J Glob Optim (2012) 53:573–586
DOI 10.1007/s10898-011-9762-6

Bilevel problems over polyhedra with extreme point optimal solutions

Herminia I. Calvete · Carmen Galé · Stephan Dempe · Sebastian Lohse

Received: 20 July 2010 / Accepted: 25 July 2011 / Published online: 9 August 2011

Abstract Bilevel programming involves two optimization problems where the constraint region of the upper level problem is implicitly determined by another optimization problem. In this paper we focus on bilevel problems over polyhedra with upper level constraints involving lower level variables. On the one hand, under the uniqueness of the optimal solution of the lower level problem, we prove that the fact that the objective functions of both levels are quasiconcave characterizes the property of the existence of an extreme point of the polyhedron defined by the whole set of constraints which is an optimal solution of the bilevel problem. An example is used to show that this property is in general violated if the optimal solution of the lower level problem is not unique. On the other hand, if the lower level objective function is not quasiconcave but convex quadratic, assuming the optimistic approach we prove that the optimal solution is attained at an extreme point of an 'enlarged' polyhedron.

This research work has been partially funded by the Spanish Ministry of Education and Science under grant MTM2007-66893.

H. I. Calvete (✉)
Dpto. de Métodos Estadísticos, IUMA, Universidad de Zaragoza, Pedro Cerbuna 12,
50009 Zaragoza, Spain
e-mail: herminia@unizar.es

C. Galé
Dpto. de Métodos Estadísticos, IUMA, Universidad de Zaragoza, María de Luna 3,
50018 Zaragoza, Spain
e-mail: cgale@unizar.es

S. Dempe · S. Lohse
Department of Mathematics and Computer Science, Technical University Bergakademie, Freiberg,
Germany
e-mail: dempe@tu-freiberg.de

S. Lohse
e-mail: sebastian.lohse@math.tu-freiberg.de

Keywords Bilevel programming · Quasiconcave · Quadratic · Extreme point

Mathematics Subject Classification (2000) 90C30

1 Introduction

Bilevel programming has been proposed for modeling hierarchical decision processes with two decision makers, the leader or upper level decision maker and the follower or lower level decision maker. All the variables are controlled either by the leader or by the follower, both of whom have their own objective function and constraints. The follower optimizes his objective function under the given parameters from the leader. In return, having complete information on the possible reactions of the lower level decision maker, the leader selects the parameters so as to optimize his own objective function. Bilevel problems can be formulated as:

$$\min_{x_1,x_2} \quad f_1(x_1,x_2) \tag{1a}$$

$$\text{s.t.} \quad (x_1,x_2) \in R \tag{1b}$$

where x_2 solves

$$\min_{x_2} \quad f_2(x_1,x_2) \tag{1c}$$

$$\text{s.t.} \quad (x_1,x_2) \in S \tag{1d}$$

where $x_1 \in \mathbb{R}^{n_1}$ and $x_2 \in \mathbb{R}^{n_2}$ are the upper level and lower level variables, respectively; $f_1, f_2 : \mathbb{R}^n \longrightarrow \mathbb{R}, n = n_1 + n_2$ are the upper level and lower level objective functions, respectively; and $R, S \subseteq \mathbb{R}^n$ are the sets defined by the upper level and the lower level constraints, respectively. Usually, upper level constraints involving upper and lower level variables are called coupling constraints. The set defined by all constraints $T = R \cap S$ is called the constraint region. Let R_1, S_1 and T_1 be the projection of R, S and T onto $\mathbb{R}^{n_1}$, respectively. Due to their structure, bilevel programs are nonconvex and quite difficult to deal with and solve. Bard [2], Dempe [8], Migdalas and Pardalos [14], Migdalas et al. [15] and Shimizu et al. [18] are good general references on this topic. Additionally, Dempe [9] and Vicente and Calamai [19] provide surveys which cover applications as well as major theoretical developments. Chinchuluun et al. [6] discuss some algorithmic and theoretical results on multilevel programming.

Where all the functions involved are linear (LB problem), Savard [17] has proved that there is an extreme point of the polyhedron T which solves the problem. Xu [20] has achieved similar results under weaker assumptions via a penalty function approach. The purpose of this paper is to analyze bilevel problems over polyhedra which verify the extreme point property and point out the differences with the linear case. Similarly to the case when there is only one level of decision making, we will prove that the fact that f_1 and f_2 are quasiconcave functions characterizes the property of the existence of an extreme point of the polyhedron T which solves the bilevel problem, under the uniqueness of the optimal solution of the lower level problem. Finally, we will prove that exchanging the quasiconcave lower level objective function for a convex quadratic one allows us to conclude that an optimal solution occurs at an extreme point of an 'enlarged' polyhedron that takes into account complementary constraints of Karush-Kuhn-Tucker conditions, without assuming that the optimal solution of the lower level problem is a singleton. The paper is organized as follows. Section 2 states the problem and provides some insight into its behavior and approaches regarding the optimal solution. In Sect. 3 the characterization of bilevel problems with an optimal solution at an extreme point of T is given. Section 4 provides the main theoretical result on the optimal solution of bilevel problems with quasiconcave upper level and convex quadratic lower level objective functions. Finally, Sect. 5 concludes the paper with some final remarks.

2 Existence of an optimal solution of the bilevel problem

For a given x_1, the follower solves the lower level problem (1c)–(1d). Let $S(x_1) = \{x_2 : (x_1, x_2) \in S\}$ be its feasible region and $M(x_1)$ be the set of optimal solutions, also called the follower rational reaction set. The feasible region of the bilevel problem, called the inducible (or induced) region, is:

$$\text{IR} = \{(x_1, x_2) : (x_1, x_2) \in T,\ x_2 \in M(x_1)\} \tag{2}$$

Any point of IR is a bilevel feasible solution. A point $x_1 \in \mathbb{R}^{n_1}$ is called permissible if an $x_2 \in \mathbb{R}^{n_2}$ exists so that $(x_1, x_2) \in \text{IR}$. Let P be the set of permissible points. Taking into account previous definitions, the bilevel problem formulated in (1) can be equivalently written as:

$$\begin{aligned} \min_{x_1, x_2} \ & f_1(x_1, x_2) \\ \text{s.t.} \ & (x_1, x_2) \in \text{IR}. \end{aligned} \tag{3}$$

Two main difficulties arise when looking for the existence of an optimal solution of bilevel programs. On the one hand, IR could be an empty set although T is a nonempty set. The following example in $\mathbb{R}^2$ illustrates this fact.

Example 1

$$\begin{aligned} \min_{x_1, x_2} \quad & \frac{x_1 + 2x_2 + 2}{x_1 + x_2 + 4} \\ \text{s.t.} \quad & -x_1 + 10x_2 \le 36 \\ & 5x_1 + 13x_2 \ge 18 \\ & -6x_1 + 22x_2 \ge 32 \end{aligned}$$

where x_2 solves

$$\begin{aligned} \max_{x_2} \quad & x_2 \\ \text{s.t.} \quad & x_1 - x_2 \le 0 \\ & -x_1 - x_2 \le 0 \\ & 0 \le x_2 \le 5 \end{aligned}$$

Figure 1 displays the constraint region T which is the shaded nonempty compact polyhedron ABCDE. The thick line shows the optimal solution of the lower level problem for all $x_1 \in S_1$. Therefore, IR $= \emptyset$.

On the other hand, complications arise when there are multiple optima in the lower level problem, that is to say, $M(x_1)$ is not a singleton for some permissible x_1. In fact, some authors use the term min in (1a) in quotation marks to express a certain ambiguity in the problem formulation in the case of non-unique lower level solutions [8,13]. If the upper level objective function is sensitive to different values of $x_2 \in M(x_1)$, it is necessary to give a rule to select

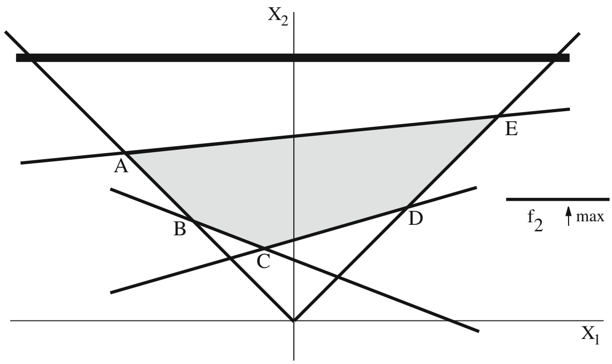

Fig. 1 The inducible region of Example 1 is empty

$x_2^* \in M(x_1)$ in order to evaluate f_1. Notice, however, that taking into account the intrinsic noncooperative nature of the model, there is no reason why both decision makers should collaborate, so it is not certain that the upper level decision maker can influence the selection of the lower level one.

Several assumptions have been proposed in the literature to make sure that the bilevel problem is well posed [8,13]. The most common are:

1. To assume that the lower level decision maker always selects the optimal decision which gives the worst value of f_1. This is the pessimistic or strong approach, which is used when the leader is not able to influence the follower and is forced to choose an approach bounding the damage resulting from an unfavorable selection by the follower. The resulting problem is: $\min_{x_1 \in P} \quad \phi(x_1)$ where $\phi(x_1) = \max_{x_2 \in M(x_1)} f_1(x_1, x_2)$.
2. To assume that the upper level decision maker is able to influence the lower level one so that the latter always selects the variables x_2 to provide the best value of f_1. This results in the so-called optimistic or weak bilevel problem $\min_{x_1 \in P} \psi(x_1)$ where $\psi(x_1) = \min_{x_2 \in M(x_1)} f_1(x_1, x_2)$. Due to its important properties, we will analyze this case in depth in the following sections.

The following example in $\mathbb{R}^2$ allows us to illustrate pessimistic and optimistic approaches.

Example 2

$$
\begin{aligned}
\min_{x_1, x_2} \quad & -x_1 + x_2, \quad \text{where } x_2 \text{ solves} \\
\min_{x_2} \quad & -(x_2 - 1)^2 \\
\text{s.t.} \quad & 4x_1 + x_2 \geq 5 \\
& x_1 - 2x_2 \leq 8 \\
& x_1 + 4x_2 \leq 20 \\
& x_1 \leq 7
\end{aligned}
$$

In Fig. 2, the shaded region is the constraint region and the thick line displays the optimal solution of the lower level problem for all permissible x_1. If we take the optimistic approach then (4, −2) is the optimal solution of the bilevel problem and the best value of f_1 is −6. But if we take the pessimistic approach, then the infimum of f_1 is not attained. Indeed, if the leader selected $x_1 = 4$, then the follower would choose $x_2 = 4$ since points $x_2 = -2$ and

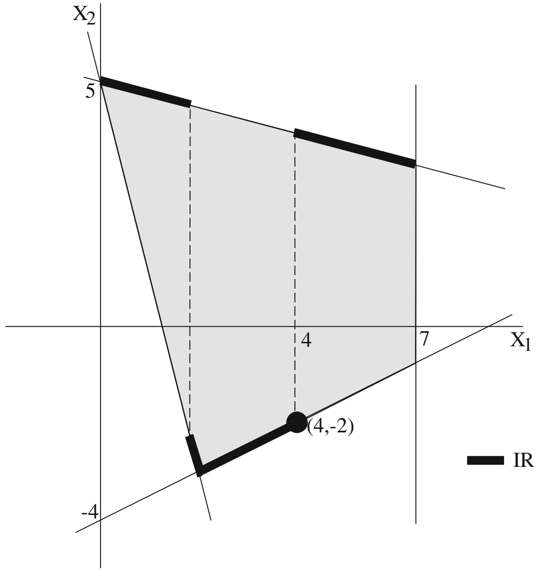

Fig. 2 Inducible region and optimistic optimal solution of Example 2

$x_2 = 4$ are optimal solutions of the lower level problem and $f_1(4, 4) > f_1(4, -2)$. Hence, the pessimistic bilevel problem has no optimal solution.

3 Bilevel problems over polyhedra

From now on, we restrict our attention to bilevel problems defined over polyhedra. We assume that S and R are nonempty polyhedra and $T \neq \emptyset$. $S(x_1)$ and T are assumed to be bounded in order to guarantee the existence of an optimal solution in the lower level and upper level problems, respectively. We also assume that f_1 and f_2 are continuous functions.

Theorem 1 *Assuming the optimistic approach, IR is closed.*

Proof It directly follows from Theorem I.2.2 in [7]. □

Theorem 2 *Assuming the optimistic approach, if IR is nonempty, then problem* (3) *has an optimal solution.*

Proof Problem (3) minimizes a continuous function over a nonempty and compact region. Hence, by applying Weierstrass's theorem the conclusion follows. □

Note that the case in which $M(x_1)$ is a singleton for all permissible x_1 can be considered a particular case of the optimistic approach.

In the particular case of LB problems, if IR is nonempty, it has been proved that at least one optimal solution is obtained at an extreme point of the constraint region T [13,17]. This important property allows us, amongst other things, to develop enumerative algorithms which search amongst extreme points to solve LB problems [5,17]. The following example in which we have exchanged max for min in the lower level problem of Example 1 suggests that this property is valid for more general problems.

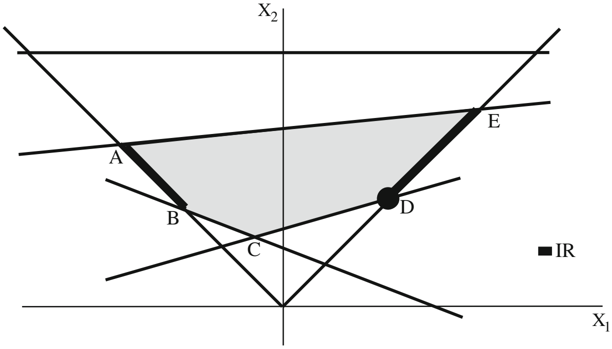

Fig. 3 Inducible region and optimal solution of Example 3

Example 3

$$\begin{array}{ll} \min\limits_{x_1,x_2} & \dfrac{x_1 + 2x_2 + 2}{x_1 + x_2 + 4} \\ \text{s.t.} & -x_1 + 10x_2 \leq 36 \\ & 5x_1 + 13x_2 \geq 18 \\ & -6x_1 + 22x_2 \geq 32 \end{array}$$

where x_2 solves

$$\begin{array}{ll} \min\limits_{x_2} & x_2 \\ \text{s.t.} & x_1 - x_2 \leq 0 \\ & -x_1 - x_2 \leq 0 \\ & 0 \leq x_2 \leq 5 \end{array}$$

The shaded region in Fig. 3 is the constraint region T. The two faces of T with thick lines are the feasible region IR. Vertex D is the optimal solution of Example 3. However, Example 2 shows that this property is not true in general. There, the lower level objective function is strongly (quasi)concave and an optimal solution of the lower level problem is a vertex of the feasible set of this problem for each fixed x_1. But the solution of the bilevel problem is not a vertex of the constraint region.

For the purpose of characterizing bilevel problems over polyhedra for which there exists an extreme point of T that solves the problem, we need two additional assumptions:

(A1) $M(x_1)$ is a singleton for all $x_1 \in S_1$.
(A2) f_1 is quasiconcave on T and $f_2(x_1, \cdot)$ is quasiconcave on $S(x_1)$, $x_1 \in T_1$.

We will call this the quasiconcave bilevel (QB) problem. Recall that a real-valued function h defined on a convex subset D of $\mathbb{R}^n$ is quasiconcave on D if and only if $d^1, d^2 \in D$, $\lambda \in (0, 1)$ and $h(d^1) \leqslant h(d^2)$ imply $h(d^1) \leqslant h[(1-\lambda)d^1 + \lambda d^2]$.

In the first part of this Section we will extend to the more general problem considered in this paper some results on the geometry of IR obtained in [3] for the quasiconcave bilevel problem without coupling constraints. For this purpose, let us introduce the following bilevel problem in which upper level constraints (1b) have been shifted from the upper to the lower level:

$$P_T: \quad \min_{x_1,x_2} \; f_1(x_1,x_2), \text{ where } x_2 \text{ solves} \\ \min_{x_2} \; f_2(x_1,x_2) \\ \text{s.t.} \; (x_1,x_2) \in T \tag{4}$$

The feasible set for the follower is $T(x_1) = \{x_2 : (x_1, x_2) \in T\}$, $x_1 \in T_1$. Let IR_T denote the inducible region of problem P_T. Let $\widehat{S} = \{(x_1, x_2) : x_1 \in T_1, (x_1, x_2) \in S\}$. In like manner, we consider:

$$P_{\widehat{S}}: \min_{x_1,x_2} \; f_1(x_1,x_2), \quad \text{where } x_2 \text{ solves} \\ \min_{x_2} \; f_2(x_1,x_2) \\ \text{s.t.} \; (x_1,x_2) \in \widehat{S} \tag{5}$$

In this case, the feasible set for the follower is $S(x_1)$, $x_1 \in T_1$. Let $\mathrm{IR}_{\widehat{S}}$ denote the inducible region of problem $P_{\widehat{S}}$.

Lemma 1 *A point $\tilde{x}_1 \in T_1$ is permissible for problem* (1) *if and only if an $\tilde{x}_2 \in \mathbb{R}^{n_2}$ exists so that*

$$(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_T \cap \mathrm{IR}_{\widehat{S}}. \tag{6}$$

Proof If $\tilde{x}_1$ is permissible then an $\tilde{x}_2 \in \mathbb{R}^{n_2}$ exists so that $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}$. Since $(\tilde{x}_1, \tilde{x}_2) \in T$ and $\tilde{x}_2 = \mathrm{argmin}_{y_2}\{f_2(\tilde{x}_1, y_2) : y_2 \in S(\tilde{x}_1)\}$, then $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_{\widehat{S}}$.

Taking into account that $T(\tilde{x}_1) \subseteq S(\tilde{x}_1)$, we obtain that $\tilde{x}_2 = \mathrm{argmin}_{y_2}\{f_2(\tilde{x}_1, y_2) : y_2 \in T(\tilde{x}_1)\}$ and so $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_T$. Therefore, $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_T \cap \mathrm{IR}_{\widehat{S}}$.

Similarly, from (6) we obtain that $\tilde{x}_2$ is the optimal solution to the lower level problem of problem (1) for $x_1 = \tilde{x}_1$ and $(\tilde{x}_1, \tilde{x}_2) \in T$. Therefore, $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}$, and $\tilde{x}_1$ is permissible.

□

In words, Lemma 1 establishes that a point $\tilde{x}_1 \in T_1$ is permissible if and only if an $\tilde{x}_2 \in \mathbb{R}^{n_2}$ exists so that $\tilde{x}_2$ is an optimal solution of the lower level problem of problem (4) (i.e. $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_T$) and $\tilde{x}_2$ is an optimal solution of the lower level problem of problem (5) (i.e. $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_{\widehat{S}}$). Note that, for x_1 permissible, the optimal solution of the lower level problem of problems (5) and (1) coincide.

Lemma 2 *Assuming hypothesis* (A1) *and* (A2)*, the feasible regions of problems* (1), (4) *and* (5) *are the union of faces of the corresponding constraints regions.*

Proof Since problems (4) and (5) do not include coupling constraints, the assertion of the lemma regarding these problems directly follows from Lemmas 2.1–2.5 in [3]. That is to say, there exist J and I finite index sets so that

$$\mathrm{IR}_T = \bigcup_{j \in J} T_j \tag{7a}$$

$$\mathrm{IR}_{\widehat{S}} = \bigcup_{i \in I} \widehat{S}_i \tag{7b}$$

where T_j, $j \in J$, and $\widehat{S}_i$, $i \in I$, denote nonempty faces of T and $\widehat{S}$, respectively.

Now, let us prove that there exists a finite index set K so that the feasible region of problem (1) can be expressed as follows:

$$\mathrm{IR} = \bigcup_{k \in K} T_k \tag{8}$$

where T_k, $k \in K$, denote nonempty faces of T.

We recall here that every face of a polyhedron is itself a polyhedron and the fact that the collection of all relative interiors of nonempty faces of a nonempty convex set is a partition of that set (Theorem 18.2 in [16]).

Let $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}$. As a consequence of Lemma 1, we can write $\mathrm{IR} = \mathrm{IR}_T \cap \mathrm{IR}_{\widehat{S}}$. Hence, $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR}_T$ and, by applying (7a) and previous remarks, we conclude that there is a face T_k of T, $k \in J$, so that $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{ri}\ T_k$, where ri denotes relative interior. We associate face T_k to the point $(\tilde{x}_1, \tilde{x}_2)$. Let $K \subset J$ be the index set of the nonempty faces of T associated to points in IR. Hence, $\mathrm{IR} \subset \bigcup_{k \in K} T_k$.

In order to prove the other direction of the proof, we will show that for all $k \in K$, $T_k \subset \mathrm{IR}$.

According with the selection of faces T_k, $k \in K$, there is $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR} \cap \mathrm{ri}\ T_k$. Since $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{IR} = \mathrm{IR}_T \cap \mathrm{IR}_{\widehat{S}}$, by applying (7b) and previous remarks, we conclude that there exists $i \in I$ so that $(\tilde{x}_1, \tilde{x}_2) \in \mathrm{ri}\ \widehat{S}_i$.

Bearing in mind that $\widehat{S}$ is a convex set, $\widehat{S}_i$ is a face of $\widehat{S}$ and T_k is a convex set in $\widehat{S}$ such that ri T_k meets $\widehat{S}_i$, we conclude that the face $T_k \subset \widehat{S}_i$ (Theorem 18.1 in [16]). Taking into account that T_k is a face in IR_T and $\widehat{S}_i$ is a face in $\mathrm{IR}_{\widehat{S}}$, it directly follows that $T_k \subseteq \mathrm{IR}_T \cap \mathrm{IR}_{\widehat{S}} = \mathrm{IR}$. □

Remark 1 It is worth pointing out that, unlike quasiconcave problems without coupling constraints, when such constraints exist IR is not necessarily connected. The Example 3 shows this fact. The feasible region IR is formed by the segments AB and DE (see Fig. 3) and so is a non-connected union of faces of T. Bilevel problems are very sensitive to the existence of upper level constraints including upper level and lower level variables. Mersha and Dempe [13] and Audet et al. [1] have investigated the consequences of shifting upper level constraints to the lower level for linear bilevel problems. Calvete and Galé [4] have done the same for linear bilevel problems with several followers.

Let us now concentrate on optimality. For one level problems, $\min_{x \in T} f(x)$ where f is continuous and T is a compact polyhedron, it is well-known that f attains its minimum at an extreme point of T and in all its polyhedra subsets if and only if f is quasiconcave on T [12]. In fact, this is an alternative definition of quasiconcavity. The following two theorems extend this characterization to bilevel problems.

Theorem 3 *If IR is nonempty and* (A1) *and* (A2) *are verified, then there is an extreme point of T which is an optimal solution of the bilevel problem* (1).

Proof Taking into account Theorem 2, problem (1) can be reformulated as:

$$\begin{aligned} \min_{x_1,x_2}\ & f_1(x_1,x_2) \\ \text{s.t.}\ & (x_1,x_2) \in \bigcup_{k \in K} T_k. \end{aligned} \tag{9}$$

By Theorem 2, a minimizing solution $(\tilde{x}_1, \tilde{x}_2)$ exists. If it is an extreme point of T, the proof finishes. Otherwise, $k \in K$ exists so that $(\tilde{x}_1, \tilde{x}_2)$ is an optimal solution to the problem:

$$\begin{aligned} \min_{x_1,x_2}\ & f_1(x_1,x_2) \\ \text{s.t.}\ & (x_1,x_2) \in T_k. \end{aligned} \tag{10}$$

Since f_1 is quasiconcave and T_k is a nonempty compact polyhedron, there is an extreme point of T_k (therefore an extreme point of T) which solves problem (10). Since this extreme point gives the same value of f_1 as $(\tilde{x}_1, \tilde{x}_2)$, this extreme point of T solves problem (1). □

Theorem 4 *Let $C \subseteq \mathbb{R}^{n_1+n_2}$ be a convex set. Let f_1 and f_2 be continuous functions on C. If for each nonempty polyhedra $R \subseteq C$ and $S \subseteq C$, with $S(x_1)$ and $R \cap S$ bounded, assumption* (A1) *holds and the bilevel problem* (1) *attains its optimal value at an extreme point of $R \cap S$, then f_1 is quasiconcave on C and $f_2(x_1, \cdot)$ is quasiconcave on $C(x_1) = \{x_2 : (x_1, x_2) \in C\}$, $x_1 \in C_1$ the projection of C onto $\mathbb{R}^{n_1}$.*

Proof Let (x_1^1, x_2^1), $(x_1^2, x_2^2) \in C$ so that $f_1(x_1^1, x_2^1) \le f_1(x_1^2, x_2^2)$. Let

$$R = S = \left\{(x_1, x_2) : (x_1, x_2) = \lambda \left(x_1^1, x_2^1\right) + (1 - \lambda) \left(x_1^2, x_2^2\right), \lambda \in [0, 1]\right\}$$

Notice that, $R, S \subseteq C$ are nonempty compact polyhedra whose extreme points are (x_1^1, x_2^1) and (x_1^2, x_2^2). By hypothesis, the associated bilevel problem (1) attains its optimal value at an extreme point of $R \cap S = R = S$. Hence

$$f_1 \left(x_1^1, x_2^1\right) \le f_1 \left[\lambda \left(x_1^1, x_2^1\right) + (1 - \lambda) \left(x_1^2, x_2^2\right)\right], \quad \lambda \in [0, 1]$$

and f_1 is quasiconcave on C.

To continue, let $\tilde{x}_1 \in C_1$ and x_2^1, $x_2^2 \in C(\tilde{x}_1)$ so that $f_2(\tilde{x}_1, x_2^1) \le f_2(\tilde{x}_1, x_2^2)$. Let

$$R = S = \left\{(x_1, x_2) : (x_1, x_2) = \lambda \left(\tilde{x}_1, x_2^1\right) + (1 - \lambda) \left(\tilde{x}_1, x_2^2\right), \lambda \in [0, 1]\right\}$$

As before, we conclude that either $(\tilde{x}_1, x_2^1)$ or $(\tilde{x}_1, x_2^2)$ is an optimal solution of the corresponding bilevel problem. If $(\tilde{x}_1, x_2^1)$ is the optimal solution, then, in particular, it is a bilevel feasible solution. Hence, taking into account the lower level problem associated with $x_1 = \tilde{x}_1$, $f_2(\tilde{x}_1, x_2^1) \le f_2(\tilde{x}_1, y_2)$, $\forall y_2 \in S(\tilde{x}_1)$. Therefore

$$f_2 \left(\tilde{x}_1, x_2^1\right) \le f_2 \left[\lambda \left(\tilde{x}_1, x_2^1\right) + (1 - \lambda) \left(\tilde{x}_1, x_2^2\right)\right], \quad \lambda \in [0, 1].$$

Similarly, if $(\tilde{x}_1, x_2^2)$ is the optimal solution, then $f_2(\tilde{x}_1, x_2^2) \le f_2(\tilde{x}_1, y_2)$, $\forall y_2 \in S(\tilde{x}_1)$. Since $x_2^1 \in S(\tilde{x}_1)$ then

$$f_2 \left(\tilde{x}_1, x_2^1\right) = f_2 \left(\tilde{x}_1, x_2^2\right) \le f_2 \left[\lambda \left(\tilde{x}_1, x_2^1\right) + (1 - \lambda) \left(\tilde{x}_1, x_2^2\right)\right], \quad \lambda \in [0, 1].$$

Therefore, in both cases the quasiconcavity of $f_2(x_1, \cdot)$ on $C(x_1)$, $x_1 \in C_1$ follows. □

Remark 2 It is worth mentioning that if the assumption (A1) is omitted these theorems are no longer valid. Example 2 shows that (A1) is necessary.

4 Bilevel problems with quadratic lower level problems

The focus is now on a slightly different problem where it is not possible to prove that an optimal solution can be found at an extreme point of the set T but at an extreme point of a related single-level problem. In this Section assumption (A1) is omitted and the optimistic approach of the bilevel programming problem is taken.

We consider the quasiconcave quadratic bilevel (QQB) problem in which f_1 is quasiconcave, f_2 is convex quadratic and R and S are polyhedra. It can be formulated as:

$$\min_{x_1, x_2} \quad f_1(x_1, x_2), \tag{11a}$$

$$\text{s.t.} \quad (x_1, x_2) \in R \tag{11b}$$

where x_2 solves

$$\min_{x_2} \quad \frac{1}{2} x_2^\top Q x_2 + c(x_1)^\top x_2 \tag{11c}$$
$$\text{s.t.} \quad Ax_2 \leq b - Bx_1 \tag{11d}$$

where Q is an $n_2 \times n_2$ matrix which is assumed to be positive semidefinite, $c(x_1) : \mathbb{R}^{n_1} \longrightarrow \mathbb{R}^{n_2}$ is a linear function, A is a $p \times n_2$-matrix, B is a $p \times n_1$-matrix and b is a p-dimensional column vector. Here clearly we have $S = \{(x_1, x_2) \in \mathbb{R}^n : Bx_1 + Ax_2 \leq b\}$. As before, $S(x_1)$ and $R \cap S$ are assumed to be bounded in order to guarantee the existence of an optimal solution in the lower level and upper level respectively.

By using the (in this case necessary and sufficient) Karush-Kuhn-Tucker optimality conditions applied to the lower level problem (11c)–(11d), we can replace problem (11) by:

$$\begin{aligned} \min_{x_1,x_2,u} \quad & f(x_1, x_2) \\ & (x_1, x_2) \in R \\ & Qx_2 + c(x_1) + A^\top u = 0 \\ & u \geq 0 \\ & Ax_2 \leq b - Bx_1 \\ & u^\top (Ax_2 + Bx_1 - b) = 0 \end{aligned} \tag{12}$$

Both problems (11) and (12) are equivalent if global optimal solutions of the bilevel programming problem are considered [10].

Let $\Lambda(x_1, x_2)$ denote the set of Lagrange multipliers related to problem (11):

$$\Lambda(x_1, x_2) := \{u \geq 0 : Qx_2 + c(x_1) + A^\top u = 0,\ u^\top (Ax_2 + Bx_1 - b) = 0\}.$$

This set is a convex polyhedron [11] having a finite number of vertices. The following theorem characterizes local optimal solutions to QQB problems in terms of local optimal solutions of problem (12).

Theorem 5 *The point $(\widetilde{x}_1, \widetilde{x}_2) \in T$ is a local optimal solution of problem* (11) *if and only if $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{u})$ is a local optimal solution of problem* (12) *for each vertex $\widetilde{u} \in \Lambda(\widetilde{x}_1, \widetilde{x}_2)$.*

Proof If $(\widetilde{x}_1, \widetilde{x}_2) \in T$ is a local optimal solution of problem (11) the result is obvious.

Assume now that the point $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{u})$ is a local optimal solution of problem (12) for each vertex $\widetilde{u} \in \Lambda(\widetilde{x}_1, \widetilde{x}_2)$ and that, arguing by contradiction, $(\widetilde{x}_1, \widetilde{x}_2)$ is not a local optimal solution of problem (11). Then, there exists a sequence (x_1^k, x_2^k) of feasible points to the bilevel programming problem converging to $(\widetilde{x}_1, \widetilde{x}_2)$ with $f_1(x_1^k, x_2^k) < f_1(\widetilde{x}_1, \widetilde{x}_2)$ for all k.

Then, since the Karush-Kuhn-Tucker conditions are necessary and sufficient optimality conditions for problem (11c)–(11d), there exists a sequence of vertices $u^k \in \Lambda(x_1^k, x_2^k)$ such that the triple (x_1^k, x_2^k, u^k) is feasible to problem (12) with a smaller objective function value than $(\widetilde{x}_1, \widetilde{x}_2)$ for each k. Hence, each accumulation point $(\widehat{x}_1, \widehat{x}_2, \widehat{u})$ of the sequence (x_1^k, x_2^k, u^k) cannot be local optimal for this problem. The existence of accumulation points is a result of finiteness of the number of variables and constraints: for each vertex u^k of the set $\Lambda(x_1^k, x_2^k)$ there exists a subset $I^k \subseteq \{1, \ldots, p\}$ such that the system

$$\begin{aligned} A^\top u &= -Qx_2^k - c(x_1^k) \\ u_i &= 0, \quad i \in I^k \end{aligned} \tag{13}$$

has the unique solution u^k. Since $\{1, \ldots, p\}$ is finite we can consider the finite family of all the sets I^k which appear infinitely often in the family of all the sets $\{I^k\}_k$ corresponding to the sequence $\{x_1^k, x_2^k, u^k\}_k$. For each of these sets, the coefficient matrix in system (13) is a regular quadratic matrix having an inverse matrix which is independent of k. Hence, for $k \to \infty$, all accumulation points of the sequence $\{u^k\}_k$ are solutions $\widehat{u}$ of the systems (13) for the different selections of I^k and (x_1^k, x_2^k) being replaced with $(\widetilde{x}_1, \widetilde{x}_2)$. Since it is easy to see that $\widehat{u}$ is a vertex of $\Lambda(\widetilde{x}_1, \widetilde{x}_2)$, a contradiction to our assumption is revealed and the proof is concluded. □

To proceed, let us consider the complementarity condition in problem (12) and note that, for each feasible solution (x_1, x_2, u) of this problem there exist sets $I, J \subseteq \{1, \ldots, p\}$ with

$$\begin{aligned} u_i \geq 0,\ (Ax_2 + Bx_1 - b)_i = 0, \quad i \in I,\\ u_i = 0,\ (Ax_2 + Bx_1 - b)_i \leq 0, \quad i \in J. \end{aligned}$$

Then, problem (12) can be replaced by a patchwork of problems (P_{IJ}) for each possible selection of the sets $I, J, I \cap J = \emptyset, I \cup J = \{1, \ldots, p\}$:

$$\begin{aligned} \min_{x_1, x_2, u}\ & f(x_1, x_2)\\ & (x_1, x_2) \in R\\ & Qx_2 + c(x_1) + A^\top u = 0\\ & u \geq 0\\ & Ax_2 \leq b - Bx_1\\ & u_i \geq 0,\ (Ax_2 + Bx_1 - b)_i = 0, \quad i \in I\\ & u_i = 0,\ (Ax_2 + Bx_1 - b)_i \leq 0, \quad i \in J\ . \end{aligned} \tag{P_{IJ}}$$

Computing the best optimal solution among all these problems (P_{IJ}) yields the optimal solution of (12). Note that each of the problems (P_{IJ}) has a polyhedral feasible set and that, due to quasiconcavity of the function f_1, an optimal solution of each of the problems (P_{IJ}) can be found at an extreme point of this feasible set. Hence, local optimal solutions of the problem (12) are located at extreme points of the union of the finite number of polyhedral feasible sets of the problems (P_{IJ}) and, hence, at extreme points of the problem (12), too, even if the complementary slackness condition is dropped. This allows us to conclude the following corollary:

Corollary 1 *Considering problem* (12)*, an optimal solution* $(\overline{x}_1, \overline{x}_2)$ *can be found at a vertex* $(\overline{x}_1, \overline{x}_2, \overline{u})$ *of the set*

$$\begin{aligned} & (x_1, x_2) \in R\\ & Qx_2 + c(x_1) + A^\top u = 0\\ & u \geq 0\\ & Ax_2 \leq b - Bx_1\ . \end{aligned} \tag{14}$$

As a consequence of this result, it is possible to develop enumerative algorithms considering vertices of (14) to solve problem (11) globally.

The next example in $\mathbb{R}^2$ illustrates that the bilevel optimal solution of a QQB problem is a vertex of an ‘enlarged’ polyhedron and not a vertex of $R \cap S$.

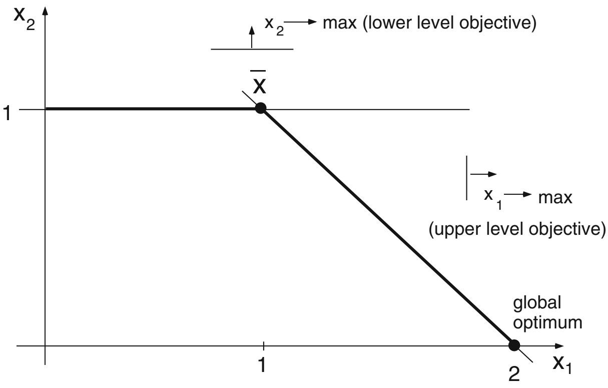

Fig. 4 Lower level problem of Example 5

Example 4 Let us consider the problem

$$\min_{x_1,x_2} -(x_1^2+x_2^2)$$
$$1 \le x_1 \le 2,$$

where x_2 solves

$$\begin{aligned}\min_{x_2}\ & x_2^2 - 2x_2\\ \text{s.t.}\ & -x_1 + x_2 \le 2\\ & -x_1 - x_2 \le 1\ .\end{aligned}$$

For each parameter x_1 the solution of the lower level problem is $x_2 = 1$. Hence, the global optimal solution of the bilevel problem is $(x_1, x_2) = (2, 1)$, which is not a vertex of the polyhedron T defined by all the constraints.

Considering the feasible sets of the problems (P_{IJ}), it is easy to see that there are only feasible solutions if $u_1 = u_2 = 0$. Then, $x_2 = 1$ and x_1 can be arbitrarily chosen in $[1, 2]$. The optimal solution of this subproblem is $(x_1, x_2, u_1, u_2) = (2, 1, 0, 0)$, which is a vertex of the 'enlarged' polyhedron:

$$\{(x_1, x_2, u_1, u_2) : 1 \le x_1 \le 2,\ 2x_2 - 2 + u_1 - u_2 = 0,$$
$$u_1 \ge 0,\ u_2 \ge 0,\ -x_1 + x_2 \le 2,\ -x_1 - x_2 \le 1\}\ .$$

In [10] it is shown that in Theorem 5 it is really necessary to verify local optimality of $(\widetilde{x}_1, \widetilde{x}_2, \widetilde{u})$ to problem (12) for all vertices of $\Lambda(\widetilde{x}_1, \widetilde{x}_2)$. The following example in $\mathbb{R}^2$, displayed in Fig. 4, shows this fact.

Example 5

$$\min_{x_1,x_2} x_1$$
$$0 \le x_1 \le 10,$$

where x_2 solves

$$\begin{aligned}\min_{x_2}\ & x_2\\ \text{s.t.}\ & x_1 + x_2 \le 2\\ & 0 \le x_2 \le 1\end{aligned}$$

Consider the point $(\overline{x}_1, \overline{x}_2) = (1, 1)$. Then, there are two active constraints in the lower level problem with only one variable. The multiplier set is $\Lambda(\overline{x}_1, \overline{x}_2) = \{u = (u_1, u_2, u_3) : u_1, u_2 \geq 0, u_1 + u_2 = 1, u_3 = 0\}$. Take $\overline{u} = (1, 0, 0)$. Then, in an open neighborhood of the point $(\overline{x}_1, \overline{x}_2, \overline{u}_1, \overline{u}_2, \overline{u}_3) = (1, 1, 1, 0, 0)$ we have $u_1 > 0$ and hence, in problem (12), by complementarity slackness, $x_2 = 1$. Solving problem (12) under these conditions we obtain the optimal solution $(\overline{x}_1, \overline{x}_2, \overline{u}_1, \overline{u}_2, \overline{u}_3) = (1, 1, 1, 0, 0)$. On the contrary, if we take $\overline{u} = (0, 1, 0)$, in an open neighborhood of the point $(\overline{x}_1, \overline{x}_2, \overline{u}_1, \overline{u}_2, \overline{u}_3) = (1, 1, 0, 1, 0)$ we have $u_2 > 0$ and hence, in (10), $x_1 = 2 - x_2$. Therefore, in this open neighborhood, we can choose points $(x_1, x_2, u_1, u_2, u_3)$ with better values of f_1 than 1.

5 Conclusions

In this paper we have analyzed bilevel problems over polyhedra with extreme point optimal solutions. We have proved that the fact that both objective functions are quasiconcave characterizes the property of the existence of an extreme point of the polyhedron defined by the whole set of constraints which solves the bilevel problem. This property allows us to consider enumerative methods which search amongst extreme points to solve the problem. In order to obtain this property it is necessary to assume that the optimal solution of the lower level problem is a singleton. Quasiconcave functions include as important particular cases linear, linear fractional or linear multiplicative functions.

If the upper level objective function is quasiconcave and the lower level one is convex quadratic, by replacing the lower level problem with the Karush-Kuhn-Tucker conditions we have proved that an optimal solution can be found at an extreme point of a related polyhedron which takes into account complementary constraints. In this case, no assumption is required for the set of optimal solutions of the lower level problem and the optimistic approach is taken. If the lower level function is an arbitrary convex parametric optimization problem, we can replace it with the Karush-Kuhn-Tucker conditions but, since these conditions are not polyhedral, we cannot conclude that optimal solutions can be found at the vertices.

Acknowledgments The authors gratefully acknowledge the anonymous referees for their valuable suggestions to improve the presentation of the paper.

References

1. Audet, C., Haddad, J., Savard, G.: A note on the definition of a linear bilevel programming solution. Appl. Math. Comput. **181**, 351–355 (2006)
2. Bard, J.: Practical Bilevel Optimization. Algorithms and applications. Kluwer Academic Publishers, Dordrecht (1998)
3. Calvete, H., Galé, C.: On the quasiconcave bilevel programming problem. J. Optim. Theory Appl. **98**(3), 613–622 (1998)
4. Calvete, H., Galé, C.: Linear bilevel multi-follower programming with independent followers. J. Global Optim. **39**(3), 409–417 (2007)
5. Calvete, H., Galé, C., Mateo, P.: A new approach for solving linear bilevel problems using genetic algorithms. To appear in Eur. J. Oper. Res. (2007). doi:10.1016/j.ejor.2007.03.034
6. Chinchuluun, A., Huang, H., Pardalos, P. : Multilevel (hierarchical) optimization: complexity issues, optimality conditions, algorithms. In: Gao, D., Sherali, H. (eds.) Advances in Applied Mathematics and Global Optimization, pp. 197–221. Springer, New York (2009)
7. Dantzig, G., Folkman, J., Shapiro, N.: On the continuity of the minimum set of a continuous function. J. Math. Anal. Appl. **17**, 519–548 (1967)

8. Dempe, S.: Foundations of Bilevel Programming. Kluwer Academic Publishers, Dordrecht Boston London (2002)
9. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**, 333–359 (2003)
10. Dempe, S., Dutta, J.: Is bilevel programming a special case of a mathematical program with complementary constraints? Technical report, Department of Mathematics and Computer Science, TU Bergakademie Freiberg (2008)
11. Gauvin, J.: A necessary and sufficient regularity condition to have multipliers in nonconvex programming. Math. Program. **12**, 136–139 (1977)
12. Martos, B.: The direct power of adjacent vertex programming problems. Manag. Sci. **12**(3), 241–252 (1965)
13. Mersha, A., Dempe, S.: Linear bilevel programming with upper level constraints depending on the lower level solution. Appl. Math. Comput. **180**, 247–254 (2006)
14. Migdalas, A., Pardalos, P.M. (eds): Hierarchical and bilevel programming. J. Global Optim. **8**(3), 209–215 (1996)
15. Migdalas, A., Pardalos, P., Värbrand, P. (eds.): Multilevel Optimization Algorithms and Applications. Kluwer Academic Publishers, Dordrecht (1998)
16. Rockafellar, R.: Convex Analysis. Princeton University Press, Princeton, New Jersey (1970)
17. Savard, G.: Contribution à la programmation mathématique à deux niveaux.Ph.D. thesis, Ecole Polytechnique de Montréal, Université de Montréal, Montréal, QC, Canada (1989)
18. Shimizu, K., Ishizuka, Y., Bard, J.: Nondifferentiable and Two-level Mathematical Programming. Kluwer Academic Publishers, Boston (1997)
19. Vicente, L., Calamai, P.: Bilevel and multilevel programming: a bibliography review. J. Global Optim. **5**, 291–306 (1994)
20. Xu, Z.: Deriving the properties of linear bilevel programming via a penalty function approach. J. Optim. Theory Appl. **103**(2), 441–456 (1999)