J Glob Optim (2010) 47:233–245
DOI 10.1007/s10898-009-9468-1

Second order optimality conditions for bilevel set optimization problems

S. Dempe · N. Gadhi

Received: 30 April 2008 / Accepted: 12 August 2009 / Published online: 3 September 2009

Abstract In this work, we use the notion of the support function to the feasible set mapping to establish second order necessary and sufficient optimality conditions for the optimistic case of bilevel optimization problems. The main tools we exploit are approximate Jacobians, approximate Hessians, second order approximations and second order contingent sets.

Keywords Approximate Jacobian · Approximate Hessian · Recession matrices · Bilevel optimization · Second order approximation · Necessary optimality conditions · Regularity condition · Set valued mappings · Support function

Mathematics Subject Classification (2000) 90C29 · 49J52 · 90C30 · 49K99

1 Introduction

Bilevel programming problems have been considered e.g. in the papers [6,7,9,10,17,22,23,25–27]. Necessary and sufficient optimality conditions without assuming that the lower level problem is a convex one and without assuming uniqueness of optimal lower level solutions can be found in the paper Ye and Zhu [26]. Under semi-Lipschitz property, Zhang [27] extends the classical approach to allow nonsmooth problem data; he derives existence and optimality conditions for problems in terms of the graph of the solution multifunction to the lower-level problem using the coderivative of Mordukhovich [21]. Dempe et al. [13] derive necessary optimality conditions for using the reformulation (P^*) below under a partial calmness assumption.

S. Dempe (✉)
Department of Mathematics and Computer Sciences, Technical University Bergakademie Freiberg, Freiberg, Germany
e-mail: dempe@tu-freiberg.de; dempe@math.tu-freiberg.de

N. Gadhi
Dhar El Mehrez, Department of Mathematics, Sidi Mohamed Ben Abdellah University, Fes, Morocco
e-mail: ngadhi@hotmail.com

Second order optimality conditions are commonly used in mathematical programming. They are also often very helpful in convergence analyses for solution algorithms. Quite a few number of publications exist on second order conditions, among which we cite the papers [1,8,14,15] for problems with C^2 and $C^{1.1}$ data, and [16,19] for problems with only C^1 data.

In this paper, we investigate the bilevel optimization problem

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

where, for each $x \in \mathbb{R}^{n_1}$, $S(x)$ is the solution set of the following parametric optimization problem (the lower level problem)

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

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. Since the upper level problem (P) is formulated using the point-to-set mapping S, it belongs to the class of set-valued optimization problems. Another approach to see this is the one given in [11].

A pair $(\overline{x}, \overline{y})$ is said to be a local optimal solution to (P) if it is a local optimal solution to the following problem: $\min_{(x,y)\in\overline{S}} f(x, y)$ where

$$\overline{S} = \left\{(x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \; 0 \in F(x, y) \text{ and } y \in S(x)\right\}.$$

This is the optimistic approach in bilevel programming; for more details, see [11] and the references therein.

Problem (P) is a sequence of two optimization problems in which the feasible region of the upper-level problem is determined implicitly by the solution set mapping of the parametric lower-level problem.

The rest of the paper is written as follows: Sect. 2 contains basic definitions and preliminary results. Sections 3 and 4 are devoted to second order necessary and sufficient optimality conditions.

2 Preliminaries

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

$$C_H\left(y^*, x\right) := \inf_{y\in H(x)} \langle y^*, y\rangle,$$

where $\langle ., .\rangle$ is the inner product.

Let S be an arbitrary nonempty subset of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. The contingent cone to S at $\overline{u}$ is

$$K(S, \overline{u}) := \left\{ \begin{array}{l} d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \exists (t_n) \searrow 0, \; \exists (d_n) \to d \text{ such that} \\ \overline{u} + t_n d_n \in S, \; \forall n \end{array} \right\}.$$

The second order contingent set to S at $\overline{u}$ in the direction $d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ is

$$K^2(S, \overline{u}, d) := \left\{ \begin{array}{l} w \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \exists (t_n) \searrow 0, \; \exists (w_n) \to w \text{ such that} \\ \overline{u} + t_n d + t_n^2 w_n \in S, \; \forall n \end{array} \right\}.$$

Remark 2.1 Let S be an arbitrary nonempty subset of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $\overline{u} \in S$ and d, $w \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. Then,

$$w \in K^2(S, \overline{u}, d) \implies d \in K(S, \overline{u}).$$

Remark 2.2 The set $K^2(S, \overline{u}, d)$ is in general not a cone, and it might be empty when S is not a polyhedral set, see [24].

The following proposition gives geometric properties of second order tangents. For more details, see Propositions 13.12 and 13.13 in Rockafellar and Wets [24].

Proposition 2.3 [24] *For any $\overline{u} \in S$ and any $d \in K(S, \overline{u})$, the set $K^2(S, \overline{u}, d)$ is closed. Moreover, if S is closed and convex, one has*

$$K^2(S, \overline{u}, d) + v \subset K^2(S, \overline{u}, d) \quad \text{for all } v \in K(S, \overline{u}).$$

If S is a polyhedral set, then

$$K^2(S, \overline{u}, d) = K(K(S, \overline{u}), d) = K(S, \overline{u}) + Rd.$$

In the sequel we need the following definitions. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $\overline{u} \in \mathbb{R}^n$. Then, $L(\mathbb{R}^n, \mathbb{R})$ denotes the set of all continuous linear operators mapping $\mathbb{R}^n$ to $\mathbb{R}$, and $B(\mathbb{R}^n, \mathbb{R}^n, \mathbb{R})$ is the set of all continuous bi-linear operators mapping $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. $\mathbb{B}_{\mathbb{R}^n}$ denotes the closed unit ball of $\mathbb{R}^n$ centered at the origin, and $\mathbb{R}^n_+$ is the nonnegative orthant of $\mathbb{R}^n$.

Definition 2.4 [2–4] The set $A_f(\overline{u}) \subset L(\mathbb{R}^n, \mathbb{R})$ is a first order approximation of the function f at $\overline{u}$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$f(u) - f(\overline{u}) \in A_f(\overline{u})(u - \overline{u}) + \varepsilon \|u - \overline{u}\| \mathbb{B}_{\mathbb{R}}$$

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

Definition 2.5 [2,3] The couple $\left(A_f(\overline{u}), B_f(\overline{u})\right) \subset L(\mathbb{R}^n, \mathbb{R}) \times B(\mathbb{R}^n, \mathbb{R}^n, \mathbb{R})$ is a second order approximation of f at $\overline{u}$ if

(1) $A_f(\overline{u})$ is a first order approximation of f at $\overline{u}$, and
(2) for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$f(u) - f(\overline{u}) \in A_f(\overline{u})(u - \overline{u}) + B_f(\overline{u})(u - \overline{u})(u - \overline{u}) + \varepsilon \|u - \overline{u}\|^2 \mathbb{B}_{\mathbb{R}}$$

for all $u \in \overline{u} + \delta \mathbb{B}_{\mathbb{R}^n}$, i.e. for all $\varepsilon > 0$ and for all $u \in \overline{u} + \delta \mathbb{B}_{\mathbb{R}^n}$ there exist $\delta > 0$, $A \in A_f(\overline{u})$ and $B \in B_f(\overline{u})$ such that

$$f(u) - f(\overline{u}) = \langle A, u - \overline{u} \rangle + \langle B(u - \overline{u}), u - \overline{u} \rangle + \varepsilon \|u - \overline{u}\|^2.$$

Definition 2.6 [18]

(1) Let $f : \mathbb{R}^n \to \mathbb{R}^p$ be a function. A closed subset $\partial f(\overline{u})$ of $L(\mathbb{R}^n, \mathbb{R}^p)$ is called an approximate Jacobian (generalized subdifferential) of f at $\overline{u}$ if for every $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^p_+$ one has

$$\liminf_{t \searrow 0} \frac{\langle v, f(\overline{u} + tu) \rangle - \langle v, f(\overline{u}) \rangle}{t} \leq \sup_{M \in \partial f(\overline{u})} \langle v, M(u) \rangle.$$

(2) For a local C^1 function $f : \mathbb{R}^n \to \mathbb{R}$ around $\overline{u}$, a closed set $\partial^2 f(\overline{u}) \subset B(\mathbb{R}^n, \mathbb{R}^n, \mathbb{R})$ is said to be an approximate Hessian of f at $\overline{u}$ if it is an approximate Jacobian of ∇f at $\overline{u}$.

Remark 2.7 For continuous functions, approximate Jacobians are examples of first order approximations.

Proposition 2.8 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a local C^1 function around $\overline{u}$ and let, for every u close to $\overline{u}$, $\partial^2 f(u)$ denote an approximate Hessian. Suppose that $\partial^2 f(.)$ is upper semi-continuous in an open neighborhood of $\overline{u}$. Then, f admits $\left(\nabla f(\overline{u}), \frac{1}{2}\, cl\ co\ \partial^2 f(\overline{u})\right)$ as a second order approximation at $\overline{u}$.*

Here, $cl\ coA$ denotes the closure of the convex hull of a set A.

Proof Let $\pi > 0$ be such that f is C^1 on $\overline{u} + \pi\mathbb{B}_{\mathbb{R}^n}$ and let $\varepsilon > 0$. Since $\partial^2 f(.)$ is upper semi-continuous, there exists $\pi > \delta > 0$ satisfying

$$\partial^2 f(u) \subset \partial^2 f(\overline{u}) + \varepsilon\mathbb{B}_{\mathbb{R}^n \times \mathbb{R}^m} \tag{2.1}$$

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

Let u be an arbitrary element of $\overline{u} + \delta\mathbb{B}_{\mathbb{R}^n}$. From [18], taking $v := u - \overline{u}$ and using the generalized Taylor's expansions for C^1 functions, there exist $c \in]\overline{u}, u[$ and $M \in cl\ co\ \partial^2 f(c)$ such that

$$f(u) - f(\overline{u}) = \langle \nabla f(\overline{u}), v\rangle + \frac{1}{2}\langle Mv, v\rangle .$$

Using (2.1), we derive

$$f(u) - f(\overline{u}) \in \langle \nabla f(\overline{u}), v\rangle + \frac{1}{2} cl\ co\ \partial^2 f(\overline{u})(v, v) + \frac{1}{2}\varepsilon \|v\|^2 \mathbb{B}_{\mathbb{R}}.$$

□

The recession cone A_∞ of a non empty set $A \subset \mathbb{R}^n$ consists of all limit points $\lim_{i\to\infty} t_i a_i$, where $\{a_i\} \subset A$ and $\{t_i\}$ is a sequence of positive numbers converging to 0. It is important to note that a set is bounded if and only if its recession cone is trivial.

3 Second order necessary optimality conditions

As in [26], we give optimality conditions without any convexity assumption on the lower level problem and without the assumption that the solution set $S(x)$ is a singleton. In this case, according to [12], problem (P) can be replaced by

$$(P^*) : \quad \begin{cases} \text{Minimize } f(x, y) \\ \text{subject to : } 0 \in F(x, y),\ 0 \in G(x, y), \\ g(x, y) - V(x) \le 0, \end{cases}$$

provided that (P^*) has an optimal solution [20], where

$$V(x) := \min_y \left\{ g(x, y) : 0 \in G(x, y),\ y \in \mathbb{R}^{n_2} \right\}$$

denotes the optimal value function of the lower level problem.

Remark 3.1 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 feasible set of problem (P^*) is nonempty and bounded.

Let C be the feasible set of problem (P^*) and let

$$\overline{V}(x, y) := V(x) \text{ and}$$
$$H(x, y) := (F(x, y), G(x, y), g(x, y) - V(x) + \mathbb{R}_+).$$

For $t^* = (\mu^*, \upsilon^*, \gamma^*) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}_+$, we consider the function

$$\begin{aligned} L\left(t^*, (x, y)\right) &:= f(x, y) + C_H\left(t^*, (x, y)\right) \\ &= f(x, y) + C_F\left(\mu^*, (x, y)\right) + C_G\left(\upsilon^*, (x, y)\right) + \gamma^*\left(g(x, y) - V(x)\right) \end{aligned}$$

and the set

$$C_{t^*} := \left\{u := (x, y) \in C \text{ such that } 0 = C_H\left(t^*, u\right)\right\}.$$

Let $\overline{u} := (\overline{x}, \overline{y}) \in C$. In the following theorem, we will need

$$J(\overline{u}) = \left\{t^* \in \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}_+ \text{ such that } \left\|t^*\right\| \le 1 \text{ and } 0 = C_H\left(t^*, \overline{u}\right)\right\}$$

and

$$T^{Lin}(\overline{u}) := \left\{d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \forall t^* \in J(\overline{u})\ \exists p^* \in A_{C_H(t^*,.)}(\overline{u}) \text{ with } \left\langle p^*, d\right\rangle \le 0\right\}$$

Remark 3.2 In general, we have $K(C, \overline{u}) \subseteq T^{Lin}(\overline{u})$. Using a suitable regularity assumption, one can get equality of both sets (see [14] and [5]).

Theorem 3.3 *Let $t^* \in J(\overline{u})$ and let $\overline{u} := (\overline{x}, \overline{y})$ be a local optimal solution of problem (P). Assume that f, $C_F(\mu^*, x)$, $C_G(\upsilon^*, x)$, g and $\overline{V}$ admit compact first order approximations $A_f(\overline{u})$, $A_{C_F(\mu^*,.)}(\overline{u})$, $A_{C_G(\upsilon^*,.)}(\overline{u})$, $A_g(\overline{u})$ and $A_{\overline{V}}(\overline{u})$ at $\overline{u}$. Then, for all $d \in K(C_{t^*}, \overline{u})$, there exist $A_1 \in A_f(\overline{u})$, $A_2 \in A_{C_F(\mu^*,.)}(\overline{u})$, $A_3 \in A_{C_G(\upsilon^*,.)}(\overline{u})$, $A_4 \in A_g(\overline{u})$ and $A_5 \in A_V(\overline{x}) \times \{0\}$ such that*

$$\left\langle A_1 + A_2 + A_3 + \gamma^*(A_4 - A_5), d\right\rangle \ge 0.$$

Proof Let $\overline{u} := (\overline{x}, \overline{y})$ be a local optimal solution to (P). Then, by definition of the mapping H, $(\overline{x}, \overline{y})$ is also a local optimal solution of the problem

$$\left(P^*\right) : \begin{cases} \text{Minimize } f(x, y) \\ \text{subject to : } 0 \in H(x, y). \end{cases}$$

Let $\varepsilon > 0$ and $d := \left(d^1, d^2\right) \in K(C_{t^*}, \overline{u})$ be arbitrarily chosen. Then, there exist sequences $(t_n) \searrow 0$, $(d_n) := \left(d^1_n, d^2_n\right) \to d$ such that $\overline{u} + t_n d_n \in C_{t^*}$. Hence, for n large enough,

$$0 \in H(\overline{u} + t_n d_n),\ C_H\left(t^*, \overline{u} + t_n d_n\right) = 0 \quad \text{and} \quad f(\overline{u} + t_n d_n) - f(\overline{u}) \ge 0.$$

Thus,

$$L\left(t^*, \overline{u} + t_n d_n\right) - L\left(t^*, \overline{u}\right) \ge 0.$$

Consequently, by the definition of first order approximations, there exist $A_n \in A_{L(t^*,.)}(\overline{u})$ and $b_n \in [-1; 1]$ such that

$$L\left(t^*, \overline{u} + t_n d_n\right) - L\left(t^*, \overline{u}\right) = A_n(t_n d_n) + \varepsilon\left\|t_n d_n\right\| b_n \ge 0.$$

Due to compactness of the first order approximations, extracting a subsequence if necessary, and passing to the limit for n to infinity, we derive the existence of $A \in A_{L(t^*,.)}(\overline{u})$ and $b \in [-1; 1]$ such that

$$\langle A, d\rangle + \varepsilon \|d\| b \geq 0.$$

Then, for $\varepsilon \to 0$, we derive the desired inequality

$$\langle A, d\rangle \geq 0. \qquad \square$$

Remark 3.4 Assume that g and the support functions of F, G are continuous (i.e. u.s.c. and l.s.c.) functions and that they admit compact first order approximations. Then, due to [6], the set $J(\overline{u})$ is nonempty and there exists $(\mu^*, \upsilon^*) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+$ such that

$$\left\langle \mu^*, F(\overline{u})\right\rangle = 0 \quad and \quad \left\langle \upsilon^*, G(\overline{u})\right\rangle = 0.$$

For more details see [6, Theorem 7].

Remark 3.5 Let

$$S(\overline{x}) = \left\{y \in \mathbb{R}^{n_2} : 0 \in G(\overline{x}, y) \text{ and } g(\overline{x}, y) = V(\overline{x})\right\}$$

denote the set of optimal solutions of the lower level problem. Then,

$$co\left\{A_g(., y)(\overline{x}) : y \in S(\overline{x})\right\}$$

can be taken as a first order approximation of V at $\overline{x}$, see [3,6].

Theorem 3.6 *Let $t^* \in J(\overline{u})$ and $\overline{u} := (\overline{x}, \overline{y})$ be a local optimal solution of (P). Assume that f and $C_H(t^*, .)$ admit second order approximations $\left(A_f(\overline{u}), B_f(\overline{u})\right)$ and $\left(A_{C_H(t^*,.)}(\overline{u}), B_{C_H(t^*,.)}(\overline{u})\right)$ at $\overline{u}$.*

(1) *Suppose that $A_f(\overline{u})$, $A_{C_H(t^*,.)}(\overline{u})$, $B_f(\overline{u})$ and $B_{C_H(t^*,.)}(\overline{u})$ are bounded sets. Then, for all $d \in N_{L(t^*,.)}(\overline{u})$, $w \in K^2(C_{t^*}, \overline{u}, d)$, there exist $A \in A_{L(t^*,.)}(\overline{u})$ and $B \in B_{L(t^*,.)}(\overline{u})$ such that*

$$\langle A, w\rangle + \langle Bd, d\rangle \geq 0.$$

(2) *Suppose that $A_f(\overline{u})$ and $A_{C_H(t^*,.)}(\overline{u})$ are bounded and that $B_f(\overline{u})$ and $B_{C_H(t^*,.)}(\overline{u})$ are not bounded. Then, for all $d \in N_{L(t^*,.)}(\overline{u})$ there exist $B \in \left[B_{L(t^*,.)}(\overline{u})\right]_\infty \setminus \{0\}$ such that*

$$\langle Bd, d\rangle \geq 0.$$

Here,

$$N_{L(t^*,.)}(\overline{u}) := \left\{d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \max_{A \in A_{L(t^*,.)}(\overline{u})} \langle A, d\rangle = 0\right\}.$$

Proof Let $\overline{u} := (\overline{x}, \overline{y})$ be a local optimal solution of problem (P). Let $d \in N_{L(t^*,.)}(\overline{u})$ and $w \in K^2(C_{t^*}, \overline{u}, d)$. Then, there exists $(t_n, w_n) \to \left(0^+, w\right)$ such that

$$\overline{u} + t_n d + t_n^2 w_n \in C_{t^*}.$$

Thus, for n large enough,

$$0 \in H\left(\overline{u} + t_n d + t_n^2 w_n\right), \; C_H\left(t^*, \overline{u} + t_n d + t_n^2 w_n\right) = 0$$

and

$$f\left(\overline{u}+t_nd+t_n^2w_n\right)-f\left(\overline{u}\right)\geq 0.$$

Consequently,

$$L\left(t^*,\overline{u}+t_nd+t_n^2w_n\right)-L\left(t^*,\overline{u}\right)\geq 0.$$

Let $\varepsilon>0$ be arbitrarily chosen. Then, for n large enough, there exist $A_n\in A_{L(t^*,.)}\left(\overline{u}\right)$, $B_n\in B_{L(t^*,.)}\left(\overline{u}\right)$ and $b_n\in[-1;1]$ such that

$$\left\langle A_n,t_nd+t_n^2w_n\right\rangle+\left\langle B_n\left(t_nd+t_n^2w_n\right),t_nd+t_n^2w_n\right\rangle+\varepsilon\left\|t_nd+t_n^2w_n\right\|^2b_n\geq 0.$$

That is ,

$$t_n\langle A_n,d\rangle+t_n^2\langle A_n,w_n\rangle+t_n^2\langle B_n\left(d+t_nw_n\right),d+t_nw_n\rangle+t_n^2\varepsilon\left\|d+t_nw_n\right\|^2b_n\geq 0.$$

Since $d\in N_{L(t^*,.)}\left(\overline{u}\right)$, one has

$$\langle A_n,d\rangle\leq 0.$$

Then,

$$\langle A_n,w_n\rangle+\langle B_n\left(d+t_nw_n\right),d+t_nw_n\rangle+\varepsilon\left\|d+t_nw_n\right\|^2b_n\geq 0.$$

Consequently,

$$\begin{aligned}&\langle A_n,w_n\rangle+\langle B_nd,d\rangle+t_n\langle B_nd,w_n\rangle+t_n\langle B_nw_n,d\rangle+t_n^2\langle B_nw_n,w_n\rangle\\&\quad+\quad\varepsilon\left\|d+t_nw_n\right\|^2b_n\geq 0.\end{aligned}$$

- If the sequences $\{A_n\}$ and $\{B_n\}$ are bounded, then we may assume that they converge to some $A\in A_{L(t^*,.)}\left(\overline{u}\right)$ and $B\in B_{L(t^*,.)}\left(\overline{u}\right)$. Moreover, by compactness of $[-1;1]$, extracting a subsequence if necessary, one may assume that there exist $b\in[-1;1]$ such that

$$\langle A,w\rangle+\langle Bd,d\rangle+\varepsilon\left\|d\right\|^2b\geq 0.$$

  Letting $\varepsilon\to 0$, we get

$$\langle A,w\rangle+\langle Bd,d\rangle\geq 0.$$

- If the sequence $\{B_n\}$ is not bounded and the sequence $\{A_n\}$ is bounded, we may assume that

$$\lim_{n\to\infty}\left\|B_n\right\|=\infty$$

  and

$$\lim_{n\to\infty}\frac{B_n}{\left\|B_n\right\|}=B_0\in\left[B_{L(t^*,.)}\left(\overline{u}\right)\right]_\infty\setminus\{0\}.$$

  Consequently,

$$\langle B_0d,d\rangle\geq 0.$$

□

4 Second order sufficient condition

Let $\overline{u} \in C$. Assume that for all $t^* \in \mathbb{R}_+^p \times \mathbb{R}_+^q \times \mathbb{R}_+$, the functions f and $C_H(t^*,.)$ admit second order approximations $\left(A_f(\overline{u}), B_f(\overline{u})\right)$ and $\left(A_{C_H(t^*,.)}(\overline{u}), B_{C_H(t^*,.)}(\overline{u})\right)$ at $\overline{u}$ such that $A_f(\overline{u})$ and $A_{C_H(t^*,.)}(\overline{u})$ are compact sets.

Theorem 4.1 *Suppose that there exists $t^* \in J(\overline{u})$ such that for all directions $d \in K(C,\overline{u})\backslash\{0\}$ and $A \in A_{L(t^*,.)}(\overline{u})$ one has*

$$\langle A, d\rangle > 0.$$

Then, $\overline{u}$ is a strict local optimal solution of (P).

Proof Assume that $\overline{u}$ is not a strict local optimal solution of (P). Then, there exists a sequence (u_n) such that

$$\begin{cases} u_n := \overline{u} + t_n d_n \in C,\ u_n \to \overline{u},,\ t_n := \|u_n - \overline{u}\|,\ t_n \searrow 0, \\ d_n := \dfrac{u_n - \overline{u}}{\|u_n - \overline{u}\|},\ d_n \to d,\ f(u_n) \le f(\overline{u}) \text{ for all } n. \end{cases} \tag{4.1}$$

Since u_n is feasible, we obtain $C_H(t^*, u_n) \le 0$ and, since $t^* \in J(\overline{u})$, we have

$$0 = C_H\left(t^*, \overline{u}\right).$$

This implies

$$L\left(t^*, u_n\right) - L\left(t^*, \overline{u}\right) \le 0$$

Let $\varepsilon > 0$ be arbitrarily chosen. Then, for n large enough, there exist $A_n \in A_{L(t^*,.)}(\overline{u})$ and $b_n \in [-1; 1]$ such that

$$\langle A_n, t_n d_n\rangle + \varepsilon \|t_n d_n\| b_n \le 0$$

or equivalently

$$t_n \langle A_n, d_n\rangle + t_n\varepsilon \|d_n\| b_n \le 0.$$

By compactness, extracting a subsequence if necessary and passing to the limit, we obtain that there exist $A \in A_{L(t^*,.)}(\overline{u})$ and $b \in [-1; 1]$ such that

$$\langle A, d\rangle + \varepsilon \|d\| b \le 0.$$

For $\varepsilon \to 0$, we derive $\langle A, d\rangle \le 0$, which is a contradiction to the assumptions of the theorem. □

Define $dist(d, Z)$ to be the Euclidean distance of a vector d from a set Z.

Theorem 4.2 *Assume that $\overline{u}$ is feasible and that there exist $t^* \in J(\overline{u})$ and $\delta > 0$ such that for all $A \in A_{L(t^*,.)}(\overline{u})$, for all $d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that*

$$dist\,(d, K(C, \overline{u})) < \delta$$

we have

$$\langle A, d\rangle \ge 0.$$

If for all $d \in T^{Lin}(\overline{u})$ and for all $B \in B_{L(t^*,.)}(\overline{u}) \cup \left(\left[B_{L(t^*,.)}(\overline{u})\right]_\infty \setminus \{0\}\right)$ we have

$$\langle Bd, d\rangle > 0,$$

then $\overline{u}$ is a strict local optimal solution of (P).

Proof Proving the theorem by contradiction, assume that $\overline{u}$ is not a strict local optimal solution of (P). Then, there exists a sequence (u_n) of feasible points satisfying conditions (4.1). Without loss of generality, we can assume that $d_n := \|u_n - \overline{u}\|^{-1}(u_n - \overline{u})$ converges to a vector d_1. Obviously,

$$d_1 \in K(C, \overline{u}) \setminus \{0\} \subset T^{Lin}(\overline{u})$$

by Remark 3.2. Put $u_n := \overline{u} + t_n d_n$ for all n. Since u_n is feasible and $t^* \in J(\overline{u})$, we obtain

$$C_H\left(t^*, u_n\right) \leq 0 \quad \text{and} \quad C_H\left(t^*, \overline{u}\right) = 0.$$

Hence,

$$f(u_n) + C_H\left(t^*, u_n\right) \leq f(\overline{u}) + C_H\left(t^*, \overline{u}\right).$$

Take an arbitrary $\varepsilon > 0$. Then, this inequality implies the existence of $A_n \in A_{L(t^*,.)}(\overline{u})$, $B_n \in B_{L(t^*,.)}(\overline{u})$ and $b_n \in [-1; 1]$ such that

$$\langle A_n, d_n\rangle + t_n \langle d_n, B_n d_n\rangle + \varepsilon t_n \|d_n\|^2 b_n \leq 0.$$

By the assumption of the theorem, due to $d_1 \in K(C, \overline{u}) \setminus \{0\}$, we have $\langle A_n, d_n\rangle \geq 0$. Inserting this into the last inequality, we derive

$$t_n \langle d_n, B_n d_n\rangle + \varepsilon t_n \|d_n\|^2 b_n \leq 0.$$

and, hence,

$$\langle d_n, B_n d_n\rangle + \varepsilon \|d_n\|^2 b_n \leq 0.$$

We proceed investigating the two possible cases that the sequence $\{B_n\}$ is bounded or not.

- If the sequence $\{B_n\}$ is bounded, then we may assume that it converges to some $B \in B_{L(t^*,.)}(\overline{u})$. If, without loss of generality, the sequence $\{b_n\}$ converges to a vector b, this implies

  $$\langle d_1, B d_1\rangle + \varepsilon \|d_1\|^2 b \leq 0.$$

  For $\varepsilon \to 0$, we get

  $$\langle d_1, B d_1\rangle \leq 0,$$

  which is a contradiction.
- If the sequence $\{B_n\}$ is not bounded, we may assume that

  $$\lim_{n\to\infty} \|B_n\| = \infty$$

  and

  $$\lim_{n\to\infty} \frac{B_n}{\|B_n\|} = B_0 \in \left[B_{L(t^*,.)}(\overline{u})\right]_\infty \setminus \{0\}.$$

  Consequently,

  $$\langle d_1, B_0 d_1\rangle \leq 0,$$

  which again is a contradiction. □

Remark 4.3 Similar results can be obtained using the notion of approximate Hessians when $f,\ g,\ C_F(\mu^*, .),\ C_G(\upsilon^*, .)$ are local C^1 functions around $\overline{u}$.

With the following example, we illustrate the usefulness of our optimality conditions (in Theorem 3.3). Let $k > 1$ be a given real number. We consider

$$\begin{cases} f(x,y) := (x_1-2)^2 + (x_2-2)^2 - ky, \\ g(x,y) := -y, \\ F(x,y) = 0, \\ G(x,y) = \left\{ \begin{array}{l} (x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} : \\ y \le 1,\ y^2 \le 3 - x_1^2 - x_2^2 \\ \text{and } (y-1.5)^2 \le 0.75 - (x_1-0.5)^2 - (x_2-0.5)^2 \end{array} \right\}. \end{cases}$$

Under these assumptions, we investigate the optimization problem

$$\left(P^{\otimes}\right): \quad \begin{cases} \text{Minimize } f(x,y) \\ \text{subject to} : 0 \in F(x,y),\ y \in S(x), \end{cases}$$

where, for each $x \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$, $S(x)$ is the solution set of the following parametric optimization problem

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

This is a special case of the general type (P). In this example, the optimal solution of $\left(P^{\otimes}\right)$ is

$$y(x) = \begin{cases} y^1 = 1, & x \in \text{Supp}(y, y^1), \\ y^2 = \sqrt{3 - x_1^2 - x_2^2}, & x \in \text{Supp}(y, y^2), \\ y^3 = 1.5 - \sqrt{0.75 - (x_1-0.5)^2 - (x_2-0.5)^2}, & x \in \text{Supp}(y, y^3), \end{cases}$$

where

$$\begin{aligned} &\text{Supp}(y, y^1) = \{x : x_1^2 + x_2^2 \le 2,\ (x_1-0.5)^2 + (x_2-0.5)^2 \ge 0.5\}, \\ &\text{Supp}(y, y^2) = \{x : 2 \le x_1^2 + x_2^2 \le 3\}, \\ &\text{Supp}(y, y^3) = \{x : (x_1-0.5)^2 + (x_2-0.5)^2 \le 0.5\}. \end{aligned}$$

In addition, the optimal value function is

$$V(x) = -\min\left\{1, \sqrt{3 - x_1^2 - x_2^2}, 1.5 - \sqrt{0.75 - (x_1-0.5)^2 - (x_2-0.5)^2}\right\}.$$

and the feasible set C reduces to

$$C = \left\{ \begin{array}{l} (x,y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R} : \\ y \le 1, \\ y^2 \le 3 - x_1^2 - x_2^2, \\ (y-1.5)^2 \le 0.75 - (x_1-0.5)^2 - (x_2-0.5)^2, \\ \min\left\{1,\ \sqrt[2]{3 - x_1^2 - x_2^2},\ 1.5 - \sqrt[2]{0.75 - (x_1-0.5)^2 - (x_2-0.5)^2}\right\} \le y \end{array} \right\}.$$

Remarking that $(x_1, x_2, y) = (1, 1, 1)$ is a local optimal solution of $\left(P^{\otimes}\right)$, we apply our results at $\overline{u} = (\overline{x}, \overline{y}) = (1, 1, 1)$.

$$K(C, \overline{u}) = \{(d_1, d_2, 0) : d_1 + d_2 \ge 0\} \cup \{(d_1, d_2, d_1 + d_2) : -(d_1 + d_2) \le 0\}.$$

In this case,

$$J(1,1,1) = \mathbb{B}^4.$$

Let $t^* = (\alpha, \beta, \gamma, \upsilon) \in \mathbb{B}^4$, $A \in A_{L(t^*,.)}(\overline{u})$ and $d \in K(C, \overline{u})$.Bear in mind that, $A \in A_{L(t^*,.)}(\overline{u}) \Longleftrightarrow \exists \mu \in [-1, 1]$ such that

$$A = (-2 + 2\beta + \gamma, -2 + 2\beta + \gamma, -k + \alpha + 2\beta - \gamma - \upsilon) + \mu (1, 1, 0).$$

Since $t^* = (\alpha, \beta, \gamma, \upsilon) \in \mathbb{B}^4$, one has

$$|\alpha| + |\beta| + |\gamma| + |\upsilon| \leq 1.$$

Then

$$-2 + 2\beta + \gamma + \mu \leq 0.$$

Since $d_1 + d_2 \leq 0$, k sufficiently large, and $A \in A_{L(t^*,.)}(\overline{u})$, there exist $\mu \in [-1, 1]$ such that

$$\begin{cases} A = (-2 + 2\beta + \gamma, -2 + 2\beta + \gamma, -k + \alpha + 2\beta - \gamma - \upsilon) + \mu (1, 1, 0), \\ \quad (-2 + 2\beta + \gamma + \mu)(d_1 + d_2) + (-k + \alpha + 2\beta - \gamma - \upsilon) d_3 \geq 0. \end{cases}$$

Then,

$$\langle A, d \rangle \geq 0.$$

5 Application

If $G = 0_{\mathbb{R}^q}$ and $g = 0$, then the problem

$$\left(P^{\triangleleft}\right): \quad \begin{cases} \text{Minimize } f(x, y) \\ \text{subject to} : 0 \in F(x, y) \end{cases}$$

is obtained, where $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \longrightarrow \mathbb{R}$ is a continuous function and $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightrightarrows \mathbb{R}^p$ is a given set valued mapping; $n_1 \geq 1$ and $n_2 \geq 1$ are integers. In this case,

$$H(x, y) := F(x, y)$$

and

$$C := \left\{u \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : 0 \in H(u)\right\} = \left\{u = (x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : 0 \in F(x, y)\right\}.$$

Consider the function

$$L\left(\mu^*, .\right) := f(.) + C_F\left(\mu^*, .\right)$$

for $\mu^* \in \mathbb{R}_+^p$, and the set

$$C_{\mu^*} := \left\{u \in C \text{ such that } 0 = C_F\left(\mu^*, u\right)\right\}.$$

For $\overline{u} \in C$, the sets $J(\overline{u})$ and $T^{Lin}(\overline{u})$ become

$$J(\overline{u}) = \left\{\mu^* \in \mathbb{R}_+^p \text{ such that } \left\|\mu^*\right\| \leq 1 \text{ and } 0 = C_F\left(\mu^*, \overline{u}\right)\right\}$$

and

$$T^{Lin}(\overline{u}) = \left\{d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : \forall \mu^* \in J(\overline{u}) \ \exists p^* \in A_{C_F(\mu^*,.)}(\overline{u}) \text{ with } \left\langle p^*, d\right\rangle \leq 0\right\}.$$

The following corollaries give second order necessary and sufficient optimality conditions for $\left(P^{\triangleleft}\right)$.

Corollary 5.1 [4] *Let $\mu^* \in J(\overline{u})$ and let $\overline{u} := (\overline{x}, \overline{y})$ be a local optimal solution of problem $\left(P^{\triangleleft}\right)$. Assume that f and $C_F(\mu^*, .)$ admit second order approximations $\left(A_f(\overline{u}), B_f(\overline{u})\right)$ and $\left(A_{C_F(\mu^*,.)}(\overline{u}), B_{C_F(\mu^*,.)}(\overline{u})\right)$ at $\overline{u}$ such that $A_f(\overline{u})$ and $A_{C_F(\mu^*,.)}(\overline{u})$ are compact sets. Then,*

(1) *for all $d \in K(C_{t^*}, \overline{u})$, there exist $A_1 \in A_f(\overline{u})$ and $A_2 \in A_{C_F(\mu^*,.)}(\overline{u})$ such that*

$$\langle A_1 + A_2, d\rangle \geq 0.$$

(2) *for all $d \in N_{L(\mu^*,.)}(\overline{u})$, $w \in K^2\left(C_{\mu^*}, \overline{u}, d\right)$, there exist $A \in A_{L(t^*,.)}(\overline{u})$ and $B \in B_{L(t^*,.)}(\overline{u}) \cup \left(\left[B_{L(t^*,.)}(\overline{u})\right]_\infty \setminus \{0\}\right)$ such that*

$$\langle A, w\rangle + \langle Bd, d\rangle \geq 0.$$

Corollary 5.2 *Let $\overline{u} \in C$. Assume that for all $\mu^* \in \mathbb{R}^p_+$, the functions $C_F(\mu^*, .)$ and f admit second order approximations $\left(A_{C_F(\mu^*,.)}(\overline{u}), B_{C_F(\mu^*,.)}(\overline{u})\right)$ and $\left(A_f(\overline{u}), B_f(\overline{u})\right)$ at $\overline{u}$ such that $A_f(\overline{u})$ and $A_{C_F(\mu^*,.)}(\overline{u})$ are compact sets.*

The feasible point $\overline{u}$ is strict local optimal solution of $\left(P^{\triangleleft}\right)$ if one of the following conditions is satisfied:

(1) *There exists $\mu^* \in J(\overline{u})$ such that for all $d \in K(C, \overline{u}) \setminus \{0\}$ and $A \in A_{L(\mu^*,.)}(\overline{u})$ one has*

$$\langle A, d\rangle > 0.$$

(2) *There exists $\mu^* \in J(\overline{u})$ and $\delta > 0$ such that for all $A \in A_{L(\mu^*,.)}(\overline{u})$, for all $d \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ such that*

$$dist\,(d, K(C, \overline{u})) < \delta$$

one has

$$\langle A, d\rangle \geq 0,$$

and for all $d \in T^{Lin}(\overline{u})$ and for all $B \in B_{L(t^*,.)}(\overline{u}) \cup \left(\left[B_{L(t^*,.)}(\overline{u})\right]_\infty \setminus \{0\}\right)$, one has

$$\langle Bd, d\rangle > 0.$$

Remark 5.3 If f and $C_F(\mu^*, .)$ are continuously differentiable, the sufficient optimality conditions given by Amahroq and Gadhi [4] are obtained.

6 Conclusion

In this paper, optimistic bilevel programming problems with generalized equations as constraints have been investigated. Using the notions of first and second order approximations [2,3] necessary and sufficient optimality conditions of first and second order have been described. For this, the bilevel programming problem is reformulated into a one-level problem using the optimal value function of the lower level. An approximation $K(C_{t^*,\overline{u}})$ of the contingent cone to the feasible set of the bilevel problem is described using the support function to an auxiliary set-valued mapping. Then, the first order necessary optimality condition reduces to nonexistence of a direction of descent for some Lagrangian function to problem (P^*) within $K(C_{t^*,\overline{u}})$. Under some boundedness assumptions on the second order approximations, the second order necessary optimality condition means that there is no descent direction in the second order contingent set to the Lagrange function that is also a critical direction to the Lagrangian. The sufficient optimality conditions can be interpreted similarly.

Acknowledgment This work has been supported by the Alexander-von-Humboldt foundation.

References

1. Aghezzaf, B., Hachimi, M.: Second order optimality conditions in multiobjective optimization problems. J. Optim. Theory Appl. **102**, 37–50 (1999)
2. Allali, K., Amahroq, T.: Second order approximations and primal and dual necessary optimality conditions. Optimization **40**, 229–246 (1997)
3. Amahroq, T., Gadhi, N.: On the regularity condition for vector programming problems. J. Glob. Optim. **21**, 435–443 (2001)
4. Amahroq, T., Gadhi, N.: Second order optimality conditions for an extremal problem under inclusion constraints. J. Math. Anal. Appl. **285**, 74–85 (2003)
5. Amahroq, T., Gadhi, N., Riahi, H.: Epi-differentiability and optimality conditions for an extremal problem under inclusion constraints. J. Inequal. Pure Appl. Math. **4**(2) Article. 41, (2003).
6. Babahadda, H., Gadhi, N.: Necessary optimality conditions for bilevel optimization problems using convexificators. J. Glob. Optim. **34**, 535–549 (2003)
7. Bard, J.F.: Some properties of the bilevel programming problem. J. Optim. Theory Appl. **68**, 371–378 (1991)
8. Bolintineau, S., El maghri, M.: Second order efficiency conditions and sensitivity of efficient points. J. Optim. Theory Appl. **98**, 569–592 (1998)
9. 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, dg-jogo (1992)
10. Dempe, S.: A necessary and a sufficient optimality condition for bilevel programming problems. Optimization **25**, 341–354 (1992)
11. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**, 333–359 (2003)
12. Dempe, S.: First-order necessary optimality conditions for general bilevel programming problems. J. Optim. Theory Appl. **95**, 735–739 (1997)
13. Dempe, S., Dutta, J., Mordukhovich, B.: New necessary optimality conditions in optimistic bilevel programming. Optimization **56**, 577–604 (2007)
14. Dien, P.H., Sach, P.H.: Second order optimality conditions for the extremal problem under inclusion constraints. Appl. Math. Optim. **20**, 71–80 (1989)
15. Guerraggio, A., Luc, D.T.: Optimality conditions for $C^{1,1}$ vector optimization problems. J. Optim. Theory Appl. **109**, 615–629 (2001)
16. Guerraggio, A., Luc, D.T., Minh, N.B.: Second order optimality conditions for C^1 multiobjective programming problems. Acta. Math. Vietnam **26**, 257–268 (2001)
17. Huang, H.X., Pardalos, P.M.: A multivariate partition approach to optimization problems. Cybern. Syst. Anal. **38**, 265–275 (2002)
18. Jeyakumar, V., Luc, D.T.: Approximate Jacobian matrices for continuous maps and C^1 optimization. SIAM J. Control Optim. **36**, 1815–1832 (1998)
19. Luc, D.T.: Second order optimality conditions for problems with continuously differentiable data. Optimization **51**, 497–510 (2002)
20. Luderer,B.: Über der Äquivalenz nichtlinearer Optimierungsaufgaben, Technical Report, Technische Universität Karl-Marx-Stadt, Germany (1983)
21. Mordukhovich, B.S.: Variational Analysis and Generalized Differentiation: Basic Theory, vol. 1 . Springer, Berlin (2006)
22. Outrata, J.V.: On necessary optimality conditions for Stackelberg problems. J. Optim. Theory Appl. **76**, 306–320 (1993)
23. Outrata, J.V.: Optimality problems with variational inequality constraints. SIAM J. Optimization **4**, 340–357 (1993)
24. Rockafellar, R.T., Wets, R.J.B: Variational Analysis, vol. 317. Springer, Heidelberg (2004)
25. Wang, S., Wang, Q., Romano-Rodriguez, S.: Optimality conditions and an algorithm for linear-quadratic bilevel programs. Optimization **4**, 521–536 (1993)
26. Ye, J.J., Zhu, D.L.: Optimality conditions for bilevel programming problems. Optimization **33**, 9–27 (1995)
27. Zhang, R.: Problems of hierarchical optimization in finite dimensions. SIAM J. Optim. **4**, 521–536 (1993)