Wuhan University Journal of Natural Sciences

2017, Vol.22 No.1, 063-067

Article ID 1007-1202(2017)01-0063-05

DOI 10.1007/s11859-017-1217-6

Maximum Entropy Approach for Solving Pessimistic Bilevel Programming Problems

□ ZHENG Yue, ZHUO Xiangzhi, CHEN Jingxiao

School of Management, Huaibei Normal University, Huaibei 235000, Anhui, China

Abstract: Bilevel programming problems are of growing interest both from theoretical and practical points of view. In this paper, we study a pessimistic bilevel programming problem in which the set of solutions of the lower level problem is discrete. We first transform such a problem into a single-level optimization problem by using the maximum-entropy techniques. We then present a maximum entropy approach for solving the pessimistic bilevel programming problem. Finally, two examples illustrate the feasibility of the proposed approach.

Key words: bilevel programming; pessimistic formulation; maximum entropy approach

CLC number: O 221

Received date: 2016-02-25

Foundation item: Supported by the National Natural Science Foundation of China (11501233), the Key Project of Anhui Province University Excellent Youth Support Plan (gxyqZD2016102)

Biography: ZHENG Yue, male, Associate professor, research direction: optimization theory and application. E-mail: zhengyuestone@126.com

0 Introduction

Bilevel programming is a sequence of two optimization problems where the constraint region of the upper level problem is determined implicitly by the lower level problem. It plays exceedingly important roles in different application fields, such as transportation, economics, ecology, and engineering[1]. The recent monographs and surveys can refer to Refs. [2-7].

A general formulation of bilevel programming problem can be written as follows:

$$\begin{aligned} \text{“}\min \quad & F(x,y)\text{”} \\ \text{s.t.} \quad & G(x) \leqslant 0 \end{aligned} \tag{1}$$

where y is the solution of the lower level problem,

$$\begin{aligned} \min_{y} \quad & f(x,y) \\ \text{s.t.} \quad & h(x,y) \leqslant 0 \end{aligned}$$

where $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$, $F, f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, $G : \mathbf{R}^n \to \mathbf{R}^q$, $h : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^p$.

Let $X = \{x \in \mathbf{R}^n \mid G(x) \leqslant 0\}$ be a compact set of $\mathbf{R}^n$. For each $x \in X$, denote by $\psi(x)$ the set of solutions of the lower level problem. When the set $\psi(x)$ is not a singleton, the upper level objective function value cannot be predicted in general without knowledge of the response of the lower level problem. In this case, the definition of problem (1) is unclear. That is the reason why we use the quotation marks in (1). To overcome this ambiguity, most of the authors used either the optimistic formulation or pessimistic formulation[3]. The corresponding problems are called strong and weak bilevel programming problem, respectively[8]. For the optimistic bilevel programming problem, the leader can affect the follower's decision, so that the follower always selects a strategy in $\psi(x)$ that suits the leader best. In a competitive world, on one hand, the cooperation between both the leader and the follower may not be allowed. In this case, the leader is not able to influence the follower's choice. On the other hand, the leader may be risk-averse. Thus, the leader wishes to bound the damage resulting from an undesirable selection of the follower in $\psi(x)$, i.e. he minimizes the function $\max F(x,y)$ which subjects to the constraint $y \in \psi(x)$. This leads to the following pessimistic bilevel programming problem:

$$\min_{x\in X} \sup_{y\in\psi(x)} F(x,y) \tag{2}$$

Several papers have been contributed to pessimistic bilevel programming problems from different subjects. For example, Refs. [9-14] discussed the existing results and approximations results of solutions. Using a variety of techniques from convex, nonconvex and nonsmooth analysis, Dassanayaka[15] presented first order necessary and sufficient optimality conditions for pessimistic bilevel programming problems. Employing advanced tools of variational analysis and generalized differentiation, Dempe *et al*[16] derived lower (i.e. conventional) subdifferential optimality conditions for pessimistic bilevel programming problem via the LLVF and KKT approaches. Wiesemann *et al* [17] analyzed the structural properties for the independent pessimistic bilevel programming problem, and developed a solvable ε-approximation algorithm. Cervinka *et al* [18] presented a new numerical method to compute approximate and the so-called relaxed pessimistic solutions to mathematical programs with equilibrium constraints which is a generalized bilevel programming problem. Based on an exact penalty function, Zheng *et al* [19] proposed an algorithm which can find at least a local solution for the weak linear bilevel programming problem. For a class of pessimistic semivectorial bilevel programming, Liu *et al*[20] established the first-order necessary optimality conditions by using the scalarization method and the generalized differentiation calculus of Mordukhovich [16]. Recently, Zheng *et al* [21] developed a modified *K*th-Best algorithm for the weak linear bilevel programming problem.

It is worthwhile noting that Mallozzi *et al* [22] presented an intermediate method for solving static non-zero-sum game, and supposed that the leader could obtain or know a density function $p(x)$ on $\psi(x)$ for each $x\in X$. Before using this method for all practical purposes, however, we need to know the density function $p(x)$. In general, the leader is difficult to obtain an exact description of such a function. In order to overcome this dilemma, the maximum entropy approach is considered since it can be used to estimate the probability distribution of the set $\psi(x)$. So, in this paper, the pessimistic bilevel programming problem is solved by using the maximum entropy approach. To the best knowledge of the authors, this is the first time that such a problem is solved via maximum entropy approach.

The paper is organized as follows. We consider the case in which the set $\psi(x)$ is discrete for any $x\in X$ in Section 1, and then give some examples to illustrate the feasibility of the maximum entropy approach in Section 2. In Section 3, we give a conclusion.

1 Maximum Entropy Approach for Discrete Reaction Set

We first give the related definition of pessimistic solution.

Definition 1

(a) Constraint region of problem (2):

$$S=\{(x,y)\,|\,x\in X, h(x,y)\leqslant 0\}.$$

(b) Projection of S onto the leader's decision space:

$$S(X)=\{x\in X\,|\,\exists y, \text{such that } (x,y)\in S\}.$$

(c) Feasible set for the follower:

$$Y(x)=\{y\,|\,h(x,y)\leqslant 0\}.$$

(d) The follower's rational reaction set:

$$\psi(x)=\{y\,|\,y\in \operatorname{Arg\,min}(f(x,y)\,|\,y\in Y(x))\}.$$

(e) Inducible region or feasible region of the leader:

$$\mathrm{IR}=\{(x,y)\,|\,(x,y)\in S, y\in\psi(x)\}.$$

(f) Define a function $\theta(x)$ as follows:

$$\theta(x)=\sup_{y\in\psi(x)} F(x,y).$$

A point $x^*\in X$ with $y^*\in\psi(x^*)$ is called a pessimistic solution to (2), if

$$\theta(x^*)=F(x^*,y^*)$$
$$\theta(x^*)\leqslant\theta(x), \forall (x,y)\in \mathrm{IR}.$$

Next, suppose that the set $\psi(x)$ has a finite number of elements. More precisely, let

$$\psi(x)=\{y_i(x)\,|\,i=1,2,\cdots,m\}.$$

Let

$$\phi(x)=\sup_{1\leqslant i\leqslant m} F(x,y_i(x)).$$

Then, the pessimistic bilevel programming problem (2) with discrete reaction set can be written as follows:

$$\min_{x\in X} \phi(x) \tag{3}$$

Note that, it is difficult to solve problem (3), because $\phi(x)$ may be nondifferentiable at some points. To avoid this, a natural idea is to find a differentiable approximation of $\phi(x)$ to replace $\phi(x)$.

Now, suppose that the leader would attribute probability distribution $p_i(x)$ on $y_i(x)$ ($i=1,2,\cdots,m$), where $p_i(x)\geqslant 0$, and $\sum_{i=1}^{m} p_i(x)=1$. Note particularly that, the density function $p_i(x)$ ($i=1,2,\cdots,m$) only play important roles in deriving the formulation of $\phi_\rho(x)$ below, but they have no link to the function $\phi_\rho(x)$.

Based on the maximum entropy approach solving constrained optimization[23], we consider the following problem:

$$\begin{aligned} \max & \sum_{i=1}^{m}\left[p_i(x)F(x,y_i(x))-\frac{1}{\rho}p_i(x)\ln p_i(x)\right] \\ \text{s.t.} \ & p_i(x)\geqslant 0, \\ & \sum_{i=1}^{m} p_i(x)=1, \end{aligned} \tag{4}$$

where $\rho > 0$ is a real parameter, and $H=-\sum_{i=1}^{m} p_i(x)$ $\ln p_i(x)$ is the Shannon information entropy.

It follows from Templeman and Li[23] that the solution of problem (4) is written as:

$$p_i(x)=\frac{\exp\left(\rho F(x,y_i(x))\right)}{\sum_{i=1}^{m}\exp\left(\rho F(x,y_i(x))\right)}, i=1,2,\cdots,m,$$

and its optimal value is $\frac{1}{\rho}\ln\sum_{i=1}^{m}\exp\left(\rho F(x,y_i(x))\right)$.

For any $\rho > 0$, define:

$$\phi_\rho(x)=\frac{1}{\rho}\ln\sum_{i=1}^{m}\exp\left(\rho F(x,y_i(x))\right).$$

Then, it is easy to obtain the following result. For details, see Ref.[23].

Theorem 1 For any $x\in X$, we have

$$\lim_{\rho\to+\infty}\phi_\rho(x)=\phi(x).$$

The result above implies that we can use the function $\phi_\rho(x)$ to approximate $\phi(x)$. Thus, the problem (3) can be approximated by the following problem:

$$\min_{x\in X}\phi_\rho(x) \tag{5}$$

Theorem 2 Let x_ρ be a solution of problem (5). If $x_\rho \to x^*$ as $\rho\to+\infty$, then x^* is a solution of problem (3).

Proof By the definition of x_ρ, for all $x\in X$, we find that

$$\phi_\rho(x_\rho)\leqslant\phi_\rho(x).$$

As $\rho\to+\infty$, for all $x\in X$, it follows that

$$\phi(x^*)\leqslant\phi(x)$$

which implies that x^* is a solution of problem (3). This completes the proof.

2 Maximum Entropy Approach and Numerical Examples

In this section, we first present the following maximum entropy approach based on the mathematical results in Section 1.

Maximum entropy approach

Step 1 Let $\varepsilon > 0$. Choose a sequence $\{\rho_k\}$ such that $\rho_k\to+\infty$ as $k\to+\infty$, and let $k=1$.

Step 2 Compute the solution set $\psi(x)$ of the lower level problem.

Step 3 Solve problem (5) for $\rho=\rho_k$, and denote its solution by x_k.

Step 4 If $\left|\phi_{\rho_k}(x_k)-\phi(x_k)\right| < \varepsilon$, then the algorithm terminates and x_k is an approximate solution to problem (2). Otherwise, let $k=k+1$, and go to Step 3.

To illustrate the feasibility of the proposed approach, we consider the following two examples which are adapted from Ref. [22].

Example 1

$$\begin{aligned} & \min_{x}\max_{y\in\psi(x)} x+y \\ & \text{s.t.} \ -1\leqslant x\leqslant 1, \end{aligned}$$

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

$$\begin{aligned} & \min_{y}\left|y^2-x^4\right| \\ & \text{s.t.} \ -1\leqslant y\leqslant 1. \end{aligned}$$

For this example, $\forall x\in[-1,1]$, $\psi(x)=\{x^2,-x^2\}$, and the pessimistic solution is $x^*=-0.5$.

Example 2

$$\begin{aligned} & \min_{x}\max_{y\in\psi(x)} xy \\ & \text{s.t.} \ -1\leqslant x\leqslant 1 \end{aligned}$$

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

$$\begin{aligned} & \min_{y}\max\left\{0, y\left(y-\frac{x}{2}+\frac{1}{2}\right)\right\} \\ & \text{s.t.} \ -1\leqslant y\leqslant 1. \end{aligned}$$

The set of solutions of the lower level problem in Example 2 is $\psi(x)=\left[\frac{x-1}{2},0\right]$.

Any point x^* in [0,1] is the pessimistic solution.

The numerical tests were run on a PC(1.3GHz Inter Core i5, 4GB RAM). All nonlinear programming problems in Step 3 of the proposed maximum entropy approach are solved by GAMS/BARON[24]. BARON is one of the recently updated solvers on the NEOS server. In particular, it is commonly used for the global solution of nonconvex nonlinear programming problems. In our test, the parameters are set as follows: $\varepsilon=10^{-8}$, $\rho_1=1$, and $\rho_{k+1}=10\rho_k$. Let $\phi_0(x)=\phi(x)$. Furthermore, we can obtain the following results.

(i) Example 1

From Table 1, it can be seen that when $\rho=100$, $x_\rho=-0.5$, and $\phi_\rho(x_\rho)-\phi(x_\rho)=0$ which shows that $x^*=-0.5$ is the pessimistic solution for this example. Moreover, $\phi_\rho(x)$ is almost identical to $\phi_0(x)$ as $\rho=100$ in Fig. 1.

Table 1 Computational results of Example 1

ρ	x_ρ	$\phi_\rho(x_\rho)$	$\phi(x_\rho)$
1	–0.833	0.084	–0.139
10	–0.506	–0.249	–0.25
100	–0.5	–0.25	–0.25

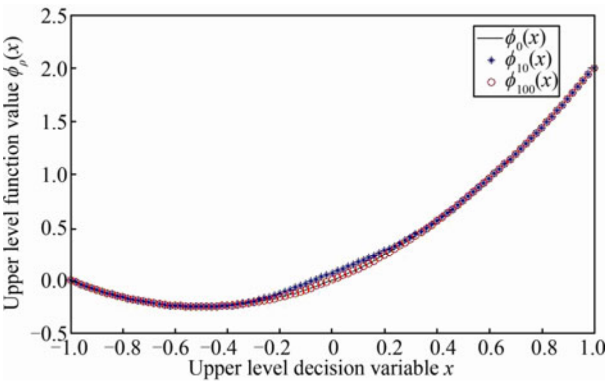

Fig. 1 Comparison of Example 1 between $\phi_\rho(x)$ and $\phi(x)$

(ii) Example 2

From Table 2, it can be seen that $\phi_\rho(x_\rho)$ will be close to $\phi(x_\rho)$ with sufficiently large ρ. Obviously, $x^*=0$ is an approximately pessimistic solution to Example 2. Comparing the curves of upper level objective values of $\phi_\rho(x)$ in Fig. 2, we observe that the objective values achieved by the proposed maximum entropy approach are almost identical while $\rho=1\times10^{10}$. Note that, this example has an infinite number of pessimistic solutions. Our proposed approach only obtains an approximately pessimistic solution.

Table 2 Computational results of Example 2

ρ	x_ρ	$\phi_\rho(x_\rho)$	$\phi(x_\rho)$
1	0.5	–0.062	0
10	0.5	–0.056	0
100	0.5	–0.025	0
10^3	0.5	–0.005	0
10^4	0.497	-7.131^{-4}	0
10^5	0.507	-9.433^{-5}	0
10^{10}	0	6.931^{-11}	0

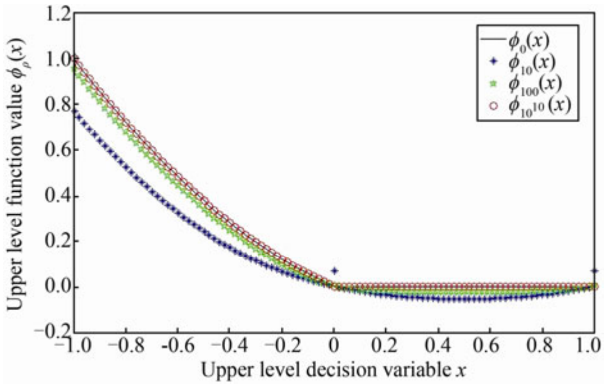

Fig. 2 Comparison of Example 2 between $\phi_\rho(x)$ and $\phi(x)$

These two examples show the feasibility of maximum entropy approach. With sufficiently large ρ, we could find an approximation solution of the pessimistic bilevel programming problem. Moreover, this method is easy to implement. Therefore, it will provide us with a new way to solve the pessimistic bilevel programming problems.

3 Conclusion

In this paper, we present a maximum entropy approach for solving pessimistic bilevel programming problems. Two examples show that the proposed approach is feasible. In fact, this approach would not be viable for an intricate pessimistic bilevel programming problem, because it needs to obtain the set of solutions of the lower level problem in advance. Clearly, it allows us to use the constrained nonlinear programming tools to solve the pessimistic bilevel programming problems. Therefore, it may provide us with a new way to discuss the pessimistic bilevel programming problems.

References

[1] Dempe S. Annottated bibliography on bilevel programming and mathematical problems with equilibrium constraints [J]. *Optimization*, 2003, **52**: 333-359.

[2] Bard J F. *Practical Bilevel Optimization: Algorithms and Applications* [M]. Dordrecht: Kluwer Academic, 1998.

[3] Dempe S. *Foundations of Bilevel Programming ,Nonconvex Optimization and Its Applications Series*[M]. Dordrecht: Kluwer Academic, 2002.

[4] Colson B, Marcotte P, Savard G. An overview of bilevel optimization [J]. *Annals of Operations Research*, 2007, **153**: 235-256.

[5] Shimizu K, Ishizuka Y, Bard J F. *Nondifferentiable and Two-Level Mathematical Programming* [M]. Dordrecht: Kluwer Academic, 1997.

[6] Wang G, Wan Z, Wang X. Bibliography on bilevel programming [J]. *Advances in Mathematics*, 2007, **36**: 513-529(Ch).

[7] Lu J, Han J, Hu Y, *et al*. Multilevel decision-making: A survey [J]. *Information Sciences*, 2016, **346**: 463-487.

[8] Loridan P, Morgan J. Weak via strong Stackelberg problem: New results [J]. *Journal of Global Optimization*, 1996, **8**: 263-287.

[9] Aboussoror A, Mansouri A. Weak linear bilevel programming problems: Existence of solutions via a penalty method [J]. *Journal of Mathematical Analysis and Applications*, 2005, **304**: 399-408.

[10] Aboussoror A, Mansouri A. Existence of solutions to weak nonlinear bilevel problems via MinSup and d.c. problems [J]. *RAIRO Operations Research*, 2008, **42**: 87-103.

[11] Aboussoror A, Adly S, Jalby V. Weak nonlinear bilevel problems: existence of solutions via reverse convex and convex maximization problems [J]. *Journal of Industrial and Management Optimization*, 2011, **7**: 559-571.

[12] Lignola M B, Morgan J. Topological existence and stability for Stackelberg problems [J]. *Journal of Optimization Theory Applications*, 1995, **84**: 145-169.

[13] Loridan P, Morgan J. New results on approximate solutions in two-level optimization [J]. *Optimization*, 1989, **20**: 819-836.

[14] Loridan P, Morgan J. ε-regularized two-level optimization problems: approximation and existence results [C]// *Proceeding of the Fifth French-German Optimization Conference*. New York: Springer -Verlag, 1989: 99-113.

[15] Dassanayaka S. *Methods of Variational Analysis in Pessimistic Bilevel Programming* [D]. Detroit: Wayne State University, 2010.

[16] Dempe S, Mordukhovich B S, Zemkoho A B. Necessary optimality conditions in pessimistic bilevel programming [J]. *Optimization*, 2014, **63**: 505-533.

[17] Wiesemann W, Tsoukalas A, Kleniati P, *et al*. Pessimistic bi-level optimisation [J]. *SIAM Journal on Optimization*, 2013, **23**: 353-380.

[18] Cervinka M, Matonoha C, Outrata J V. On the computation of relaxed pessimistic solutions to MPECs [J]. *Optimization Methods and Software*, 2013, **28**: 186-206.

[19] Zheng Y, Wan Z, Sun K, *et al*. An exact penalty method for weak linear bilevel programming problem [J]. *Journal of Applied Mathematics and Computing*, 2013, **42**: 41-49.

[20] Liu B, Wan Z, Chen J, *et al*. Optimality conditions for pessimistic semivectorial bilevel programming problems [J]. *Journal of Inequalities and Applications*, 2014, **1**: 1-26.

[21] Zheng Y, Fang D, Wan Z. A solution approach to the weak linear bilevel programming problems [J]. *Optimization*, 2016, **65**: 1437-1449.

[22] Mallozzi L, Morgan J. Hierarchical systems with weighted reaction set [C] // *Nonlinear Optimization and Applications*. New York:Springer-Verlag , 1996 :271-282.

[23] Templeman A B, Li X. A maximum entropy approach to constrained non-linear programming [J]. *Engineering Optimization*, 1987, **12**: 191-205.

[24] Tawarmalani M, Sahinidis N V. A polyhedral branch-and-cut approach to global optimization [J]. *Mathematical Programming*, 2005, **103**: 225-249.

□