

RESEARCH

Open Access



A note on the optimality condition for a bilevel programming

Jie Zhang^{1,2*}, Huan Wang¹ and Yue Sun¹

*Correspondence:

zhangjie04212001@163.com

¹School of Mathematics, Liaoning Normal University, Huanghe Road 850, Dalian, 116029, China

²Research Center of Information and Control, Dalian University of Technology, Hongling Road, Dalian, 116024, China

Abstract

The equality type Mordukhovich coderivative rule for a solution mapping to a second-order cone constrained parametric variational inequality is derived under the constraint nondegenerate condition, which improves the result published recently. The rule established is then applied to deriving a necessary and sufficient local optimality condition for a bilevel programming with a second-order cone constrained lower level problem.

MSC: 90C33; 90C46

Keywords: coderivative; second-order cone; parametric variational inequality; constraint nondegenerate condition; bilevel programming

1 Introduction

In this paper, we focus on the following *bilevel programming* (BP):

$$\begin{aligned} \min & f(x, y) \\ \text{s.t. } & y \in S(x), \end{aligned} \quad (1)$$

where $f(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ is continuously differentiable and $S(x)$ is the optimal solution set of the following problem:

$$\begin{aligned} \min & \psi(x, y) \\ \text{s.t. } & A(x)y + b \in \mathcal{K}^p \end{aligned} \quad (2)$$

with $\psi(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ being a continuously differentiable convex mapping, $b \in \mathbb{R}^p$, $A(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^{p \times m}$, and $\mathcal{K}^p \subseteq \mathbb{R}^p$ being the *second-order cone* (SOC), also called the Lorentz cone, defined by

$$\mathcal{K}^p := \{z = (z_0, z_1) \in \mathbb{R} \times \mathbb{R}^{p-1} : z_0 \geq \|z_1\|\},$$

where $\|\cdot\|$ stands for the Euclidean norm. If $p = 1$, then \mathcal{K}^p is the set of nonnegative reals \mathbb{R}_+ , in this case, problem (1) is the bilevel programming studied by [1] and [2]. If the lower level problem of BP is replaced by its KKT condition, this is a mathematical program with a second-order cone complementarity problem among the constraints [3].

Since for fixed $x \in \mathbb{R}^n$, problem (2) is a convex problem, the solution mapping $S(\cdot)$ in (1) can be rewritten as

$$S(x) := \{y \in \mathbb{R}^m : \langle F(x, y), y' - y \rangle \geq 0, \forall y' \in \Omega(x)\}, \quad (3)$$

where $F(x, y) := \nabla_y \psi(x, y)$ and $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a convex-valued multifunction defined by

$$\Omega(x) = \{y \in \mathbb{R}^m : A(x)y + b \in \mathcal{K}^p\}. \quad (4)$$

For fixed x , $S(x)$ denotes the solution set of a variational inequality problem, which has been intensively studied by [4–8].

To establish the necessary and sufficient local optimality condition for bilevel programming (1), a crucial step is to compute generalized differentiation for the solution mapping $S(\cdot)$ defined by (3). The generalized differentiation in our study is Mordukhovich's coderivative [4], which plays an important role in characterizations of metric regularity and openness properties of set-valued mappings; see [9] and the references therein.

Mordukhovich and Outrata [5] has established upper estimations of the coderivatives for the solution mapping (3) with \mathcal{K}^p being a closed convex set under appropriate calmness assumptions and constraint qualifications. However, the equality type calculus rules of the coderivatives of a solution mapping S (3) are not mentioned. Recently, Zhang *et al.* [10] has established equality type calculus rules of the coderivatives of a solution mapping S (3) under the constraint nondegenerate condition and applied the results obtained to deriving necessary and sufficient condition of the Lipschitz-like property [4] of the solution mapping S (3).

In this paper, the equality type representation of the coderivative of a solution mapping S (3) is established under conditions weaker than [10], Theorem 3.2, and it then is used to obtain a necessary and sufficient local optimality conditions for the bilevel programming (1). This is done on the basis of an exact description of the coderivative of the normal cone operator onto the second-order cone.

This paper is organized as follows: Section 2 gives preliminaries needed throughout the paper. In Section 3, the main results are established, *i.e.*, the equality type calculus rule of the coderivatives of a solution mapping S (3) is established and then used to derive the optimality condition of bilevel programming (1). Some examples are provided.

2 Preliminaries

Throughout this paper we use the following notations. For an extended real-valued function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}$, $\nabla\varphi(x)$ denotes its the gradient of φ at x . For a continuously differentiable mapping $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathcal{J}\phi(x)$ denotes the Jacobian of ϕ at x . We use \mathbb{B}_n , $\|\cdot\|$ and \mathbb{R}_+ to stand for the closed unit ball in \mathbb{R}^n , the Euclidean norm and the nonnegative reals, respectively. $[|x|] = \{tx : t \in \mathbb{R}\}$, $S^\perp = \{\eta \in \mathbb{R}^n : \langle \eta, x \rangle = 0, \forall x \in S\}$, $\text{Sp}(S) = \mathbb{R}_+(S - S)$ and $\text{lin}(C)$ denote the linear space generated by vector x , the orthogonal complement of the set $S \subseteq \mathbb{R}^n$, the linear space generated by S and the linearity subspace of the convex cone C , respectively.

Given a closed set $\Xi \subset \mathbb{R}^n$ and a point $\bar{x} \in \Xi$, the *Mordukhovich limiting normal cone* to Ξ at \bar{x} is defined by

$$N_\Xi(\bar{x}) := \limsup_{x \xrightarrow{\Xi} \bar{x}} \widehat{N}_\Xi(x),$$

see for instance [11] and [4], where the cone

$$\widehat{N}_{\Xi}(\bar{x}) := \left\{ x^* \in \mathbb{R}^n \mid \limsup_{x \xrightarrow{\Xi} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\}$$

is called the *regular normal cone* to Ξ at \bar{x} with ‘lim sup’ being the outer limit of a set-valued mapping or the upper limit of a real-valued function; see [11]. It follows from the definition that $\widehat{N}_{\Xi}(\bar{x}) \subseteq N_{\Xi}(\bar{x})$. If the above inclusion becomes equality, we say that Ξ is *normally regular* at \bar{x} (or Clarke regular by [11]). According to [11], Theorem 6.9, each convex set is normally regular at all its points.

For set-valued maps, the definition of the coderivative was introduced by Mordukhovich in [12] based on the Mordukhovich limiting normal cone.

Definition 2.1 Consider a mapping $S: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $\bar{x} \in \text{dom } S$. The coderivative of S at \bar{x} for any $\bar{u} \in S(\bar{x})$ is the mapping $D^*S(\bar{x}, \bar{u}): \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$D^*S(\bar{x}, \bar{u})(y) = \{v: (v, -y) \in N_{\text{gph } S}(\bar{x}, \bar{u})\}.$$

The notation $D^*S(\bar{x}, \bar{u})$ is simplified to $D^*S(\bar{x})$ when S is single-valued at \bar{x} , $S(\bar{x}) = \{\bar{u}\}$. Similarly, and with the same provision for simplified notation, the regular coderivative $\widehat{D}^*S(\bar{x}, \bar{u}): \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is defined by

$$\widehat{D}^*S(\bar{x}, \bar{u})(y) = \{v: (v, -y) \in \widehat{N}_{\text{gph } S}(\bar{x}, \bar{u})\}.$$

Next we give the following proposition to show the description of the coderivative of some special set-valued mappings.

Proposition 2.1 ([10], Proposition 2.1) *For any $(\bar{x}, \bar{y}) \in \text{gph } N_{\mathcal{K}^p}$, let $\bar{z} = \bar{x} + \bar{y}$.*

(1) *In the case when $\bar{x} \neq 0, \bar{y} \neq 0$, we have*

$$\widehat{D}^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = D^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = \begin{cases} [|\eta| + \frac{\bar{z}_0 - \|\bar{z}_1\|}{\bar{z}_0 + \|\bar{z}_1\|} \begin{pmatrix} \bar{z}_1^T y_1^* \\ -y_1^* \end{pmatrix}], & \eta^T y^* = 0, \\ \emptyset, & \text{otherwise,} \end{cases}$$

$$\text{where } \eta = (1, -\frac{\bar{z}_1^T}{\|\bar{z}_1\|})^T.$$

(2) *In the case when $\bar{z} \in \text{int}(\mathcal{K}^p)^-$, we have*

$$\widehat{D}^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = D^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = \begin{cases} \mathbb{R}^p, & y^* = 0, \\ \emptyset, & y^* \neq 0. \end{cases}$$

(3) *In the case when $\bar{z} \in \text{int } \mathcal{K}^p$, we have*

$$\widehat{D}^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = D^*N_{\mathcal{K}^p}(\bar{x}, \bar{y})(y^*) = \{0\}_p$$

for any $y^* \in \mathbb{R}^p$.

We need the following stability notations; see [4].

Definition 2.2 Consider the multifunction $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$.

- (a) (*Lipschitz-like property*) We say F has Lipschitz-like property at $(\bar{y}, \bar{x}) \in \text{gph } F$, if there exist some $\kappa > 0$ and some neighborhoods U of \bar{x} and V of \bar{y} such that

$$F(y') \cap U \subset F(y) + \kappa \|y' - y\| \mathbb{B}_n \quad \text{for all } y, y' \in V.$$

- (b) (*Calmness*) We say F is calm at $(\bar{y}, \bar{x}) \in \text{gph } F$ if there exist some $k > 0$ and some neighborhoods U of \bar{x} and V of \bar{y} such that

$$d(x, F(\bar{y})) \leq k \|y - \bar{y}\| \quad \text{for all } y \in V, x \in F(y) \cap U.$$

We know from the definition that the calmness property is weaker than the Lipschitz-like property. As shown in [11], Theorem 9.43, F has Lipschitz-like property at $(\bar{y}, \bar{x}) \in \text{gph } F$ if and only if the coderivative condition

$$D^*F(\bar{y}, \bar{x})(0) = \{0\}, \quad (5)$$

see [13], Proposition 2.8. This condition is the famous Mordukhovich criterion [11], Theorem 9.40.

Under the calmness condition, when the constraint set is structured, the normal cones can be estimated or calculated.

Proposition 2.2 ([14], Theorem 4.1) Assume the multifunction $M : \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^{n_1}$, defined by

$$M(q) := \{z \in Z : G(z) + q \in K\}$$

for closed sets $Z \subseteq \mathbb{R}^{n_1}$ and $K \subseteq \mathbb{R}^{n_2}$ and a C^1 mapping $G : \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2}$, is calm at $(0, \bar{z}) \in \text{gph } M$. Then one has

$$N_{M(0)}(\bar{z}) \subseteq \mathcal{J}G(\bar{z})^T N_K(G(\bar{z})) + N_Z(\bar{z}). \quad (6)$$

We know from [11], Theorem 6.14 that

$$\widehat{N}_{M(0)}(\bar{z}) \supseteq \mathcal{J}G(\bar{z})^T \widehat{N}_K(G(\bar{z})) + \widehat{N}_Z(\bar{z}).$$

Thus if, in addition, Z is normally regular at \bar{z} and K is normally regular at $G(\bar{z})$, then C is normally regular at \bar{z} and inclusion in (6) becomes equality.

We know from [4], Theorem 4.32 that M defined in Proposition 2.2 is Lipschitz-like around $(0, \bar{z}) \in \text{gph } M$ if the following constraint qualification holds:

$$\left. \begin{array}{l} 0 \in \mathcal{J}G(\bar{z})^T \eta + N_Z(\bar{z}), \\ \eta \in N_K(G(\bar{z})) \end{array} \right\} \Rightarrow \eta = 0. \quad (7)$$

3 Main results

In this section, we provide conditions ensuring the equality type calculus rule of the coderivatives of a solution mapping S (3), which is an improvement of [10], Theorem 3.2.

The result obtained is used to derive a necessary and sufficient condition for the bilevel programming (1).

We know from the definition of normal cone in convex analysis that the solution mapping S (3) can be rewritten as

$$S(x) = \{y \in \mathbb{R}^m : 0 \in F(x, y) + N_{\Omega(x)}(y)\},$$

where $N_{\Omega(x)}(y)$ denotes the normal cone of $\Omega(x)$ at y . For a parameter $\bar{x} \in \mathbb{R}^n$, if the following *Slater constraint qualification* (SCQ) is satisfied at \bar{x} :

$$\exists \bar{y} \in \mathbb{R}^m \quad \text{such that} \quad A(\bar{x})\bar{y} + b \in \text{int } K^p, \quad (8)$$

then, by [11], Theorem 10.49, we can compute the normal cone $N_{\Omega(\bar{x})}(y)$ at $y \in \Omega(\bar{x})$ and obtain

$$N_{\Omega(\bar{x})}(y) = A(\bar{x})^T N_{K^p}(A(\bar{x})y + b). \quad (9)$$

We need the following conditions, which are popularly used conditions in SOCP.

Definition 3.1 Let $\bar{x} \in \mathbb{R}^n$, $\bar{y} \in \Omega(\bar{x})$ and $\bar{v} \in N_{\Omega(\bar{x})}(\bar{y})$.

(a) We say that the constraint nondegenerate condition ($\mathcal{CN}\mathcal{C}$) holds true at \bar{y} for \bar{x} , if

$$A(\bar{x})\mathbb{R}^m + \text{lin}(T_{K^p}(A(\bar{x})\bar{y} + b)) = \mathbb{R}^p. \quad (10)$$

(b) We say the strict complementarity (\mathcal{SC}) condition holds at $(\bar{x}, \bar{y}, \bar{v})$, if

$$\lambda \in \text{ri } N_{K^p}(A(\bar{x})\bar{y} + b)$$

for all λ satisfying $\lambda \in N_{K^p}(A(\bar{x})\bar{y} + b)$ and $A(\bar{x})^T \lambda = \bar{v}$.

We introduce the *Lagrangian mapping* $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^m$ defined by

$$\mathcal{L}(x, y, \lambda) := F(x, y) + A(x)^T \lambda \quad (11)$$

and the *Lagrangian multiplier mapping* $\Lambda : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^p$ defined by

$$\Lambda(x, y) := \{\lambda \in \mathbb{R}^p \mid \mathcal{L}(x, y, \lambda) = 0\}.$$

In [10], Theorem 3.2, an equality type representation of the coderivative of a solution mapping S (3) has been established under some constraint qualifications, we cite it as a lemma.

Lemma 3.1 Assume the SCQ (8) holds for \bar{x} and the multifunction $P : \mathbb{R}^m \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ defined by

$$P(\gamma, q) := \{(x, y, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \mid \mathcal{L}(x, y, \lambda) + \gamma = 0\} \cap M(q) \quad (12)$$

is calm at the points $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$ with $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$, where the multifunction $M : \mathbb{R}^{2p} \rightrightarrows \mathbb{R}^{n+m+p}$ is defined by

$$M(q) := \left\{ (x, y, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \mid q + \begin{pmatrix} A(x)y + b \\ \lambda \end{pmatrix} \in \text{gph } N_{\mathcal{K}^p} \right\}. \quad (13)$$

Then:

- (a) In the case when $\bar{z}_0 > \|\bar{z}_1\|$, where $\bar{z} := A(\bar{x})\bar{y} + b$, we have for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$,

$$\begin{aligned} D^*S(\bar{x}, \bar{y})(y^*) &= \{ (\mathcal{J}_x \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + (\mathcal{J}_x(A(\bar{x})\bar{y}))^T w \mid \\ &\quad 0 \in y^* + (\mathcal{J}_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + A(\bar{x})^T w, \\ &\quad w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u) \} \end{aligned} \quad (14)$$

holds for all $y^* \in \mathbb{R}^m$.

- (b) In the case when $\bar{z}_0 = \|\bar{z}_1\|$, if the mapping $M(\cdot)$ (13) is calm at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$, CNC (10) holds at \bar{y} for \bar{x} and SC condition holds at $(\bar{x}, \bar{y}, -F(\bar{x}, \bar{y}))$. Then the equality (14) holds for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$.

Under conditions weaker than the ones in (b) of Theorem 3.2 in [10], we obtain the same equality type coderivative rule as follows.

Theorem 3.1 Assume:

- (a) SCQ (8) holds for \bar{x} and $P(\gamma, q)$ (12) is calm at the points $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$ with $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$.
 (b) CNC (10) holds at \bar{y} for \bar{x} and SC condition holds at $(\bar{x}, \bar{y}, -F(\bar{x}, \bar{y}))$.

Then in the case when $\bar{z}_0 = \|\bar{z}_1\|$, the equality (14) holds for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$.

Proof According to Lemma 3.1(b), we need to show under conditions (a) and (b) that the mapping $M(\cdot)$ (13) is calm at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ for any $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$. We know from Definition 2.2 that the calmness of $M(\cdot)$ at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ is ensured by the Lipschitz-like property of $M(\cdot)$ at $(0, \bar{x}, \bar{y}, \bar{\lambda})$, which holds under the condition

$$\left. \begin{aligned} 0 &= \mathcal{J}(A(x)y)^T|_{(x,y)=(\bar{x},\bar{y})} \eta, \\ \eta &\in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(0) \end{aligned} \right\} \Rightarrow \eta = 0. \quad (15)$$

Indeed, notice that, by the Mordukhovich criterion (5), we only need to verify

$$D^*M(0, \bar{x}, \bar{y}, \bar{\lambda})(0) = \{0\} \quad (16)$$

under condition (15). Let $y^* \in D^*M(0, \bar{x}, \bar{y}, \bar{\lambda})(0)$, by Definition 2.1, we have

$$\begin{pmatrix} y^* \\ 0 \end{pmatrix} \in N_{\text{gph } M} \begin{pmatrix} 0 \\ \bar{x} \\ \bar{y} \\ \bar{\lambda} \end{pmatrix}. \quad (17)$$

Since

$$\text{gph } M = \left\{ (q, x, y, \lambda) \in \mathfrak{N}^{2p} \times \mathfrak{N}^n \times \mathfrak{N}^m \times \mathfrak{N}^p \mid q + \begin{pmatrix} A(x)y + b \\ \lambda \end{pmatrix} \in \text{gph } N_{\mathcal{K}^p} \right\},$$

we know from Proposition 2.2 that if the condition

$$\left. \begin{aligned} 0 &= \mathcal{J}_{(q,x,y,\lambda)} \left[q + \begin{pmatrix} A(x)y + b \\ \lambda \end{pmatrix} \right]^T \Big|_{(q,x,y,\lambda)=(0,\bar{x},\bar{y},\bar{\lambda})} \xi, \\ \xi &\in N_{\text{gph } N_{\mathcal{K}^p}} \left(\begin{pmatrix} A(\bar{x})\bar{y} + b \\ \bar{\lambda} \end{pmatrix} \right) \end{aligned} \right\} \Rightarrow \xi = 0 \quad (18)$$

holds, then

$$N_{\text{gph } M} \begin{pmatrix} 0 \\ \bar{x} \\ \bar{y} \\ \bar{\lambda} \end{pmatrix} \subseteq \mathcal{J}_{(q,x,y,\lambda)} \left[q + \begin{pmatrix} A(x)y + b \\ \lambda \end{pmatrix} \right]^T \Big|_{(q,x,y,\lambda)=(0,\bar{x},\bar{y},\bar{\lambda})} N_{\text{gph } N_{\mathcal{K}^p}} \begin{pmatrix} A(\bar{x})\bar{y} + b \\ \bar{\lambda} \end{pmatrix}. \quad (19)$$

Notice that

$$\mathcal{J}_{(q,x,y,\lambda)} \left[q + \begin{pmatrix} A(x)y + b \\ \lambda \end{pmatrix} \right]^T \Big|_{(q,x,y,\lambda)=(0,\bar{x},\bar{y},\bar{\lambda})} = \begin{pmatrix} I_p & 0 & \mathcal{J}(A(x)y)|_{(x,y)=(\bar{x},\bar{y})} & 0 \\ 0 & I_p & 0 & I_p \end{pmatrix},$$

we have (18), then (19) holds and hence by (17), we have

$$\begin{pmatrix} y^* \\ 0 \end{pmatrix} \in \begin{pmatrix} I_p & 0 & \mathcal{J}(A(x)y)|_{(x,y)=(\bar{x},\bar{y})} & 0 \\ 0 & I_p & 0 & I_p \end{pmatrix}^T N_{\text{gph } N_{\mathcal{K}^p}} \begin{pmatrix} A(\bar{x})\bar{y} + b \\ \bar{\lambda} \end{pmatrix},$$

which, by (15) and Definition 2.1, means that $y^* = 0$. Therefore (16) holds.

Next we show \mathcal{CNC} condition implies (15). In the case when $\bar{z}_0 = \|\bar{z}_1\| = 0$, \mathcal{CNC} condition means that $A(\bar{x})\mathfrak{N}^m = \mathfrak{N}^p$, which is equivalent to

$$0 = A(\bar{x})^T \eta \Rightarrow \eta = 0$$

and hence condition (15) holds. In the case when $\bar{z}_0 = \|\bar{z}_1\| \neq 0$, we proceed in the proof in two main steps.

Step 1. Taking the orthogonal complements on both sides of (10), the \mathcal{CNC} condition can be rewritten as

$$(A(\bar{x})\mathfrak{N}^m)^\perp \cap \text{lin}(T_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b))^\perp = \{0\}. \quad (20)$$

We know from [15], Proposition 4.73 that

$$\text{lin}(T_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b))^\perp = \text{Sp}\{N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)\},$$

which, by (20), means that the \mathcal{CNC} condition is equivalent to

$$\text{Sp}\{N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)\} \cap \text{Ker } A(\bar{x})^T = \{0\}. \quad (21)$$

Step 2. We next show

$$\text{Sp}\{N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)\} = D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(0). \quad (22)$$

Since $\bar{\lambda} \in N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$, we have $\bar{\lambda} = k(-\bar{z}_0, \bar{z}_1)$ with $k \in \mathfrak{N}_+$, where $\bar{z} = A(\bar{x})\bar{y} + b$. Then we know from Proposition 2.1 that

$$\begin{aligned} D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(0) &= \left[\left(-1, \frac{(\bar{z} + \bar{\lambda})_1^T}{\|(\bar{z} + \bar{\lambda})_1\|} \right)^T \right] \\ &= \left[\left(-1, \frac{(k+1)\bar{z}_1^T}{\|(k+1)\bar{z}_1\|} \right)^T \right] = \left[\left(-1, \frac{\bar{z}_1^T}{\|\bar{z}_1\|} \right)^T \right], \end{aligned}$$

which, by $\|\bar{z}_1\| = \bar{z}_0$, means that (22) holds. Combining with (21) and (22), the $\mathcal{CN}\mathcal{C}$ condition is equivalent to

$$\left. \begin{aligned} 0 &= A(\bar{x})^T \eta, \\ \eta &\in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(0) \end{aligned} \right\} \Rightarrow \eta = 0,$$

which implies (15). We complete the proof. \square

Remark 3.1 We know from the proof of Theorem 3.1 that the calmness of $M(\cdot)$ at $(0, \bar{x}, \bar{y}, \bar{\lambda})$ is ensured by the $\mathcal{CN}\mathcal{C}$ condition, which means the condition in Theorem 3.1 is weaker than the conditions in [10], Theorem 3.2.

In the following, we apply the results obtained to derive a necessary and sufficient optimality condition for the bilevel programming (1).

Theorem 3.2 Suppose the function f in BP (1) is convex, $\nabla_y \psi(x, y)$ is a linear function and the conditions in Theorem 3.1 hold at (\bar{x}, \bar{y}) with the involved function $F(x, y) := \nabla_y \psi(x, y)$. Then (\bar{x}, \bar{y}) is a locally optimal solution of BP (1) if and only if there exists $(w, u) \in \mathfrak{N}^p \times \mathfrak{N}^m$ satisfying $w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u)$ for some $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$ such that

$$0 = \nabla f(\bar{x}, \bar{y}) + (\mathcal{J}_{x,y} \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + (\mathcal{J}G(\bar{x}, \bar{y}))^T w, \quad (23)$$

where $G(x, y) := A(x)y + b$, $\mathcal{L}(x, y, \lambda) = \nabla_y \psi(x, y) + A(x)^T \lambda$ and $\Lambda(x, y) = \{\lambda : \mathcal{L}(x, y, \lambda) = 0\}$.

Proof Since $S(x)$ is the optimal solution set of the parametric problem (2) and for any $x \in \mathfrak{N}^n$, (2) is a convex optimization problem, $S(x)$ can be written as

$$S(x) = \{y : 0 \in \nabla_y \psi(x, y) + Q(x, y)\}, \quad (24)$$

where $\Omega(x) = \{y : A(x)y + b \in \mathcal{K}^p\}$ and $Q(x, y) = N_{\Omega(x)}(y)$. As a result, the bilevel problem can be reformulated as

$$\begin{aligned} \min & f(x, y) \\ \text{s.t.} & (x, y) \in \text{gph } S, \end{aligned}$$

where

$$\text{gph } S = \left\{ (x, y) \in \mathfrak{N}^n \times \mathfrak{N}^m \mid \begin{bmatrix} x \\ y \\ -\nabla_y \psi(x, y) \end{bmatrix} \in \text{gph } Q \right\}.$$

We next show that $\text{gph } Q$ is normally regular at $(\bar{x}, \bar{y}, -\nabla \psi(\bar{x}, \bar{y}))$. We know from the proof of [10], Theorem 3.1 that, under conditions (a) and (b) in Theorem 3.1,

$$D^*Q(\bar{x}, \bar{y}, \bar{v})(u) = (\mathcal{J}_{x,y}(A(\bar{x})^T \lambda))^T u + D^*(N_{\mathcal{K}^p} \circ G)(\bar{x}, \bar{y}, \lambda)(A(\bar{x})u)$$

holds for any $\lambda \in \Lambda(\bar{x}, \bar{y})$, which, by Definition 2.1, means that

$$\begin{aligned} \begin{bmatrix} w \\ -u \end{bmatrix} &\in N_{\text{gph } Q}(\bar{x}, \bar{y}, \bar{v}) \\ \iff w &\in D^*Q(\bar{x}, \bar{y}, \bar{v})(u) \\ \iff w &\in (\mathcal{J}_{x,y}(A(\bar{x})^T \lambda))^T u + D^*(N_{\mathcal{K}^p} \circ G)(\bar{x}, \bar{y}, \lambda)(A(\bar{x})u) \\ \iff \begin{bmatrix} w - (\mathcal{J}_{x,y}(A(\bar{x})^T \lambda))^T u \\ -A(\bar{x})u \end{bmatrix} &\in N_{\text{gph } N_{\mathcal{K}^p} \circ G}(\bar{x}, \bar{y}, \lambda) \\ \iff \begin{bmatrix} I_{n+m} & (\mathcal{J}_{x,y}(A(\bar{x})^T \lambda))^T \\ 0 & A(\bar{x}) \end{bmatrix} \begin{bmatrix} w \\ -u \end{bmatrix} &\in N_{\text{gph } N_{\mathcal{K}^p} \circ G}(\bar{x}, \bar{y}, \lambda) \end{aligned} \quad (25)$$

holds for any $\lambda \in \Lambda(\bar{x}, \bar{y})$. Under conditions (a) and (b) in Theorem 3.1, we know from the proof of [16], Lemma 3.3 that

$$\widehat{D}^*Q(\bar{x}, \bar{y}, \bar{v})(u) = (\mathcal{J}_{x,y}(A(\bar{x})^T \lambda))^T u + \widehat{D}^*(N_{\mathcal{K}^p} \circ G)(\bar{x}, \bar{y}, \lambda)(A(\bar{x})u). \quad (26)$$

Under the \mathcal{SC} condition, by Proposition 2.1, we have

$$N_{\text{gph } N_{\mathcal{K}^p} \circ G}(\bar{x}, \bar{y}, \lambda) = \widehat{N}_{\text{gph } N_{\mathcal{K}^p} \circ G}(\bar{x}, \bar{y}, \lambda). \quad (27)$$

Consequently, combining with (25), (26), and (27), we see that $\text{gph } Q$ is normally regular at $(\bar{x}, \bar{y}, -\nabla \psi(\bar{x}, \bar{y}))$. We know from the proof of [5], Theorem 4.3 that if the set-valued mapping P (12) is calm at the points $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$ with $\bar{\lambda} \in \Lambda(\bar{x}, \bar{y}) \cap N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b)$, then the set-valued mapping $\Psi : \mathfrak{N}^n \times \mathfrak{N}^m \times \mathfrak{N}^m \rightrightarrows \mathfrak{N}^n \times \mathfrak{N}^m$ defined by

$$\Psi(\zeta) := \left\{ (x, y) \in \mathfrak{N}^n \times \mathfrak{N}^m \mid \begin{bmatrix} x \\ y \\ -\nabla_y \psi(x, y) \end{bmatrix} + \zeta \in \text{gph } Q \right\}$$

is calm at $(0, \bar{x}, \bar{y})$, which, by Proposition 2.2, implies that

$$N_{\text{gph } S}(\bar{x}, \bar{y}) \subseteq \begin{bmatrix} I_n & 0 & -\mathcal{J}_x(\nabla_y \psi(\bar{x}, \bar{y}))^T \\ 0 & I_m & -\mathcal{J}_y(\nabla_y \psi(\bar{x}, \bar{y}))^T \end{bmatrix} \circ N_{\text{gph } Q}(\bar{x}, \bar{y}, -\nabla_y \psi(\bar{x}, \bar{y})). \quad (28)$$

On the other hand, we know from [11], Theorem 6.14 that

$$\widehat{N}_{\text{gph } S}(\bar{x}, \bar{y}) \supseteq \begin{bmatrix} I_n & 0 & -\mathcal{J}_x(\nabla_y \psi(\bar{x}, \bar{y}))^T \\ 0 & I_m & -\mathcal{J}_y(\nabla_y \psi(\bar{x}, \bar{y}))^T \end{bmatrix} \circ \widehat{N}_{\text{gph } Q}(\bar{x}, \bar{y}, -\nabla_y \psi(\bar{x}, \bar{y})). \quad (29)$$

Notice that

$$\widehat{N}_{\text{gph } S}(\bar{x}, \bar{y}) \subseteq N_{\text{gph } S}(\bar{x}, \bar{y}).$$

Then combining (28) and (29), the normal regularity of $\text{gph } S$ at (\bar{x}, \bar{y}) is directly from the normal regularity of $\text{gph } Q$ at $(\bar{x}, \bar{y}, -\nabla \psi(\bar{x}, \bar{y}))$. Therefore, (\bar{x}, \bar{y}) is a locally optimal solution if and only if (\bar{x}, \bar{y}) satisfying $0 \in \nabla f(x, y) + N_{\text{gph } S}(\bar{x}, \bar{y})$, i.e.,

$$0 \in \nabla_x f(\bar{x}, \bar{y}) + D^*S(\bar{x}, \bar{y})(\nabla_y f(\bar{x}, \bar{y})). \quad (30)$$

Under the conditions in Theorem 3.1, we have

$$\begin{aligned} D^*S(\bar{x}, \bar{y})(y^*) &= \{ (\mathcal{J}_x \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + (\mathcal{J}_x(A(\bar{x})\bar{y}))^T w \mid \\ &\quad 0 \in y^* + (\mathcal{J}_y \mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + A(\bar{x})^T w, \\ &\quad w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u) \}. \end{aligned} \quad (31)$$

Consequently, the conclusion is directly from (30) and (31). \square

In [10], Theorem 5.1, a necessary and sufficient global optimality condition for the bilevel programming (1) has been derived under some strong condition such as $G(x, y) + \lambda \in (\text{int } \mathcal{K}^p) \cup (\text{int } (\mathcal{K}^p)^-)$. In the case when one of the conditions in [10], Theorem 5.1 is not satisfied at a point, we do not know whether the point is a global optimal solution. However, by Theorem 3.2, we may verify that the point is a local optimal solution for the bilevel programming (1). We next give an example to show this.

Example 3.1 Consider

$$\begin{aligned} \min f(x_1, x_2, y_1, y_2) &:= e^{x_1} + 2x_2 + y_1^2 - 3y_1 + y_2^4 - 2y_2 \\ \text{s.t. } y &\in S(x), \end{aligned} \quad (32)$$

where $S(x)$ is the optimal solution set of the following problem:

$$\begin{aligned} \min \psi(x_1, x_2, y_1, y_2) &:= y_1^2 - 2y_2 + e^{x_1} + x_2 \\ \text{s.t. } G(x, y) &:= \begin{pmatrix} x_1 + 1 & 0 \\ 0 & x_2 + 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathcal{K}^2, \end{aligned}$$

where $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Consider a point $(\bar{x}, \bar{y}) = (0, 0, 1, 1)^T \in \mathbb{R}^4$. By simple computing, we have the multiplier set $\Lambda(\bar{x}, \bar{y}) = \{\bar{\lambda}\} = \{(-2, 2)^T\}$. Then we have $G(\bar{x}, \bar{y}) + \bar{\lambda} = (-1, 3)^T \notin (\text{int } \mathcal{K}^2) \cup (\text{int } (\mathcal{K}^2)^-)$, which means that one of the conditions in [10], Theorem 5.1 is not satisfied at (\bar{x}, \bar{y}) and hence we do not know whether it is a global solution to problem

(32). Next by Theorem 3.2, we verify that $(\bar{x}, \bar{y}) = (0, 0, 1, 1)^T$ is a locally optimal solution of problem (32). (i) Since there exists $\hat{y} = (1, 0)^T$ such that $A(\bar{x})\hat{y} = \hat{y} \in \text{int } \mathcal{K}^2$, the SCQ (8) holds for \bar{x} . (ii) Since $A(\bar{x}) = I_2$, the \mathcal{CN} (10) holds at (\bar{x}, \bar{y}) . (iii) The multifunction $P(\cdot)$ defined by (12) is calm at the points $(0, 0, \bar{x}, \bar{y}, \bar{\lambda})$. In fact, by a simple computation, we obtain

$$\begin{aligned} & (\mathcal{J}_{x,y}\mathcal{L}(\bar{x}, \bar{y}, \bar{\lambda}))^T u + (\mathcal{J}_{x,y}(A(\bar{x})\bar{y}))^T w \\ &= \begin{pmatrix} -2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix}^T u + \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}^T w, \end{aligned}$$

which means that

$$\left. \begin{aligned} 0 &\in (\mathcal{J}_{x,y}\mathcal{L}(\bar{x}, \bar{y}, \lambda))^T u + (\mathcal{J}_{x,y}(A(\bar{x})\bar{y}))^T w, \\ w &\in D^*N_{\mathcal{K}^p}(G(\bar{x}, \bar{y}), \lambda)(A(\bar{x})u) \end{aligned} \right\} \Rightarrow w = 0, \quad u = 0.$$

This, by the Mordukhovich criterion (5), is a condition ensuring the Lipschitz-like property of P at $(0, 0, \bar{x}, \bar{y}, \lambda)$, which ensures the calmness of P at $(0, 0, \bar{x}, \bar{y}, \lambda)$. (iv) By simple computing, the SC condition holds. (v) Next we show there exists $(w, u) \in \mathbb{R}^2 \times \mathbb{R}^2$ satisfying $w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u)$ such that (23) holds. If the equality (23) holds for $(\bar{x}, \bar{y}, \bar{\lambda})$ and (w, u) , then we have

$$0 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 0 & 2 \\ 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \quad (33)$$

We know from Proposition 2.1 that $w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u)$ means that

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in [u]^\perp + \begin{pmatrix} -6u_2 \\ 2u_2 \end{pmatrix}. \quad (34)$$

Then, combining (33) and (34), we have $w = (0, -2)^T$ and $u = (1/2, 0)^T$ satisfying $w \in D^*N_{\mathcal{K}^p}(A(\bar{x})\bar{y} + b, \bar{\lambda})(A(\bar{x})u)$ such that (23) holds. Therefore, by Theorem 3.2, (\bar{x}, \bar{y}) is a locally optimal solution of problem (32).

Remark 3.2 By a similar computation, we can infer the point $(\bar{x}, \bar{y}) = (0, 0, 1, 1)^T$ is a locally optimal solution of problem (32) with $f(x, y) := x_1^2 + x_2^2 + (y_1 - 1)^2 + (y_2 - 1)^2$ and it is not a locally optimal solution of problem (32) with $f(x, y) := e^{x_1} + x_2^2 - x_2 + 2y_1 + 1/2y_2^2 - 2y_2$.

4 Conclusion

In this paper, an equality type representation of the coderivative of the solution mapping S (3) is obtained, which is an improvement of [10], Theorem 3.2. The result obtained is then used to develop a necessary and sufficient local optimality condition for a bilevel programming with SOC as its lower level problem.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Project No. 11201210, CPSF grant 2014M560200 and Program for Liaoning Excellent Talents in University No. LJQ2015059.

Received: 29 June 2015 Accepted: 1 November 2015 Published online: 14 November 2015

References

- Dempe, S: Foundations of Bilevel Programming. Kluwer Academic, Dordrecht (2002)
- Ye, J: Constraint qualifications and KKT conditions for bilevel programming problems. *Math. Oper. Res.* **31**, 811-824 (2006)
- Outrata, JV, Sun, DF: On the coderivative of the projection operator onto the second-order cone. *Set-Valued Anal.* **16**, 999-1014 (2008)
- Mordukhovich, BS: Variational Analysis and Generalized Differentiation. I: Basic Theory, II: Applications. Springer, Berlin (2006)
- Mordukhovich, BS, Outrata, JV: Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. *SIAM J. Optim.* **18**, 389-412 (2007)
- Shen, J, Pang, LP: A bundle-type auxiliary problem method for generalized variational-like inequality. *Comput. Math. Appl.* **55**, 2993-2998 (2008)
- Shen, J, Pang, LP: An approximate bundle-type auxiliary problem method for generalized variational inequality. *Math. Comput. Model.* **48**, 769-775 (2008)
- Shen, J, Pang, LP: A proximal analytic center cutting plane algorithm for solving variational inequality problems. *J. Appl. Math.* **2012**, Article ID 503242 (2012)
- Mordukhovich, BS: Coderivatives of set-valued mappings: calculus and applications. *Nonlinear Anal.* **30**, 3059-3070 (1997)
- Zhang, J, Li, Y, Zhang, L: On the coderivative of the solution mapping to a second-order cone constrained parametric variational inequality. *J. Glob. Optim.* **61**, 379-396 (2015)
- Rockafellar, RT, Wets, RJB: Variational Analysis. Springer, Berlin (1998)
- Mordukhovich, BS: Metric approximations and necessary optimality conditions for general classes of extremal problems. *Sov. Math. Dokl.* **22**, 526-530 (1980)
- Mordukhovich, BS: Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.* **183**, 250-288 (1994)
- Henrion, R, Jourani, A, Outrata, JV: On the calmness of a class of multifunctions. *SIAM J. Optim.* **13**, 603-618 (2002)
- Bonnans, JF, Shapiro, A: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
- Zhang, J, Zhang, L, Pang, L: On the convergence of coderivative of SAA solution mapping for a parametric stochastic variational inequality. *Set-Valued Var. Anal.* **20**, 75-109 (2012)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com