda, \\ 0 \leq \lambda \perp g(\bar{x}), \quad \lambda \perp g'(\bar{x}; d) \end{array} \right. \implies \lambda = 0$$

  then the KKT type directional optimality condition holds.

- We call the above condition the first order sufficient condition for metric subregularity (or the directional MFCQ).

Directional neighborhood

$$\mathcal{V}_{\varepsilon,\delta}(d) = \begin{cases} \mathbb{B}_\varepsilon(0), & d = 0, \\ \{0\} \cup \{z \in \mathbb{B}_\varepsilon(0) \,|\, \left\| \frac{z}{\|z\|} - \frac{d}{\|d\|} \right\| \leq \delta\}, & d \neq 0. \end{cases}$$



Directional Lipschitz continuity

Definition

We say that a single-valued map $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous at $\bar{x}$ in direction d if there exists $L > 0$ and a directional neighborhood $\mathcal{V}_{\varepsilon,\delta}(d)$ of d such that

$$\|\varphi(x) - \varphi(x')\| \le L\|x - x'\| \quad \forall x, x' \in \bar{x} + \mathcal{V}_{\varepsilon,\delta}(d).$$

• When $d = 0$, the directional Lipschitz continuity reduces to the classical Lipschitz continuity.

Directional Clarke calmness condition

Definition

Suppose $\bar{x}$ is a local optimal solution of (P). We say that (P) is (Clarke) calm at $\bar{x}$ in direction d if $\bar{x}$ also solves

$$\begin{aligned} &\min\ f(x) + \rho\|g_+(x)\| \\ &s.t.\ x \in \bar{x} + \mathcal{V}_{\epsilon,\delta}(d), \end{aligned}$$

for some $\rho, \epsilon, \delta \geq 0$.

• When $d = 0$, the directional calmness reduces to the calmness introduced by Clarke. Directional Clarke calmness with $d \neq 0$ is weaker than the Clarke calmness.

Directional KKT condition under directional calmness

Theorem (Bai and JY 2021)

Let $\bar{x}$ be a local minimizer of (P). Suppose $f(x)$ is continuously differentiable at $\bar{x}$ and $g(x)$ is directionally Lipschitz and directionally differentiable at $\bar{x}$ in direction $d \in C(\bar{x})$. Suppose that (P) is Clarke calm at $\bar{x}$ in direction d. Then there exists a vector $\lambda \in \mathbb{R}^m$ such that the directional KKT condition holds at $\bar{x}$ in direction d:

$$
\begin{aligned}
&0 \in \nabla f(\bar{x}) + \partial g(\bar{x}; d)^T \lambda \\
&0 \le \lambda \perp g(\bar{x}), \ \lambda \perp g'(\bar{x}; d).
\end{aligned}
$$

- When $d \neq 0$, directional KKT condition is sharper than the (nondirectional) KKT condition.

Definition (Directional metric subregularity; Gfrerer, 2013)

Given $d \in \mathbb{R}^n$. *The set-valued map* $G(x) := -g(x) + \mathbb{R}^m_-$ *is metrically subregular at* $(\bar{x}, 0) \in gphG$ *in direction* d *if there are positive reals* $\epsilon > 0, \delta > 0$, *and* $\kappa > 0$ *such that*

$$\mathrm{dist}(x, G^{-1}(0)) \leq \kappa \|g_+(x)\| \quad \forall x \in \bar{x} + \mathcal{V}_{\epsilon,\delta}(d),$$

where $G^{-1}(0) := \{x | g(x) \leq 0\}$.

- When $d = 0$, the set-valued map $G(x) := -g(x) + \mathbb{R}^m_-$ is metrically subregular at $(\bar{x}, 0)$ or the metric subregularity constraint qualification (MSCQ) holds at $\bar{x}$ or the perturbed set-valued map $M(v) := \{x | g(x) \leq v\}$ is calm at $(0, \bar{x}) \in gphM$.

Directional FOSCMS/MFCQ

Definition (Directional first order sufficient condition for metric subregularity (FOSCMS) by Gfrerer)

Let $g(\bar{x}) \leq 0$ and $0 \neq d \in L(\bar{x})$. Let g be directionally Lipschitz continuous and directionally differentiable at $\bar{x}$ in direction d. We say FOSCMS/directional MFCQ at $\bar{x}$ in direction d if

$$0 \in \partial g(\bar{x}; d)^T \lambda, 0 \leq \lambda \perp g(\bar{x}), \lambda \perp g'(\bar{x}; d) \quad \Longrightarrow \quad \lambda = 0$$

Directional quasi-normality

Definition (Directional quasi-normality; Bai, Ye and Zhang, 2019)

Let $g(\bar{x}) \leq 0$. Let g be directionally Lipschitz continuous and directionally differentiable. We say the directional quasi-normality holds at $\bar{x}$ in direction $0 \neq d \in L(\bar{x})$ if

$$\left.\begin{array}{l} 0 \in \partial g(\bar{x}; d)^T \lambda, \quad 0 \leq \lambda \perp g(\bar{x}), \quad \lambda \perp g'(\bar{x}; d) \\ \exists x^k \to \bar{x} \text{ s.t. } g_i(x^k) > 0 \text{ if } \lambda_i > 0 \end{array}\right. \Longrightarrow \quad \lambda = 0$$

Directional constraint qualifications and their relationships

FOSCMS/MFCQ in direction d

$\Longrightarrow$ quasi-normality in direction d

$\Longrightarrow$ metric subregularity in direction d

$\Longrightarrow$ Clarke calmness in direction d

$\Longrightarrow$ KKT condition in direction d for a local minimum.

Directional KKT condition for bilevel program

Theorem (Bai and JY 2021)

Let $(\bar{x}, \bar{y})$ be a local minimizer of (BP). Suppose the value function $V(x)$ is Lipschitz continuous and directionally differentiable at $\bar{x}$ in direction u and $(u, v) \in C(\bar{x}, \bar{y})$.

$$\begin{aligned} L(\bar{x}, \bar{y}) &:= \{(u, v) | \nabla f(\bar{x}, \bar{y})(u, v) \leq V'(\bar{x}; u), \nabla g(\bar{x}, \bar{y})(u, v) \leq 0\} \\ C(\bar{x}, \bar{y}) &:= \{(u, v) \in L(\bar{x}, \bar{y}) | F(\bar{x}, \bar{y})(u, v) \leq 0\} \end{aligned}$$

Suppose that (VP) is calm at $(\bar{x}, \bar{y})$ in direction (u, v). Then there exists $(\lambda, \mu) \geq 0$ such that

$$\begin{aligned} 0 \in \nabla F(\bar{x}, \bar{y}) &+ \lambda (\nabla f(\bar{x}, \bar{y}) - \partial^c V(\bar{x}; u) \times \{0\}) \\ &+ \nabla g(\bar{x}, \bar{y})^T \mu, \\ \mu \perp g(\bar{x}, \bar{y}), &\ \ \mu \perp \nabla g(\bar{x}, \bar{y})(u, v). \end{aligned}$$

Difficulties

Problem (VP) is a constrained optimization problem with the difficult constraint

$$\varphi(x,y) := f(x,y) - V(x) \leq 0.$$

There are two important issues.

- Which directional constraint qualification is applicable for bilevel programs?
- How to calculate or give the upper estimate for the directional Clarke subdifferential $\partial^c V(\bar{x}; u)$ and the directional derivative $V'(\bar{x}; u)$?

Recall that

MFCQ/NNAMCQ
$\Longrightarrow$ MFCQ in direction d
$\Longrightarrow$ quasi-normality in direction d
$\Longrightarrow$ Metric subregularity in direction d
$\Longrightarrow$ calmness in direction d.

Question We know that MFCQ/NNAMCQ fails for (VP). Directional MFCQ is weaker. Does directional MFCQ hold for (VP)?
We will show that the answer is "No"!

Directional FOSCMS/MFCQ fails for VP

Proposition (Ye and Zhu, 1995)

The nonsmooth MFCQ fails at any feasible point of (VP).

Proposition (Bai and JY 2021)

FOSCMS/MFCQ fails at any feasible solution of (VP) in any critical direction.

Denote the feasible mapping by $\mathcal{F}(x) := \{y \in \mathbb{R}^m | g(x,y) \leq 0\}$ and the active index set $I_g(x,y) := \{i = 1, \dots, p | g_i(x,y) = 0\}$.

Definition (RCR regularity;Minchenko and Stakhovski, 2011)

We say that the set-valued map $\mathcal{F}(x)$ is relaxed constant rank (RCR) regular at $(\bar{x}, \bar{y}) \in \mathrm{gph}\mathcal{F}$ if there exists $\delta > 0$ such that for any index subset $K \subseteq I_g(\bar{x}, \bar{y})$, the family of gradient vectors $\nabla_y g_j(x,y), j \in K$, has the same rank for all $(x,y) \in \mathbb{B}_\delta(\bar{x}, \bar{y})$.

- RCR regularity is weaker then MFCQ.

Definition (Directional Inner Semicontinuity)

Given $\bar{y} \in S(\bar{x})$, we say that the set-valued map $S(x) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is inner semicontinuous at $(\bar{x}, \bar{y})$ in direction u, if for any sequences $x^k \xrightarrow{u} \bar{x}$, there exists a sequence $y^k \in S(x^k)$ converging to $\bar{y}$.

Formula for directional derivative, directional inner semicontinuous case

Theorem (Bai and JY 2021)

Suppose that there exists $\bar{y} \in S(\bar{x})$ such that $S(x)$ is inner semicontinuous at $(\bar{x}, \bar{y})$ in direction u. Moreover assume that $\mathcal{F}(x)$ is RCR-regular at $(\bar{x}, \bar{y})$. Then the value function $V(x)$ is directionally differentiable at $\bar{x}$ in direction u and

$$V'(\bar{x}; u) = \max_{\lambda \in \mathrm{KT}(\bar{x}, \bar{y})} \nabla_x \mathcal{L}(\bar{x}, \bar{y}; \lambda) u,$$

where $\mathcal{L}(\bar{x}, \bar{y}; \lambda) := f(\bar{x}, \bar{y}) + \langle g(\bar{x}, \bar{y}), \lambda \rangle$ is the Lagrange function.

Estimate for $\partial^c V(\bar{x}; u)$, directional inner semicontinuous case

Theorem (Bai and JY 2021)

Suppose that there exists $\bar{y} \in S(\bar{x})$ such that $S(x)$ is inner semicontinuous at $(\bar{x}, \bar{y})$ in direction u and the set-valued map $\mathcal{F}(x) := \{y | g(x, y) \le 0\}$ is RCR-regular at $(\bar{x}, \bar{y})$, then $V(x)$ is Lipschitz at $\bar{x}$ in direction u. Moreover

$$\partial^c V(\bar{x}; u) \subseteq \left\{ \nabla_x \mathcal{L}(\bar{x}, \bar{y}; \lambda) | \lambda \in \mathrm{KT}(\bar{x}, \bar{y}) \cap \{\nabla g(\bar{x}, \bar{y})(u, v)\}^{\perp} \right\},$$

where v satisfies $\nabla g(\bar{x}, \bar{y})(u, v) \le 0$, $V'(\bar{x}; u) = \nabla f(\bar{x}, \bar{y})(u, v)$.

Directional KKT condition for (VP), directional IS case

Theorem (Bai and JY 2021)

Let $(\bar{x}, \bar{y})$ be a local minimizer of (BP). Suppose $\mathcal{F}(x)$ is RCR-regular at $(\bar{x}, \bar{y})$ and $S(x)$ is inner semicontinuous at $(\bar{x}, \bar{y})$ in direction u. Suppose that there exists v such that the directional quasi-normality holds at $(\bar{x}, \bar{y})$ in direction $(u, v) \in C(\bar{x}, \bar{y})$. Then there exists a vector $(\lambda, \mu, \mu_g) \in \mathbb{R}_+^{1+p+p}$ satisfying

$$0 = \nabla_x F(\bar{x}, \bar{y}) + \nabla_x g(\bar{x}, \bar{y})^T (\mu - \lambda \mu_g),$$

$$0 = \nabla_y F(\bar{x}, \bar{y}) + \lambda \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^T \mu,$$

$$0 = \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^T \mu_g, \mu_g \perp g(\bar{x}, \bar{y}), \mu_g \perp \nabla g(\bar{x}, \bar{y})(u, v),$$

$$\mu \perp g(\bar{x}, \bar{y}), \ \mu \perp \nabla g(\bar{x}, \bar{y})(u, v).$$

Example

$$
\begin{aligned}
\text{(BP)} \qquad \min \quad & F(x,y) = (\sqrt{3}x - y - \sqrt{3})^2 + x + \sqrt{3}y + 3 \\
s.t. \quad & y \text{ minimizes } f(x,y) = -(x-y)^2 + 1, \\
& s.t. \; g_1(x,y) = (x-1)^2 + y^2 - 4 \leq 0, \\
& \qquad g_2(x,y) = -\sqrt{3}x - y - \sqrt{3} \leq 0.
\end{aligned}
$$

In this example, we have shown that

- $S(x)$ is not inner semi-continuous but it is inner semi-continuous in a critical direction.
- The quasi-normality holds in a critical direction but the classical quasi-normality fails.
- The directional KKT condition holds.

Reference

- JY and D.L. Zhu 1995, *Optimality conditions for bilevel programming problems*, Optimization **33**, 9-27.
- B. S. Mordukhovich 2018, *Variational Analysis and Applications*, Springer 2018.
- K. Bai, J. J. Ye and J. Zhang 2019, *Directional quasi-/pseudo-normality as sufficient conditions for metric subregularity*, SIAM J. Optim., 29(2019), pp. 2625-2649.
- K. Bai and J. J. Ye 2021, *Directional necessary optimality conditions for bilevel programs*, to appear in Math. Oper. Res., arXiv (2007.07933).

- Thank You !-