On Global Partial Calmness for Bilevel Programming Problems with Linear Lower Level Problem

Leonid I. Minchenko
Belarusian State University of Informatics and Radioelectronics
6 Brovki str, 220013 Minsk, Belarus, e-mail: inform@bsuir.by
Daniil E. Berezhnov
Belarusian State University of Informatics and Radioelectronics
6 Brovki str, 220013 Minsk, Belarus, e-mail: inform@bsuir.by

Abstract

Bilevel programming problems occupy an important place in optimization theory and its various applications. We consider bilevel programs such that their lower level problem is linear with respect to the lower level variable. It is known that such programs are partially calm and, therefore, can be reduced to a one level optimization problem by means of local penalization. One of the goals of our paper is to show that these programs can be reduced to a one level optimization problem via a global exact penalization. In the paper we also prove sufficient conditions of partial calmness for bilevel programs and study the pseudo-Lipschitzian continuity (Aubin property) of the solution mapping of the lower level problem.

Keywords: bilevel programming, partial calmness, optimal value function, penalization, solution mapping, pseudo-Lipschitzian continuity.

2000 AMS subject classifications: 90C30, 91A65, 90B06

1 Introduction

A bilevel programming problem is a hierarchical optimization problem with two players, where the upper level player (leader) pursues the goal to minimize his objective function while taking into account the reaction of the lower level player (follower).

Let $x \in R^n, y \in R^m$. Consider the following bilevel programming problem (BLPP):

$$G(x,y) \to \min, \tag{1.1}$$

$$x \in X = \{x \in R^n \,|g_j(x) \le 0 \; j \in J\}, \tag{1.2}$$

$$y \in S(x) \overset{\Delta}{=} Arg\min\{f(x,y)\,|y \in F(x)\}, \tag{1.3}$$

In: Yu. G. Evtushenko, M. Yu. Khachay, O. V. Khamisov, Yu. A. Kochetov, V.U. Malkova, M.A. Posypkin (eds.): Proceedings of the OPTIMA-2017 Conference, Petrovac, Montenegro, 02-Oct-2017, published at http://ceur-ws.org

where $F(x) = \{y \in R^m \,|h_i(x,y) \leq 0 \; i \in I\}$, functions $G(x,y)$, $f(x,y)$, $g_j(x)$ and $h_i(x,y)$ are continuously differentiable, $I = \{1,...,p\}$, $J = \{1,...,s\}$.

Inequalities $g_j(x) \leq 0 \; j \in J$ and $h_i(x,y) \leq 0 \; i \in I$ are usually referred to as the upper level and the lower level constraints, variables x and y are called the upper level and the lower level variables, respectively.

Thus, the problem (BLPP) consists in minimizing the upper level (*leader's*) objective function $G(x,y)$ subject to the upper level constraints and to solutions $y(x) \in S(x)$ of the lower level (*follower's*) problem (1.3). The solution $y(x) \in S(x)$ of the lower level problem (1.3) is called the *rational reaction* of the follower on the leader's choice x. The point (x,y) is said to be a *feasible point* in the problem (BLPP) if $x \in X$, $y \in S(x)$. A feasible point (x^0,y^0) is called a *solution* (*local solution*) of the problem (BLPP) if $G(x^0,y^0) \leq G(x,y)$ for all feasible points (x,y) (for all feasible points from some neighborhood of (x^0,y^0)).

Consider the multivalued mappings $F: x \mapsto F(x)$ and $S: x \mapsto S(x)$. Their graphs and domains are denoted by $grF, domF$ and $grS, domS$.

Note that the problem (BLPP) has been mostly investigated in [Dempe, 2002]-[Henrion, 2011], sometimes it is referred to as a classical bilevel programming problem or simply a classical bilevel program.

In [Outrata, 1988] Outrata proposed the following reformulation of the problem (BLPP) into the equivalent one level problem:

$$G(x,y) \to \min_{x,y}, \; x \in X, \; y \in S(x) = \{y \in F(x) \,|f(x,y) \leq \varphi(x)\}, \tag{1.4}$$

where $\varphi(x)$ is the optimal value function of the lower level problem (1.3), that is,

$$\varphi(x) = \min\{f(x,y) \,|y \in F(x)\}.$$

The problem (BLPP) in the form (1.4) was studied in numerous works [Dempe, 2002], [Ye, 1997]-[Dempe, 2013]. The main difficulty in solving (1.4) is a nonsmooth constraint involving the value function $\varphi(x)$. In [Ye, 1995] Ye and Zhu introduced the now well-known concept of partial calmness and a new method which allowed to move the nonsmooth constraint from the feasible set to the objective function in (1.4).

Let (x^0,y^0) be a feasible point of the problem (BLPP). Recall [Ye, 1995] that the problem (BLPP) in the form (1.4) is called *partial calm at* (x^0,y^0) if there exist a number $\mu > 0$ and a neighborhood V of the point $(x^0,y^0,0)$ in $R^n \times R^m \times R$ such that $G(x,y) - G(x^0,y^0) + \mu\,|u| \geq 0$ for all $(x,y,u) \in V$ such that $x \in X$, $y \in S(x)$, $f(x,y) - \varphi(x) + u = 0$.

In [Ye, 1995] Ye and Zhu proved that the problem (BLPP) in the form (1.4) is partial calm at its local solution (x^0,y^0) if and only if there exists a number $\mu > 0$ such that (x^0,y^0) is a local solution of the partially penalized problem

$$G(x,y) + \mu(f(x,y) - \varphi(x)) \to \min, \; x \in X, \; y \in F(x). \tag{1.5}$$

In their seminal paper [Ye, 1995] where Ye and Zhu introduced the partial calmness, they proved that the problem (BLPP) with a lower level problem linear in (x,y) is partially calm. Dempe and Zemkoho showed in [Dempe, 2013] that this proof can be adapted to the case where the lower level problem is linear only in y. Since the partial exact penalization via partial calmness allows to obtain necessary optimality conditions for (BLPP) and calculate stationary points for partially calm bilevel programs, partial calmness has drawn a lot of attention in literature (see e.g. [Ye, 2010, Dempe, 2007, Henrion, 2011]). However, an approach involving a global penalty function can be also interesting.

In Section 2 we consider a bilevel program (BLPP) with the lower level problem linear in y and prove that its solutions (global) *are global solutions* of the penalized problem (1.5). Section 3 is devoted to so-called extended error bound property for multivalued mappings. In Section 4 we prove some sufficient conditions of the pseudo-Lipschitz continuity for the solution mapping S and the partial calmness for the problem (BLPP).

In our paper we consider the bilevel program (BLPP) under the following assumption:

$$X \cap domS = X \cap domF. \tag{H1}$$

In other words, (H1) is a quite natural assumption that there exists a rational reaction $y(x) \in S(x)$ of the follower on every leader's choice $x \in X \cap domF$.

Denote by $d(v,C)$ the Euclidean distance between a point $v \in R^m$ and a set $C \subset R^m$. Let $|v|$ be the Euclidean norm of a vector v, B be an open unit ball centered at 0 in R^m or R^n, $V(x)$, $V(y)$ and $V_\varepsilon(y) = y + \varepsilon B$, $V_\varepsilon(x) = x + \varepsilon B$ $(\varepsilon > 0)$ be neighborhoods of x and y. Also denote $I(x,y) = \{i \in I \,|h_i(x,y) = 0\}$, $h_0(x,y) = f(x,y) - \varphi(x)$.

2 Global Penalization in Bilevel Programs

In this section we consider the problem (BLPP) with an additional assumption

$$f(x,y) = \langle a_0, y\rangle,\ h_i(x,y) = \langle a_i, y\rangle + b_i(x), \tag{H2}$$

where $a_0, a_i \in R^m$, $b_i(x)$ are scalar functions.

It is known that under assumption (H2) the tangent cone $T_{F(x)}(y)$ to the set $F(x)$ at a point $y \in F(x)$ and a normal cone $N_{S(x)}(y)$ to the set $S(x)$ at a point $y \in S(x)$ can be defined as

$$T_{F(x)}(y) = clcon(F(x) - y) = \{\bar{y} \in R^m \,|\langle a_i, \bar{y}\rangle \leq 0\ i \in I(x,y)\}$$

and

$$N_{S(x)}(y) = \{\sum_{i\in\{0\}\cup I(x,y)} \lambda_i a_i \,|\lambda_i \geq 0\ i \in \{0\} \cup I(x,y)\},$$

respectively.

Lemma 2.1. *Let assumptions (H1) and (H2) hold. Then there exists a number $M > 0$ such that*

$$d(v, S(x)) \leq M \max\{0, \langle a_0, v\rangle - \varphi(x)\} \tag{2.1}$$

for all $v \in F(x)$ and all $x \in domS$.

Proof. Set $h(x,y) = \max\{h_i(x,y)\,|i = 0,1,...,p\}$ where $h_0(x,y) = \langle a_0, y\rangle + b_0(x)$, $b_0(x) = -\varphi(x)$. Then $S(x) = \{y \in R^m\,|h(x,y) \leq 0\}$. Consider any point $x \in domS$. If $v \in S(x)$, the inequality (2.1) holds for all $v \in F(x)$ provided that $M > 0$.

In case $v \in F(x)\backslash S(x)$ denote as $y = y(x,v)$ the point from $S(x)$ closest to v. Then $d(v,S(x)) = |v-y|$. Let $l = (v-y)\,|v-y|^{-1}$, $v = y(t) = y + tl$, where $t = |v-y|$. In this case $l \in N(x,y) = \{l \in R^m \,\big|l \in N_{S(x)}(y) \cap T_{F(x)}(y),\ |l| = 1\}$.

Taking into consideration that $\langle a_i, v-y\rangle \leq 0$ for $i \in I(x,y)$, obtain

$$h(x,v) = h(x,y(t)) \geq \max_{i\in\{0\}\cup I(x,y)} h_i(x,y(t)) = \max_{i\in\{0\}\cup I(x,y)} \{\langle a_i, y\rangle + t\langle a_i, l\rangle + b_i(x)\} =$$

$$= t \max_{i\in\{0\}\cup I(x,y)} \langle a_i, l\rangle = t\langle a_0, l\rangle \geq t \min_{l\in N(x,y)} \langle a_0, l\rangle = t\delta(x,y) \tag{2.2}$$

for all $l \in N(x,y)$.

Suppose that $\delta(x,y) = \langle a_0, l\rangle \leq 0$. Since $h_i(x,y) < 0$ for $i \notin I(x,y)$ and $h_i(x,y)$ are continuous in y, there exists a number $\varepsilon_0 > 0$ such that $h_i(x, y+\varepsilon l) = \langle a_i, y\rangle + b_i(x) + \varepsilon\langle a_i, l\rangle < 0$ for all $i \notin I(x,y)$ and all positive numbers $\varepsilon \leq \varepsilon_0$.

On the other hand, $y + \varepsilon l \notin S(x)$ for all $\varepsilon_0 \geq \varepsilon > 0$ and, therefore, the inequality

$$0 < h(x, y+\varepsilon l) = \max_{i\in\{0\}\cup I(x,y)} h_i(x, y+\varepsilon l) = \max_{i\in\{0\}\cup I(x,y)} \{\langle a_i, y\rangle + \varepsilon\langle a_i, l^0\rangle + b_i(x)\} =$$

$$= \varepsilon \max_{i\in\{0\}\cup I(x,y)} \langle a_i, l\rangle = \varepsilon\langle a_0, l\rangle = \varepsilon\delta(x,y)$$

holds for all positive $\varepsilon \leq \varepsilon_0$. This contradicts the assumption that $\delta(x,y) \leq 0$. Therefore, $\delta(x,y) = \langle a_0(x), (v-y)\,|v-y|^{-1}\rangle > 0$ for all $v \in F(x)\backslash S(x)$.

The vectors a_i don't depend on x, consequently, the set $N(x,y)$ and the value of $\delta(x,y)$ are fully determined by a choice of subsets $I(x,y)$ in the set $I = \{1,...,p\}$. Since there exists only a finite number of subsets $I(x,y)$ in the set I, positivity of $\delta(x,y)$ implies that $\delta(x,y) \geq \delta > 0$ as y ranges over all boundary points in $S(x)$ and x ranges over all points in $domS$.

Since any point $v \in F(x)\backslash S(x)$ can be represented as $v = y + tl$ where $l \in N_{S(x)}(y) \cap T_{F(x)}(y)$, $|l| = 1$, $t > 0$, y is a boundary point in $S(x)$, then from (2.2) follows $h(x,v) \geq \delta t = \delta d(v, S(x))$ and, therefore,

$$d(v, S(x)) \leq \frac{1}{\delta} \max\{0, \langle a_i, v\rangle + b_i(x)\,|i = 0,1,...,p\} = \frac{1}{\delta} \max\{0, \langle a_0, v\rangle - \varphi(x)\}.$$

Introduce the sets

$$D = \{(x,y)\,|h_i(x,y) \le 0 \; i \in I, \; g_j(x) \le 0 \; j \in J\}, \; C = \{(x,y) \in D\,|h_0(x,y) \le 0\}$$

and their corresponding multivalued mappings

$$D(\cdot) : x \mapsto D(x) = \{y \in R^m \,|(x,y) \in D\}, C(\cdot) : x \mapsto C(x) = \{y \in R^m \,|(x,y) \in C\}. \tag{2.3}$$

Theorem 2.1. *Suppose that assumptions (H1) and (H2) hold. Let a feasible point (x^0, y^0) be a solution (global) of the problem (BLPP) and the function G be Lipschitz continuous on the set D with Lipschitz constant $l_0 > 0$. Then there exists a number $\mu_0 > 0$ such that for any $\mu > \mu_0$ the point (x^0, y^0) is a global solution of the problem $G(x,y) + \mu(f(x,y) - \varphi(x)) \to \min, \; (x,y) \in D$.*

Proof. First of all, note that $domD(\cdot) = domC(\cdot)$ due to (H1). Moreover, from (H1) follows $d(y, C(x)) = d((x,y), \{x\} \times C(x)) \ge d((x,y), C)$ for all $(x,y) \in D$.

In virtue of Proposition 2.4.3 [Clarke, 1983] for all $\alpha > l_0$ the point (x^0, y^0) is a solution of the problem $G(x,y) + \alpha d((x,y), C) \to \min, \; (x,y) \in D$. Then $d((x,y), C) \le d(y, C(x))$ for all $(x,y) \in D$ and, therefore, $G(x,y) + \alpha d(y, C(x)) \ge G(x^0, y^0) + \alpha d(y^0, C(x^0)) = G(x^0, y^0)$.

Then, applying Lemma 2.1 to the set $C(x)$, obtain $G(x,y) + \alpha M \max\{0, h_0(x,y)\} \ge G(x^0, y^0)$ for all $(x,y) \in D$. The last inequality is equivalent to the assertion of the theorem with $\mu_0 = l_0 M$.

3 Extended Error Bound Property and Its Application to Bilevel Programming

Let $\Phi(x) = \{y \in R^m \,|q_i(x,y) \le 0 \; i \in K\}$ where $K = \{1, ..., r\}$. In this section we consider a multivalued mapping $\Phi : x \mapsto \Phi(x)$ assuming that the functions $q_i : R^n \times R^m \to R$ are continuous together with their partial gradients $\nabla_y q_i(x,y)$. Following [Fedorov, 1979, Luderer et al., 2002] introduce the so-called *extended error bound property* (named R-regularity in [Luderer et al., 2002, Minchenko, 2015, Minchenko, 2011]).

We say the multivalued mapping Φ has the *extended error bound property (EEBP)* at a point $(x^0, y^0) \in gr\Phi$ (relative to $X \subset R^n$) if there exist a number $M > 0$ and neighborhoods $V(x^0)$ and $V(y^0)$ such that

$$d(y, \Phi(x)) \le M \max\{0, q_i(x,y) \; i \in K\} \tag{3.1}$$

for all $y \in V(y^0)$ and all $x \in V(x^0)$ $(x \in V(x^0) \cap X)$.

We say that the multivalued mapping Φ has the *global extended error bound property (GEEBP)* if there exists a number $M > 0$ such that (3.1) is valid for all $y \in R^m$ and all $x \in domF$.

It is known [Luderer et al., 2002, Borwein, 1986] that in the case of continuously differentiable constraints $q_i : R^n \times R^m \to R$ EEBP holds at $(x^0, y^0) \in gr\Phi$ if y^0 satisfies the Mangasarian-Fromovitz constraint qualification (MFCQ) [Mangasarian, 1967] for the set $\Phi(x^0)$ and EEBP also holds at $(x^0, y^0) \in gr\Phi$ relative to $dom\Phi$ if (x^0, y^0) satisfies the constant rank constraint qualification (CRCQ) [Luderer et al., 2002, Rockafellar, 1998].

Recall some definitions (see, e.g. [Rockafellar, 1998, Klatte, 2015]).

We say that a mapping Φ *is pseudo-Lipschitzian* or it *has the Aubin property* relative to $X \subset R^n$ at a point $(x^0, y^0) \in gr\Phi$ (where $x^0 \in X$) if there exist a number $l_\Phi > 0$ and neighborhoods $V_\delta(x^0)$ and $V_\varepsilon(y^0)$ such that

$$\Phi(x^1) \cap V_\varepsilon(y^0) \subset \Phi(x^2) + l_\Phi \left|x^2 - x^1\right| B \tag{3.2}$$

for all $x^1, x^2 \in V_\delta(x^0)$ (all $x^1, x^2 \in V_\delta(x^0) \cap X$).

Note that in (3.2) $\Phi(x) \cap V_\varepsilon(y^0) \ne \emptyset$ if $\delta < \varepsilon / l_\Phi$.

A mapping Φ is *uniformly bounded* at $x^0 \in dom\Phi$ if there exist a neighborhood $V(x^0)$ and a bounded set Y_0 such that $\Phi(x) \subset Y_0$ for all $x \in V(x^0)$.

A mapping Φ is called *lower semicontinuous (lsc)* at a point $(x^0, y^0) \in gr\Phi$ (relative to $X \subset R^n$) if for any neighborhood $V(y^0)$ there is a neighborhood $V(x^0)$ such that $\Phi(x) \cap V(y^0) \ne \emptyset$ for all $x \in V(x^0)$ (for all $x \in V(x^0) \cap X$).

A mapping Φ is called *Lipschitz lower semicontinuous* at a point $(x^0, y^0) \in gr\Phi$ if there exist positive numbers l and δ such that $d(y^0, \Phi(x)) \le l \left|x - x^0\right| \; \forall x \in V_\delta(x^0)$.

Let $v \in R^m$, $v \notin \Phi(x)$. Denote by $\Pi_{\Phi(x)}(v)$ the set of points from $\Phi(x)$ closest to v. Let $y \in \Pi_{\Phi(x)}(v)$. Then $(v - y)\left|v - y\right|^{-1}$ is a proximal normal [Rockafellar, 1998] to $\Phi(x)$ at y.

Set

$$\Lambda_v(x,y)=\{\lambda\in R^r\left|\frac{y-v}{|y-v|}+\sum_{i=1}^{r}\lambda_i\nabla_y q_i(x,y)=0,\ \lambda_i\ge 0\ and\ \lambda_i q_i(x,y)=0\ i\in K\right.\}.$$

The next assertion generalizes results [Minchenko, 2011, Guo, 2013].

Lemma 3.1. *Assume that Φ is lsc at $(x^0,y^0)\in gr\Phi$ relative to $dom\Phi$. Then the following assertions are equivalent:*

(a) Φ has EEBP at (x^0,y^0) relative to $dom\Phi$;

(b) there exists a number $M>0$ such that for any sequences

$$x^k\to x^0\ (x^k\in dom\Phi),\ v^k\to y^0\ (v^k\notin\Phi(x^k)),\ y^k\in\Pi_{\Phi(x^k)}(v^k)$$

the condition

$$\Lambda^M_{v^k}(x^k,y^k)=\{\lambda\in\Lambda_{v^k}(x^k,y^k)\left|\sum_{i=1}^{r}|\lambda_i|\le M\right.\}\ne\emptyset \tag{3.3}$$

holds starting with some $k=k_0$;

(c) for every pair of sequences $x^k\to x^0$ $(x^k\in dom\Phi)$, $v^k\to y^0$ $(v^k\notin\Phi(x^k))$ there exists a number $M>0$ such that, for all k starting with some $k=k_0$, the condition (3.3) holds for some sequence $y^k\in\Pi_{\Phi(x^k)}(v^k)$.

Proof. The implication $(a)\Rightarrow(b)$ follows immediately from the proof of Theorem 2 [Minchenko, 2011]. The implication $(b)\Rightarrow(c)$ is obvious. Prove that $(c)\Rightarrow(a)$. Suppose to the contrary that EEBP does not hold at (x^0,y^0) relative to $dom\Phi$. This means that there exist sequences $x^k\to x^0$, $x^k\in dom\Phi$ and $v^k\to y^0$, $v^k\notin\Phi(x^k)$ such that for all $k=1,2...$

$$d(v^k,\Phi(x^k))>k\max\{0,h_i(x^k,v^k)\ i\in K\}. \tag{3.4}$$

Let $y^k\in\Pi_{\Phi(x^k)}(v^k)$. Then there exists a vector $\lambda^k\in R^r$ such that $\sum\limits_{i=1}^{r}\left|\lambda_i^k\right|\le M$ and

$$\frac{y^k-v^k}{|y^k-v^k|}+\sum_{i=1}^{r}\lambda_i^k\nabla_y q_i(x^k,y^k)=0,\ \lambda_i^k\ge 0\ i\in K(x^k,y^k),\ \lambda_i^k=0\ i\in K\backslash K(x^k,y^k) \tag{3.5}$$

where $K(x,y)=\{i\in K\,|q_i(x,y)=0\}$.

Since Φ is lsc at (x^0,y^0), there exists a sequence $q^k\in\Phi(x^k)$ such that $q^k\to y^0$ and $\left|v^k-y^k\right|\le\left|v^k-q^k\right|$. This means that $y^k\to y^0$.

Then from (3.5) follows

$$\left|y^k-v^k\right|=\langle\sum_{i=1}^{r}\lambda_i^k\nabla_y q_i(x^k,y^k),v^k-y^k\rangle\le\sum_{i=1}^{r}\lambda_i^k(q_i(x^k,v^k)-q_i(x^k,y^k)+o(\left|v^k-y^k\right|))=$$

$$=\sum_{i=1}^{r}\lambda_i^k q_i(x^k,v^k)+\sum_{i=1}^{r}\lambda_i^k o(\left|v^k-y^k\right|)\le\sum_{i=1}^{r}\lambda_i^k q_i(x^k,v^k)+\frac{1}{2}\left|v^k-y^k\right|.$$

Therefore, $d(v^k,\Phi(x^k))=\left|y^k-v^k\right|\le 2M\max\{0,q_i(x^k,v^k)\ i\in K\}$.

This contradicts (3.4). Thus $(c)\Rightarrow(a)$.

Lemma 3.2. *Let $(x^0,y^0)\in gr\Phi$ and $|q_i(x,y)-q_i(\tilde{x},y)|\le l_i\,|x-\tilde{x}|$ for all $x,\tilde{x}\in V(x^0)$ and $y\in V(y^0)$ where $l_i=const>0$ for all $i\in K$. Assume that the extended error bound property holds at the point (x^0,y^0) relative to $dom\Phi$. Then Φ is pseudo-Lipschitzian at this point relative to $dom\Phi$.*

Proof. If Φ has EEBP at (x^0,y^0) relative to $dom\Phi$, it means that there exist numbers $M>0$, $\delta>0$, $\varepsilon>0$ such that

$$d(y,\Phi(x))\le M\max\{0,q_i(x,y)\ i\in K\} \tag{3.6}$$

for all $y\in V_\varepsilon(y^0)$ and all $x\in V_\delta(x^0)\cap dom\Phi$.

Denote $l=\max\{l_i\,|i=1,...,r\}$. Choose numbers δ and ε such that $V_\delta(x^0)\subset V(x^0), V_\varepsilon(y^0)\subset V(y^0)$. Then

$$d(y^0,\Phi(x))\le M\max\{0,q_i(x,y^0)\ i\in K\}\le$$

$$\leq M \max\{0, q_i(x, y^0) - q_i(x^0, y^0)\ i \in K\} \leq l \left| x - x^0 \right|$$

for all $x \in V_\delta(x^0) \cap dom\Phi$.

This means that Φ is lower Lipschitz continuous at (x^0, y^0) relative to $dom\Phi$ and, consequently, $\Phi(x) \cap V_\varepsilon(y^0) \neq \emptyset$ for all $x \in V_{\delta_0}(x^0) \cap dom\Phi$ and any $\delta_0 < \min\{\delta, \varepsilon/l\}$. Let $x, \tilde{x} \in V_{\delta_0}(x^0) \cap dom\Phi$ and let $\tilde{y} \in \Phi(\tilde{x}) \cap V_\varepsilon(y^0)$. Then from (3.6) follows

$$d(\tilde{y}, \Phi(x)) \leq M \max\{0, q_i(x, \tilde{y})\ i \in K\} \leq$$

$$\leq M \max\{0, q_i(x, \tilde{y}) - q_i(\tilde{x}, \tilde{y})\ i \in K\} \leq l \left| x - \tilde{x} \right|.$$

This is equivalent to $\Phi(\tilde{x}) \cap V_\varepsilon(y^0) \subset \Phi(x) + l \left| x - \tilde{x} \right| B$ for all $x, \tilde{x} \in V_{\delta_0}(x^0) \cap dom\Phi$.

4 Pseudo-Lipschitz Continuity of Solution Mapping and Partial Calmness in BLPP

In this section we consider the problem BLPP which was stated in Section 1.

It is well known that the pseudo-Lipschitz continuity of a multivalued mapping at some point of its graph is closely related to constraint qualifications that hold at that point. However, it is easy to see that if the Kuhn-Tucker necessary condition holds for a solution of the lower level problem, then such classical constraint qualifications as the linear independence of the gradients of all active constraints and the Mangasarian-Fromovitz condition [Mangasarian, 1967] can't be fulfilled for the mapping S. It means that we need to involve some other constraint qualifications.

Following [Janin, 1984] we say that the *constant rank constraint qualification condition* (CRCQ) holds for the mapping S at a point $(x^0, y^0) \in C$ if there exist neighborhoods $V(x^0)$ and $V(y^0)$ such that $rank\{\nabla_y h_i(x, y)\ i \in K\} = const$ for all $x \in V(x^0)$, $y \in V(y^0)$ and any index set $K \subset \{0\} \cup I(x^0, y^0)$.

It is easy to see that $C \subset grS$ and CRCQ for the mapping $C(\cdot)$ at a point $(x^0, y^0) \in C$ is equivalent to CRCQ for the mapping S at this point. Before we prove the next lemma we also note that $domC(\cdot) \subset domS$ and the lower semicontinuity of S relative to $domS$ implies the lower semicontinuity of $C(\cdot)$ relative to $domC(\cdot)$.

Lemma 4.1. *Let $(x^0, y^0) \in C$. Assume that φ is continuous at x^0 relative to domS, S is lsc at (x^0, y^0) relative to domS and CRCQ holds for the mapping S at the point (x^0, y^0). Then S and $C(\cdot)$ have EEBP at (x^0, y^0).*

Proof. Suppose that $C(\cdot)$ does not have EEBP at (x^0, y^0). Then, due to Lemma 3.1 there exist sequences $x^k \to x^0$, $v^k \to y^0$ and $\{y^k\}$ such that $C(x^k) \neq \emptyset$, $v^k \notin C(x^k)$, $y^k \in \Pi_{C(x^k)}(v^k)$, and either $\Lambda_{v^k}(x^k, y^k) = \emptyset$ or $d(0, \Lambda_{v^k}(x^k, y^k)) \to \infty$ where

$$\Lambda_v(x, y) = \{\lambda \in R^{p+1} \left| \frac{y - v}{|y - v|} + \sum_{i=0}^{p} \lambda_i \nabla_y h_i(x, y) = 0, \lambda_i \geq 0 \text{ and } \lambda_i h_i(x, y) = 0\ i \in \{0\} \cup I \right. \}.$$

Lower semicontinuity of $C(\cdot)$ implies that $y^k \to y^0$. Then, without loss of generality, one can assume that CRCQ holds at all points (x^k, y^k). This means that $\Lambda_{v^k}(x^k, y^k) \neq \emptyset$ and, therefore, $\left| \lambda^k \right| \to \emptyset$ for each $\lambda^k \in \Lambda_{v^k}(x^k, y^k)$.

Since there is only a finite number of possible index sets $I(x^k, y^k)$, by working with a subsequence if necessary, we may assume that these index sets are the same for all $k = 1, 2, \ldots$. In other words, $I(x^k, y^k) = I^* \subset I(x^0, y^0)$, where I^* doesn't depend on $k = 1, 2, \ldots$, and $h_i(x^k, y^k) = 0\ i \in \{0\} \cup I^*$, $h_i(x^k, y^k) < 0\ i \in (I \backslash I^*)$.

It is known that if $\Lambda_{v^k}(x^k, y^k) \neq \emptyset$, then there exists a maximal linearly independent subfamily $\{\nabla_y h_i(x^k, y^k)\ i \in I^*(x^k, y^k) \subset \{0\} \cup I^*\}$ in the family $\{\nabla_y h_i(x^k, y^k)\ i \in \{0\} \cup I^*\}$ and a vector $\lambda^k \in \Lambda_{v_k}(z_k)$ such that $\lambda_i^k = 0$ for all $i \notin I^*(x^k, y^k)$. By choosing a subsequence if necessary, we may assume that the index set $I^*(x^k, y^k)$ is constant and denote it as $I^*(x^k, y^k) = I^0$. Then, for all $k = 1, 2, \ldots$ $\{\nabla_y h_i(x^k, y^k)\ i \in I^0\}$ is a maximal linearly independent subfamily in $\{\nabla_y h_i(x^k, y^k)\ i \in \{0\} \cup I^*\}$ and there exists a vector λ^k such that

$$\frac{y^k - v^k}{|y^k - v^k|} + \sum_{i=0}^{p} \lambda_i^k \nabla_y h_i(x^k, y^k) = 0\ \lambda_i^k \geq 0\ i \in \{0\} \cup I,\ \lambda_i^k = 0\ i \notin I^0. \tag{4.1}$$

Since $\lambda^k \to \infty$ in (4.1), without loss of generality one may assume that $\lambda^k \left| \lambda^k \right|^{-1} \to \bar{\lambda}$. Then from (4.2) follows

$$\sum_{i\in I^0} \bar\lambda_i \nabla_y h_i(x^0,y^0) = 0,\ \bar\lambda_i \geq 0\ i \in \{0\}\cup I,\ \bar\lambda_i = 0\ \forall i \notin I^0,\ |\bar\lambda| = 1.$$

This means that the vectors $\{\nabla_y h_i(x^0,y^0)\ i\in I^0\}$ are linearly dependent and we arrive to the contradiction with CRCQ for the mapping $C(\cdot)$ at the point (x^0,y^0).

Lemma 4.2. *Let the solution mapping S be uniformly bounded at x^0 and S be lsc relative to $domS$ at some point (x^0,y^0) where $y^0\in S(x^0)$. Then φ is continuous at x^0 relative to $domS$.*

Proof. Consider an arbitrary sequence $\{x^k\}\subset domS$ such that $x^k\to x^0$. Lower semicontinuity of S provides that there exists $y^k\in S(x^k)$ such that $y^k\to y^0$. Then

$$\lim_{k\to\infty}\sup\varphi(x^k) = \lim f(x^k,y^k) = f(x^0,y^0) = \varphi(x^0).$$

On the other hand, $\lim\limits_{k\to\infty}\inf\varphi(x^k)\geq\varphi(x^0)$ due to Lemma 3.18 [Luderer et al., 2002].

Lemma 4.3. *Assume that:*

1) the solution mapping S is uniformly bounded at x^0 and S is lsc at $(x^0,y^0)\in grS$ relative to $domS$;

2) CRCQ holds for the mapping S at (x^0,y^0).

Then the mapping S has EEBP at (x^0,y^0).

Proof. According to Lemma 4.2, the function φ is continuous at x^0 relative to *domS*. Then due to Lemma 4.1 EEBP holds at (x^0,y^0).

Theorem 4.1. *Let $(x^0,y^0)\in grS$. Assume that:*

1) the solution mapping S is uniformly bounded at x^0 and S is lsc at (x^0,y^0) relative to $domS$;

2) CRCQ holds for the mapping S at (x^0,y^0).

Then the mapping S is pseudo-Lipschitzian at (x^0,y^0) relative to $domS$.

Proof. Lemma 4.2 implies that CRCQ holds for the mapping S at (x^0,y^0). Then due to Lemma 3.2 S is pseudo-Lipschitzian at (x^0,y^0).

Example 4.1. Let $y_1-y_2\to min,\ y_2^2-y_1+x\leq 0,\ (y_1-x)^2+y_2\leq 0,\ 0\leq y_3\leq 1,\ 0\leq x\leq 1.$

Let $x^0=0$, $y^0=(0,0,0)$. CRCQ and other conditions of Theorem 4.1 hold at (x^0,y^0), therefore, S is pseudo-Lipschitzian at (x^0,y^0). On the other hand, it is easy to check that $\varphi(x)=x$ and $S(x)=\{y\in R^3\,|y_1=x,y_2=0,x\leq y_3\leq 1\}$.

Theorem 4.2. *Let a feasible point (x^0,y^0) be a local solution of the problem BLPP. Assume that:*

1) the function G is Lipschitz continuous on the set D with Lipschitz constant $l_0>0$;

2) the solution mapping S is uniformly bounded at x^0 and S is lsc at (x^0,y^0) relative to $domS$;

3) CRCQ holds for the mapping S at (x^0,y^0).

Then there exists a number $\mu_0>0$ such that for any $\mu>\mu_0$ the point (x^0,y^0) is a local solution of the problem $G(x,y)+\mu(f(x,y)-\varphi(x))\to\min$, $(x,y)\in D$.

The proof is similar to the proof of Theorem 2.1.

References

[Dempe, 2002] Dempe S. Foundations of bilevel programming. Kluwer Acad. Publishers, Dordrecht, 2002.

[Bard, 1998] Bard J.F. Practical bilevel optimization. Kluwer Acad. Publishers, Dordrecht, 1998.

[Ye, 1995] Ye J.J., Zhu D.L. Optimality conditions for bilevel programming problems. Optimization, 33, 1995. 9–27.

[Ye, 1997] Ye J.J., Zhu D.L., Zhu Q.J. Exact penalization and necessary optimality conditions for generalized bilevel programming problems. SIAM J. Optim. 7 (1997), no. 2. 481–507.

[Ye, 2010] Ye J.J., Zhu D.L. New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches. SIAM J. Optim. 20 (2010), no. 4. 1885–1905.

[Dempe, 2007] Dempe S., Dutta J., Mordukhovich B.S. New necessary optimality conditions in optimistic bilevel programming. Optimization 56 (2007), no. 5-6. 577–604.

[Dempe, 2013] Dempe S., Zemkoho A.B. Bilevel programming: reformulations, constraint qualifications and optimality conditions. Math. Programming 138, 2013. 447–473.

[Henrion, 2011] Henrion R., Surowiec T.M. On calmness conditions in convex bilevel programming. Appl. Anal. 90 (2011), no. 6. 951–970.

[Outrata, 1988] Outrata J.V. A note on the usage of nondifferentiable exact penalties in some special optimization problems. Kybernetika 24 (1988), no. 4. 251–258.

[Clarke, 1983] Clarke F. Optimization and nonsmooth analysis. Wiley, New York, 1983.

[Fedorov, 1979] Fedorov V.V. Numerical methods of maxmin. Nauka, Moscow, 1979.

[Luderer et al., 2002] Luderer B., Minchenko L., Satsura T. Multivalued analysis and nonlinear programming problems with perturbations. Kluwer Acad. Publishers, Dordrecht, 2002.

[Minchenko, 2015] Minchenko L.I., Tarakanov A.N. On second order derivatives of value functions. Optimization 64 (2015). 389–407.

[Minchenko, 2011] Minchenko L., Stakhovski S. Parametric nonlinear programming problems under the relaxed constant rank condition. SIAM J.Optim. 21 (2011), no. 1. 314–332.

[Borwein, 1986] Borwein J.M. Stability and regular points of inequality systems. J. Optimiz. Theory and Appl. 48 (1986). 9–52.

[Mangasarian, 1967] Mangasarian O. L., Fromovitz S. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. Journal of Mathem. Analysis and Applications 17 (1967). 37–47.

[Rockafellar, 1998] Rockafellar R.T., Wets R.J.-B. Variational analysis. Springer, Berlin, 1998.

[Klatte, 2015] Klatte D., Kummer B. On calmness of the argmin mapping in parametric optimization problems. J. Optimiz. Theory and Appl. 165 (2015). 708–719.

[Guo, 2013] Guo L., Lin G.H. Notes on some constraint qualifications for mathematical programs with equilibrium. J. Optimiz. Theory and Appl. 156 (2013). 600–616.

[Janin, 1984] Janin R. Directional derivative of the marginal function in nonlinear programming. Mathematical Programming Studies 21 (1984). 110–126.