J Appl Math Comput (2013) 42:41–49
DOI 10.1007/s12190-012-0620-6

JAMC

APPLIED MATHEMATICS

An exact penalty method for weak linear bilevel programming problem

Yue Zheng · Zhongping Wan · Kangtai Sun · Tao Zhang

Received: 21 April 2012 / Published online: 15 November 2012

Abstract In this paper, we present an exact penalty method, which is different from the existing penalty method, for solving weak linear bilevel programming problem. Then, we establish an existence result of solutions for such a problem. Finally, we propose an algorithm and give two examples to illustrate its feasibility.

Keywords Linear bilevel programming · Weak linear bilevel programming · Penalty method

Mathematics Subject Classification 90C05 · 90C26

1 Introduction

In this paper, we consider the following weak linear bilevel programming problem:

$$\min_{x\in X}\sup_{y\in\Psi(x)} c_1^T x + d_1^T y, \tag{1}$$

where $\Psi(x)$ is the set of solutions of the lower level problem,

Supported by the National Science Foundation of China (71171150, 71201146) and the Ph.D. Fund of Huanggang Normal University to Yue Zheng.

Y. Zheng (✉)
College of Mathematics and Computer Sciences, Huanggang Normal University,
Huanggang 438000, China
e-mail: zhengyuestone@126.com

Z. Wan · K. Sun
School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

T. Zhang
State Key Laboratory of Water Resource and Hydropower Engineering Science, Wuhan University,
Wuhan 430072, China

$$\begin{aligned} &\min_{y \geq 0} d_2^T y \\ &\text{s.t.} \quad Ax + By \leq b. \end{aligned} \tag{2}$$

Here, $x, c_1 \in R^n$, $y, d_1, d_2 \in R^m$, $A \in R^{q \times n}$, $B \in R^{q \times m}$, $b \in R^q$, X is a closed subset of R^n and T stands for transpose.

Remark 1 In the definition of the objective function of problem (2), we have ignored a term of the form $c_2^T x$, since for a given x, $c_2^T x$ is a constant.

Problem (1) is also called pessimistic formulation or weak linear Stackelberg problem. If the solution of the lower level problem is not unique, i.e. for at least $x \in X$, the set $\Psi(x)$ contains of more than one point, it is rational that the leader provides himself/herself against the possible worst choice of the follower in $\Psi(x)$. Several papers have been devoted to weak bilevel programming problem from different subjects, for example, optimality conditions [11, 14], existence of solutions [1–3, 17], regularization [18], approximation [19, 21], and so on. The reader can also be referred to at least two monographs [7, 12] and the annotated bibliography [10, 13, 22, 23] on bilevel programming problem.

It is worthwhile noting that, Aboussoror and Mansouri [2] presented an exact penalty method for weak linear bilevel programming problem. Unfortunately, no numerical results were reported. In this paper, we develop a new variant of the penalty method of [2], for solving weak linear bilevel programming problem. The proposed penalty method, which is inspired from [2, 5, 8, 24], is also exact. Moreover, we give theoretical results on the existence of solution. Finally, we develop an algorithm and give some numerical experiments.

The organization of the paper is as follows. In Sect. 2, we propose a penalty method and establish theoretical results on the existence of solution in Sect. 3. In Sect. 4, an algorithm is proposed, and then two examples are given to illustrate its feasibility. Finally, we finish with a conclusion section.

2 Penalty method

In order to establish theoretical results, we state the main assumptions throughout the paper.

Assumptions

(A1) For any $x \in X$, $Y(x) = \{y \in R^m \mid By \leq b - Ax,\ y \geq 0\} \neq \emptyset$, and there exists a compact subset Z of R^m such that $Y(x) \subset Z$ for all $x \in X$.
(A2) The set X is a bounded polyhedron.

For each $x \in X$, it follows from assumption (A1) and Theorem 3.1 in [12] that $v(x) = \sup_{y \in \Psi(x)} d_1^T y$ is attained. Then, problem (1) can also be written as:

$$\min_{x \in X} \left[c_1^T x + v(x)\right]. \tag{3}$$

Denote the optimal value of problem (2) by $\varphi(x)$. Then,

$$\Psi(x) = \left\{ y \in Y(x) \,\middle|\, d_2^T y = \varphi(x) \right\}.$$

Now, we consider the penalized problem of (3) by the penalty parameter k:

$$\min_{x \in X} \left\{ c_1^T x + \max_{y \in Y(x)} \left[d_1^T y + k\left(\varphi(x) - d_2^T y\right)\right] \right\}, \tag{4}$$

which is also rewritten as:

$$\min_{x \in X} \left\{ c_1^T x + k\varphi(x) + \max_{y \in Y(x)} \left[d_1^T y - k d_2^T y \right] \right\}.$$

For each $x \in X$, the dual of the following problem

$$\max_{y \in Y(x)} \left[d_1^T y - k d_2^T y \right] \tag{5}$$

can be written as:

$$\begin{aligned} &\min_{u \geq 0} (b - Ax)^T u \\ &\text{s.t.} \;\; -B^T u \leq k d_2 - d_1. \end{aligned} \tag{6}$$

Note that, for any $x \in X$, it follows from (A1) that problem (5) has at least one solution. Then, problem (6) also has a solution, and these two problems have the common optimal value. Hence, problem (4) is equivalent to the following problem:

$$\min_{x \in X} \left[c_1^T x + k\varphi(x) + \psi(x) \right], \tag{7}$$

where $\psi(x)$ is the optimal value of problem (6).

For $k > 0$, define the following function:

$$v_k(x) = k\varphi(x) + \max_{y \in Y(x)} \left[d_1^T y - k d_2^T y \right]. \tag{8}$$

It follows from the definition of $\psi(x)$ that $v_k(x) = k\varphi(x) + \psi(x)$. Then, problems (4) and (7) are equivalent to the following problem:

$$\min_{x \in X} \left[c_1^T x + v_k(x) \right]. \tag{9}$$

For the sake of simplicity, we denote

$$\begin{aligned} S &= \left\{ (x, y) \,\middle|\, x \in X, \;\; y \in Y(x) \right\}, \\ U_k &= \left\{ u \,\middle|\, -B^T u \leq k d_2 - d_1, \;\; u \geq 0 \right\}. \end{aligned}$$

Now, we consider the following problem:

$$\begin{aligned} &\min_{x, y, u} \left[c_1^T x + k d_2^T y + (b - Ax)^T u \right] \\ &\text{s.t.} \quad (x, y) \in S, \\ &\qquad\;\; u \in U_k. \end{aligned} \tag{10}$$

Then, we have the following result which shows the relations between (9) and (10).

Lemma 1 *For a fixed value of* k, *if* $(\bar{x}, \bar{y}, \bar{u})$ *is a solution of problem* (10), *then* $\bar{x}$ *is a solution of problem* (9). *Furthermore,* $\varphi(\bar{x}) = d_2^T \bar{y}$ *and* $\psi(\bar{x}) = (b - A\bar{x})^T \bar{u}$.

Proof Suppose that $\tilde{x}$ solves problem (9). Then, there exist $\tilde{y} \in Y(\tilde{x})$ and $\tilde{u} \in U_k$ such that $\varphi(\tilde{x}) = d_2^T \tilde{y}$ and $\psi(\tilde{x}) = (b - A\tilde{x})^T \tilde{u}$. Thus, it follows that $(\tilde{x}, \tilde{y}, \tilde{u})$ is a feasible point of problem (10).

Since $(\bar{x}, \bar{y}, \bar{u})$ is a solution of problem (10), we have

$$c_1^T \bar{x} + kd_2^T \bar{y} + (b - A\bar{x})^T \bar{u} \leq c_1^T \tilde{x} + kd_2^T \tilde{y} + (b - A\tilde{x})^T \tilde{u},$$

which implies that,

$$c_1^T \bar{x} + kd_2^T \bar{y} + (b - A\bar{x})^T \bar{u} \leq c_1^T \tilde{x} + k\varphi(\tilde{x}) + \psi(\tilde{x}). \tag{11}$$

On the other hand, it follows from $d_2^T \bar{y} \geq \varphi(\bar{x})$ and $(b - A\bar{x})^T \bar{u} \geq \psi(\bar{x})$ that

$$c_1^T \bar{x} + kd_2^T \bar{y} + (b - A\bar{x})^T \bar{u} \geq c_1^T \bar{x} + k\varphi(\bar{x}) + \psi(\bar{x}). \tag{12}$$

Combining (11) and (12), then we have

$$c_1^T \bar{x} + v_k(\bar{x}) = c_1^T \bar{x} + k\varphi(\bar{x}) + \psi(\bar{x}) \leq c_1^T \tilde{x} + k\varphi(\tilde{x}) + \psi(\tilde{x}) = c_1^T \tilde{x} + v_k(\tilde{x}),$$

which implies that, $\bar{x}$ is a solution of problem (9).

Replacing $\tilde{x}$ of the right part of (11) with $\bar{x}$, we can easily obtain that $\varphi(\bar{x}) = d_2^T \bar{y}$ and $\psi(\bar{x}) = (b - A\bar{x})^T \bar{u}$. This completes the proof. □

3 Main results

Denote by $V(\cdot)$ the set of vertices of the set to be concerned. For a fixed value of k, define the following function:

$$\theta_k(u) = \inf_{(x,y)\in S} \left[c_1^T x + kd_2^T y + (b - Ax)^T u\right].$$

Then, we have the following result.

Theorem 1 *If assumptions* (A1) *and* (A2) *are satisfied,*

$$\min_{u\in U_k} \theta_k(u) \tag{13}$$

has at least one solution in $V(U_k)$.

Proof It is easy to verify that $\theta(\cdot)$ is a concave function. Then,

$$\begin{aligned}\inf_{u\in U_k} \theta_k(u) &= \inf_{x\in X} \inf_{\substack{y\in Y(x),\\ u\in U_k}} \left[c_1^T x + kd_2^T y + (b - Ax)^T u\right] \\ &\geq \inf_{x\in X} \left\{c_1^T x + k\varphi(x) + \sup_{y\in Y(x)} \left[d_1^T y - kd_2^T y\right]\right\} \\ &\geq \inf_{x\in X} \left[c_1^T x + d_1^T y + k\big(\varphi(x) - d_2^T y\big)\right],\end{aligned}$$

for any $y \in Y(x)$.

In particular, for any $x \in X$, if $\tilde{y}(x)$ solves problem (2), we have $\varphi(x) - d_2^T \tilde{y}(x) = 0$, and then

$$\inf_{u \in U_k} \theta_k(u) \geq \inf_{x \in X} \left[c_1^T x + d_1^T \tilde{y}(x)\right] \geq \inf_{(x,y) \in S} \left[c_1^T x + d_1^T y\right] = \min_{(x,y) \in S} \left[c_1^T x + d_1^T y\right].$$

Therefore, it follows from (A1) and (A2) that the function $\theta_k(u)$ is bounded from below. Using Corollary 32.3.4 of Rockafellar [20], we deduce that problem (13) has at least one solution in $V(U_k)$. This completes the proof. □

On the solution of the penalized problem (12), we can obtain the following result.

Theorem 2 *Under assumptions* (A1) *and* (A2), *problem* (10) *has at least one solution.*

Proof It follows from Theorem 1 that problem (13) has a solution $u^* \in V(U_k)$. Then, under assumptions (A1) and (A2), the following linear programming problem

$$\min_{(x,y) \in S} \left[c_1^T x + k d_2^T y + (b - Ax)^T u^*\right]$$

has at least one solution $(x^*, y^*) \in V(S)$. Hence, (x^*, y^*, u^*) is a solution of (10). □

Theorem 3 *Let assumptions* (A1) *and* (A2) *be satisfied. For any* $x \in X$, *there exists* $k^* > 0$ *such that* $v_k(x) = v(x)$ *for all* $k > k^*$.

Proof For each $x \in X$, it is clear that the solution of the linear programming problem (5) is attained at $V(Y(x))$. Then, if any $y \in V(Y(x))$ satisfies $y \in \Psi(x)$, we have $v_k(x) = v(x)$.

Without loss of generality, we can assume that, for at least $x \in X$, the set

$$D(x) = \left\{y \in V\big(Y(x)\big) \mid \varphi(x) - d_2^T y < 0\right\}$$

is not empty. Then, consider the following problems:

$$g_1 = \max_{x \in X, y \in D(x)} \left[\varphi(x) - d_2^T y\right]$$

and

$$g_2(x) = \max_{y \in Y(x)} d_1^T y.$$

Now, let $k^*(x) = -\frac{g_2(x) - v_k(x)}{g_1}$. Then, $k^*(x) > 0$. For any $x \in X$ and $k > k^*(x)$, we have

$$\begin{aligned}
\max_{y \in D(x)} \left\{d_1^T y + k\left[\varphi(x) - d_2^T y\right]\right\} &< \max_{y \in D(x)} \left\{d_1^T y + k^*(x)\left[\varphi(x) - d_2^T y\right]\right\} \\
&\leq \max_{y \in Y(x)} d_1^T y + k^*(x) \max_{y \in D(x)} \left[\varphi(x) - d_2^T y\right] \\
&\leq g_2(x) + k^*(x) \cdot g_1 \\
&= v_k(x),
\end{aligned}$$

which implies that, for any $x \in X$ and $k > k^*(x)$, $v_k(x)$ must be attained at $y \in \{y \in V(Y(x)) \mid \varphi(x) = d_2^T y\}$, i.e. $y \in \Psi(x)$, and then $v_k(x) = v(x)$.

Moreover, it follows from the definitions of $v_k(x)$ and $g_2(x)$ that they are continuous, and then $k^*(x)$ is also continuous over the bounded polyhedron X. Hence, there exists $k^* > 0$, i.e. $k^* = \max_{x \in X} k^*(x)$, such that $v_k(x) = v(x)$ for all $k > k^*$. □

Theorem 4 *Let assumptions* (A1) *and* (A2) *be satisfied, and* $\{(x_k, y_k, u_k)\}$ *be a sequence of solutions of problem* (10). *Then, for all* $k > k^*$, *which is defined in Theorem* 3, x_k *is a solution of problem* (3).

Proof It follows from Lemma 1 that x_k is a solution of problem (9). Then,

$$c_1^T x_k + v_k(x_k) \le c_1^T x + v_k(x), \quad \forall x \in X.$$

For all $k > k^*$, it follows from Theorem 3 and the above inequality that

$$c_1^T x_k + v(x_k) \le c_1^T x + v(x), \quad \forall x \in X,$$

which implies that, for all $k > k^*$, x_k is also a solution of problem (3). □

4 The algorithm and numerical experiment

According to the results in Sect. 3, we can propose a simple algorithm for solving problem (3) as follows.

Algorithm 1

Step 0. Choose $k > 0$, $\lambda > 0$ and set $i = 1$.
Step 1. Solve problem (10), and denote its solution by (x^i, y^i, u^i).
Step 2. For $x = x^i$, solve the following problem:

$$\max_{y \in Y(x)} \left[d_1^T y - k d_2^T y\right],$$

and denote its solution by y^{i*}.
Step 3. If $d_2^T y^i = d_2^T y^{i*}$, then x^i is a solution of problem (3), stop. Otherwise, set $k = k + \lambda$, $i = i + 1$ and go to step 1.

Remark 2 Obviously, solving problem (10), which is a disjoint bilinear programming problem [15], is the main compute tasks of the above algorithm. Fortunately, however, many authors proposed algorithms to solve such a problem, for example, see [4, 6, 16].

Remark 3 It follows from Lemma 1 that $\varphi(x^i) = d_2^T y^i$, and from step 3 of the above algorithm that $v_k(x^i) = k\varphi(x^i) + d_1^T y^{i*} - k d_2^T y^{i*}$. Furthermore, if $d_2^T y^i = d_2^T y^{i*}$, then $y^{i*} \in \Psi(x^i)$. Hence, we have $v_k(x^i) = v(x^i)$, and x^i is a solution of problem (3) by using Theorem 4.

To illustrate the feasibility of the proposed algorithm, we consider the following examples which are adapted from [9].

Table 1 Numerical results of Algorithms 1 and 2

Example	Algorithm 1		Algorithm 2	
	$(x^*, y^*)^T$	$f(x^*, y^*)$	$(x^*, y^*)^T$	$f(x^*, y^*)$
1	$(0, 10, 0, 10)^T$	−90	$(0, 10, 0, 10)^T$	−90
2	$(10, 0, 0, 0, 0, 0)^T$	−80	$(10, 0, 0, 0, 0, 0)^T$	−80

Example 1

$$\min_{x\in X}\max_{y\in\Psi(x)} -8x_1 - 10x_2 - 2y_1 + y_2,$$

where $X = \{x \mid x = (x_1, x_2)^T,\ x_1 + x_2 \le 10,\ x_1, x_2 \ge 0\}$, and $\Psi(x)$ is the set of solutions of the lower level problem,

$$\begin{aligned} &\min_y -y_1 - y_2 \\ &\text{s.t.}\ \ y_1 + y_2 \le 20 + x_1 - x_2, \\ &\qquad y = (y_1, y_2)^T \ge 0. \end{aligned}$$

Example 2

$$\min_{x\in X}\max_{y\in\Psi(x)} -8x_1 - 6x_2 - 25y_1 - 30y_2 + 2y_3 + 16y_4,$$

where $X = \{x \mid x = (x_1, x_2)^T,\ x_1 + x_2 \le 10,\ x_1, x_2 \ge 0\}$, and $\Psi(x)$ is the set of solutions of the lower level problem,

$$\begin{aligned} &\min_y -10y_1 - 10y_2 - 10y_3 - 10y_4 \\ &\text{s.t.}\ \ y_1 + y_2 + y_3 + y_4 \le 10 - x_1 - x_2, \\ &\qquad -y_1 + y_4 \le 0.8x_1 + 0.8x_2, \\ &\qquad y_2 + y_4 \le 4x_2, \\ &\qquad y = (y_1, y_2, y_3, y_4)^T \ge 0. \end{aligned}$$

In our experiment, we first choose $k = 10$, $\lambda = 10$, and then use the algorithm proposed by Alarie et al. [4] to solve disjoint bilinear programming problem (10). The numerical results are reported in Table 1 where $(x^*, y^*)^T$ and $f(x^*, y^*)$ denote the solution and the optimal value, respectively. Moreover, for comparison purposes, results of the penalty method in [2], which is denoted by Algorithm 2, are given in Table 1 as well. As shown in Table 1, the results obtained by Algorithm 1 is the same as that of Algorithm 2.

Now, to better illustrate our algorithm, we consider Example 1, and present the steps in solving this example as follows.

Iteration 1

Step 0: Let $k = 10$ and $\lambda = 10$.

Step 1: Using (10), we can get the penalty problem of Example 1 as follows:

$$\begin{aligned}
\min_{x,y,u}\ & -8x_1 - 10x_2 + 10(-y_1 - y_2) + (20 + x_1 - x_2)u \\
\text{s.t.}\ \ & x_1 + x_2 \le 10, \\
& y_1 + y_2 \le 20 + x_1 - x_2, \\
& -u \le -8, \\
& -u \le -11, \\
& x, y, u \ge 0,
\end{aligned}$$

and obtain its solution $(x^1, y^1, u^1)^T = (0, 10, 5, 5, 11)^T$.

Step 2: Solve the following linear programming problem:

$$\max_{y \in Y(x^1)} -2y_1 + y_2 + 10(y_1 + y_2),$$

and get its solution $y^{1*} = (0, 10)^T$.

Step 3: $(x^1, y^{1*})^T = (0, 10, 0, 10)^T$ is a solution of this example because of $d_2^T y^1 = d_2^T y^{1*} = -10$.

5 Conclusion

In this paper, we consider the weak linear bilevel programming problem where the solution of the lower level problem is not unique. We then present a simple algorithm based on an exact penalty method, which is different from the existing penalty method, for such problems. Finally, two numerical examples illustrate its feasibility. It is interesting and useful to transfer this method to weak nonlinear bilevel programming problem in our future research.

References

1. Aboussoror, A.: Weak bilevel programming problems: existence of solutions. Adv. Math. Res. **1**, 83–92 (2002)
2. Aboussoror, A., Mansouri, A.: Weak linear bilevel programming problems: existence of solutions via a penalty method. J. Math. Anal. Appl. **304**, 399–408 (2005)
3. Aboussoror, A., Mansouri, A.: Existence of solutions to weak nonlinear bilevel problems via MinSup and d.c. problems. RAIRO. Rech. Opér. **42**, 87–103 (2008)
4. Alarie, S., Audet, C., Jaumard, B., Savard, G.: Concavity cuts for disjoint bilinear programming. Math. Program. **90**, 373–398 (2001)
5. Anandalingam, G., White, D.J.: A solution for the linear static Stackelberg problem using penalty function. IEEE Trans. Autom. Control **35**, 1170–1173 (1990)
6. Audet, C., Hansen, P., Jaumard, B., Savard, G.: A symmetrical linear maxmin approach to disjoint bilinear programming. Math. Program. **85**, 573–592 (1999)
7. Bard, J.F.: Practical Bilevel Optimization: Algorithms and Applications. Kluwer Academic, Dordrecht (1998)
8. Campelo, M., Dantas, S., Scheimberg, S.: A note on a penalty function approach for solving bi-level linear programs. J. Glob. Optim. **16**, 245–255 (2000)
9. Cao, D., Leung, L.C.: A partial cooperation model for non-unique linear two-level decision problems. Eur. J. Oper. Res. **140**, 134–141 (2002)

10. Colson, B., Marcotte, P., Savard, G.: An overview of bilevel optimization. Ann. Oper. Res. **153**, 235–256 (2007)
11. Dassanayaka, S.: Methods of variational analysis in pessimistic bilevel programming. PhD thesis, Wayne State University (2010)
12. Dempe, S.: Foundations of Bilevel Programming. Nonconvex Optimization and Its Applications. Kluwer Academic, Dordrecht (2002)
13. Dempe, S.: Annotated bibliography on bilevel programming and mathematical problems with equilibrium constraints. Optimization **52**, 333–359 (2003)
14. Dempe, S., Mordukhovich, B.S., Zemkoho, A.B.: Necessary optimality conditions in pessimistic bilevel programming. Optimization (2012). doi:10.1080/02331934.2012.696641
15. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches, 3rd edn. Springer, Berlin (1996)
16. Konno, H.: A cutting plane algorithm for solving bilinear programs. Math. Program. **11**, 14–27 (1976)
17. Lignola, M.B., Morgan, J.: Topological existence and stability for Stackelberg problems. J. Optim. Theory Appl. **84**, 145–169 (1995)
18. Loridan, P., Morgan, J.: On strict ϵ-solutions for a two-level optimization problem. In: Proceedings of the International Conference on Operations Research, pp. 165–172. Springer, Berlin (1992)
19. Loridan, P., Morgan, J.: Weak via strong Stackelberg problem: new results. J. Glob. Optim. **8**, 263–287 (1996)
20. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
21. Tsoukalas, A., Wiesemann, W., Rustem, B.: Global optimisation of pessimistic bi-level problems. Fields Inst. Commun. **55**, 1–29 (2009)
22. Vicente, L.N., Calamai, P.H.: Bilevel and multilevel programming: a bibliography review. J. Glob. Optim. **5**, 1–23 (1994)
23. Wen, U.P., Hsu, S.T.: Linear bilevel programming problems—a review. J. Oper. Res. Soc. **42**, 125–133 (1991)
24. White, D.J., Anandalingam, G.: A penalty function for solving bi-level linear programs. J. Glob. Optim. **3**, 397–419 (1993)