DEPENDENCE OF BILEVEL PROGRAMMING ON IRRELEVANT DATA

STEPHAN DEMPE, SEBASTIAN LOHSE

ABSTRACT. In 1997, Macal and Hurter [7] have found that adding a constraint to the lower level problem, which is not active at the computed global optimal solution, can destroy global optimality. In this paper this property is reconsidered and it is shown that this solution remains locally optimal under inner semicontinuity of the original solution set mapping. In the second part of the paper we prove that adding a variable in the linear lower level problem can also destroy global optimality. But here the solution remains locally optimal, provided the optimal solution in the lower level was dual non-degenerated.

1. Introduction

Bilevel programming problems are hierarchical optimization problems where the feasible set of the so-called upper level or leader's problem is restricted in part by the graph of the solution set mapping of a second optimization problem. This latter problem is the follower's or lower level problem.

To formulate the bilevel programming problem formally, consider the parametric follower's problem first:

(1.1)
$$\min_{y} \{ f(x,y) : g(x,y) \leq 0 \},$$
 where $f : \mathbb{R}^{n} \times \mathbb{R}^{m} \to \mathbb{R}, \ g : \mathbb{R}^{n} \times \mathbb{R}^{m} \to \mathbb{R}^{p}.$ Let
$$\Psi(x) := \text{Argmin } \{ f(x,y) : g(x,y) \leq 0 \}$$

be its solution set mapping. Then, the leader's problem is given as

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

where $\operatorname{gph} \Psi := \{(x,y) : y \in \Psi(x)\}$ denotes the graph of the solution set mapping of the problem (1.1), $X \subseteq \mathbb{R}^n$ is a closed set and $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$. This problem has been investigated in the monographs [1, 4]. It has many applications, [5] is an annotated bibliography on bilevel programming problems, [3] gives an overview of the problem.

Problem (1.2) is a \mathcal{NP} -hard problem [6], which makes its solution difficult, especially for large problems. Hence, reduction of the dimension of the problem is desirable. Two such reductions are used: either only some part of the constraints in the lower level problem are used and more constraints are added if necessary. This is done e.g. in cutting plane methods for solving discrete linear optimization problems. Or, variables can be dropped and added only if necessary. This is e.g. the column generation approach for solving large linear optimization problems. We will investigate the implications of such approaches to bilevel programming problems where the lower level problem is reduced.

Bilevel programming problems have some surprising properties. When solving integer optimization problems one often starts with solving its linear relaxation where the integer variables are replaced with continuous ones. Then, if the optimal solution is integer, too, i.e. if it satisfies all constraints of the original problem, it is the optimal solution of the integer problem. This is in general not true for integer bilevel programming problems [1, Section 6.3].

Adding constraints to the lower level which are not active at the computed global optimum of a bilevel programming problem will modify the feasible set of the problem. This can imply that global optimality at the computed solution is lost [7].

The outline of the paper is as follows. In Section 2 we reconsider the result in [7]. Here we will see that a global optimal solution of a bilevel programming problem need not to be a local optimal solution after adding an irrelevant constraint to the lower level problem. To guarantee that it remains locally optimal we need inner semicontinuity of the solution set mapping of the original lower level problem. In Section 3 we investigate the question if global optimality of a solution is maintained if variables are added to the lower level problem in case this is a linear optimization problem perturbed in the objective function, see Subsection 3.1, and in the right-hand side vector, see Subsection 3.2. We will see that this property also fails in general. But here, if the new variable is irrelevant, local optimality is maintained.

2. Irrelevant constraints

Definition 2.1. A point-to-set mapping $\Gamma: \mathbb{R}^n \to 2^{\mathbb{R}^m}$ is upper semicontinuous at $\overline{x} \in \mathbb{R}^n$, if for all open sets $U \supset \Gamma(\overline{x})$ there exists an open set W with $\overline{x} \in W$ such that $\Gamma(x) \subset U$ for all $x \in W$. It is inner semicontinuous at $(\overline{x}, \overline{y}) \in \operatorname{gph} \Gamma$, if for each sequence $\{x^k\}_{k=1}^{\infty}$ with $\Gamma(x^k) \neq \emptyset$ converging to \overline{x} , there is a sequence $\{y^k\}_{k=1}^{\infty}, y^k \in \Gamma(x^k)$ for all k, converging to \overline{y} .

Theorem 2.2. Let (x^0, y^0) be a global optimal solution of the problem (1.2). Let Ψ be inner semicontinuous at (x^0, y^0) . Then, (x^0, y^0) is a local optimal solution of the problem

(2.1)
$$\min_{x,y} \{ F(x,y) : x \in X, \ (x,y) \in \mathrm{gph} \, \Psi^1 \}$$

with

$$\Psi^{1}(x) := \underset{y}{\operatorname{Argmin}} \{ f(x,y) : g(x,y) \leq 0, \ h(x,y) \leq 0 \}$$

with $h: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ provided that $h(x^0, y^0) < 0$ and that the function $h: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is continuous.

Proof. Since $y \in \Psi(x)$ is an optimal solution of the lower level problem to (2.1) provided it is feasible for that problem, $y^0 \in \Psi^1(x^0)$. Hence, the point (x^0, y^0) is feasible for (2.1).

Assume that (x^0,y^0) is not a local optimum of problem (2.1). Then, there exists a sequence $\{(x^k,y^k)\}_{k=1}^\infty$ converging to (x^0,y^0) such that $x^k\in X,\,y^k\in \Psi^1(x^k)$ and $F(x^k,y^k)< F(x^0,y^0)$. Note that (x^k,y^k) is feasible for problem (1.2).

Since Ψ is inner semicontinuous at (x^0, y^0) there exists a sequence $\widehat{y}^k \in \Psi(x^k)$ converging to y^0 . By continuity of the function h, $h(x^k, \widehat{y}^k) < 0$ and $\widehat{y}^k \in \Psi^1(x^k)$.

Hence,
$$f(x^k, \widehat{y}^k) = f(x^k, y^k)$$
,

$$\begin{split} \Psi^1(x^k) = & \{ y : g(x,y) \le 0, \ h(x,y) \le 0, \ f(x^k,y) = f(x^k, \widehat{y}^k) \} \\ \subseteq & \{ y : g(x,y) \le 0, \ f(x^k,y) = f(x^k, \widehat{y}^k) \} = \Psi(x^k) \end{split}$$

and, hence,

$$\min_y \{F(x^k,y): y \in \Psi(x^k)\} \leq \min_y \{F(x^k,y): y \in \Psi^1(x^k)\} \leq F(x^k,y^k) < F(x^0,y^0)$$

for sufficiently large k. This contradicts global optimality of (x^0, y^0) .

The assumption of inner semicontinuity is restrictive. But if it is not satisfied, the assertion of the Theorem is in general not correct. This can be seen in the following example.

Example 2.3. Consider the lower level problem

$$\min_{y} \{ -y_2 : y_2 \le \cos y_1 - 1 + xy_1, -3\pi \le y_1 \le 3\pi \}.$$

For x=0 this problem has three global minima at $(y_1,y_2), y_2=0, y_1 \in \{0,2\pi,-2\pi\}$. For $x \neq 0$ the problem has only one global optimum at $y_1>0$ for x>0 and $y_1<0$ if x<0. Let the upper level objective be $F(x,y)=(y_1-\pi)^2+(y_2+5)^2$. The global optima of problem (1.2) are then $(x^0;y^0)=(0;(0,0))$ and $(x^1;y^1)=(0;(2\pi,0))$. Take $(x,y)=(x^1,y^1)$ and add the constraint $y_1\geq \pi$ to the lower level problem. Then, for x<0 sufficiently close to zero, the lower level problem has a unique optimal solution $(\widehat{x};\widehat{y})$ with $F(\widehat{x},\widehat{y})< F(x^1,y^1)$ converging to $F(x^1,y^1)$ for \widehat{x} converging to x^1 . Hence, (x^1,y^1) is not a local optimum of (2.1).

3. Irrelevant variables

3.1. Parameter in the objective function of the lower level problem. Consider now the bilevel programming problem

(3.1)
$$\min_{x,y} \{ F(x,y) : x \in X, \ (x,y) \in gph \, \Psi_L \},$$

where $X \subseteq \mathbb{R}^n$ is a closed set, $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and the linear lower level problem parameterized in the objective function

(3.2)
$$\Psi_L(x) := \underset{y}{\operatorname{Argmin}} \{ x^\top y : Ay = b, \ y \ge 0 \}.$$

Let $(\overline{x}, \overline{y})$ be a global optimal solution of this problem. Now, add one new variable y_{n+1} to the lower level problem with objective function coefficient x_{n+1} and a new column A_{n+1} in the coefficient matrix of the lower level problem, i.e. replace the lower level problem with

(3.3)
$$\Psi_{NL}(x) := \underset{y}{\operatorname{Argmin}} \{x^{\top}y + x_{n+1}y_{n+1} : Ay + A_{n+1}y_{n+1} = b, \ y, y_{n+1} \ge 0\}$$

and consider the problem

$$(3.4) \quad \min_{x,y} \{ \widetilde{F}(x, x_{n+1}, y, y_{n+1}) : (x, x_{n+1}) \in \widetilde{X}, \ (x, x_{n+1}, y, y_{n+1}) \in gph \ \Psi_{NL} \}.$$

In the next example it is shown that, adding one new variable, global optimality of a solution can be destroyed, but strict local optimality can be maintained.

Example 3.1. Consider the following bilevel programming problem with the lower level problem

(3.5)
$$\Psi_L(x) := \underset{y}{\operatorname{Argmin}} \{x_1y_1 + x_2y_2 : y_1 + y_2 \le 2, -y_1 + y_2 \le 0, y \ge 0\}$$

and upper level problem

$$(3.6) \qquad \min\{(x_1 - 0.5)^2 + (x_2 - 0.5)^2 - 3y_1 - 3y_2 : (x, y) \in \operatorname{gph} \Psi_L\}.$$

Then, the unique global optimum is $\overline{x} = (0.5; 0.5)$, $\overline{y} = (1; 1)$ with optimal objective function value -6. Now, adding one variable to the lower level problem (3.7)

$$\Psi_{NL}(x) := \underset{y}{\operatorname{Argmin}} \{x_1y_1 + x_2y_2 + x_3y_3 : y_1 + y_2 + y_3 \le 2, -y_1 + y_2 \le 0, y \ge 0\}$$

and investigating the bilevel programming problem

(3.8)
$$\min\{(x_1 - 0.5)^2 + (x_2 - 0.5)^2 + x_3^2 - 3y_1 - 3y_2 - 6y_3 : (x, y) \in \operatorname{gph} \Psi_{NL}\}$$

the point $x=(0.5;0.5;0.5),\ y=(0;0;2)$ has objective function value -11.75. Hence, global optimality of $(\overline{x},\overline{y})$ is destroyed. But, the point $((\overline{x},0),(\overline{y},0))$ remains feasible and it is a strict local minimum.

Theorem 3.2. Consider problems (3.1) and (3.4). Let $(\overline{x}, \overline{y})$ be a global optimal solution of problem (3.1). Then, if $(\widetilde{x}, \widetilde{x}_{n+1})$ is a local minimum of the problem

$$\min_{x} \{ \widetilde{F}(x, x_{n+1}, \overline{y}, 0) : (x, x_{n+1}) \in \widetilde{X} \},$$

and $(\overline{y},0)$ is a unique optimal solution of (3.3) for $(x,x_{n+1})=(\widetilde{x},\widetilde{x}_{n+1})$ then $(\widetilde{x},\widetilde{x}_{n+1},\overline{y},0)$ is a local minimum of (3.4).

Proof. By the assumptions and parametric linear programming there is an open neighborhood V of $(\widetilde{x}, \widetilde{x}_{n+1})$ such that $\{(\overline{y}, 0)\} = \Psi_{NL}(x)$ for all $(x, x_{n+1}) \in V$. Since $(\widetilde{x}, \widetilde{x}_{n+1})$ is a local minimum of $\widetilde{F}(x, x_{n+1}, \overline{y}, 0)$ there is an open neighborhood V_1 of $(\widetilde{x}, \widetilde{x}_{n+1})$ with $\widetilde{F}(x, x_{n+1}, \overline{y}, 0) \geq \widetilde{F}(\widetilde{x}, \widetilde{x}_{n+1}, \overline{y}, 0)$ for all $(x, x_{n+1}) \in V_1$. Moreover,

$$\{(x, x_{n+1}, y, y_{n+1}) : (x, x_{n+1}) \in \widetilde{X} \cap V \cap V_1, (y, y_{n+1}) \in \Psi_{NL}(x, x_{n+1})\}$$

= $(\widetilde{X} \cap V \cap V_1) \times \{(\overline{y}, 0)\}.$

This implies the proof.

Corollary 3.3. Under the assumptions of Theorem 3.2, if $(\overline{x}, \overline{y})$ is a global optimal solution of problem (3.1), $\{(\overline{y}, 0)\} = \Psi_{NL}(\overline{x}, 0)$ and

$$(\overline{x},0) \in \underset{(x,x_{n+1})}{\operatorname{Argmin}} \{\overline{F}(x,x_{n+1},\overline{y},0)\},\$$

then, $(\overline{x}, 0, \overline{y}, 0)$ is a local optimal solution of problem (3.4).

The most restrictive assumption in this Corollary is the uniqueness of the optimal solution of the lower level problem. We will drop the uniqueness assumption restricting to a subclass of bilevel programming problems.

If the objective function $F(\cdot, \cdot)$ in the upper level problem is a linear or concave one and the set X is a polyhedron, then at least one (global) optimal solution of the problem (3.1) can be found at a vertex of the set [1, 2]

$$\{(x,y): x \in X, Ay = b, y > 0\}.$$

Then, there exists a basic matrix B of the matrix A such that

$$(3.9) y = (y_B, y_N), x = (x_B, x_N), y_B = B^{-1}b, y_N = 0, x_B^{\top}B^{-1}A - x^{\top} \le 0.$$

Note that the basic matrix does not need to be unique. Call a basic matrix satisfying the conditions (3.9) a basic matrix for y and x.

Theorem 3.4. Let (x^0, y^0) be a global optimal solution for problem (3.1) and assume that the functions F, \widetilde{F} are concave, X, \widetilde{X} are polyhedra. Let

(3.10)
$$x^{0\top}B^{-1}A_{n+1} < 0$$
 for each basic matrix for y^0 and x

and $(x^0,0)$ be a local minimum of the problem

$$\min\{F((x,x_{n+1}),(y,0)):(x,x_{n+1})\in\widetilde{X},\ y\in\Psi_L(x^0)\}.$$

Then, the point $((x^0,0),(y^0,0))$ is a local optimal solution of problem (3.4).

Proof. Assume that $((x^0,0),(y^0,0))$ is not a local optimum. Then, there exists a sequence $((x^k,x_{n+1}^k),(y^k,y_{n+1}^k))$ converging to $((x^0,0),(y^0,0))$ with

$$F((x^k, x_{n+1}^k), (y^k, y_{n+1}^k)) < F((x^0, 0), (y^0, 0)) \text{ for all } k.$$

Since $((x^k, x_{n+1}^k), (y^k, y_{n+1}^k))$ is feasible for (3.4) and gph Ψ_{NL} equals the union of faces of the set [4]

$$\{(x,y): x \in \widetilde{X}, Ay + A_{n+1}y_{n+1} = b, y, y_{n+1} \ge 0\},\$$

then, since $((x^k, x_{n+1}^k), (y^k, y_{n+1}^k))$ converges to $((x^0, 0), (y^0, 0))$ there exists, without loss of generality, one facet M of this set with $((x^k, x_{n+1}^k), (y^k, y_{n+1}^k)) \in M$ for all k. Moreover, by upper semicontinuity of $\Psi_{NL}(\cdot)$, $(y^0, 0) \in M$. By [8] there exists $c \in \mathbb{R}^{n+1}$ such that M equals the set of optimal solutions of the problem

$$\min\{c^{\top}(y,y_{n+1})^{\top}:Ay+A_{n+1}y_{n+1}=b,\ y,y_{n+1}\geq 0\}.$$

Since $(y^0,0) \in M$ there exists a basic matrix for $(y^0,0)$ and c. Then, the assumptions of the theorem imply that $(x^0,0) \neq c$ if x_{n+1} is a basic variable in (y^k,y^k_{n+1}) (since this implies that $c_B^\top B^{-1}A_{n+1}-c_{n+1}=0$ by linear optimization). This implies that there is an open neighborhood V of $(x^0,0)$ such that $\Psi_{NL}(x,x_{n+1})\subseteq\{(y,y_{n+1}):y_{n+1}=0\}.$

Hence, $y_{n+1}^k = 0$ for sufficiently large k.

By parametric linear programming, $\Psi_L(x) \subseteq \Psi_L(x^0)$ for x sufficiently close to x^0 . Hence, the assertion follows.

Corollary 3.5. If the upper level objective function is separable, i.e. if

$$F((x, x_{n+1}), (y, y_{n+1})) = F_1(x, x_{n+1}) + F_2(y, y_{n+1}),$$

then (3.10) together with the assumption that $(x^0,0)$ is a local minimum of

$$\min\{F_1(x, x_{n+1}) : (x, x_{n+1}) \in \widetilde{X}\}\$$

will guarantee local minimality of $((x^0, 0), (y^0, 0))$ for (3.4).

3.2. Parameter in the right-hand side of the lower level problem. We consider now the linear bilevel program

(3.11)
$$\min_{x,y} \{ x^{\top} d^1 + y^{\top} d^2 : x \in X, y \in \Psi_L(x) \},$$

where X is a polyhedron and the lower level is linear with right hand side parameter x

(3.12)
$$\Psi_L(x) = \underset{y}{\operatorname{Argmin}} \{ c^\top y : A_1 y = b - A_2 x, y \ge 0 \}.$$

Like in the last chapter, we are interested in the behavior if we add a new variable x_{n+1} . At first, it could happen that our problem becomes infeasible, because of different reasons.

Example 3.6. Consider the lower level problem

(3.13)
$$\Psi_L(x) = \underset{y}{\operatorname{Argmin}} \{y_1 - y_2 : y_1 + 4y_3 = 1 + x, y_2 + 5y_3 = 7 - 2x, y \ge 0\}$$

and

(3.14)
$$\min\{-y_1 - y_2 - 2x : x \in [1, 2], y \in \Psi_L(x)\}\$$

as upper level. With $\Psi_L(x) = \{(1+x, 7-2x, 0)\}$ for all $x \in [1, 2]$, we get the global solution $\overline{x} = 2$, $\overline{y} = (3, 3, 0)$. If we add a variable y_4 with a column (-1, -1) in the coefficient matrix, lower level cost coefficient c_4 and upper level cost coefficient $d_4^2 = 0$, we get the problem

$$(3.15) \quad \min\{-y_1 - y_2 - 2x : x \in [1, 2], y \in \Psi_{NL}(x)\}\$$

with

(3.16)

$$\Psi_{NL}(x) = \underset{y}{\operatorname{Argmin}} \left\{ y_1 - y_2 + c_4 y_4 : y_1 + 4y_3 - y_4 = 1 + x, \ y_2 + 5y_3 - y_4 = 7 - 2x, \ y \ge 0 \right\}.$$

If we choose $c_4 = -1$ our problem becomes infeasible, because the lower level objective function is unbounded over the feasible set for each parameter $x \in [1, 2]$. If we choose on the other hand $c_4 = 0$, we get

$$\Psi_{NL}(x) = \{(1+x+y_4, 7-2x+y_4, 0, y_4): y_4 > 0\}$$

and therefore the upper level objective function is unbounded over $\Psi_{NL}(x)$ for each parameter $x \in [1, 2]$.

Remark 3.7. A similar example can also be constructed to verify the results in Subsection 3.1.

In Theorem 3.2 it was shown, that the uniqueness of the lower level solution is needed to guarantee at least local optimality of a former global solution. Next theorem states the same fact, but we need an additional assumption.

Theorem 3.8. Consider problem (3.11) and let $(\overline{x}, \overline{y})$ a global optimal solution. The lower level problem is now replaced by

$$\Psi_{NL}(x) := \underset{y}{\operatorname{Argmin}} \{c^{\top}y + c_{n+1}y_{n+1} : A_1y + A_{n+1}y_{n+1} = b - A_2x, y, y_{n+1} \ge 0\}$$

and assume that $(\overline{y},0)$ is an optimal solution of (3.17) for \overline{x} . Furthermore, suppose that all optimal basic solutions of (3.17) for \overline{x} are non-degenerate and the reduced cost coefficient of y_{n+1} is negative. Then, $(\overline{x},\overline{y},0)$ is a local optimal solution of

(3.18)
$$\min_{x,y} \{ x^{\top} d^1 + y^{\top} d^2 + d_{n+1}^2 y_{n+1} : x \in X, (y, y_{n+1}) \in \Psi_{NL}(x) \}.$$

Proof. Under the assumptions of the theorem, we have $y_{n+1} = 0$ for all $(y, y_{n+1}) \in \Psi_{NL}(x)$ where x is in a small neighborhood of \overline{x} .

Example 3.9. This example will show, that we need both non-degeneracy and uniqueness of the lower level solution. Let us consider Example 3.6 again. We introduce a new variable y_4 and get

(3.19)

$$\Psi_{NL}(x) = \underset{y}{\text{Argmin}} \{y_1 - y_2 : y_1 + 4y_3 + y_4 = 1 + x, y_2 + 5y_3 + y_4 = 7 - 2x, y \ge 0\}$$

as lower level and

$$(3.20) \quad \min\{-y_1 - y_2 - 2x - 3y_4 : x \in [1, 2], y \in \Psi_{NL}(x)\}\$$

as upper level. Since we have

$$\Psi_{NL}(x) = \text{conv}\{(1+x, 7-2x, 0, 0), (0, 6-3x, 0, 1+x)\},\$$

 $(\overline{x}, \overline{y}, 0) = (2, 3, 3, 0, 0)$ is not a local solution of (3.22).

Now, replace in (3.13) the upper level constraint $x \in [1, 2]$ by $x \in [1, 4]$. Then, we get the new global solution $(\overline{x}, \overline{y}) = (\frac{7}{2}, \frac{9}{2}, 0, 0)$. After introducing an additional variable

(3.21)

$$\Psi_{NL}(x) = \underset{y}{\operatorname{Argmin}} \{ y_1 - y_2 + y_4 : y_1 + 4y_3 - y_4 = 1 + x, \ y_2 + 5y_3 - y_4 = 7 - 2x, \ y \ge 0 \}$$

and considering

$$(3.22) \quad \min\{-y_1 - y_2 - 3y_4 - 2x : x \in [1, 4], y \in \Psi_{NL}(x)\}\$$

as new upper level, we see that the optimal solution of the lower level is given by (3x - 6, 0, 0, 2x - 7) for $x > \frac{7}{2}$. Hence, $(\overline{x}, \overline{y}, 0)$ is not a local optimal solution of (3.22).

References

- J.F. Bard, Practical bilevel optimization: Algorithms and applications, Kluwer Academie Publishers, Dordrecht, 1998.
- H.I. Calvete and C. Gale, On the quasiconcave bilevel programming problem, 98 (1998), 613–622.
- 3. B. Colson, P. Marcotte, and G. Savard, A trust-region method for nonlinear bilevel programming: algorithm and computational experience, Computational Optimization and Applications 30 (2005), 211–227.
- 4. S. Dempe, Foundations of bilevel programming, Kluwer Academie Publishers, Dordrecht, 2002.
- Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints, Optimization 52 (2003), 333–359.
- X. Deng, Complexity issues in bilevel linear programming, Multilevel Optimization: Algorithms and Applications (A. Migdalas, P.M. Pardalos, and P. Värbrand, eds.), Kluwer Academic Publishers, Dordrecht, 1998, pp. 149–164.
- C.M. Macal and A.P. Hurter, Dependence of bilevel mathematical programs on irrelevant constraints, 24 (1997), 1129–1140.
- Alexander Schrijver, Theory of linear and integer programming. Repr., Chichester: Wiley, 1998.

Department of Mathematics and Computer Science, TU Bergakademie Freiberg $E\text{-}mail\ address:}$ dempe@tu-freiberg.de