J Glob Optim (2007) 39:529–542
DOI 10.1007/s10898-007-9154-0

Necessary optimality conditions for bilevel set optimization problems

S. Dempe · N. Gadhi

Received: 10 November 2005 / Accepted: 26 February 2007 / Published online: 9 June 2007

Abstract Bilevel programming problems are hierarchical optimization problems where in the upper level problem a function is minimized subject to the graph of the solution set mapping of the lower level problem. In this paper necessary optimality conditions for such problems are derived using the notion of a convexificator by Luc and Jeyakumar. Convexificators are subsets of many other generalized derivatives. Hence, our optimality conditions are stronger than those using e.g., the generalized derivative due to Clarke or Michel-Penot. Using a certain regularity condition Karush-Kuhn-Tucker conditions are obtained.

Keywords Bilevel optimization · Convexificator · Karush-Kuhn-Tucker multipliers · Necessary Optimality conditions · Regularity condition · Set valued mappings · Support function

Mathematics Subject Classification (2000) Primary 49J52 · 90C29 · Secondary 49K99

1 Introduction

The bilevel programming problem (P) considered in this paper is a sequence of two optimization problems in which the feasible region of the upper-level problem is determined implicitly by the solution set of the lower-level problem. It is given as problem (P):

$$(P): \quad \begin{cases} \text{Minimize } f(x, y) \\ \text{subject to : } F(x, y) \cap \left(-\mathbb{R}_+^p\right) \neq \emptyset, \;\; y \in S(x), \end{cases}$$

S. Dempe
Department of Mathematics and Computers Sciences, Technical University Bergakademie Freiberg, Freiberg, Germany
e-mail: dempe@tu-freiberg.de

N. Gadhi (✉)
Department of Mathematics, Sidi Mohamed Ben Abdellah University, Dhar Al Mehrez, B.P. 1796 Atlas, Fez, Marokko
e-mail: ngadhi@math.net

where, for each $x \in \mathbb{R}^n$, $S(x)$ is the solution set of the following parametric optimization problem (the lower level problem)

$$\begin{cases} \underset{y}{\text{Minimize}}\ g(x, y) \\ \text{subject to} : G(x, y) \cap \left(-\mathbb{R}^q_+\right) \neq \emptyset, \end{cases}$$

where $f,\ g : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \longrightarrow \mathbb{R}$ are continuous functions, $G : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^q$ and $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^p$ are given set-valued mappings; $n_1 \geq 1$ and $n_2 \geq 1$ are integers. A pair $(\overline{x}, \overline{y})$ is said to be a local optimal solution of (P) if it is a local optimal solution to the following problem:

$$\min_{(x,y)\in\overline{S}} f(x, y)$$

where

$$\overline{S} = \left\{(x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} :\ F(x, y) \cap \left(-\mathbb{R}^p_+\right) \neq \emptyset \text{ and } y \in S(x)\right\}.$$

Note that we use the optimistic approach in bilevel programming here.

A lot of research has been carried out in bilevel optimization problems [4,7,13,21,37,40,41,46,47,49], see also the monograph [14] and the annotated bibliography [15]. Ye and Zhu [47] give optimality conditions without convexity assumption on the lower level problem and without the assumption that the solution set $S(x)$ is a singleton. Under semi-Lipschitz property, Zhang [49] extends the classical approach to allow nonsmooth problem data; he derives existence and optimality conditions for problems using the coderivative due to Mordukhovich applied to the graph of the solution multifunction to the lower-level problem. Ye [48] first reformulates the bilevel programming problem as a one-level one and then uses a generalized derivative to derive necessary optimality conditions.

Due to the formulation of the constraints and the use of the solution set mapping of the lower level programming problem in the constraints of the upper level one, problem (P) is a special type of a set-valued optimization problem. Generally speaking set-valued optimization means set-valued analysis and its application to optimization, and it is an extension of continuous optimization to the set-valued case. In this research area one investigates optimization problems with constraints and/or an objective function described by set-valued maps, or investigations in set-valued analysis are applied to standard optimization problems. In the last decade there has been an increasing interest in set-valued optimization [20,24,28–30,32,33]. General optimization problems with set-valued constraints or a set-valued objective function are closely related to problems in stochastic programming [34], interval programming [6], vector optimization [23] and optimal control [8]. If the values of a given function vary in a specified region, this fact could be described by using a membership function in the theory of fuzzy sets or using information on the distribution of the function values. Optimal control problems with differential inclusions belong to this class of set-valued optimization problems as well.

It is our aim to develop sharp necessary optimality conditions for problem (P). For this we need a generalized derivative of continuous functions. In recent years, a great deal of research in nonsmooth analysis has focused on the development of generalized subdifferentials that provide sharp extremality conditions and good calculus rules for nonsmooth functions [8,10,12,39,42,45]. Very recently, as an extension of the notion of subdifferentials, the idea of convexificators has been used to extend, unify, and sharpen various results in nonsmooth analysis and optimization [12,25,26]. In Ref. [27], Jeyakumar and Luc gave a revised version of convexificators by introducing the notion of a convexificator which is a closed set but is not necessarily bounded or convex. Such a new notion will allow applications of convexificators to the above bilevel optimization problem under continuous data. For a locally Lipschitz function, most known subdifferentials such as the subdifferentials of Clarke [8], Michel-Penot [35,36], Ioffe-Mordukhovich [38] and Treiman [45] are convexificators. For more details, see [27] and the references therein. Moreover, for a continuous function, the symmetric subdifferential is a convexificator (see Proposition 1).

Our approach consists of using a support function [1,2,17,18,44] for the study of necessary optimality conditions for bilevel optimization problems. In [17], Dien gave a characterization of a set-valued mapping by its support function. Applying the support function, the bilevel programming problem is transformed into an equivalent problem for which, by use of the convexificator, necessary optimality conditions can be derived. Throughout, the data are assumed to be continuous but not necessarily locally Lipschitz. Therefore, the result we establish by means of convexificators is not only valid for locally Lipschitz optimization problems. Convexificators are subsets of many other generalized derivatives. Hence, our optimality conditions are stronger than those using e.g., the generalized derivative due to Clarke or Michel-Penot.

If additionally a certain appropriate regularity condition is satisfied we are able to detect necessary optimality conditions in terms of Karush-Kuhn-Tucker multipliers. Some examples that illustrate the usefulness of convexificators are also given.

The rest of the paper is written as follows: Section 2 contains basic definitions and preliminary results. Section 3 is devoted to the optimality conditions.

2 Preliminaries

Let $f : \mathbb{R}^p \to \mathbb{R} \cup \{+\infty\}$ be an extended real valued function. The expressions

$$f_d^-(x, v) := \liminf_{t \searrow 0} [f(x + tv) - f(x)]/t,$$
$$f_d^+(x, v) := \limsup_{t \searrow 0} [f(x + tv) - f(x)]/t$$

signify, respectively, the lower and upper Dini directional derivatives of f at x in the direction of v.

Definition 1 [27] The function $f : \mathbb{R}^p \to \mathbb{R} \cup \{+\infty\}$ is said to have a convexificator $\partial^* f(x)$ at x if $\partial^* f(x) \subset \mathbb{R}^p$ is a closed set and, for each $v \in \mathbb{R}^p$,

$$f_d^-(x, v) \le \sup_{x^* \in \partial^* f(x)} \langle x^*, v\rangle \text{ and } f_d^+(x, v) \ge \inf_{x^* \in \partial^* f(x)} \langle x^*, v\rangle .$$

Note that convexificators are not necessarily compact or convex [12]. These relaxations allow applications to a large class of nonsmooth continuous functions.

The Clarke generalized subdifferential [8] $\partial^c f(x)$ of f at x defined by

$$\partial^c f(x) := \left\{ x^* \in \mathbb{R}^p : \limsup_{u \to x,\ t \searrow 0} \frac{f(u + tv) - f(u)}{t} \ge \langle x^*, v\rangle \ \forall\, v \in \mathbb{R}^p \right\}$$

is a convexificator of f at x when f is locally Lipschitz. However, the convex hull of a convexificator of a locally Lipschitz function may be strictly contained in the Clarke subdifferential.

To progress, we need the following definition.

Definition 2 [11] A set valued mapping $F : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ is upper semicontinuous (u.s.c.) at x, if for each $\varepsilon > 0$, there exists $\delta > 0$ such that, for each $x' \in x + \delta\mathbb{B}_{\mathbb{R}^p}$, we have

$$F\left(x'\right) \subset F(x) + \varepsilon\mathbb{B}_{\mathbb{R}^q},$$

where $\mathbb{B}_{\mathbb{R}^p}$ and $\mathbb{B}_{\mathbb{R}^q}$ are the closed unit balls in $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively.

An example of a convexificator for a function which is not locally Lipschitz continuous is given as follows.

Definition 3 [39] Let $f : \mathbb{R}^p \to \overline{\mathbb{R}} := [-\infty, +\infty]$ be an extended real valued function and $\overline{x} \in \mathrm{dom}(f)$. The symmetric subdifferential of f at $\overline{x}$ is defined by

$$\partial^0 f(\overline{x}) := \partial f(\overline{x}) \cup [-\partial(-f)(\overline{x})]$$

where $\partial f(\overline{x}) := \limsup\limits_{x \xrightarrow{f} \overline{x},\ \varepsilon \searrow 0} \widehat{\partial}_\varepsilon f(x)$ and $\widehat{\partial}_\varepsilon f(x)$ is the $\varepsilon-$ Fréchet subdifferential of f at x defined by

$$\widehat{\partial}_\varepsilon f(x) := \left\{x^* \in \mathbb{R}^p : \liminf_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x\rangle}{\|u - x\|} \geq -\varepsilon\right\},\ \varepsilon \geq 0.$$

Here, $x \xrightarrow{f} \overline{x}$ is an abbreviation of $x \to \overline{x}$ and $f(x) \to f(\overline{x})$. For more details see [38,39].

Note that sufficient conditions for upper semicontinuity of $\partial^0 f(.)$ can be found in [22] and [31].

Proposition 1 *Let $f : \mathbb{R}^p \to \overline{\mathbb{R}} := [-\infty, +\infty]$ be continuous and $\overline{x} \in dom(f)$. Suppose that $\partial^0 f(\overline{x})$ is closed and that $\partial^0 f(.)$ is upper semicontinuous at $\overline{x}$. Then $\partial^0 f(\overline{x})$ is a convexificator of f at $\overline{x}$.*

Proof Let $\varepsilon > 0$. By the upper semicontinuity of $\partial^0 f(.)$, there exists $\delta > 0$ such that

$$\partial^0 f(x) \subset \partial^0 f(\overline{x}) + \varepsilon\mathbb{B}_{\mathbb{R}^q},$$

for all $x \in \overline{x} + \delta\mathbb{B}_{\mathbb{R}^p}$.

Using Theorem 2.3 of [26] (the mean value theorem), there exists $c \in]x, \overline{x}[$ such that

$$f(x) - f(\overline{x}) \in \partial^0 f(c)(x - \overline{x}) \subset \partial^0 f(\overline{x})(x - \overline{x}) + \varepsilon\|x - \overline{x}\|\mathbb{B}_{\mathbb{R}}.$$

Now, let $v \in \mathbb{R}^p$. Since $\mathbb{B}_{\mathbb{R}}$ is compact,

$$f_d^-(\overline{x}, v) \in \partial^0 f(\overline{x})(v) + \varepsilon\|v\|\mathbb{B}_{\mathbb{R}} \text{ and } f_d^+(\overline{x}, v) \in \partial^0 f(\overline{x})(v) + \varepsilon\|v\|\mathbb{B}_{\mathbb{R}}.$$

Consequently, there exist $x_1^*,\ x_2^* \in \partial^0 f(\overline{x})$ and $b_1,\ b_2 \in \mathbb{B}_{\mathbb{R}}$ such that

$$f_d^-(\overline{x}, v) = \left\langle x_1^*, v\right\rangle + \varepsilon\|v\| b_1 \text{ and } f_d^+(\overline{x}, v) = \left\langle x_2^*, v\right\rangle + \varepsilon\|v\| b_2.$$

Then,

$$f_d^-(\overline{x}, v) \leq \sup_{x^* \in \partial^0 f(\overline{x})} \left\langle x^*, v\right\rangle + \varepsilon\|v\| \text{ and } f_d^+(\overline{x}, v) \geq \inf_{x^* \in \partial^0 f(\overline{x})} \left\langle x^*, v\right\rangle - \varepsilon\|v\|.$$

Letting $\varepsilon \to 0$, one gets

$$f_d^-(\overline{x}, v) \leq \sup_{x^* \in \partial^0 f(x)} \left\langle x^*, v\right\rangle \text{ and } f_d^+(\overline{x}, v) \geq \inf_{x^* \in \partial^0 f(x)} \left\langle x^*, v\right\rangle.$$

The proof is finished. □

Now, we recall the chain rule for composite functions in terms of convexificators established by Jeyakumar and Luc in [27].

Proposition 2 [27] *Let $f = (f_1, \ldots, f_n)$ be a continuous function from $\mathbb{R}^p$ to $\mathbb{R}^n$, and g be a continuous function from $\mathbb{R}^n$ to $\mathbb{R}$. Suppose that, for each $i = 1, 2, \ldots, n$, f_i admits a bounded convexificator $\partial^* f_i(\overline{x})$ and that g admits a bounded convexificator $\partial^* g(f(\overline{x}))$ at $f(\overline{x})$. For each $i = 1, \ldots, n$, if $\partial^* f_i$ is u.s.c. at $\overline{x}$ and $\partial^* g$ is u.s.c. at $f(\overline{x})$, then the set $\partial^*(g \circ f)(\overline{x}) := \left\{ \sum_{i=1}^{n} a_i\, \partial^* f_i(\overline{x}) : (a_1, \ldots, a_n) \in \partial^* g(f(\overline{x})) \right\}$ is a convexificator of the function $g \circ f$ at $\overline{x}$.*

Since the convex hull of a convexificator of a locally Lipschitz function may be strictly contained in the Clarke subdifferential, Corollary 1 is an extension of Proposition 2.3.12 [8].

Corollary 1 *Let $f = (f_1, \ldots, f_n)$ be a continuous function from $\mathbb{R}^p$ to $\mathbb{R}^n$. Suppose that for $i = 1, 2, \ldots, n$, the function f_i admits a bounded convexificator $\partial^* f_i(\overline{x})$ at $\overline{x}$. Let*

$$h(x) = \max\{f_i(x) : i = 1, 2, \ldots, n\}$$

and $I(\overline{x}) = \{i : f_i(\overline{x}) = h(\overline{x})\}$. Then $co \left\{ \bigcup_{i \in I(\overline{x})} \partial^* f_i(\overline{x}) \right\}$ is a convexificator of h at $\overline{x}$, where, "co" denotes the convex hull.

The following result is a variant of Theorem 2.8.2 [8]. See also [9,44].

Corollary 2 *Let $\mathbb{T}$ be a subset of $\mathbb{R}^p$, $\overline{x} \in \mathbb{R}^p$, $t \in \mathbb{T}$, $f_t : \mathbb{R}^p \to \overline{\mathbb{R}}$ and*

$$h(x) = \sup_{t \in \mathbb{T}} \{f_t(x)\} \text{ and } J(\overline{x}) = \{t \in \mathbb{T} : f_t(\overline{x}) = h(\overline{x})\}.$$

Suppose that there exists a neighborhood U of $\overline{x}$ in $\mathbb{R}^p$ such that for each $t \in \mathbb{T}$, the function f_t is finite on U and admits a bounded convexificator on U. If in addition $t \longmapsto f_t$ is upper semicontinuous then, $clco\{\partial^* f_t(\overline{x}) : t \in J(\overline{x})\}$ is a convexificator of h at $\overline{x}$.

Proof It suffices to repeat (with very slight modification) the argument of the first part of the proof of Theorem 2.8.2 in Clarke [8]. □

Let $H : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ be a set-valued mapping. For every $y^* \in \mathbb{R}^q$, the support function of H at x is defined by

$$C_H(y^*, x) := \sup_{y \in H(x)} < y^*, y >,$$

where $\langle ., . \rangle$ is the dual pairs.

Suppose that for all $x \in \mathbb{R}^p$, $H(x)$ is a non empty, closed and convex set. The distance function of H to $y \in \mathbb{R}^q$,

$$d(y, H(x)) = \inf\{\|y - z\| : z \in H(x)\}$$

is related to the support function of H by the relation

$$d(y, H(x)) = \max_{y^* \in Y_H^* \cap \mathbb{B}^q} \left\{ \langle y^*, y \rangle - C_H(y^*, x) \right\},$$

where Y_H^* denotes the barrier cone of H defined by

$$Y_H^* := \left\{ y^* \in \mathbb{R}^q : \sup_{y \in H(x)} < y^*, y > < +\infty \right\}.$$

If $d(y, H(x)) > 0$ then there is a unique $y^* \in Y_H^* \cap \mathbb{B}_{\mathbb{R}^q}$ satisfying $\|y^*\| = 1$ and $d(y, H(x)) = \langle y^*, y\rangle - C_H(y^*, x)$, see [17] and [44].

Using Corollary 2, one can deduce the following result which is an extension of Proposition 2.2 [17].

Corollary 3 *Suppose that there exists a neighborhood U of $\overline{x}$ such that for each $x \in U$ and $y^* \in Y_H^* \cap \mathbb{B}_{\mathbb{R}^q}$, the support function $C_H(y^*, .)$ is continuous on U and admits a bounded convexificator $\partial^* C_H(y^*, .)(x) := \partial_x^* C_H(y^*, x)$. Then, for all $x \in X$ and $y \in Y$, the distance function $d(y, H(x))$ admits*

$$cl\ co \bigcup_{y^* \in J(x,y)} \left\{-\partial_x^* C_H\left(y^*, x\right) \times \left\{y^*\right\}\right\}$$

as a convexificator at (x, y). Here,

$$J(x, y) = \left\{y^* \in Y_H^* : \left\|y^*\right\| \leq 1 \text{ and } d(y, H(x)) = \left\langle y^*, y\right\rangle - C_H\left(y^*, x\right)\right\}.$$

If, in addition, $d(y, H(x)) > 0$, then $J(x, y)$ consists of only one single element y^* with $\|y^*\| = 1$ and the symbol "co" can be deleted.

In what follows, the set valued mappings F and G are assumed to have the following property (cl-property)

(i) If $x_n^* \in \partial^* C_F\left(y_n^*, \cdot\right)(x_n)$ where $x_n^* \to x^*, y_n^* \to y^*$ and $x_n \to x$, then $x^* \in \partial^* C_F(y^*, \cdot)(x)$.

(ii) If $x_n^* \in \partial^* C_G\left(y_n^*, \cdot\right)(x_n)$ where $x_n^* \to x^*, y_n^* \to y^*$ and $x_n \to x$, then $x^* \in \partial^* C_G(y^*, \cdot)(x)$.

Remark 4 1. The above property was first introduced by Dien [17,18] for locally Lipschitz set-valued mappings. He calls it the cl-property.
2. In some cases, the cl-property can be established without difficulty. See the following example.
3. The cl-property can be seen as the sequentially u.s.c. of $\partial^* C_H(y^*, .)$.

Example 1 Let $y^* \in \mathbb{B}_{\mathbb{R}^q}^+$ and $F(x) = f(x) - \mathbb{B}_{\mathbb{R}^q}^+$, where $f\colon \mathbb{B}_{\mathbb{R}^p} \to \mathbb{B}_{\mathbb{R}^q}$ is a locally Lipschitz mapping.

Suppose that $x_n^* \in \partial^* C_F\left(y_n^*, \cdot\right)(x_n)$ where $x_n^* \to x^*$, $y_n^* \to y^*$ and $x_n \to x$. Remarking that $C_F\left(y_n^*, x_n\right) = \left\langle y_n^*, f(x_n)\right\rangle$, one has $x_n^* \in \partial^* \left\langle y_n^*, f\right\rangle(x_n)$. Then, $x^* \in \partial^* \langle y^*, f\rangle(x)$.

3 Optimality conditions

According to [15,47], (P) can be replaced by

$$\left(P^*\right): \quad \begin{cases} \displaystyle\operatorname*{Minimize}_{x,y} f(x, y) \\ \text{subject to :} \\ \quad F(x, y) \cap \left(-\mathbb{R}_+^p\right) \neq \emptyset,\ \ G(x, y) \cap \left(-\mathbb{R}_+^q\right) \neq \emptyset, \\ \quad g(x, y) - V(x) \leq 0, \\ \quad (x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \end{cases}$$

where

$$V(x) := \min_y \left\{g(x, y) : G(x, y) \cap \left(-\mathbb{R}_+^q\right) \neq \emptyset,\ \ y \in \mathbb{R}^{n_2}\right\}.$$

Remark 5 Under the following hypotheses (H_1), (H_2) and (H_3), the optimization problem (P) has at least one optimal solution.

(H_1): $f(.,.)$, $g(.,.)$ are continuous, $F(.,.)$ and $G(.,.)$ are continuous on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$;
(H_2): $V(.)$ is upper semicontinuous (u.s.c.) on $\mathbb{R}^{n_1}$;
(H_3): The problem (P^*) has at least one feasible solution (i.e. the infimal value v^* of the function $f(.,.)$ on the feasible set of this problem is less than infinity), there exists $v^* < c < \infty$ such that

$$M := \{(x,y) : G(x,y) \cap \left(-\mathbb{R}^q_+\right) \neq \emptyset,\ F(x,y) \cap \left(-\mathbb{R}^p_+\right) \neq \emptyset,\ f(x,y) \leq c\}$$

is not empty and bounded.

The following regularity assumption will be used to get Karush-Kuhn-Tucker multipliers.

Definition 4 $(\overline{x}, \overline{y})$ is said to be a regular point of (P) if $\overline{z}_1 \in F(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^p_+\right)$, $\overline{z}_2 \in G(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^q_+\right)$, $y^* = (\mu^*, \nu^*, \gamma)$, $\|y^*\| = 1$ and

$$C_F\left(\mu^*, (\overline{x}, \overline{y})\right) + C_G\left(\nu^*, (\overline{x}, \overline{y})\right) = \left\langle \mu^*, \overline{z}_1 \right\rangle + \left\langle \nu^*, \overline{z}_2 \right\rangle$$

imply

$$0 \notin \partial^* C_F(\mu^*, .)(\overline{x}, \overline{y}) + \partial^* C_G\left(\nu^*, .\right)(\overline{x}, \overline{y}) + \gamma \partial^* g(\overline{x}, \overline{y}) - \gamma\left(\partial^* V(\overline{x}) \times \{0\}\right),$$

where

$$\partial^* V(\overline{x}) \subseteq co\left\{\partial^* g(.,y)(\overline{x})\ \forall y : G(\overline{x}, y) \cap \left(-\mathbb{R}^q_+\right) \neq \emptyset,\ g(\overline{x}, y) = V(\overline{x})\right\}. \tag{1}$$

Theorem 6 *Let $(\overline{x}, \overline{y})$ be a local optimal solution of (P) such that*

$$(\overline{x}, \overline{y}) \notin Arg\min\left\{f(x,y) : (x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}\right\}.$$

Suppose that there exists a neighborhood U of $(\overline{x}, \overline{y})$ such that the functions f and g are continuous on U and admit bounded convexificators $\partial^* f(\overline{x}, \overline{y})$ and $\partial^* g(\overline{x}, \overline{y})$, the set valued mappings $F$and G have the cl-properties (i) and (ii) and satisfy assumptions (H_1), (H_2), (H_3), and the support functions of F and G admit at the point $(\overline{x}, \overline{y})$ bounded convexificators $\partial^* C_F(\mu^*, .)(\overline{x}, \overline{y})$ and $\partial^* C_G(\upsilon^*, .)(\overline{x}, \overline{y})$. Also, assume that $\partial^* f$, $\partial^* g$, $\partial^* C_F(\mu^*, .)$ and $\partial^* C_G(\upsilon^*, .)$ are upper semicontinuous at $(\overline{x}, \overline{y})$.

Then, for all $\overline{z}_1 \in F(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^p_+\right)$ and $\overline{z}_2 \in G(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^q_+\right)$ there exist scalars $\lambda_1, \lambda_2 \geq 0$ and vectors $t^* = (\mu^*, \upsilon^*, \gamma^*) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}_+$ such that

$$\left\|\left(\mu^*, \upsilon^*, \gamma^*\right)\right\| = 1 \text{ and } \lambda_1 + \lambda_2 = 1,$$

$$C_F\left(\mu^*, (\overline{x}, \overline{y})\right) = \left\langle \mu^*, \overline{z}_1 \right\rangle \text{ and } C_G\left(\upsilon^*, (\overline{x}, \overline{y})\right) = \left\langle \upsilon^*, \overline{z}_2 \right\rangle, \tag{2}$$

$$\begin{aligned}(0,0) \in \lambda_1\, co\partial^* f(\overline{x}, \overline{y}) & \\ -\lambda_2 & \left[\partial^* C_F\left(\mu^*, .\right)(\overline{x}, \overline{y}) + \partial^* C_G\left(\upsilon^*, .\right)(\overline{x}, \overline{y})\right] \\ -\lambda_2 & \,\gamma^* \left[\partial^* g(\overline{x}, \overline{y}) - \left(\partial^* V(\overline{x}) \times \{0\}\right)\right],\end{aligned} \tag{3}$$

where $\partial^* V(\overline{x})$ satisfies (1).

If in addition to the above assumptions, the problem (P) is regular at $(\overline{x}, \overline{y})$, one has

$$\lambda_1 > 0.$$

Proof The proof of this theorem consists of several steps.

Let $(\overline{x}, \overline{y})$ is be a local optimal solution to (P), $\overline{z}_1 \in F(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^p_+\right)$ and $\overline{z}_2 \in G(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}^q_+\right)$. According to [15], it is also a local optimal solution of

$$\left(P^*\right): \begin{cases} \text{Minimize } f(x, y) \\ \text{subject to : } H(x, y) \cap \left(-Z^+\right) \neq \emptyset, \\ (x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, \end{cases}$$

where

$$Z := \mathbb{R}^{p+q+1} \text{ and } H(x, y) := (F(x, y), G(x, y), g(x, y) - V(x)).$$

First Step Let

$$\Psi_1(x, y, z) := f(x, y) + \delta_{-Z^+}(z) - f(\overline{x}, \overline{y}) + \frac{1}{n},$$
$$\Psi_2(x, y, z) := d(z, H(x, y))$$

and

$$h_n(x, y, z) := \max(\Psi_1(x, y, z), \Psi_2(x, y, z)).$$

Set $\overline{z} := (\overline{z}_1, \overline{z}_2, 0)$. Then we have

$$h_n(\overline{x}, \overline{y}, \overline{z}) \leq \frac{1}{n} + \inf_{(x,y,z) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times Z} h_n(x, y, z).$$

By using Ekeland's Variational Principle [19], there exists a sequence $(x_n, y_n, z_n) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times Z$ such that $z_n := (z_{1n}, z_{2n}, z_{3n})$ and

$$\begin{cases} \|(x_n, y_n, z_n) - (\overline{x}, \overline{y}, \overline{z})\| \leq \dfrac{1}{\sqrt{n}} \\ h_n(x_n, y_n, z_n) \leq h_n(x, y, z) + \dfrac{1}{\sqrt{n}} \|(x, y, z) - (x_n, y_n, z_n)\|, \end{cases}$$

for all $(x, y, z) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times Z$. Hence (x_n, y_n, z_n) is a local minimum of $h_n(x, y, z) + \frac{1}{\sqrt{n}} \|(x, y, z) - (x_n, y_n, z_n)\|$ and we get

$$0 \in cl\ co\ \partial^* \left(h_n + \frac{1}{\sqrt{n}} \|. - (x_n, y_n, z_n)\|\right)(x_n, y_n, z_n).$$

Consequently,

$$0 \in cl\ co\ \partial^* h_n(x_n, y_n, z_n) + \frac{1}{\sqrt{n}} \mathbb{B}_{\mathbb{R}^{n_1+n_2} \times Z}.$$

In view of Corollary 1, it follows that

$$\partial^* h_n \subset co\left\{\partial^* \Psi_i : i \in I(x_n, y_n, z_n)\right\},$$

where $I(x_n, y_n, z_n) := \{i : h_n(x_n, y_n, z_n) = \Psi_i(x_n, y_n, z_n)\}$.

Consequently, there exist $\lambda_{n,1}, \lambda_{n,2} \in [0, 1]$ such that $\lambda_{n,1} + \lambda_{n,2} = 1$ and

$$0 \in \lambda_{n,1} co \partial^* \Psi_1(x_n, y_n, z_n) + \lambda_{n,2} co \partial^* \Psi_2(x_n, y_n, z_n) + \frac{1}{\sqrt{n}} \mathbb{B}_{\mathbb{R}^{n_1+n_2} \times Z} \quad (4)$$

where

$$\begin{aligned} &\lambda_{n,1} = 0 && \text{if } \Psi_1(x_n, y_n, z_n) < \Psi_2(x_n, y_n, z_n),\\ &\lambda_{n,2} = 0 && \text{if } \Psi_2(x_n, y_n, z_n) < \Psi_1(x_n, y_n, z_n),\\ &0 \le \lambda_{n,1} \le 1,\ 0 \le \lambda_{n,2} \le 1 && \text{if } \Psi_1(x_n, y_n, z_n) = \Psi_2(x_n, y_n, z_n). \end{aligned}$$

Second Step We have $\max(\Psi_1(x_n, y_n, z_n), \Psi_2(x_n, y_n, z_n)) > 0$, since otherwise

$$\begin{cases} d(z_n, H(x_n, y_n)) = 0, \\ f(x_n, y_n) - f(\overline{x}, \overline{y}) + \delta_{-Z^+}(z_n) + \dfrac{1}{n} = 0, \end{cases}$$

so that $z_n \in H(x_n, y_n)$ and $f(x_n, y_n) - f(\overline{x}, \overline{y}) = -\delta_{-Z^+}(z_n) - \dfrac{1}{n}$. Since $(\overline{x}, \overline{y})$ is an optimal solution of the problem (P), one has $f(x_n, y_n) - f(\overline{x}, \overline{y}) \ge 0$, a contradiction.

Moreover, $\Psi_2(x_n, y_n, z_n) > 0$, since otherwise $\lambda_{n,2} = 0, \lambda_{n,1} = 1$ and $0 \in co\partial^* f(x_n, y_n) + \frac{1}{\sqrt{n}} \mathbb{B}_{\mathbb{R}^{n_1+n_2}}$. Then, $0 \in co\partial^* f(\overline{x}, \overline{y})$, which contradicts the assumption.

From (4), and using Corollary 3, there exist $t_n^* = \left(\mu_n^*, \upsilon_n^*, \gamma_n\right) \in \mathbb{R}_+^{m_1} \times \mathbb{R}_+^{m_2} \times \mathbb{R}_+$ such that $\left\|t_n^*\right\| = 1$ and

$$\begin{cases} 0 \in \lambda_{n,1}\, co\partial^* f(x_n, y_n) - \lambda_{n,2}\, co\partial^* C_H\left(t_n^*, .\right)(x_n, y_n) + \frac{1}{\sqrt{n}} \mathbb{B}_{\mathbb{R}^{n_1+n_2} \times Z}, \\ 0 \in \lambda_{n,1} N_{-Z^+}(z_n) + \lambda_{n,2} t_n^* + \frac{1}{\sqrt{n}} \mathbb{B}_Z, \\ d(z_n, H(x_n, y_n)) = \left\langle t_n^*, z_n\right\rangle - C_H\left(t_n^*, (x_n, y_n)\right), \end{cases} \tag{5}$$

with

$$\begin{aligned} \partial^* C_H\left(t_n^*, .\right)(x_n, y_n) = {} & \partial^* C_F\left(\mu_n^*, .\right)(x_n, y_n) \\ & + \partial^* C_G\left(\upsilon_n^*, .\right)(x_n, y_n) + \gamma_n \partial^* g(x_n, y_n) - \gamma_n\left(\partial^* V(x_n) \times \{0\}\right). \end{aligned}$$

Taking a subsequence if necessary, we can assume that for n tending to $+\infty$: $(\lambda_{n,1}) \to \lambda_1 \in [0, 1]$, $(\lambda_{n,2}) \to \lambda_2 \in [0, 1]$, $\mu_n^* \to \mu^* \geqslant 0$, $\upsilon_n^* \to \upsilon^* \geqslant 0$, $\gamma_n^* \to \gamma^* \ge 0$, $t_n^* \to t^* = (\mu^*, \upsilon^*, \gamma^*) \in \mathbb{R}_+^{m_1} \times \mathbb{R}_+^{m_2} \times \mathbb{R}_+$ and $\|(\mu^*, \upsilon^*, \gamma^*)\| = 1$.

Then, we get Eq. 3:

$$\begin{aligned} 0 \in \lambda_1\, & co\partial^* f(\overline{x}, \overline{y}) - \lambda_2[\partial^* C_F\left(\mu^*, .\right)(\overline{x}, \overline{y}) + \partial^* C_G\left(\upsilon^*, .\right)(\overline{x}, \overline{y}) \\ & + \gamma^* \partial^* g(\overline{x}, \overline{y}) - \gamma^*(\partial^* V(\overline{x}) \times \{0\})]. \end{aligned}$$

Using Corollary 2, $co\{\partial^* g(., y)(\overline{x}) : y \in J(\overline{x})\}$ can be taken as a convexificator of V at $\overline{x}$. We remind the reader that

$$J(\overline{x}) = \left\{y \in \mathbb{R}^{n_2} : G(\overline{x}, y) \cap \left(-\mathbb{R}_+^q\right) \neq \emptyset \text{ and } g(\overline{x}, y) = V(\overline{x})\right\}.$$

Third Step At last we have to show Eq. 2:

On the one hand, since $\overline{z} \in H(\overline{x}, \overline{y})$, we have $C_H(t^*, (\overline{x}, \overline{y})) \ge \langle t^*, \overline{z}\rangle$ for each $t^* \in \mathbb{R}_+^{m_1} \times \mathbb{R}_+^{m_2} \times \mathbb{R}_+$. That is, for all $\overline{z}_1 \in F(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}_+^p\right)$ and $\overline{z}_2 \in G(\overline{x}, \overline{y}) \cap \left(-\mathbb{R}_+^q\right)$ we have

$$C_F\left(\mu^*, (\overline{x}, \overline{y})\right) + C_G\left(\upsilon^*, (\overline{x}, \overline{y})\right) + \gamma^* g(\overline{x}, \overline{y}) - \gamma^* V(\overline{x}) \ge \left\langle \mu^*, \overline{z}_1\right\rangle + \left\langle \upsilon^*, \overline{z}_2\right\rangle, \tag{6}$$

for each $t^* = (\mu^*, \upsilon^*, \gamma^*) \geqslant 0$.

Since $F(., .)$ and $G(., .)$ are continuous, one have

$$\begin{aligned} C_F\left(\mu^*, (\overline{x}, \overline{y})\right) &= \liminf_{n\to\infty} C_F\left(\mu_n^*, (x_n, y_n)\right) \\ &= \liminf_{n\to\infty} \left[\left\langle \mu_n^*, z_{1n}\right\rangle - d(z_{1n}, F(x_n, y_n))\right] \\ &\le \left\langle \mu^*, \overline{z}_1\right\rangle \end{aligned} \tag{7}$$

and

$$C_G\left(\upsilon^*, (\overline{x}, \overline{y})\right) = \liminf_{n\to\infty} C_G\left(\upsilon_n^*, (x_n, y_n)\right) \tag{8}$$
$$= \liminf_{n\to\infty} \left[\left\langle \upsilon_n^*, z_{2n}\right\rangle - d\left(z_{2n}, G\left(x_n, y_n\right)\right)\right]$$
$$\leq \left\langle \mu^*, \overline{z}_2\right\rangle.$$

Using the sum of (7) and (8) together with (6) and the fact that

$$g\left(\overline{x}, \overline{y}\right) - V\left(\overline{x}\right) = 0,$$

one can deduce

$$C_F\left(\mu^*, (\overline{x}, \overline{y})\right) = \left\langle \mu^*, \overline{z}_1\right\rangle \text{ and } C_G\left(\upsilon^*, (\overline{x}, \overline{y})\right) = \left\langle \upsilon^*, \overline{z}_2\right\rangle. \tag{9}$$

Under the regularity assumption of (P) at $(\overline{x}, \overline{y})$, one can prove that $\lambda_1 > 0$. □

With the following example, we illustrate the usefulness of our necessary optimality conditions. This example show that optimality conditions using convexificators are stronger and more general than those using the Clarke subdifferential.

Example 2 We consider the following bilevel optimization problem $(P^{\rhd})$

$$\begin{cases} \underset{x,y}{\text{Minimize}}\ y \\ \text{subject to}: |x| + |y| = 0,\ \ y \in S(x), \end{cases}$$

where, for each $x \in R$, $S(x)$ is the solution set of the following parametric optimization problem (parameterized in x)

$$\begin{cases} \underset{y}{\text{Maximize}}\ x \\ \text{subject to}: x + |y| \leq 0. \end{cases} \tag{10}$$

This is a special case of the general type (P). The example problem is illustrated in Fig. 1. Observing that $(0, 0)$ is not a local minimum of $(P^{\rhd})$, the usefulness of our Theorem becomes clear when we remark that the optimality conditions using convexificators are not satisfied and the optimality conditions using the Clarke subdifferential are satisfied.

On the one hand, using the Clarke generalized derivative, necessary optimality conditions reduces to the existence of scalars $\lambda_0,\ \lambda_1,\ \lambda_2 \geq 0$ satisfying

$$(0, 0) \in \lambda_0\,(0, -1) + \lambda_1\mathbb{R}^2 + \lambda_2 co\left\{(1, -1), (1, 1)\right\}.$$

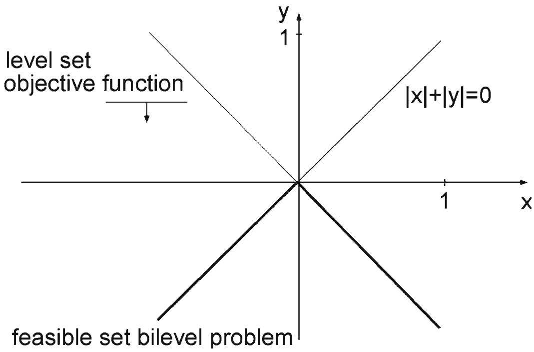

Fig. 1 Illustration of Example 2: The feasible set of the bilevel programming problem is drawn with thick lines. A level set of the objective function of the upper-level problem is also given. The picture shows that the origin is not a locally optimal solution

On the other hand, using convexificators,

$$(0,0) \notin \lambda_0 (0,-1) + \lambda_1 \{(1,-1),(-1,1)\} + \lambda_2 co \{(1,-1),(1,1)\}$$

for all $\lambda_0,\ \lambda_1,\ \lambda_2 \geq 0$.

Remark 7 Problem (10) can be reformulated as a parametric linear programming problem. Properties of such problems can be found e.g. in [5]. If the optimal solution of the lower level problem is uniquely determined and directionally differentiable [43], the bilevel programming problem can be transformed into an equivalent one-level problem and solved by descent algorithms [16]. In this case, paper [15] describes necessary and sufficient optimality conditions for the bilevel programming problem. Since the assumptions in [15] are very restrictive, the results developed here are better suited for bilevel programming problems than those using parametric optimization. This concerns especially the uniqueness assumption for an optimal solution in the lower-level problem which is not used here.

Remark 8 Optimality conditions established by means of convexificators remain valid for locally Lipschitz optimization problems when the Clarke subdifferential, the Michel-Penot subdifferential, the Ioffe-Mordukhovich subdifferential or the Treiman subdifferential is used.

4 Special cases

Consider the following bilevel optimization problem (P^*)

$$(P_1): \quad \begin{cases} \text{Minimize } f(x,y) \\ \text{subject to : } F(x,y) \leq 0, \ \ y \in S(x), \end{cases}$$

where, for each $x \in X$, $S(x)$ is the solution set of the following parametric optimization problem (parameterized in x) [5]

$$(P_2): \quad \begin{cases} \underset{y}{\text{Minimize}}\ g(x,y) \\ \text{subject to : } G(x,y) \leq 0, \end{cases}$$

where $f,\ g,\ F,\ G : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \longrightarrow \mathbb{R}$ are given; n_1 and n_2 are integers with $n_i \geq 1$.

Using Theorem 6, it is easy to deduce necessary optimality conditions for the bilevel optimization problem (P^*), since for functions F, the condition $F(x,y) \cap \left(-\mathbb{R}^p_+\right) \neq \emptyset$ is equivalent to $F(x,y) \leq 0$.

Corollary 9 [3] *Let $(\overline{x},\overline{y})$ be a solution of (P) such that*

$$(\overline{x},\overline{y}) \notin Arg \min \left\{ f(x,y) : (x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \right\}.$$

Suppose that the assumptions (H_1), (H_2), (H_3) are satisfied and that there exists a neighborhood U of $(\overline{x},\overline{y})$ such that the functions $F,\ f,\ g,\ G$ are continuous on U and admit bounded convexificators $\partial^* F(\overline{x},\overline{y})$, $\partial^* f(\overline{x},\overline{y})$, $\partial^* g(\overline{x},\overline{y})$ and $\partial^* G(\overline{x},\overline{y})$ at $(\overline{x},\overline{y})$. Also, suppose that $\partial^* F,\ \partial^* f,\ \partial^* g$ and $\partial^* G$ are upper semicontinuous at $(\overline{x},\overline{y})$.

Then, there exist scalars $\lambda_1,\ \lambda_2,\ \mu^*,\ \upsilon^*,\ \gamma^* \geq 0$ such that

$$\|(\mu^*,\upsilon^*,\gamma^*)\| = 1,\ \lambda_1 + \lambda_2 = 1,\ \mu^* F(\overline{x},\overline{y}) = 0,\ \upsilon^* G(\overline{x},\overline{y}) = 0 \ \textit{and}$$
$$(0,0) \in \lambda_1\, co\partial^* f(\overline{x},\overline{y}) - \lambda_2 \mu^* \partial^* F(\overline{x},\overline{y}) - \lambda_2 \upsilon^* \partial^* G(\overline{x},\overline{y})$$
$$-\lambda_2\, \gamma^* \left[\partial^* g(\overline{x},\overline{y}) - (\partial^* V(\overline{x}) \times \{0\})\right],$$

where

$$\begin{cases} \partial^* V(\overline{x}) \subseteq co\{\partial^* g(.,y)(\overline{x}) : y \in J(\overline{x})\}, \\ J(\overline{x}) = \{y \in \mathbb{R}^{n_2} : G(\overline{x}, y) \leq 0 \text{ and } g(\overline{x}, y) = V(\overline{x})\}. \end{cases}$$

If in addition to the above assumptions, the problem (P) *is regular at* $(\overline{x}, \overline{y})$, *one has*

$$\lambda_1 > 0.$$

5 Conclusion

In this work, we used convexificators and the support function to feasible set mapping to establish necessary optimality conditions for a set valued bilevel optimization problem with inclusion constraints. We assumed that all data are continuous but not necessary Lipschitz. We used an intermediate set-valued optimization problem together with an appropriate regularity condition to detect necessary optimality conditions in terms of Karush-Kuhn-Tucker multipliers. The approach used hier is more general than those using the Clarke subdifferential, the Mordukhovich subdifferential and the symmetric subdifferential. Optimality conditions using the above subdifferentials can be deduced from our result.

Acknowledgement Our sincere acknowledgements to the anonymous referees and the Associate editor for their insightful remarks and suggestions. This work has been supported by the Alexander-von-Humboldt foundation.

References

1. Amahroq, T., Gadhi, N.: Second order optimality conditions for an extremal problem under inclusion constraints. J. Math. Anal. Appl. **285**, 74–85 (2003)
2. Amahroq, T., Gadhi, N., Riahi, H.: Epi-differentiability and optimality conditions for an extremal problem. J. Inequal. Pure Appl. Math., Victoria University, ISSN electronic 1443–5756 **4**(2) Article 41 (2003)
3. Babahadda, H., Gadhi, N.: Necessary optimality conditions for bilevel optimization problems using convexificators (To appear) in J. Global. Optim.
4. Bard, J.F.: Some properties of the bilevel programming problem. J. Optim. Theory Appl. **68**, 371–378 (1991)
5. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Nonlinear Parametric Optimization. Akademie-Verlag, Berlin (1982)
6. Chanas, S., Delgado, M., Verdegay, J.L., Vila, M.A.: Interval and fuzzy extensions of classical transportation problems. Transport. Plan. Technol. **17**, 203–218 (1993)
7. Chen, Y., Florian, M.: On the geometry structure of linear bilevel programs: A dual approach, Technical Report CRT-867, Centre de Recherche sur les transports. Université de Montreal, Montreal, Quebec, Canada (1992)
8. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley-Interscience (1983)
9. Clarke, F.H.: Necessary conditions for a general control problem in calculus of variations and control. In: Russel D. (ed.) Mathematics research center, Pub.36. University of Wisconsin, pp. 259–278, New York Academy (1976)
10. Craven, B.D., Ralph, D., Glover, B.M.: Small convex-valued subdifferentials in mathematical programming. Optimization **32**, 1–21 (1995)
11. Dantzig, G.B., Folkman, J., Shapiro, N.: On the continuity of the minimum set of a continuous function. J. Math. Anal. Appl. **17**, 512–548 (1967)
12. Demyanov, V.F., Jeyakumar, V.: Hunting for a smaller convex subdifferential. J. Global Optim. **10**, 305–326 (1997)
13. Dempe, S.: A necessary and a sufficient optimality condition for bilevel programming problems. Optimization **25**, 341–354 (1992)
14. Dempe, S.: Foundations of Bilevel Programming. Kluwer Academie Publishers, Dordrecht (2002)

15. Dempe, S.: First-order necessary optimality conditions for general bilevel programming problems. J. Optim. Theory Appl. **95**, 735–739 (1997)
16. Dempe, S., Schmidt, H.: On an algorithm solving two-level programming problems with nonunique lower level solutions. Comput. Optim. Appl. **6**, 227–249 (1996)
17. Dien, P.H.: Locally Lipschitzian set-valued maps and general extremal problems with inclusion constraints. Acta. Math. Vietnam. **1**, 109–122 (1983)
18. Dien, P.H.: On the regularity condition for the extremal problem under locally Lipschitz inclusion constraints. Appl. Math. Appl. **13**, 151–161 (1985)
19. Ekeland, I.: On the variational principle. J. Math. Anal. Appl. **47**, 324–353 (1974)
20. Gadhi, N.: Necessary optimality conditions for Lipschitz multiobjective optimization problems. Georgian Math. J. **12**, 65–74 (2005)
21. Huang, H.X., Pardalos, P.M.: A multivariate partition approach to optimization problems. Cybernet. Syst. Anal. **38**, 265–275 (2002)
22. Ioffe, A.D.: Approximate subdifferential and applications.. III : the Metric Theory. Mathematika **36**, 1–38 (1989)
23. Jahn, J.: Vector optimization. Springer, Berlin (2004)
24. Jahn, J., Rauh, R.: Contingent epiderivatives and set-valued optimization. Math. Method. Oper. Res. **46**, 193–211 (1997)
25. Jeyakumar, V., Luc, D.T.: Approximate Jacobian matrices for continuous maps and C^1-Optimization. SIAM J. Control Optim. **36**, 1815–1832 (1998)
26. Jeyakumar, V., Luc, D.T., Schaible, S.: Characterizations of generalized monotone nonsmooth continuous maps using approximate Jacobians. J. Convex Anal. **5**, 119–132 (1998)
27. Jeyakumar, V., Luc, D.T.: Nonsmooth calculus, minimality, and monotonicity of convexificators. J. Optim. Theory Appl. **101**, 599–621 (1999)
28. Klose, J.: Sensitivity analysis using the tangent derivative. Numer. Funct. Anal. Optimiz. **13**, 143–153 (1992)
29. Kuroiwa, D.: Natural criteria of set-valued optimization, Manuscript. Shimane University, Japan (1998)
30. Li, Z.: A theorem of the alternative and its application to the optimization of set-valued maps. J. Optim. Theory Appl. **100**, 365–375 (1999)
31. Loewen, P.D.: Limits of Frechet normals in nonsmooth analysis. Optimization and Nonlinear Analysis. Pitman Research Notes Math, Ser. **244**, 178–188 (1992)
32. Luc, D.T.: Contingent derivatives of set-valued maps and applications to vector optimization. Math. Program. **50**, 99–111 (1991)
33. Luc, D.T., Malivert, C.: Invex optimization problems. Bull. Austral. Math. Soc. **46**, 47–66 (1992)
34. Marti, K.: Stochastic Optimization Methods. Springer, Berlin (2005)
35. Michel, P., Penot, J-P.: Calcul sous-differentiel pour des fonctions Lipschitziennes et non Lipschitziennes. C.R. Acad. Sc. Paris 298 (1984)
36. Michel, P., Penot, J-P.: A generalized derivative for calm and stable functions. Diff. Integral Eq. **5**(2), 433–454 (1992)
37. Migdalas, A., Pardalos, P.M., Värbrand, P.: Multilevel optimization : algorithms and applications. Kluwer Academic Publishers (1997)
38. Mordukhovich, B.S.: Variational analysis and generalized differentiation. Vol. I, II, Springer Verlag, Berlin et al. (2006)
39. Mordukhovich, B.S., Shao, Y.: On nonconvex subdifferential calculus in Banach spaces. J. Convex Anal. **2**, 211–228 (1995)
40. Outrata, J.V.: On necessary optimality conditions for Stackelberg problems. J. Optim. Theory Appl. **76**, 306–320 (1993)
41. Outrata, J.V.: Optimality problems with variational inequality constraints. SIAM J. Optim. **4**, 340–357 (1993)
42. Penot, J.P.: On the mean-value theorem. Optimization **19**, 147–156 (1988)
43. Ralph, D., Dempe, S.: Directional derivatives of the solution of a parametric nonlinear program. Math. Program. **70**, 159–172 (1995)
44. Thibault, L.: On subdifferentials of optimal value functions. SIAM J. Control Optim. **29**, 1019–1036 (1991)
45. Treiman, J.S.: The linear nonconvex generalized gradient and Lagrange multipliers. SIAM J. Optim. **5**, 670–680 (1995)
46. Wang, S., Wang, Q., Romano-Rodriguez, S.: Optimality conditions and an algorithm for linear-quadratic bilevel programs. Optimization **4**, 521–536 (1993)
47. Ye, J.J., Zhu, D.L.: Optimality conditions for bilevel programming problems. Optimization **33**, 9–27 (1995)

48. Ye, J.J.: Constraint qualifications and KKT conditions for bilevel programming problems. Math. Oper. Res. (2007, to appear)
49. Zhang, R.: Problems of hierarchical optimization in finite dimensions. SIAM J. Optim. **4**, 521–536 (1993)