Math. Program., Ser. A (2012) 131:37–48
DOI 10.1007/s10107-010-0342-1

FULL LENGTH PAPER

Is bilevel programming a special case of a mathematical program with complementarity constraints?

S. Dempe · J. Dutta

Received: 14 August 2008 / Accepted: 15 January 2010 / Published online: 25 February 2010

Abstract Bilevel programming problems are often reformulated using the Karush–Kuhn–Tucker conditions for the lower level problem resulting in a mathematical program with complementarity constraints(MPCC). Clearly, both problems are closely related. But the answer to the question posed is "No" even in the case when the lower level programming problem is a parametric convex optimization problem. This is not obvious and concerns local optimal solutions. We show that global optimal solutions of the MPCC correspond to global optimal solutions of the bilevel problem provided the lower-level problem satisfies the Slater's constraint qualification. We also show by examples that this correspondence can fail if the Slater's constraint qualification fails to hold at lower-level. When we consider the local solutions, the relationship between the bilevel problem and its corresponding MPCC is more complicated. We also demonstrate the issues relating to a local minimum through examples.

Keywords Bilevel programming · Mathematical programs with complementarity constraints · Optimality conditions · Local and global optimum

Mathematics Subject Classification (2000) 90C30

1 Introduction

Bilevel programming problems are hierarchical optimization problems combining decisions of two decision makers, the so-called leader and the so-called follower.

S. Dempe (✉)
Technical University Bergakademie Freiberg, Freiberg, Germany
e-mail: dempe@tu-freiberg.de

J. Dutta
Indian Institute of Technology, Kanpur, India

While the leader has the first choice and the follower reacts optimally on the leader's selection, the leader's aim consists in finding such a selection which, together with the follower's response, minimizes some cost function. To be more formal, denote the leader's selection by $x \in X$ for some closed set $X \subseteq \mathbb{R}^n$ and let the follower solve an optimization problem parameterized in x:

$$y \in \Psi(x) := \operatorname*{Argmin}_y \{f(x, y) : g(x, y) \leq 0\}. \tag{1.1}$$

Then, the bilevel programming problem (in its optimistic formulation, see e.g. the monograph [4] for different formulations and relations between them) consists in solving

$$\min_{x,y}\{F(x, y) : x \in X, \ \ y \in \Psi(x)\}. \tag{1.2}$$

Here, $f, F, g_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \ i = 1, \ldots, p$. Problem (1.2) is often called the leader's problem. The bilevel programming problem (1.1), (1.2) has many applications (see e.g. the annotated bibliography Dempe [5]). Many theoretical results can be found in the monographs Bard [2] and Dempe [4] and the edited volumes Migdalas et al. [13] and Dempe and Kalashnikov [6] on that topic. Note that the lower level problem is assumed to be solved globally. Hence, to keep the investigations tractable we assume throughout this paper that, for each fixed parameter value $x \in X$, the lower level problem (1.1) is a convex problem in the variables $y \in \mathbb{R}^n$, i.e. that the functions $y \mapsto f(x, y), \ y \mapsto g_i(x, y), \ i = 1, \ldots, p$, are convex in y for every fixed $x \in X$. If the functions $f(x, y)$ and $g_i(x, y), i = 1, \ldots, p$ are strongly convex or strictly convex in y for each fixed $x \in X$ then for each $x \in X$ the lower-level problem has a unique solution and the bilevel problem (in its original formulation where the leader minimizes with respect to x only) is well defined. Without these conditions in general the solution map $\Psi(\cdot)$ is a set-valued map, which makes the objective function of the upper-level problem also a set-valued one. In order to avoid the issue of handling a set-valued objective function at the upper-level the notion of optimistic formulation and pessimistic formulation of a bilevel programming problem was introduced in the literature. For more details on the issue of optimistic and pessimistic formulation see for example Dempe [4]. The major reason for considering convexity at the lower-level can be motivated as follows. Assume for the moment that each local optimal solution of problem (1.1) is strongly stable in the sense of Kojima [10] for some $\overline{x} \in X$. Then, without convexity assumption, the mapping $\Psi(\cdot)$ of global optimal solutions can contain more than one (strongly stable, isolated) optimal solution locally around $\overline{x}$ and is in general not continuous, see Jongen and Weber [9]. Hence, the function $y(x)$ with $y(x) \in \operatorname*{Argmin}_y \{F(x, y) : y \in \Psi(x)\}$ is in general also not uniquely determined locally around $\overline{x}$ and can have jumps. This makes attempts to formulate optimality conditions or solution algorithms very difficult.

Moreover in what follows we will consider an approach for investigating bilevel programming problems in which the lower level problem is replaced with its Karush–Kuhn–Tucker conditions. This results in a mathematical program with equilibrium constraints, MPCC for short. In the nonconvex case, the Karush–Kuhn–Tucker conditions are not sufficient and this approach will, hence, result is a much larger feasible set, see Mirrlees [14]. Thus, if the lower level problem is not convex, it is clear that the bilevel programming problem cannot be equivalent to the corresponding MPCC. But, are they equivalent in the case when the lower level problem is a parametric convex optimization problem? For solving a bilevel programming problem this is often reformulated in a MPCC and this is then solved using well developed approaches, see e.g. [1,3,7]. Those methods converge to (B-) stationary solutions or local optima and the question arises if the obtained solution then corresponds to a local optimal solution of the original problem. We will show below that this is unfortunately in general not the case. This makes this approach at least difficult and asks for a careful interpretation of the computed solution, respectively, a certain modification of the used solution algorithm, see [12].

The bilevel programming problem is a nonconvex programming problem with an implicitly determined feasible set. To solve it and to find (necessary and sufficient) optimality conditions for it the problem has to be reformulated. There are several different possibilities to do this.

The first and most often used attempt is to replace the lower level problem with its Karush–Kuhn–Tucker conditions. For this assume throughout this paper that the functions f, g_i are differentiable with respect to y and that these gradients are continuous with respect to both x, y. The objective function F of the upper level problem is assumed to be at least continuous with respect to both x, y. Let *Slater's constraint qualification* be satisfied for the lower level problem at any parameter value $x \in X$:

Slater's CQ: There exists $\overline{y}(x)$ such that $g_i(x, \overline{y}(x)) < 0,\ i = 1 \dots p$.
Then, the bilevel programming problem (1.1), (1.2) can be replaced with problem

$$\min_{x,y,u}\{F(x, y) : x \in X,\ \nabla_y L(x, y, u) = 0,\ g(x, y) \leq 0,\ u \geq 0,\ u^\top g(x, y) = 0\}. \tag{1.3}$$

Here $L(x, y, u) = f(x, y) + u^\top g(x, y)$ is the Lagrange function of problem (1.1) and ∇_y denotes the gradient with respect to the variables y only. This is a special case of a *mathematical program with complementarity constraints*, sometimes called *mathematical problem with equilibrium constraints*, see Luo et al. [11] and Outrata et al. [16]. This is a one-level programming problem involving additionally the Lagrange multipliers of the lower level problem as variables. It has been shown in the paper Ye et al. [20] that the Mangasarian-Fromowitz constraint qualification is violated at any feasible point of problem (1.3).

A second reformulation uses the *optimal value function* of the lower level problem

$$\varphi(x) := \min_y\{f(x, y) : g(x, y) \leq 0\} \tag{1.4}$$

and replaces the bilevel programming problem (1.1), (1.2) with

$$\min_{x,y}\{F(x, y) : x \in X,\ f(x, y) - \varphi(x) \leq 0,\ g(x, y) \leq 0\}. \tag{1.5}$$

This reformulation has been used in the paper Ye and Zhu [19] for describing necessary optimality conditions for the bilevel programming problem. See also Outrata [15] for describing a solution approach for bilevel programming problems using this reformulation. Problem (1.5) is a nonsmooth optimization problem since, even under strong assumptions, the function $\varphi(\cdot)$ is not differentiable. Moreover, the nonsmooth variant of the Mangasarian-Fromowitz constraint qualification is violated at any feasible point of this problem, see Ye and Zhu [19]. Problem (1.5) is clearly equivalent to problem (1.1), (1.2), but it is formulated using an implicitly determined, nonsmooth function $\varphi(\cdot)$ which can in general only be computed at fixed values of the parameter and approximated locally using e.g. generalized derivatives.

These two possibilities are the most widely used ones. Others reduce to replacing the lower level problem with a generalized variational inequality or to use a reformulation as a set-valued optimization problem (see e.g. [4]).

In most of the references dealing with bilevel programming problems they are quickly replaced by problem (1.3) and it is claimed that this can be done equivalently. We will investigate this topic for the global optimal solution in Sect. 2. The case of a local optimal solution will be treated in Sect. 3. We will show in Sect. 4 by an example that the linear independence constraint qualification is not generically satisfied in the lower level problem at an optimal solution of the bilevel programming problem.

2 Global optimal solution of a bilevel program

A feasible point $(\overline{x}, \overline{y})$ to the bilevel programming problem is a local optimal solution provided that there exists $\varepsilon > 0$ such that

$$F(x, y) \geq F(\overline{x}, \overline{y}) \;\forall\; (x, y) \in \operatorname{gph}\Psi, \quad x \in X \text{ and } \|(x, y) - (\overline{x}, \overline{y})\| < \varepsilon.$$

Here $\operatorname{gph}\Psi := \{(x, y) : y \in \Psi(x)\}$ denotes the graph of the solution set mapping $\Psi : \mathbb{R}^n \to 2^{\mathbb{R}^m}$ of the lower level problem defined in Eq. (1.1). A local optimal solution of problem (1.3) is defined in a similar way. A local optimal solution is a global one, if ε can be chosen arbitrarily large.

Theorem 2.1 *Let $(\overline{x}, \overline{y})$ be a global optimal solution of the bilevel programming problem and assume that the lower level problem is a convex one for which Slater's constraint qualification is satisfied at $x = \overline{x}$. Then, for each*

$$\overline{u} \in \Lambda(\overline{x}, \overline{y}) := \{u \geq 0 : \nabla_y L(\overline{x}, \overline{y}, u) = 0,\; u^\top g(\overline{x}, \overline{y}) = 0\}$$

the point $(\overline{x}, \overline{y}, \overline{u})$ is a global optimal solution of problem (1.3).

The proof of this theorem is an obvious consequence of $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ if and only if $(\overline{x}, \overline{y}) \in \operatorname{gph}\Psi$. Further it is interesting to note that for a given $(\bar{x}, \bar{y})$ the set $\Lambda(\bar{x}, \bar{y})$ is closed and the set-valued map $(x, y) \mapsto \Lambda(x, y)$ has a closed graph.

The following small example shows that, without regularity of the lower level problem, this theorem is in general not correct. Then, the bilevel programming problem can have a global optimal solution while the corresponding MPCC has no solution.

Example 2.2 Consider the convex lower level problem

$$\min_{y_1, y_2}\{y_1 : y_1^2 - y_2 \le x,\ \ y_1^2 + y_2 \le 0\} \tag{2.1}$$

Then, for $x = 0$, Slater's condition is violated, $y = 0$ is the only feasible point for $x = 0$. The optimal solution for $x \ge 0$ is

$$y(x) = \begin{cases} (0,0)^\top & \text{for } x = 0 \\ \left(-\sqrt{x/2}, -x/2\right)^\top & \text{for } x > 0 \end{cases}$$

The Lagrange multiplier is $u(x)$ with $u_1(x) = u_2(x) = \frac{1}{4\sqrt{x/2}}$ for $x > 0$. For $x = 0$, the problem is not regular and the Karush–Kuhn–Tucker conditions are not satisfied.

Consider the bilevel programming problem

$$\min\{x : x \ge 0,\ \ y \in \Psi(x)\} \tag{2.2}$$

where $\Psi(x)$ is the solution set mapping of problem (2.1).

Then, the unique (global) optimal solution of the bilevel programming problem (2.2) is $x = 0,\ \ y = 0$ and there does not exist local optimal solutions.

Consider the corresponding MPCC. Then, $(x, y(x), u(x))$ is feasible for the MPCC for $x > 0$ and the objective function value converges to zero for $x \to 0$. But, an optimal solution of this problem does not exist, since for $x = 0$ the only optimal solution of the lower level problem is $y = 0$ and there does not exist a corresponding feasible solution of the MPCC.

Note that this example can be used to show that the MPCC need not have a global optimal solution even if its feasible set is not empty and bounded.

The opposite implication of Theorem 2.1 is also true under a very mild assumption. We add the proof of this result, since despite of many authors are using it, we have not been able to find a thorough proof of it.

Theorem 2.3 *Let* $(\overline{x}, \overline{y}, \overline{u})$ *be a global optimal solution of problem* (1.3)*, let the lower level problem* (1.1) *be convex and assume that Slater's constraint qualification is satisfied for the lower level problem for each* $x \in X$*. Then,* $(\overline{x}, \overline{y})$ *is a global optimal solution of the bilevel programming problem.*

Proof If $(\overline{x}, \overline{y}, \overline{u})$ is a global optimal solution of (1.3) then $\Lambda(\overline{x}, \overline{y}) \neq \emptyset$ and, since the objective function value of problem (1.3) is independent of $u \in \Lambda(\overline{x}, \overline{y})$, each solution $(\overline{x}, \overline{y}, u),\ \ u \in \Lambda(\overline{x}, \overline{y})$ is a global optimal solution, too. Assume now, that $(\overline{x}, \overline{y})$ is not a global optimal solution of the bilevel programming problem. Hence there exists (x, y) with $x \in X$ and $y \in \Psi(x)$ such that $F(x, y) < F(\bar{x}, \bar{y})$. Since $y \in \Psi(x)$ and the Slater's constraint qualification holds at x the KKT conditions hold and thus there exists $u \in \mathbb{R}^p_+$ such that

$$\nabla_y f(x, y) + \sum_{i=1}^{p} u_i \nabla_y g_i(x, y) = 0,$$
$$u^T g(x, y) = 0$$
$$g_i(x, y) \leq 0.$$

This clearly shows that (x, y, u) is a feasible solution of the problem (1.3). This fact combined with the fact that $F(x, y) < F(\bar{x}, \bar{y})$ shows that $(\bar{x}, \bar{y}, \bar{u})$ is not a global optimal point of (1.3). This is a contradiction. Hence the result. □

Summing up, the bilevel programming problem is equivalent under some assumptions to the mathematical program with complementarity constraints if global optimal solutions are investigated.

It is important to observe that in Theorem 2.3 we need the assumption that for all $x \in X$ the Slater's constraint qualification holds for the lower-level problem. This assumption is essentially taken so that the KKT condition holds for the lower level problem for each (x, y) with $x \in X$ and $y \in \Psi(x)$. An interesting question is, if the KKT condition for the lower-level does not hold for every such point, what is the relation between the MPCC and the bilevel programming problem when we are investigating only global solutions. In the following example we present a bilevel programming problem whose lower-level problem satisfies the KKT conditions at only one point and none of the global solutions of the MPCC is a global solution of the bilevel programming problem. In this example, the MPCC has a global optimal solution but this point is not a (local or global) optimal solution of the bilevel programming problem. This example also shows that the assumption in Theorem 2.3 is essential.

Example 2.4 Consider the following optimistic bilevel programming problem where $(x, y) \in \mathbb{R} \times \mathbb{R}$:

$$\min_{x,y}\{(x-1)^2 + y^2 : x \in \mathbb{R}, y \in \Psi(x)\},$$

where Ψ denotes the solution set mapping of the following lower-level problem:

$$\min_{y}\{x^2 y : y^2 \leq 0\}.$$

It is clear that the lower-level problem has only one feasible solution i.e. $y = 0$ corresponding to any $x \in \mathbb{R}$. Further it is simple to check that $(x, y) = (0, 0)$ is the only point where the KKT conditions are satisfied. Observe that $(1, 0)$ is the only global solution of the bilevel problem. On the other hand the corresponding MPCC is given as

$$(x-1)^2 + y^2 \to \min_{x,y,\lambda}$$
$$x^2 + 2\lambda y = 0,$$
$$\lambda \geq 0,$$

$$y^2 \leq 0$$
$$\lambda y^2 = 0.$$

It is clear that the only feasible points of the MPCC problem are of the form $(0, 0, \lambda)$ where $\lambda \in \mathbb{R}$ and $\lambda \geq 0$. Thus all the feasible points of this MPCC are global solutions of MPCC. However, as we have already shown that $(0, 0)$ is a not a global solution of the bilevel problem it is clear that no global minimum of the MPCC is a global minimum of the bilevel programming problem.

It should be mentioned that, in the case when the KKT conditions for the lower level problem (1.1) are satisfied only in one point $(\overline{x}, \overline{y})$, $\overline{x} \in X$, which is a global optimal solution of the bilevel programming problem, then this point is the only feasible solution for the corresponding MPCC and hence also the unique (global) optimum. For the MPCC, there cannot exist local optimal solutions in this case. The bilevel programming problem, yet, can have more local and / or global optimal solutions.

3 Local optimal solutions

We have seen in the previous section that under a very mild condition the bilevel programming problem is equivalent to the corresponding MPCC when global solutions are investigated. Further it is important to note that it is very difficult to find a global solution and in many practical cases one needs to be happy with just obtaining a local solution. Thus from the point of view of solving the bilevel programming problem it is important to investigate the relation between local optimal solutions of the bilevel programming problem and the corresponding MPCC. The most important question in this regard is whether the equivalence that we observed in the case of global solutions still remains valid in the case of a local solution. The answer turns out to be NO as it is shown in the following counter example. In this example it is demonstrated that the local optimal solution of the MPCC need not be a local optimal solution to the bilevel programming problem. We remark that the lower level problem in this example is a parametric convex optimization problem having a unique (even strongly stable) optimal solution for all feasible (for the upper level problem) parameter values.

Example 3.1 Consider the lower level problem

$$y_1^2 + (y_2 + 1)^2 \rightarrow \min$$
$$(y_1 - x_1)^2 + (y_2 - 1 - x_1)^2 \leq 1$$
$$(y_1 + x_2)^2 + (y_2 - 1 - x_2)^2 \leq 1$$

and let $\Psi(x)$ denote its solution set. Figure 1 shows the lower level problem for fixed values of the parameters. These are the centers of the discs whose intersection is the feasible set of the lower level problem. These centers move along the dotted lines. Then, the optimal solution of the lower level problem is the feasible point with the smallest distance to the point $(0, -1)^\top$.

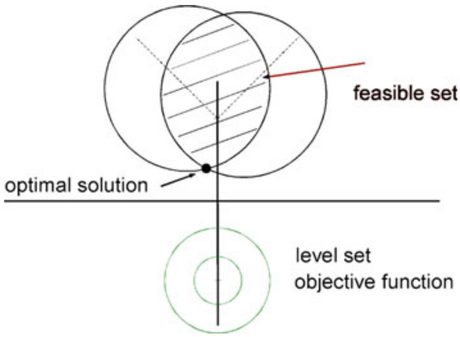

Fig. 1 The lower level problem in Example 3.1 for fixed parameter values

For $\overline{x} = (0, 0)^\top$ we have that $\overline{y} = (0, 0)^\top \in \Psi(\overline{x})$ is the unique optimal solution and

$$\Lambda(\overline{x}, \overline{y}) = \{u \in \mathbb{R}^2_+ : u_1 + u_2 = 1\}.$$

Consider the bilevel programming problem

$$\min\{-y_2 : y_1 y_2 = 0, x \geq 0, y \in \Psi(x)\}.$$

Then, since obviously $y_2 \geq 0$ for all points $y \in \Psi(x)$, $x \geq 0$ we have that each feasible point of the bilevel programming problem has an objective function value zero except at points $(\overline{x}, \overline{y})$, $\overline{y} \in \Psi(\overline{x})$ for which $\overline{y}_1 = 0, \overline{y}_2 > 0$. Those points have a negative objective function value for the bilevel programming problem.

The point $(\overline{x}, \overline{y})$, $\overline{x} = (0, 0)^\top$, $\overline{y} = (0, 0)^\top$ is feasible for the bilevel programming problem. But it is not locally optimal. We will demonstrate this fact in the following way. First observe that $(0, y_2)^T \in \Psi(x)$ with $y_2 > 0$ if and only if $x_1 = x_2$ and $u_1 = u_2$. This shows that there exists a sequence $\{(x^k, y^k)\}$ converging to $(\overline{x}, \overline{y})$ with $y^k \in \Psi(x^k)$, $y^k_1 = 0$ and $y^k_2 > 0$ for all k. For this sequence we have $x^k_1 = x^k_2$ and $u^k_1 = u^k_2$ for all k. This implies that the limit point of multipliers $u^k = (u^k_1, u^k_2)^\top$ is $\overline{u} = (0.5, 0.5)^\top$. Hence, the point $(\overline{x}, \overline{y}, \overline{u})$ is not a local minimum of the corresponding mathematical program with complementarity constraints. All other points $(\overline{x}, \overline{y}, u)$ with $u \in \Lambda(\overline{x}, \overline{y})$ are local minima.

Further it follows immediately from Theorem 2.1 that local optimal solutions $(\overline{x}, \overline{y})$ of the bilevel programming problem correspond to local optimal solutions $(\overline{x}, \overline{y}, \overline{u})$ of problem (1.3) for each $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ provided that the lower level problem (1.1) is convex and satisfies Slater's constraint qualification.

If Slater's constraint qualification is satisfied for the convex lower level problem, the set $\Lambda(\overline{x}, \overline{y})$ is a compact convex polyhedron having a finite set of vertices and being nonempty if and only if $\overline{y} \in \Psi(\overline{x})$, see Gauvin [8].

Let $(\overline{x}, \overline{y}, \overline{u})$ be a local optimal solution of problem (1.3). Then there exists $\varepsilon > 0$ such that for each (x, y, u) with $\nabla_y L(x, y, u) = 0$, $x \in X$, $g(x, y) \leq 0$, $u^\top g(x, y) = 0$, $u \geq 0$ and $\|(x, y, u) - (\overline{x}, \overline{y}, \overline{u})\| < \varepsilon$ we have $F(x, y) \geq F(\overline{x}, \overline{y})$. The complementarity conditions then imply that for each partition of $\{1, \ldots, p\}$ into two sets

$I_1, I_2 \subset \{1, \dots, p\}$, $I_1 \cup I_2 = \{1, \dots, p\}$, $I_1 \cap I_2 = \emptyset$,

$$\{j : \overline{u}_j > 0\} \subseteq I_1, \quad \{i : g_i(\overline{x}, \overline{y}) < 0\} \subseteq I_2 \tag{3.1}$$

and for each (x, y, u) satisfying $\|(x, y, u) - (\overline{x}, \overline{y}, \overline{u})\| < \varepsilon$ and

$$\nabla_y L(x, y, u) = 0, \; x \in X, \; g_i(x, y) \begin{cases} = 0 & \text{if } i \in I_1, \\ \leq 0 & \text{if } i \in I_2 \end{cases}, \; u_j \begin{cases} = 0 & \text{if } i \in I_2, \\ \geq 0 & \text{if } i \in I_1 \end{cases} \tag{3.2}$$

we have $F(x, y) \geq F(\overline{x}, \overline{y})$.

Theorem 3.2 *Let the lower level problem (1.1) be convex, Slater's constraint qualification be satisfied at the point $\overline{x}$ and $(\overline{x}, \overline{y}, \overline{u})$ be a local optimal solution for problem (1.3) for all $\overline{u} \in \Lambda(\overline{x}, \overline{y})$. Then, the point $(\overline{x}, \overline{y})$ is a local optimal solution of problem (1.1), (1.2), too.*

Proof Let $(\overline{x}, \overline{y})$ not be a local optimal solution, i.e. let there exists a sequence $\{(x^k, y^k)\} \subset \operatorname{gph} \Psi$ converging to $(\overline{x}, \overline{y})$ with $x^k \in X$ and $F(x^k, y^k) < F(\overline{x}, \overline{y})$ for all k. Since Slater's constraint qualification is satisfied at $\overline{x}$ and persistent in some open neighborhood of $\overline{x}$, there exists $u^k \in \Lambda(x^k, y^k)$ having an accumulation point $\widehat{u} \in \Lambda(\overline{x}, \overline{y})$ by upper semicontinuity of the Lagrange multiplier set mapping [17]. This means that there is a sequence (x^k, y^k, u^k) of feasible solutions to problem (1.3) converging to a feasible point $(\overline{x}, \overline{y}, \widehat{u})$ of (1.3) with $F(x^k, y^k) < F(\overline{x}, \overline{y})$. This violates local optimality of $(\overline{x}, \overline{y}, \overline{u})$ for all $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ for (1.3). Hence, $(\overline{x}, \overline{y})$ is local optimal for the bilevel programming problem.

It will be interesting to note in Example 3.1 the crucial requirement of Theorem 3.2 that $(\bar{x}, \bar{y}, \bar{u})$ has to be a local minimum of the MPCC problem (1.3) *for all* $\bar{u} \in \Lambda(\bar{x}, \bar{y})$ is violated. Note that in Example 3.1, $u = (0.5, 0.5) \in \Lambda(0, 0)$ but $(0, 0, 0.5, 0.5)$ is not a local minimum of the corresponding MPCC problem.

For checking local optimality of the bilevel programming problem the Example 3.1 indicates that it is (in the worst case) necessary to verify if all feasible solutions $(\widehat{x}, \widehat{y}, \widehat{u})$, $\widehat{x} \in X$, $g(\widehat{x}, \widehat{y}) \leq 0$, $\widehat{u} \in \Lambda(\widehat{x}, \widehat{y})$ are local optimal solutions of the MPCC. This is an infinite number of points, and possibly only one point out of them (where $\widehat{u} \in \operatorname{int} \Lambda(\widehat{x}, \widehat{y})$ is possible) is not a local optimal solution proving that $(\widehat{x}, \widehat{y})$ is not a local optimal solution for the bilevel programming problem. Under the constant rank constraint qualification below we reduce this effort to a finite number of points. Note, that $\widehat{u}$ needs not to be a vertex of $\Lambda(\overline{x}, \overline{u})$ in the proof of Theorem 3.2.

Consider the *constant rank constraint qualification*:

Constant rank CQ: There exists an open neighborhood V of $(\overline{x}, \overline{y})$ such that for each index set $I_1 \subseteq \{j : g_j(\overline{x}, \overline{y}) = 0\}$ the family of gradients $\{\nabla_y g_j(x, y) : j \in I_1\}$ has the same rank on V.

Corollary 3.3 *If the constant rank constraint qualification is added to the assumptions in Theorem 3.2 local optimality of $(\overline{x}, \overline{y}, \overline{u})$ for all vertices $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ for* *problem* (1.3) *implies local optimality of* $(\overline{x}, \overline{y})$ *for the bilevel programming problem* (1.1)*,* (1.2).

Indeed, the multipliers u^k in the proof of Theorem 3.2 can without assumptions be taken as vertices of $\Lambda(x^k, y^k)$. The constant rank constraint qualification then implies, that the sequence $\{u^k\}$ converges to a vertex of $\Lambda(\overline{x}, \overline{y})$. Hence, $\widehat{u}$ can also be taken as a vertex of $\Lambda(\overline{x}, \overline{y})$, showing the desired result.

The following example shows that, in case of multiple Lagrange multipliers and if the constant rank constraint qualification is satisfied it is not possible to weaken the assumptions of the corollary. Theoretically fixing (x, y, u) to an open neighborhood of some local optimal solution $(\overline{x}, \overline{y}, \overline{u})$ means that an inequality constraint $g_i(x, y) \leq 0$ with $\overline{u}_i > 0$ is replaced by an equation $g_i(x, y) = 0$ thus reducing the feasible set of the bilevel programming problem to a possibly proper subset. If $(\overline{x}, \overline{y}, \overline{u})$ is a local optimal solution subject to this subset it is in general not a local optimal solution subject to the feasible set itself.

Example 3.4 Consider the linear lower level problem

$$\min_y\{-y : x + y \leq 1,\ -x + y \leq 1\} \tag{3.3}$$

having the unique optimal solution $y(x)$ and the set $\Lambda(x, y)$ of Lagrange multipliers

$$y(x) = \begin{cases} x + 1 & \text{if } x \leq 0 \\ -x + 1 & \text{if } x \geq 0 \end{cases},\ \Lambda(x, y) = \begin{cases} \{(1, 0)\} & \text{if } x > 0 \\ \{(0, 1)\} & \text{if } x < 0 \\ \text{conv}\{(1, 0), (0, 1)\} & \text{if } x = 0 \end{cases}$$

where conv A denotes the convex hull of the set A. The bilevel programming problem

$$\min\{(x - 1)^2 + (y - 1)^2 : (x, y) \in \text{gph}\,\Psi\} \tag{3.4}$$

has the unique optimal solution $(\overline{x}, \overline{y}) = (0.5, 0.5)$ and no local optimal solutions. The points $(\overline{x}, \overline{y}, \overline{u}_1, \overline{u}_2) = (0.5, 0.5, 1, 0)$ and $(x^0, y^0, u_1^0, u_2^0) = (0, 1, 0, 1)$ are a local optimal solutions of problem (1.3). To see that the point $(x^0, y^0, u_1^0, u_2^0) = (0, 1, 0, 1)$ is locally optimal remark that in an open neighborhood V of this point we have $u_2 > 0$ implying that the second constraint in problem (3.3) is active, i.e. $y = x + 1$ which by feasibility implies $x \leq 0$. Substituting this into the upper level objective function gives

$$(x - 1)^2 + (y - 1)^2 = (x - 1)^2 + x^2 \geq 1$$

for each feasible point of the mathematical program with equilibrium constraints corresponding to the bilevel problem (3.4) in V. Together with $F(0, 1) = 1$ we derive that (x^0, y^0, u_1^0, u_2^0) is indeed a local optimum of the mathematical program with equilibrium constraints corresponding to (3.4). Also observe that we had mentioned before that when we have multiple Lagrange multipliers and the constant rank constraint qualification is satisfied it is not possible to weaken the assumption of Corollary 3.3.

Consider the point $(x^0, y^0) = (0, 1)$. Then it is simple to see that the vertices of $\Lambda(x^0, y^0) = \Lambda(0, 1)$ are the vectors $(1, 0)$ and $(0, 1)$. It has been shown above that $(0, 1, 0, 1)$ is a local optimal solution of the associated MPCC problem while note that $(0, 1, 1, 0)$ is not a local minimum of the MPCC problem. Also note that $(0, 1)$ is not a solution of the bilevel problem.

4 Regularity

For mathematical programs with equilibrium constraints it has been shown by Scholtes and Stöhr [18] that the *MPCC-LICQ* is generically satisfied. Closely related is the question if the partial gradients of the active constraints with respect to y in the lower level problem are generically linearly independent, i.e. if the linear independence constraint qualification is generically satisfied for the lower level problem at a local optimal solution. If this would be true, then the above examples would be exceptions and a local optimal solution of the MPCC would generically imply local optimality for the bilevel programming problem (if the lower level problem is convex in y satisfying a suitable regularity condition).

The following example shows that the linear independence constraint qualification (and hence uniqueness of a Lagrange multiplier) in the lower level problem is not a generic assumption.

Example 4.1 Consider the lower level problem

$$\Psi_L(x) := \operatorname*{Argmin}_{y,z} \{-y - z : x + y \leq 1,\ -x + y \leq 1,\ 0 \leq z \leq 1\}$$

and the bilevel programming problem

$$\min\{0.5x - y + 3z : (y, z) \in \Psi_L(x)\}.$$

Then, we derive $\Psi_L(x) = \{(y(x), z(x)) : z(x) = 1,\ x \in \mathbb{R}\}$ with

$$y(x) = \begin{cases} 1 - x & \text{if } x \geq 0 \\ 1 + x & \text{if } x \leq 0. \end{cases}$$

Substituting this solution into the upper level objective function we derive

$$0.5x - y(x) + 3z(x) = \begin{cases} 0.5x - 1 + x + 3 = 2 + 1.5x \geq 2 & \text{if } x \geq 0 \\ 0.5x - 1 - x + 3 = 2 - 0.5x \geq 2 & \text{if } x \leq 0. \end{cases}$$

Hence, $(\overline{x}, \overline{y}, \overline{z}) = (0, 1, 1)$ is the global (and unique local) optimal solution of the problem. And three constraints in the lower level problem are active. Since only two variables occur in the lower level problem, the linear independence constraint qualification is violated. Note that small smooth perturbations of the data of both the lower and the upper level problems will have no impact on this property.

References

1. Anitescu, M.: On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. SIAM J. Optim. **15**, 1203–1236 (2006)
2. Bard, J.F.: Practical Bilevel Optimization: Algorithms and Applications. Kluwer, Dordrecht (1998)
3. DeMiguel, A.-V., Friedlander, M.P., Nogales, F.J., Scholtes, S.: An interior-point method for MPECS based on strictly feasible relaxations. Technical report, Department of Decision Sciences, London Business School (2004)
4. Dempe, S.: Foundations of Bilevel Programming. Kluwer, Dordrecht (2002)
5. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**, 333–359 (2003)
6. Dempe, S., Kalashnikov, V. (eds.): Optimization with Multivalued Mappings: Theory, Applications and Algorithms. Springer, Berlin (2006)
7. Fukushima, M., Pang, J.-S.: Convergence of a smoothing continuation method for mathematical programs with complementarity constraints Ill-posed Variational Problems and Regularization Techniques, no. 477, Springer, Berlin (1999)
8. Gauvin, J.: A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. Math. Program. **12**, 136–139 (1977)
9. Jongen, H.Th., Weber, G.-W.: Nonlinear optimization: characterization of structural optimization. J. Glob. Optim. **1**, 47–64 (1991)
10. Kojima, M.: Strongly stable stationary solutions in nonlinear programs. In: Robinson, S.M. (ed.) Analysis and Computation of Fixed Points, pp. 93–138. Academic Press, New York (1980)
11. Luo, Z.-Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
12. Mersha, A.G.: Solution methods for bilevel programming problems. Ph.D. thesis, TU Bergakademie Freiberg (2008)
13. Migdalas, A., Pardalos, P.M., Värbrand, P. (eds.): Multilevel Optimization: Algorithms and Applications. Kluwer, Dordrecht (1998)
14. Mirrlees, J.A.: The theory of moral hazard and unobservable bevaviour: part I. Rev. Econ. Stud. **66**, 3–21 (1999)
15. Outrata, J.: On the numerical solution of a class of Stackelberg problems. ZOR Math. Methods Oper. Res. **34**, 255–277 (1990)
16. Outrata, J., Kočvara, M., Zowe, J.: Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer, Dordrecht (1998)
17. Robinson, S.M.: Generalized equations and their solutions, part II: applications to nonlinear programming. Math. Program. Stud. **19**, 200–221 (1982)
18. Scholtes, S., Stöhr, M.: How stringent is the linear independence assumption for mathematical programs with stationarity constraints?. Math. Oper. Res. **26**, 851–863 (2001)
19. Ye, J.J., Zhu, D.L.: Optimality conditions for bilevel programming problems. Optimization **33**, 9–27 (1995)
20. Ye, J.J., Zhu, D.L., Zhu, Q.J.: Exact penalization and necessary optimality conditions for generalized bilevel programming problems. SIAM J. Optim. **7**, 481–507 (1997)