CrossMark

ORIGINAL ARTICLE

Approaches to four types of bilevel programming problems with nonconvex nonsmooth lower level programs and their applications to newsvendor problems

Xide Zhu[1] · **Peijun Guo**[1]

Abstract This paper concentrates on solving bilevel programming problems where the lower level programs are max–min optimization problems and the upper level programs have max–max or max–min objective functions. Because these bilevel programming problems include nonconvex and nonsmooth lower level program problems, it is a challenging undone work. Giving some assumptions, we translate these problems into general single level optimization problems or min–max optimization problems. To deal with these equivalent min–max optimization problems, we propose a class of regularization methods which approximate the maximum function by using a family of maximum entropy functions. In addition, we examine the limit situations of the proposed regularization methods and show that any limit points of the global optimal solutions obtained by the approximation methods are the same as the ones of the original problems. Finally, we apply the proposed methods to newsvendor problems and use a numerical example to show their effectiveness.

Keywords Bilevel programming · Min–max optimization · Nonconvex · Nonsmooth · Newsvendor problem · One-shot decision theory

✉ Peijun Guo
guo@ynu.ac.jp

Xide Zhu
xidezhu@gmail.com

1 Faculty of Business Administration, Yokohama National University, 79-4 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan

🖄 Springer

1 Introduction

Bilevel programming problem (BLPP) is an important optimization problem, which includes optimization problems in the constraints. The general formulation of BLPP is

$$\min_{x \in X, \ y \in S(x)} u(x, y), \tag{1.1}$$

where $S(x)$ denotes the set of solutions of the following lower level program:

$$\min_{y \in Y(x)} l(x, y), \tag{1.2}$$

where $X \subset \mathbb{R}^n$, $Y(x) \subset \mathbb{R}^m$ for any $x \in X$, and u, $l : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are continuous functions.

Let x and y respectively denote the decision variables of the upper level (leader) and the lower level (follower). BLPP (1.1)–(1.2) represents an optimistic approach in which the follower is assumed to be cooperative and the leader is allowed to choose the most suitable element from the set of solutions of the follower. On the contrary, a pessimistic approach deals with the case that the follower may be noncooperative. In this case, the leader cannot decide which of the best responses is implemented by the follower so that he/she chooses a decision that performs best in the case that the worst follower response happens, that is, solving the following pessimistic BLPP:

$$\min_{x \in X} \max_{y \in S(x)} u(x, y). \tag{1.3}$$

BLPP has been always an important research area. It was initially introduced by Von Stackelberg (1952) for modeling a duopoly market. A number of contributions including theories, algorithms and applications for BLPP have been made by researchers (Allende and Still 2013; Vicente and Calamai 1994; Colson et al. 2005; Ye and Zhu 1995; Bard 1998; Dempe 2002; Dempe and Zemkoho 2013; Dempe et al. 2012; Dempe and Zemkoho 2014; Ye and Zhu 2010; Lin et al. 2014). In the case that the lower level program is a convex optimization problem and its global optimal solutions can be computed, a common approach to BLPP is to replace the lower level program by its first order optimality condition or Karush–Kuhn–Tucker (KKT) condition and then solve a mathematical program with equilibrium constraints (MPEC) or mathematical program with complementarity constraints (MPCC). However, it is difficult to solve MPEC and MPCC because their constraints fail to satisfy the standard constraint qualifications, such as the most commonly used Mangasarian–Fromovitz constraint qualification (MFCQ). Even under some convexity conditions on the function u and the set X, MPEC and MPCC are still not easy to be solved due to the nonconvexities that occur in the Lagrangean or complementarity constraints. Till now, great efforts have been made to solve MPEC and MPCC (Luo et al. 1996; Facchinei et al. 1999; Ye 2005; Fletcher et al. 2006; Guo et al. 2015; Lin and Fukushima 2005; Scholtes 2001; Zhu and Lin 2016). However, all these available methods can only find out stationary points, there is no guarantee that they are optimal.

In the case that the lower level program is not a convex optimization problem, the KKT based method may not be valid in general, BLPP is still a difficult problem to be solved. Another approach to BLPP is to reformulate it as a single level optimization problem with considering the optimal value function of the lower level program. Define the optimal value function of (1.2) as

$$v(x) := \min_{y \in Y(x)} l(x, y), \tag{1.4}$$

then BLPP (1.1)–(1.2) can be reformulated as the following single level optimization problem:

$$\begin{aligned}
\min \quad & u(x, y) \\
\text{s.t.} \quad & l(x, y) - v(x) \leq 0, \\
& x \in X, \ y \in Y(x).
\end{aligned} \tag{1.5}$$

This reformulation was first introduced by Outrata (1990) for obtaining a numerical solution and subsequently used by Ye and Zhu (1995) for obtaining necessary optimality conditions. Recently, Lin et al. (2014) used this reformulation to solve a simple BLPP where the constraint set of the lower level program does not depend on x, that is, $Y(x) \equiv Y$. Xu and Ye (2014) proposed a smoothing projected gradient algorithm for solving (1.5) by using some smooth functions to approximate the optimal value function. All these available methods need some strong assumptions and can only find stationary points, which may still not be optimal.

Although there is a rich literature on BLPP, to the best of our knowledge, there is no paper handling BLPPs with nonconvex nonsmooth lower level programs due to its inherent mathematical difficulties and the lack of an applied background. In recent years, Guo (2011) initially proposed the one-shot decision theory (OSDT) for decision making under uncertainty. The OSDT has a wide-ranging applied background in business and management (Li and Guo 2015; Guo 2010a, b; Guo and Li 2014; Guo and Ma 2014; Wang and Guo 2017). There are four decision models which are introduced as follows:

- Model I:

$$\max_{x \in X} \ \max_{y \in S_1(x)} \ f(x, y), \tag{1.6}$$

where $S_1(x)$ denotes the set of solutions of the following lower level program:

$$\max_{y \in Y} \ \min\{g(y), f(x, y)\}, \tag{1.7}$$

where $X := [x_l, x_u]$ and $Y := [y_l, y_u]$ are bounded subsets of $\mathbb{R}$, $f(x, y) : \mathbb{R} \times \mathbb{R} \to [0, 1]$ and $g(y) : \mathbb{R} \to [0, 1]$ are continuously differentiable functions.
- Model II:

$$\max_{x \in X} \ \max_{y \in S_2(x)} \ f(x, y), \tag{1.8}$$

where $S_2(x)$ denotes the set of solutions of the following lower level program:

$$\max_{y \in Y} \ \min\{1 - g(y), f(x, y)\}. \tag{1.9}$$

- Model III:

$$\max_{x \in X} \ \min_{y \in S_3(x)} \ f(x, y), \tag{1.10}$$

where $S_3(x)$ denotes the set of solutions of the following lower level program:

$$\max_{y \in Y} \ \min\{g(y), 1 - f(x, y)\}. \tag{1.11}$$

- Model IV:

$$\max_{x \in X} \ \min_{y \in S_4(x)} \ f(x, y), \tag{1.12}$$

where $S_4(x)$ denotes the set of solutions of the following lower level program:

$$\max_{y \in Y} \ \min\{1 - g(y), 1 - f(x, y)\}. \tag{1.13}$$

Let us give a brief introduction of Models I, II, III and IV in the following. In these models, x represents a decision alternative and y is a scenario. For each decision alternative x, the lower level program is to seek a suitable scenario y which has a relatively high $g(y)$ and relatively high $f(x, y)$ (Model I); or a relatively low $g(y)$ and relatively high $f(x, y)$ (Model II); or a relatively high $g(y)$ and relatively low $f(x, y)$ (Model III); or a relatively low $g(y)$ and relatively low $f(x, y)$ (Model IV). The sought scenario is called the focus point of x. For the case that there exist multiple focus points of x in the lower level program, the upper level program is to find the optimal decision alternative to make f maximize in an optimistic way (Models I, II), or in a pessimistic way (Models III, IV).

Clearly, the OSDT based decision Models I, II, III and IV are some special BLPPs which are difficult to be solved because they include nonconvex nonsmooth lower level programs. Although the optimal value function based method can generally reformulate the bilevel programming problem as a single level optimization problem, nevertheless solving the equivalent single level optimization problem (1.5) is still difficult. First, (1.5) is a nonsmooth optimization problem since the optimal value function is usually nonsmooth even when the objective function of the lower level program is smooth. Second, the commonly used nonsmooth MFCQ for single level optimization problems will never be satisfied because the inequality constraint of (1.5) is actually an equality constraint, and hence there is no guarantee that the optimal solution of (1.5) is a stationary point of (1.5). Last but the most important, (1.5) cannot be solved directly by utilizing the approaches to the general nonsmooth optimization problems since the optimal value function that occurs in the inequality constraint

usually cannot be expressed by an explicit one. In other words, (1.5) is not a traditional single level optimization problem, it remains difficulties to be solved.

This research is the first attempt to overcome some difficulties of (1.5). By taking into account the characteristics of the OSDT based models and the optimal value function based method, we succeed in translating Models I–II into general single level optimization problems so that they can be solved by the commonly used optimization approaches, and Models III–IV into min–max optimization problems with some assumptions. In order to solve the equivalent min–max optimization problems, we propose a class of regularization methods via approximating the maximum function by using a family of maximum entropy functions. Finally, we apply the proposed methods to newsvendor problems and use a numerical example to show their effectiveness.

The remainder of this paper is organized as follows. In Sect. 2, the equivalent forms of Models I, II, III and IV are proposed. In Sect. 3, newsvendor models are analyzed using the proposed methods and a numerical example is used to demonstrate the proposed approaches. Finally, we conclude our research in Sect. 4.

2 The solutions of Models I, II, III and IV

In this section, we solve Models I, II, III and IV by translating them into general single level optimization problems or min–max optimization problems. For this purpose, we first give the following definition and assumption that will be used.

Definition 2.1 (Stephen and Vandenberghe 2004) Let C be a convex set and let $f : C \to \mathbb{R}$ be a continuous function.

- f is called quasi-concave if for all x, $y \in C$ and $\lambda \in [0, 1]$, we have

$$F\big(\lambda x + (1 - \lambda)y\big) \geq \min \big\{F(x), F(y)\big\}.$$

- f is called strictly quasi-concave if for all x, $y \in C$ where $y \neq x$ and $\lambda \in (0, 1)$, we have

$$F\big(\lambda x + (1 - \lambda)y\big) > \min \big\{F(x), F(y)\big\}.$$

- f is called (strictly) quasi-convex if $-f$ is (strictly) quasi-concave.

Assumption 2.1 For the functions $f(x, y)$ and $g(y)$ given in Models I–IV, we assume that

- for any $x \in X$, $f(x, y)$ and $g(y)$ are quasi-concave for the variable y in Y;
- $g(y_l) = g(y_u) = 0$ and there exists $y_c \in (y_l, y_u)$ such that $g(y_c) = 1$.

2.1 Equivalent model of Model I

In this subsection, we reformulate Model I as a general single level optimization problem.

Theorem 2.1 *With Assumption* 2.1, *the global optimal solutions of Model I are equivalent to the ones of the following optimization problem:*

$$
\begin{aligned}
\max \quad & f(x, y) \\
\text{s.t.} \quad & f(x, y) - g(y) \leq 0, \\
& x \in X, \ y \in Y.
\end{aligned} \tag{2.1}
$$

Proof Let $\bar{x} \in X$ and suppose $\bar{y} \in S_1(\bar{x})$, that is, $\bar{y}$ is one of the global optimal solutions of the following optimization problem:

$$
\max_{y \in Y} \ \min\{g(y), f(\bar{x}, y)\}. \tag{2.2}
$$

We divide the difference of $f(\bar{x}, \bar{y})$ and $g(\bar{y})$ into two cases, that is, $g(\bar{y}) - f(\bar{x}, \bar{y}) \leq 0$ and $g(\bar{y}) - f(\bar{x}, \bar{y}) > 0$. In the first case, that is, $g(\bar{y}) \leq f(\bar{x}, \bar{y})$, we have

$$
g(\bar{y}) = \min\{g(\bar{y}), f(\bar{x}, \bar{y})\} \geq \min\{g(y), f(\bar{x}, y)\}, \quad \forall \, y \in Y, \tag{2.3}
$$

which implies that $\bar{y}$ is a global optimal solution of the following optimization problem:

$$
\begin{aligned}
\max \quad & g(y) \\
\text{s.t.} \quad & g(y) \leq f(\bar{x}, y), \ y \in Y.
\end{aligned} \tag{2.4}
$$

In fact, the inequality constraint of (2.4) is actually an equality constraint at the global optimal solutions, that is, $g(\bar{y}) = f(\bar{x}, \bar{y})$. The reason is as follows. Based on the fact that $f(x, y), \ g(y) \in [0, 1]$ and Assumption 2.1, we have

$$
g(y_c) = \max_{y \in Y} g(y) = 1 \geq f(\bar{x}, \bar{y}). \tag{2.5}
$$

Clearly, if $g(\bar{y}) < f(\bar{x}, \bar{y})$, then $\bar{y} \neq y_c$. Considering the continuities of the functions $g(y)$ and $f(\bar{x}, y)$ as well as (2.5), we know that there exists $y_{\bar{x}} \in (\bar{y}, y_c)$ or $y_{\bar{x}} \in (y_c, \bar{y})$ such that

$$
g(\bar{y}) < g(y_{\bar{x}}) = f(\bar{x}, y_{\bar{x}}) < f(\bar{x}, \bar{y}), \tag{2.6}
$$

which conflicts with the assumption that $\bar{y}$ is the global optimal solution of (2.4), so we have $g(\bar{y}) = f(\bar{x}, \bar{y})$ in the first case. Combined with the second case, that is, $g(\bar{y}) > f(\bar{x}, \bar{y})$, we can easily understand that the global optimal solutions of (2.2) must satisfy $g(\bar{y}) \geq f(\bar{x}, \bar{y})$.

Define the optimal value function of the lower level program problem (1.7) as $v_1(x)$, that is, $v_1(x) := \max_{y \in Y} \min\{g(y), f(x, y)\}$. From the above analysis, we have

$$
v_1(x) = \max_{y \in Y}\{f(x, y) \mid f(x, y) - g(y) \leq 0\}, \tag{2.7}
$$

then BLPP (1.6)–(1.7) can be reformulated as the following optimization problem:

$$\max_{x \in X} \max_{y \in Y} \{f(x, y) \mid f(x, y) - g(y) \leq 0\}. \tag{2.8}$$

Clearly, (2.8) and (2.1) are equivalent. □

In the following, we will give another equivalent form of Model I by considering its first order optimality condition. For this purpose, we give another assumption as follows.

Assumption 2.2 For the functions $f(x, y)$ and $g(y)$ given in Models I–IV, we assume that

- for any $x \in X$, $f(x, y)$ is concave and $g(y)$ is quasi-concave for the variable y in Y;
- $g(y_l) = g(y_u) = 0$ and there exists $y_c \in (y_l, y_u)$ such that $g(y_c) = 1$;
- $g'(y) > 0$ for all $y \in (y_l, y_c)$ and $g'(y) < 0$ for all $y \in (y_c, y_u)$.

It should be noted that the first order condition is necessary and sufficient for optimality for convex optimization problems. This claim still holds for quasi-convex optimization problems if the first order condition is satisfied only at the global optimal solutions [see Sects. 3.4, 4.2 of Stephen and Vandenberghe (2004)]. It follows from Assumption 2.2 that $g'(y) \neq 0$ for all $y \in (y_l, y_c) \cup (y_c, y_u)$, that is, $g(y)$ has a unique maximum. Hence solving the lower level program problem (1.7) is equivalent to solving its first order optimality condition, namely, finding out $y \in (y_l, y_u)$ such that

$$0 \in \partial_y \min\{g(y), f(x, y)\}, \tag{2.9}$$

where $\partial_y \min\{g(y), f(x, y)\}$ denotes the subdifferential of $\min\{g(y), f(x, y)\}$ at the point y (Rockafellar and Wets 1998), that is,

$$\partial_y \min\{g(y), f(x, y)\} := \begin{cases} \{g'(y)\} & \text{if } f(x, y) > g(y), \\ \{f'_y(x, y)\} & \text{if } f(x, y) < g(y), \\ \{\lambda g'(y) + (1 - \lambda) f'_y(x, y) \mid \lambda \in [0, 1]\} & \text{if } f(x, y) = g(y). \end{cases} \tag{2.10}$$

Further, we can rewrite the first order optimality condition (2.9) as

$$\begin{cases} g'(y) = 0, & \text{if } f(x, y) - g(y) > 0; \\ f'_y(x, y) = 0, & \text{if } f(x, y) - g(y) < 0; \\ g'(y) f'_y(x, y) \leq 0, & \text{if } f(x, y) - g(y) = 0. \end{cases} \tag{2.11}$$

It is trivial to verify that the first item of (2.11) does not hold under Assumption 2.2, which implies $S_1(x)$ can be denoted by

$$S_1(x) = \{y \in (y_l, y_u) \mid f'_y(x, y) = 0, \ f(x, y) - g(y) < 0\} \cup$$
$$\{y \in (y_l, y_u) \mid g'(y) f'_y(x, y) \leq 0, \ f(x, y) - g(y) = 0\}. \tag{2.12}$$

In the following, we will appropriately expand $S_1(x)$ to the following set

$$\overline{S_1}(x) := \left\{ y \in (y_l, y_u) \mid g'(y) f_y'(x, y) \leq 0, \ f(x, y) - g(y) \leq 0 \right\}. \quad (2.13)$$

Clearly, for any $x \in X$, it is easy to check that $S_1(x)$ is just a proper subset of $\overline{S_1}(x)$, and the difference of them can be given by

$$\overline{S_1}(x) - S_1(x) = \left\{ y \in (y_l, y_u) \mid g'(y) f_y'(x, y) \leq 0, \ f(x, y) - g(y) < 0, \ f_y'(x, y) \neq 0 \right\}. \quad (2.14)$$

Theorem 2.2 *With Assumption 2.2, the global optimal solutions of Model I are equivalent to the ones of the following optimization problem:*

$$\begin{aligned}
\max \quad & f(x, y) \\
\text{s.t.} \quad & g'(y) f_y'(x, y) \leq 0, \\
& f(x, y) - g(y) \leq 0, \\
& x \in X, \ y \in Y.
\end{aligned} \quad (2.15)$$

Proof Consider the following optimization problem:

$$\begin{aligned}
\max \quad & f(x, y), \\
\text{s.t.} \quad & x \in X, \ y \in \overline{S_1}(x),
\end{aligned} \quad (2.16)$$

where $\overline{S_1}(x)$ is given by (2.13). Taking the conditions $f(x, y) \leq g(y)$ and $g(y_l) = g(y_u) = 0$ into account, we know that (2.16) and (2.15) are equivalent. To prove the global optimal solutions of Model I and (2.15) are equivalent, it suffices to show that, for any $x \in X$, it holds that

$$f(x, \bar{y}) > f(x, \tilde{y}), \quad \forall \bar{y} \in S_1(x), \ \tilde{y} \in \overline{S_1}(x) - S_1(x). \quad (2.17)$$

Let us prove (2.17) in what follows. First of all, it follows from (2.12) and (2.14) that

$$f_y'(x, \bar{y}) = 0, \ g(\bar{y}) - f(x, \bar{y}) > 0; \ \text{or} \ g'(\bar{y}) f_y'(x, \bar{y}) \leq 0, \ g(\bar{y}) - f(x, \bar{y}) = 0, \quad (2.18)$$

and

$$g'(\tilde{y}) f_y'(x, \tilde{y}) \leq 0, \ f_y'(x, \tilde{y}) \neq 0, \ g(\tilde{y}) - f(x, \tilde{y}) > 0, \quad (2.19)$$

respectively. In the following, we discuss the details.

In the case that $f_y'(x, \bar{y}) = 0$ and $g(\bar{y}) - f(x, \bar{y}) > 0$, if $f_y'(x, \tilde{y}) < 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, then $f(x, y)$ is a decreasing function in variable y at the interval $[\bar{y}, \tilde{y}]$, which implies (2.17). If $f_y'(x, \tilde{y}) > 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, then $f(x, y)$ is an increasing function in variable y at the interval $[\tilde{y}, \bar{y}]$, which implies (2.17).

In the case that $g'(\bar{y}) = 0$ and $g(\bar{y}) - f(x, \bar{y}) = 0$, we can obtain $f(x, \bar{y}) = g(\bar{y}) = 1$. Combined with $f(x, \tilde{y}) < g(\tilde{y}) \leq 1$, we have (2.17).

In the case that $f'_y(x, \bar{y}) = 0$ and $g(\bar{y}) - f(x, \bar{y}) = 0$, we know $\bar{y} = \text{argmax}_{y \in Y} f(x, y)$. Combined with $f'_y(x, \tilde{y}) \neq 0$, we have (2.17).

In the case that $f'_y(x, \bar{y}) < 0$, $g'(\bar{y}) > 0$ and $g(\bar{y}) - f(x, \bar{y}) = 0$, if $f'_y(x, \tilde{y}) < 0$, $g'(\tilde{y}) \geq 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, then $f(x, y)$ is a decreasing function in variable y at the interval $[\bar{y}, \tilde{y}]$, which implies (2.17). If $f'_y(x, \tilde{y}) > 0$, $g'(\tilde{y}) \leq 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, then $f'_y(x, \tilde{y}) > 0 > f'_y(x, \bar{y})$ and $g'(\tilde{y}) \leq 0 < g'(\bar{y})$ which conflict with the definitions of these two functions.

In the case that $f'_y(x, \bar{y}) > 0$, $g'(\bar{y}) < 0$ and $g(\bar{y}) - f(x, \bar{y}) = 0$, if $f'_y(x, \tilde{y}) > 0$, $g'(\tilde{y}) \leq 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, then $f(x, y)$ is an increasing function in variable y at the interval $[\tilde{y}, \bar{y}]$, which implies (2.17). If $f'_y(x, \tilde{y}) < 0$, $g'(\tilde{y}) \geq 0$ and $g(\tilde{y}) - f(x, \tilde{y}) > 0$ hold, it is easy to verify that this result conflicts with the definitions of these two functions.

From the above analysis, we know that (2.17) holds for any $x \in X$, which implies the global optimal solutions of the BLPP (1.6)–(1.7) and (2.15) are equivalent. □

By comparing models (2.15) and (2.1), we can find that the feasible region of (2.15) is a subset of (2.1), (2.15) further indicates that the global optimal solutions of Model I must occur not only in the case that $f(x, y) \leq g(y)$ but also in the case that $g'(y) = 0$, $f'_y(x, y) = 0$ or $g'(y) f'_y(x, y) < 0$.

2.2 Equivalent model of Model II

In this subsection, we reformulate Model II as a general single level optimization problem.

Theorem 2.3 *With Assumption* 2.1, *the global optimal solutions of Model II are equivalent to the ones of the following optimization problem:*

$$
\begin{aligned}
\max \quad & f(x, y) \\
\text{s.t.} \quad & f(x, y) + g(y) - 1 \leq 0, \\
& x \in X, \ y \in Y.
\end{aligned} \tag{2.20}
$$

Proof Let $\bar{x} \in X$ and suppose $\bar{y} \in S_2(\bar{x})$, that is, $\bar{y}$ is one of the global optimal solutions of the following optimization problem:

$$
\max_{y \in Y} \ \min\{1 - g(y), f(\bar{x}, y)\}. \tag{2.21}
$$

We divide the difference of $f(\bar{x}, \bar{y})$ and $1 - g(\bar{y})$ into two cases, that is, $1 - g(\bar{y}) < f(\bar{x}, \bar{y})$ and $1 - g(\bar{y}) \geq f(\bar{x}, \bar{y})$. In fact, the first case is impossible. The reason is as follows. If the first case holds, that is, $1 - g(\bar{y}) < f(\bar{x}, \bar{y})$, then we have

$$
1 - g(\bar{y}) = \min\{1 - g(\bar{y}), f(\bar{x}, \bar{y})\} \geq \min\{1 - g(y), f(\bar{x}, y)\}, \quad \forall y \in Y, \tag{2.22}
$$

which implies that $\bar{y}$ is a global optimal solution of the following optimization problem:

$$
\begin{aligned}
\max \quad & 1 - g(y) \\
\text{s.t.} \quad & 1 - g(y) < f(\bar{x}, y), \ y \in Y.
\end{aligned}
\tag{2.23}
$$

Based on the fact that $f(x, y), \ 1 - g(y) \in [0, 1]$ and Assumption 2.1, we have

$$
1 - g(y_l) = 1 - g(y_u) = \max_{y \in Y} 1 - g(y) = 1 \geq f(\bar{x}, \bar{y}).
\tag{2.24}
$$

It is clear that $y_l < \bar{y} < y_u$. Considering the continuities of the functions $1 - g(y)$ and $f(\bar{x}, y)$ as well as (2.24), we know that there exists $y_{\bar{x}} \in (y_l, \bar{y})$ or $y_{\bar{x}} \in (\bar{y}, y_u)$ such that

$$
1 - g(\bar{y}) < 1 - g(y_{\bar{x}}) = f(\bar{x}, y_{\bar{x}}) < f(\bar{x}, \bar{y}),
\tag{2.25}
$$

which conflicts with the assumption that $\bar{y}$ is the global optimal solution of (2.23), so we can easily understand that the global optimal solutions of (2.21) must satisfy $1 - g(\bar{y}) \geq f(\bar{x}, \bar{y})$.

Define the optimal value function of the lower level program problem (1.9) as $v_2(x)$, that is, $v_2(x) := \max_{y \in Y} \min\{1 - g(y), f(x, y)\}$. From the above analysis, we have

$$
v_2(x) = \max_{y \in Y} \{f(x, y) \mid f(x, y) + g(y) - 1 \leq 0\},
\tag{2.26}
$$

then BLPP (1.8)–(1.9) can be reformulated as the following optimization problem:

$$
\max_{x \in X} \ \max_{y \in Y} \{f(x, y) \mid f(x, y) + g(y) - 1 \leq 0\}.
\tag{2.27}
$$

Clearly, (2.27) and (2.20) are equivalent.   □

2.3 Equivalent model of Model III

In this subsection, we translate Model III into a min–max optimization problem. Moreover, we propose a class of regularization methods to solve it.

Theorem 2.4 *With Assumption 2.1, the global optimal solutions of Model III are equivalent to the ones of the following continuous min–max optimization problem:*

$$
\min_{x \in X} \ \max_{y \in Y} \ \{-f(x, y) \mid 1 - f(x, y) - g(y) = 0\}.
\tag{2.28}
$$

Proof Let $\bar{x} \in X$ and suppose $\bar{y} \in S_3(\bar{x})$, that is, $\bar{y}$ is one of the global optimal solutions of the following optimization problem:

$$
\max_{y \in Y} \ \min\{g(y), 1 - f(\bar{x}, y)\}.
\tag{2.29}
$$

Under Assumption 2.1, the global optimal solution $\bar{y} \in S_3(\bar{x})$ must satisfy

$$g(\bar{y}) = 1 - f(\bar{x}, \bar{y}). \tag{2.30}$$

The reason is as follows.

If (2.30) does not hold, we divide the difference of $g(\bar{y})$ and $1 - f(\bar{x}, \bar{y})$ into two cases, that is, $g(\bar{y}) < 1 - f(\bar{x}, \bar{y})$ and $g(\bar{y}) > 1 - f(\bar{x}, \bar{y})$. If $g(\bar{y}) < 1 - f(\bar{x}, \bar{y})$ holds, then we have

$$g(\bar{y}) = \min\{g(\bar{y}), 1 - f(\bar{x}, \bar{y})\} \geq \min\{g(y), 1 - f(\bar{x}, y)\}, \quad \forall\, y \in Y, \tag{2.31}$$

which implies that $\bar{y}$ is a global optimal solution of the following optimization problem:

$$
\begin{aligned}
\max \quad & g(y) \\
\text{s.t.} \quad & g(y) < 1 - f(\bar{x}, y), \ y \in Y.
\end{aligned}
\tag{2.32}
$$

Based on the fact that $1 - f(x, y), \ g(y) \in [0, 1]$ and Assumption 2.1, we have

$$g(y_c) = \max_{y \in Y} g(y) = 1 \geq 1 - f(\bar{x}, \bar{y}). \tag{2.33}$$

It is clear that $\bar{y} \neq y_c$. Considering the continuities of functions $g(y)$ and $1 - f(\bar{x}, y)$ as well as (2.33), we know there exists $y_{\bar{x}} \in (\bar{y}, y_c)$ or $y_{\bar{x}} \in (y_c, \bar{y})$ such that

$$g(\bar{y}) < g(y_{\bar{x}}) \leq 1 - f(\bar{x}, y_{\bar{x}}) < 1 - f(\bar{x}, \bar{y}), \tag{2.34}$$

which is not consistent with the assumption that $\bar{y}$ is the global optimal solution of (2.32). On the other hand, if $g(\bar{y}) > 1 - f(\bar{x}, \bar{y})$ holds, then we have

$$1 - f(\bar{x}, \bar{y}) = \min\{g(\bar{y}), 1 - f(\bar{x}, \bar{y})\} \geq \min\{g(y), 1 - f(\bar{x}, y)\}, \quad \forall\, y \in Y, \tag{2.35}$$

which implies that $\bar{y}$ is a global optimal solution of the following problem:

$$
\begin{aligned}
\max \quad & 1 - f(\bar{x}, y) \\
\text{s.t.} \quad & 1 - f(\bar{x}, y) < g(y), \ y \in Y.
\end{aligned}
\tag{2.36}
$$

Based on the fact that $1 - f(x, y), \ g(y) \in [0, 1]$ and Assumption 2.1, we have

$$1 - f(\bar{x}, y) \geq g(y_l) = g(y_u) = 0, \quad \forall\, y \in Y. \tag{2.37}$$

It is clear that $y_l < \bar{y} < y_u$. Considering the continuities of functions $g(y)$ and $1 - f(\bar{x}, y)$ as well as (2.33), we know that there exists $y_{\bar{x}} \in (y_l, \bar{y})$ or $y_{\bar{x}} \in (\bar{y}, y_u)$ such that

$$1 - f(\bar{x}, \bar{y}) < 1 - f(\bar{x}, y_{\bar{x}}) \leq g(y_{\bar{x}}) < g(\bar{y}), \tag{2.38}$$

which is not consistent with the assumption that $\bar{y}$ is the global optimal solution of (2.36).

Define the optimal value function of the lower level program problem (1.11) as $v_3(x)$, that is, $v_3(x) := \max_{y \in Y} \min\{g(y), 1 - f(x, y)\}$. From the above analysis, we have

$$v_3(x) = \max_{y \in Y}\{1 - f(x, y) \mid 1 - f(x, y) - g(y) = 0\}, \tag{2.39}$$

then BLPP (1.10)–(1.11) can be reformulated as the following optimization problem:

$$\max_{x \in X} \min_{y \in Y}\{f(x, y) \mid 1 - f(x, y) - g(y) = 0\}. \tag{2.40}$$

Clearly, the global optimal solutions of (2.40) and (2.28) are equivalent.   □

In what follows, we translate the continuous min–max problem (2.28) as a discrete min–max problem. First of all, it follows from Assumption 2.1 that, for any $x \in X$, we have

$$1 - f(x, y_l) - g(y_l) \geq 0, \ 1 - f(x, y_c) - g(y_c) \leq 0, \ 1 - f(x, y_u) - g(y_u) \geq 0, \tag{2.41}$$

so it is easy to understand that solving the equation $1 - f(x, y) - g(y) = 0$ is equivalent to finding out $y \in \{y_1, y_2\}$ such that

$$1 - f(x, y_1) - g(y_1) = 0, \ y_1 \in [y_l, y_c]; \ 1 - f(x, y_2) - g(y_2) = 0, \ y_2 \in [y_c, y_u]. \tag{2.42}$$

Denote $Y_1 := [y_l, y_c]$ and $Y_2 := [y_c, y_u]$, then (2.28) can be rewritten as the following optimization problem

$$\begin{aligned} \min \quad & \max\{-f(x, y_1), -f(x, y_2)\} \\ \text{s.t.} \quad & x \in X, \ y_1 \in Y_1, \ y_2 \in Y_2, \\ & 1 - f(x, y_i) - g(y_i) = 0, \ i = 1, 2. \end{aligned} \tag{2.43}$$

In what follows, we design a class of regularization methods to solve (2.43) by using a smoothing function to approximate the maximum function. For a sufficiently large $k > 0$, we define the maximum entropy function (Li and Fang 1997) as

$$\varphi_k(x, y_1, y_2) := k^{-1} \ln\left(\exp\left(-kf(x, y_1)\right) + \exp\left(-kf(x, y_2)\right)\right). \tag{2.44}$$

Denote $\varphi(x, y_1, y_2) := \max\{-f(x, y_1), -f(x, y_2)\}$, it follows from the study (Li and Fang 1997) that $\{\varphi_k : k = 1, 2, \ldots\}$ is a family of smoothing approximations for the maximum function φ, and it holds that

$$0 \leq \varphi_k(x, y_1, y_2) - \varphi(x, y_1, y_2) \leq k^{-1} \ln 2. \tag{2.45}$$

Thus for a sufficiently large $k > 0$, a smoothing approximation of (2.43) can be given as:

$$\begin{aligned}
\min \quad & \varphi_k(x, y_1, y_2) \\
\text{s.t.} \quad & x \in X, \ y_1 \in Y_1, \ y_2 \in Y_2, \\
& 1 - f(x, y_i) - g(y_i) = 0, \ i = 1, 2.
\end{aligned} \tag{2.46}$$

Obviously, (2.46) is a conventional single level optimization problem which can be solved by the commonly used optimization methods. The following theorem examines the convergence of the global optimal solutions of (2.46) as $k \to \infty$.

Theorem 2.5 *Let $k = 1, 2, \dots$ and suppose (x^k, y_1^k, y_2^k) is a global optimal solution of (2.46). If (x^*, y_1^*, y_2^*) is a limit point of the sequence (x^k, y_1^k, y_2^k), then (x^*, y_1^*, y_2^*) is a global optimal solution of (2.43).*

Proof Denote by $\mathcal{F}$ the feasible set of (2.43) and (2.46). Considering the boundedness of X, Y_1 and Y_2, we know that $\mathcal{F}$ is also a closed set, which implies that the limit point of the sequence (x^k, y_1^k, y_2^k) is still feasible for (2.43) and (2.46), that is,

$$(x^*, y_1^*, y_2^*) \in \mathcal{F}.$$

Since (x^k, y_1^k, y_2^k) is the global optimal solution of (2.46) for each $k = 1, 2, \dots$, we have

$$\varphi_k(x^k, y_1^k, y_2^k) - \varphi_k(x, y_1, y_2) \le 0, \quad \forall \, (x, y_1, y_2) \in \mathcal{F}, \ \forall \, k.$$

Let $k \to \infty$, considering the continuity of $\varphi_k(x, y_1, y_2)$ and the compactness of $\mathcal{F}$, we have

$$\varphi(x^*, y_1^*, y_2^*) - \varphi(x, y_1, y_2) \le 0, \quad \forall \, (x, y_1, y_2) \in \mathcal{F},$$

which implies (x^*, y_1^*, y_2^*) is a global optimal solution of (2.43). $\qquad\square$

2.4 Equivalent model of Model IV

In this subsection, we translate Model IV into a min–max optimization problem. Moreover, we propose a class of regularization methods to solve it.

Theorem 2.6 *With Assumption* 2.1*, the global optimal solutions of Model IV are equivalent to the ones of the following continuous min–max optimization problem:*

$$\min_{x \in X} \ \max_{y \in Y} \ \{-f(x, y) \mid g(y) - f(x, y) \le 0\}. \tag{2.47}$$

Proof Let $\bar{x} \in X$ and suppose $\bar{y} \in S_4(\bar{x})$, that is, $\bar{y}$ is one of the global optimal solutions of the following optimization problem:

$$\max_{y \in Y} \ \min\{1 - g(y), 1 - f(\bar{x}, y)\}. \tag{2.48}$$

In fact, under Assumption 2.1, the global optimal solution $\bar{y} \in S_4(\bar{x})$ must satisfy:

$$1 - g(\bar{y}) \geq 1 - f(\bar{x}, \bar{y}). \tag{2.49}$$

The reason is given as follows.

If (2.49) does not hold, that is, $1 - g(\bar{y}) < 1 - f(\bar{x}, \bar{y})$, then we have

$$1 - g(\bar{y}) = \min\{1 - g(\bar{y}), 1 - f(\bar{x}, \bar{y})\} \geq \min\{1 - g(y), 1 - f(\bar{x}, y)\}, \quad \forall \, y \in Y, \tag{2.50}$$

which implies that $\bar{y}$ is a global optimal solution of the following optimization problem:

$$\begin{aligned} \max \quad & 1 - g(y) \\ \text{s.t.} \quad & 1 - g(y) < 1 - f(\bar{x}, y), \ y \in Y. \end{aligned} \tag{2.51}$$

Based on the fact that $1 - f(x, y), \ 1 - g(y) \in [0, 1]$ and Assumption 2.1, for any $y \in Y$, we have

$$\max_{y \in Y} 1 - g(y) = 1 - g(y_l) = 1 - g(y_u) = 1 \geq 1 - f(\bar{x}, y). \tag{2.52}$$

It is clear that $y_l < \bar{y} < y_u$. Considering the continuities of the functions $1 - g(y)$ and $1 - f(\bar{x}, y)$ as well as (2.52), we know that there exists $y_{\bar{x}} \in (y_l, \bar{y})$ or $y_{\bar{x}} \in (\bar{y}, y_u)$ such that

$$1 - g(\bar{y}) < 1 - g(y_{\bar{x}}) \leq 1 - f(\bar{x}, y_{\bar{x}}) < 1 - f(\bar{x}, \bar{y}), \tag{2.53}$$

which is not consistent with the assumption that $\bar{y}$ is the global optimal solution of (2.51)

Define the optimal value function of the lower level program problem (1.13) as $v_4(x)$, that is, $v_4(x) := \max_{y \in Y} \min\{1 - g(y), 1 - f(x, y)\}$. From the above analysis, we have

$$v_4(x) = \max_{y \in Y} \{1 - f(x, y) \mid g(y) - f(x, y) \leq 0\}, \tag{2.54}$$

then BLPP (1.12)–(1.13) can be reformulated as the following optimization problem:

$$\max_{x \in X} \ \min_{y \in Y} \{1 - f(x, y) \mid g(y) - f(x, y) \leq 0\}. \tag{2.55}$$

Clearly, the global optimal solutions of (2.55) and (2.47) are equivalent. $\qquad\square$

In what follows, we translate the continuous min–max problem (2.47) as a discrete min–max problem. First of all, from Assumption 2.1, for any $x \in X$, we have

$$g(y_l) = 0 \leq f(x, y_l), \quad g(y_u) = 0 \leq f(x, y_u), \tag{2.56}$$

and

$$\max\{-f(x, y_l), -f(x, y_u)\} = \max_{y \in Y} -f(x, y) \geq -f(x, y), \ \forall \, y \in Y. \tag{2.57}$$

Moreover, if $f(x, y)$ is strictly quasi-concave in variable y, we can further obtain

$$\max\{-f(x, y_l), -f(x, y_u)\} > -f(x, y), \ \forall \, y \in (y_l, y_u). \tag{2.58}$$

So it is easy to understand that (2.47) can be rewritten as the following discrete min–max problem:

$$\min_{x \in X} \ \max\{-f(x, y_l), -f(x, y_u)\}. \tag{2.59}$$

Like (2.46), we design a class of regularization methods to solve (2.59) by using the maximum entropy function to approximate the maximum function. For a sufficiently large $k > 0$, we define the maximum entropy function as

$$\phi_k(x, y_l, y_u) := k^{-1} \ln \Big(\exp \big(- kf(x, y_l) \big) + \exp \big(- kf(x, y_u) \big) \Big). \tag{2.60}$$

Denote $\phi(x, y_l, y_u) := \max\{-f(x, y_l), -f(x, y_u)\}$, it follows from the study (Li and Fang 1997) that

$$0 \leq \phi_k(x, y_l, y_u) - \phi(x, y_l, y_u) \leq k^{-1} \ln 2. \tag{2.61}$$

Thus for a sufficiently large $k > 0$, a smoothing approximation of (2.59) can be given as:

$$\min_{x \in X} \ \phi_k(x, y_l, y_u). \tag{2.62}$$

Obviously, (2.62) is a conventional single level optimization problem which can be easily solved with the commonly used optimization methods. Like Theorem 2.5, the following theorem examines the convergence of the global optimal solutions of (2.62) as $k \to \infty$. We omit the proof here.

Theorem 2.7 *Let $k = 1, 2, \ldots$ and suppose x^k is a global optimal solution of (2.62). If x^* is a limit point of the sequence x^k, then x^* is a global optimal solution of (2.59).*

3 Applications to newsvendor problems

Consider a retailer who orders a product prior to the selling season. Suppose the demand y is distributed normally with mean μ and variance σ^2. Denote the set of demand as $D := [\mu - y_0, \mu + y_0]$ with $y_0 > 0$. We have the following normalized probability density function

$$g(y) := \frac{\rho(y) - \rho_l}{\rho_u - \rho_l}, \quad y \in D, \tag{3.1}$$

where $\rho(y)$ is the original probability density function, that is,

$$\rho(y) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right), \quad y \in D, \tag{3.2}$$

ρ_l and ρ_u respectively denote the lower and upper bounds of $\rho(y)$ in D, that is,

$$\rho_u = \rho(\mu) = \frac{1}{\sqrt{2\pi}\sigma}, \quad \rho_l = \rho(y_l) = \rho(y_u) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y_0^2}{2\sigma^2}\right). \tag{3.3}$$

$g(y)$ is used to represent the relative likelihood degree of y. It is easy to check that $g(y)$ is strictly quasi-concave continuous and differentiable and satisfies

$$\max_{y \in D} g(y) = g(\mu) = 1, \quad \min_{y \in D} g(y) = g(\mu - y_0) = g(\mu + y_0) = 0. \tag{3.4}$$

Denote x as the retailer's order quantity, w is the unit wholesale price, r is the unit revenue with $r > w$. Any excess product can be salvaged at the unit salvage price $s_0 > 0$. If there is a shortage, the unit opportunity cost is $s_u > 0$. The profit function of the retailer is

$$p(x, y) := \begin{cases} ry + (x - y)s_0 - wx, & y < x; \\ (r - w)x - s_u(y - x), & y \geq x. \end{cases} \tag{3.5}$$

Because the set of uncertain demand is D, a reasonable order quantity should also lie in this region. Given an order quantity $x \in D$, the function for evaluating the order quantity with considering the regret of the retailer is given as

$$h(x, y) := -\left(p(x, y) - p_u(x)\right)^2, \tag{3.6}$$

where $p_u(x)$ denotes the highest profit for an order x, that is,

$$p_u(x) = \max_{y \in D} p(x, y) = (r - w)x. \tag{3.7}$$

Thus, we have

$$h(x, y) := \begin{cases} -(r - s_0)^2 (x - y)^2, & y < x; \\ -s_u^2 (x - y)^2, & y \geq x. \end{cases} \tag{3.8}$$

The satisfaction level of the retailer is represented by a satisfaction function, which is obtained by normalizing $h(x, y)$, that is,

$$f\big(h(x, y)\big) := \frac{h(x, y) - h_l}{h_u - h_l}, \tag{3.9}$$

where h_l and h_u are the lower and upper bounds of $h(x, y)$ in $D \times D$, respectively. Clearly, the highest value is $h_u = 0$, that is, the demand is equal to the order quantity. The lowest value is $h_l := \min\{-(r-s_0)^2(y_u-y_l)^2, \ -s_u^2(y_u-y_l)^2\}$. The satisfaction function is strictly increasing in v, and the lowest satisfaction level is 0 and the highest satisfaction level is 1. In addition, it is easy to prove that $f\big(h(x, y)\big)$ is a concave function in $D \times D$. For convenience, the satisfaction function is written as $f(x, y)$ in the following parts.

For newsvendor problems, a real demand that represents a scenario which will occur in the future and the retailer knows that there is usually only one opportunity to order products before the selling season because the procurement lead-time is usually longer than the selling season. Considering the one-time feature of newsvendor problems, it is reasonable to argue that the retailer needs to contemplate which demand ought to be taken into account before making the order. For each order quantity, the retailer chooses one demand amongst all possible ones while considering the satisfaction level caused by the occurrence of the demand and the relative likelihood degree of the demand occurring. The selected demand is called the focus point of the order quantity. We consider the following four types of focus points:

- The active focus point of an order quantity x, denoted as $y_1(x)$, is a demand that has a higher relative likelihood degree and a higher satisfaction level for an order quantity x, that is

$$y_1(x) \in D_1(x) := \text{argmax}_{y \in D} \ \min\{g(y), f(x, y)\}. \tag{3.10}$$

- The daring focus point of an order quantity x, denoted as $y_2(x)$, is a demand that has a lower relative likelihood degree and a higher satisfaction level for an order quantity x, that is

$$y_2(x) \in D_2(x) := \text{argmax}_{y \in D} \ \min\{1 - g(y), f(x, y)\}. \tag{3.11}$$

- The passive focus point of an order quantity x, denoted as $y_3(x)$, is a demand that has a higher relative likelihood degree and a lower satisfaction level for an order quantity x, that is

$$y_3(x) \in D_3(x) := \text{argmax}_{y \in D} \ \min\{g(y), 1 - f(x, y)\}. \tag{3.12}$$

- The apprehensive focus point of an order quantity x, denoted as $y_4(x)$, is a demand that has a lower relative likelihood degree and a lower satisfaction level for an order quantity x, that is

$$y_4(x) \in D_4(x) := \text{argmax}_{y \in D} \ \min\{1 - g(y), 1 - f(x, y)\}. \tag{3.13}$$

The retailer considers the focus point as his/her most appropriate scenario for each order quantity and chooses one order quantity which can bring about the highest satisfaction level with the assumption that the focus point comes true. Hence, the optimal order quantities are obtained as follows.
- The optimal active order quantity, denoted as x_1, is

$$x_1 \in \text{argmax}_{x \in D} \ \max_{y_1(x) \in D_1(x)} \ f\big(x, y_1(x)\big). \tag{3.14}$$

- The optimal daring order quantity, denoted as x_2, is

$$x_2 \in \text{argmax}_{x \in D} \ \max_{y_2(x) \in D_2(x)} \ f\big(x, y_2(x)\big). \tag{3.15}$$

- The optimal passive order quantity, denoted as x_3, is

$$x_3 \in \text{argmax}_{x \in D} \ \min_{y_3(x) \in D_3(x)} \ f\big(x, y_3(x)\big). \tag{3.16}$$

- The optimal apprehensive order quantity, denoted as x_4, is

$$x_4 \in \text{argmax}_{x \in D} \ \min_{y_4(x) \in D_4(x)} \ f\big(x, y_4(x)\big). \tag{3.17}$$

Clearly, the above newsvendor models (3.14) with (3.10), (3.15) with (3.11), (3.16) with (3.12) and (3.17) with (3.13) are special cases of Models I, II, III and IV, respectively. It follows from Sect. 2 that solving the models (3.14) with (3.10), (3.15) with (3.11), (3.16) with (3.12) and (3.17) with (3.13) are equivalent to solving the models (2.15), (2.20), (2.43) and (2.59) with $X = Y = D$, respectively.

We demonstrate the proposed approaches with the following example. A sports clothing store, located in Tokyo, is planning to order a new fashion sportswear before the selling season. The unit wholesale price w, the unit revenue r, the unit salvage price s_0 and the unit opportunity cost s_u are 6, 9, 4 and 3 (thousand JPY), respectively. The demand is distributed normally with mean 500 and variance 200^2. The range of the possible demand is [200, 800].

By using (3.1), the normalized probability density function is

$$g(y) := \frac{\exp\big(-(y - 500)^2/80000\big) - \exp(-9/8)}{1 - \exp(-9/8)}. \tag{3.18}$$

By using (3.8), we have

$$h(x, y) := \begin{cases} -25(x - y)^2, & y < x; \\ -9(x - y)^2, & y \geq x, \end{cases} \tag{3.19}$$

Table 1 Numerical results for the newsvendor example

Retailer	x	y	g(y)	f(x, y)
Active	498.9477	500.1752	1.0000	1.0000
Daring	800.0000	800.0000	0.0000	1.0000
Passive	431.8985	229.7410/770.0530	0.1135/0.1143	0.8865/0.8857
Apprehensive	424.4930	200.0000/800.0000	0.0000/0.0000	0.8600/0.8590

where $200 \leq x,\ y \leq 800$. The highest value is $h_u = 0$ and the lowest value is $h_l = -9000000$.

By using (3.9), the satisfaction function is obtained as

$$f(x, y) := \begin{cases} -\frac{1}{600^2}(x - y)^2 + 1, & y < x; \\ -\frac{1}{1000^2}(x - y)^2 + 1, & y \geq x. \end{cases} \tag{3.20}$$

In our experiments, we use the interior-point algorithm from Optimization Toolbox of MATLAB 7.10.0 to solve the models (2.15), (2.20), (2.46) and (2.62), and the starting points are the lower bound of the feasible regions. For approximation problems (2.46) and (2.62), we set the parameter k as 500. The numerical results are listed in Table 1.

The numerical results show that the optimal order quantities of the retailer satisfy $x_4 < x_3 < x_1 < x_2$. It is in perfect accordance with the situations occurred in the real world of business.

4 Conclusions

In this paper, we examine four types of bilevel programming problems where the lower level programs are max–min optimization problems and the upper level programs have max–max or max–min objective functions. The existing optimization methods may not be applicable for solving these problems because they include nonconvex nonsmooth lower level programs.

We translate these problems into conventional single level optimization problems or min–max optimization problems and prove that they have the same global optimal solutions under some assumptions. Furthermore, we propose a class of regularization methods to solve the equivalent min–max optimization problems by using a family of maximum entropy functions to approximate the maximum function. We also show that any limit points of the global optimal solutions obtained by the approximation methods are the same as the ones of the original problems. As an application, we utilize the proposed methods to the newsvendor problem and use a numerical example to show their effectiveness.

This research not only materializes such optimization problems with managerial meaning but also provides effective approaches to solve them. It is a good beginning to push forward the progress of the research on bilevel programming problems with

nonconvex nonsmooth lower level programs. In this paper, only one dimensional case is studied. How to generalize it into finite dimensional Euclidean spaces will be our future work.

References

Allende GB, Still G (2013) Solving bilevel programs with the KKT-approach. Math Program 138(1–2):309–332

Bard JF (1998) Practical bilevel optimization: algorithms and applications, vol 30. Springer, Berlin

Colson B, Marcotte P, Savard G (2005) Bilevel programming: a survey. 4OR-Q J Oper Res 3(2):87–107

Dempe S (2002) Foundations of bilevel programming. Springer, Berlin

Dempe S, Zemkoho AB (2013) The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. Math Program 138(1–2):447–473

Dempe S, Zemkoho AB (2014) KKT reformulation and necessary conditions for optimality in nonsmooth bilevel optimization. SIAM J Optim 24(4):1639–1669

Dempe S, Mordukhovich BS, Zemkoho AB (2012) Sensitivity analysis for two-level value functions with applications to bilevel programming. SIAM J Optim 22(4):1309–1343

Facchinei F, Jiang H, Qi L (1999) A smoothing method for mathematical programs with equilibrium constraints. Math Program 85(1):107–134

Fletcher R, Leyffer S, Ralph D, Scholtes S (2006) Local convergence of SQP methods for mathematical programs with equilibrium constraints. SIAM J Optim 17(1):259–286

Guo P (2010a) One-shot decision approach and its application to duopoly market. Int J Inf Decis Sci 2(3):213–232

Guo P (2010b) Private real estate investment analysis within one-shot decision framework. Int Real Estate Rev 13(3):238–260

Guo P (2011) One-shot decision theory. IEEE Trans On Syst Man Cybern A Syst Hum 41(5):917–926

Guo P, Li Y (2014) Approaches to multistage one-shot decision making. Eur J Oper Res 236(2):612–623

Guo P, Ma X (2014) Newsvendor models for innovative products with one-shot decision theory. Eur J Oper Res 239(2):523–536

Guo L, Lin GH, Ye JJ (2015) Solving mathematical programs with equilibrium constraints. J Optim Theory Appl 166(1):234–256

Li XS, Fang SC (1997) On the entropic regularization method for solving min–max problems with applications. Math Methods Oper Res 46(1):119–130

Li Y, Guo P (2015) Possibilistic individual multi-period consumption-investment models. Fuzzy Sets Syst 274:47–61

Lin GH, Fukushima M (2005) A modified relaxation scheme for mathematical programs with complementarity constraints. Ann Oper Res 133(1–4):63–84

Lin GH, Xu M, Ye JJ (2014) On solving simple bilevel programs with a nonconvex lower level program. Math Program 144(1–2):277–305

Luo ZQ, Pang JS, Ralph D (1996) Mathematical programs with equilibrium constraints. Cambridge University Press, Cambridge

Outrata JV (1990) On the numerical solution of a class of Stackelberg problems. Z Oper Res 34(4):255–277

Rockafellar RT, Wets RJB (1998) Variational analysis. Springer, Berlin

Scholtes S (2001) Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM J Optim 11(4):918–936

Stephen B, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge

Vicente LN, Calamai PH (1994) Bilevel and multilevel programming: a bibliography review. J Global Optim 5(3):291–306

Von Stackelberg H (1952) The theory of the market economy. Oxford University Press, Oxford

Wang C, Guo P (2017) Behavioral models for first-price sealed-bid auctions with the one-shot decision theory. Eur J Oper Res 261(3):994–1000

Xu M, Ye JJ (2014) A smoothing augmented Lagrangian method for solving simple bilevel programs. Comput Optim Appl 59(1–2):353–377

Ye JJ (2005) Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. J Math Anal Appl 307(1):350–369

Ye JJ, Zhu DL (1995) Optimality conditions for bilevel programming problems. Optimization 33(1):9–27
Ye JJ, Zhu DL (2010) New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches. SIAM J Optim 20(4):1885–1905
Zhu X, Lin GH (2016) Improved convergence results for a modified Levenberg–Marquardt method for nonlinear equations and applications in MPCC. Optim Methods Softw 31(4):791–804