

Pessimistic Bilevel Linear Optimization

S. Dempe¹, G. Luo², S. Franke³

¹ TU Bergakademie Freiberg, Germany, Institute of Numerical Mathematics and Optimization

² Guangdong University of Finance, China, Department of Applied Mathematics

³ TU Bergakademie Freiberg, Germany, Institute of Numerical Mathematics and Optimization

Correspondence to: S. Dempe, email: dempe@tu-freiberg.de

Abstract: *In this paper, we investigate the pessimistic bilevel linear optimization problem (PBLOP). Based on the lower level optimal value function and duality, the PBLOP can be transformed to a single-level while nonconvex and nonsmooth optimization problem. By use of linear optimization duality, we obtain a tractable and equivalent transformation and propose algorithms for computing global or local optimal solutions. One small example is presented to illustrate the feasibility of the method.*

Keywords: Pessimistic bilevel optimization, Bilevel linear optimization problem, Nonconvex optimization, Basic matrix, Optimal value function, Minimax problem, Local and global optimization, Solution algorithms, Enumeration algorithm, Descent algorithm.

1 Introduction

The bilevel optimization problem (BOP) is an optimization problem (called the upper level problem in what follows) whose constraint region is determined implicitly by the graph of the solution set mapping of another mathematical optimization problem (the lower level problem) [8]. It is a nonconvex and nonsmooth optimization problem. Therefore, solving BOP is not an easy task: even in the linear case, the problem is \mathcal{NP} -hard, see [15, 17]. However, the BOP plays an important role in different fields such as economics, transportation, engineering, supply chain planning *etc.*, see [5, 6, 7, 10, 12]. It attracts many researchers to investigate efficient solution approaches such as extreme-point search approaches, branch-and-bound approaches, descent methods, complementary pivoting methods, penalty function methods, trust region methods and so on, see *e.g.* the bibliography [9].

Referring to the work of H.v. Stackelberg [20], special cases of the bilevel optimization problem are called Stackelberg games of two players, the leader and the follower. Here, the players act according to an hierarchical order, the leader has the first choice, the follower reacts optimally on the leader's selection and the leader has to respect the follower's decision in finding an optimal decision. While the lower level problem is the follower's optimization problem, the leader solves the upper level problem.

If the lower level problem does not have a uniquely determined optimal solution, this problem is not well-posed since the objective function value of the upper level problem may depend on the optimal solution of the lower level problem, see [8]. Two ways out of this situation are the use of the optimistic or of the pessimistic approach. In the optimistic approach, the leader supposes that the follower is willing to support him, *i.e.* that the follower will select a solution out of the set of optimal solutions for the lower level problem which is best from the leader's point of view. In fact, the optimistic situation seems to be infeasible when the cooperation is not allowed or when the follower's seriousness of keeping the agreement is not granted. Then, it is necessary for the leader to consider the damage resulting from an undesirable selection of the follower, that is, considering a pessimistic situation is reasonable. There are some articles available where pessimistic bilevel optimization problems are studied, see [1, 13, 14, 21, 23, 24].

Dempe *et al.* [14] investigated several types of lower subdifferential necessary optimality conditions using the optimal value function approach, the Karush - Kuhn - Tucker representation of lower-level optimal solution maps and upper subdifferential necessary optimality conditions for pessimistic bilevel optimization. Wiesemann *et al.* [21, 23] presented existence conditions for optimal solutions of the pessimistic bilevel optimization problem and developed algorithms convergent to solutions of these problems with continuous functions. Aboussoror *et al.* [1] studied the pessimistic linear bilevel optimization problems via an exact penalty method. Recently, a combination of optimistic and pessimistic linear bilevel optimization problems have been explored in [24] which can be viewed as an extension of the approach described in [1].

In this paper, we consider the pessimistic linear bilevel optimization problem

$$\min_{x \in K} \max_{y \in \Psi(x)} \{a^\top x + b^\top y\} \quad (1.1)$$

where

$$\Psi(x) := \{z \in \mathbb{R}^m : Dz = d - Ax, z \geq 0, c^\top z = \varphi(x)\} \quad (1.2)$$

with

$$\varphi(x) := \min_y \{c^\top y : Dy = d - Ax, y \geq 0\} \quad (1.3)$$

denotes the solution-set mapping of the lower level optimization problem. The function $\varphi(\cdot)$ is called the optimal value function of the lower level problem.

Here, K is assumed to be a nonempty compact subset of \mathbb{R}^n , $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ denote upper and lower level decision variables, $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^m$, $d \in \mathbb{R}^l$ are constant vectors, $A \in \mathbb{R}^{l \times n}$, $D \in \mathbb{R}^{l \times m}$ are constant matrices and the symbol \top denotes the transposition of a vector or a matrix. Moreover,

$$\varphi_p(x) := \max_y \{b^\top y : y \in \Psi(x)\} \quad (1.4)$$

is the pessimistic two-level value function.

Then, (1.1) can be replaced by

$$\min_x \{a^\top x + \varphi_p(x) : x \in K\}. \quad (1.5)$$

In the subsequent sections, we focus on a tractable formulation and the existence of an optimal solution for problem (1.5).

The organization of the paper is as follows. Section 2 investigates the pessimistic bilevel linear optimization problem with linear constraints and obtains an equivalent single-level optimization formulation. Using this formulation, a simple algorithm for computing a global optimal solution is presented. The bilevel optimization problem is a nonconvex optimization problem. Hence, the calculation of local optimal solutions seems to be more practical. In Section 3, a descent algorithm is proposed for the computation of such a local optimum.

A small example is presented in Section 4 to explain the feasibility and reasonability of the descent algorithm. Section 5 contains some conclusions.

2 Global Optimal Solution of Pessimistic Bilevel Linear Optimization

By linear optimization sensitivity, the set valued mapping $x \mapsto \Psi(x)$ is a Lipschitz-continuous mapping [11, 16] and the function $x \mapsto \varphi(x)$ is Lipschitz-continuous, it is piecewise affine linear and convex, see [3]. This implies that the function $x \mapsto \varphi_p(x)$ is also Lipschitz-continuous, see *e.g.* [2, 18]. Hence, problem (1.1) has an optimal solution if the set K is nonempty and compact.

For fixed $x \in K$, problem (1.4) is a linear optimization problem, its dual linear optimization problem is

$$\min_{s, \gamma} \{s^\top (d - Ax) + \gamma \varphi(x) : D^\top s + \gamma c \geq b\}. \quad (2.1)$$

This problem is again a linear optimization problem, and an optimal solution $(\bar{s}, \bar{\gamma})$ can be found at a vertex of the set of feasible points $\{(s, \gamma) \in \mathbb{R}^l \times \mathbb{R} : D^\top s + \gamma c \geq b\}$. The following theorem on the relations between problems (1.4) and (2.1) can be found in all textbooks on linear optimization, see *e.g.* [22, Section 5.2].

Theorem 2.1. *Let $\hat{x} \in K$ be fixed, \hat{y} be a feasible solution for problem (1.4) and $(\hat{s}, \hat{\gamma})$ be feasible for problem (2.1). Then, the following properties are equivalent:*

1. \hat{y} is optimal for (1.4) for $x = \hat{x}$, $(\hat{s}, \hat{\gamma})$ is optimal for (2.1) with $x = \hat{x}$.
2. $b^\top \hat{y} = \hat{s}^\top (d - A\hat{x}) + \hat{\gamma} \varphi(\hat{x})$.
3. $0 = \hat{s}^\top D\hat{y} + \hat{\gamma} c^\top \hat{y} - b^\top \hat{y}$.

Let $(\bar{s}, \bar{\gamma})$ be a vertex of the set $\{(s, \gamma) : D^\top s + \gamma c \geq b\}$. Then, a point $\bar{y}(x)$ satisfying

$$0 = \bar{s}^\top D\bar{y}(x) + \bar{\gamma}c^\top \bar{y}(x) - b^\top \bar{y}(x) \quad (2.2)$$

is an optimal solution of (1.4) as long as it is feasible for this problem:

$$D\bar{y}(x) = d - Ax, \bar{y}(x) \geq 0, c^\top \bar{y}(x) = \varphi(x),$$

i.e. as long as it is an optimal solution of the lower level problem.

Using problem (2.1) instead of (1.4), we can replace (1.1) by

$$\begin{aligned} a^\top x + s^\top (d - Ax) + \gamma \varphi(x) &\rightarrow \min_{x, s, \gamma, y} \\ \text{s.t. } x &\in K, \\ D^\top s + \gamma c &\geq b, \\ Ax + Dy &= d, \\ c^\top y &= \varphi(x), \\ y &\geq 0, \end{aligned} \quad (2.3)$$

provided this problem has an optimal solution [19]. This is a nonconvex optimization problem where, moreover, the objective function as well as some constraints are only implicitly determined, for computing the value $\varphi(x)$ we have to solve a linear optimization problem.

Let $(\hat{x}, \hat{s}, \hat{\gamma}, \hat{y})$ be a local optimal solution of (2.3). Then,

- (1) $A\hat{x} + D\hat{y} = d$, $\hat{y} \geq 0$ and $c^\top \hat{y} = \varphi(\hat{x})$, *i.e.* \hat{y} is an optimal solution of the lower level problem (1.3) for $x = \hat{x}$. \hat{y} is feasible for problem (1.4).
- (2) $(\hat{s}, \hat{\gamma})$ solves the linear optimization problem (2.1) (for fixed $x = \hat{x}$). Hence, there is an optimal solution to the primal problem (1.4), that is, \hat{y} solves

$$\max_y \{b^\top y : Dy = d - Ax, y \geq 0, c^\top y = \varphi(x)\}$$

and the optimal objective function values are equal.

The objective function in problem (2.3) is in general neither convex nor concave, the feasible set of this problem is given using the implicitly determined function $\varphi(\cdot)$. The function $\varphi(\cdot)$ is piecewise affine linear. Hence, by definition, there are finitely many polyhedral sets $X^i \subset \mathbb{R}^n$, $i = 1, \dots, q$, satisfying

$$\bigcup_{i=1}^q X^i = \{x \in \mathbb{R}^n : |\varphi(x)| < \infty\}$$

and the function φ is affine linear on X^i : $\varphi(x) = \varphi_i(x)$ for all $x \in X^i$. This makes it possible to replace problem (2.3) by

$$\begin{aligned} a^\top x + s^\top (d - Ax) + \gamma \varphi_i(x) &\rightarrow \min_{x, s, \gamma, y, i} \\ \text{s.t. } x &\in K \cap X^i, \\ D^\top s + \gamma c &\geq b, \\ Ax + Dy &= d, \\ c^\top y &= \varphi_i(x), \\ y &\geq 0, \\ i &\in \{1, 2, \dots, q\}. \end{aligned} \quad (2.4)$$

Now, since the decision variables include the index $i \in \{1, 2, \dots, q\}$, problem (2.4) is a combinatorial problem. Without loss of generality, let $\Psi(x) \neq \emptyset$ for all $x \in K$, *i.e.* let the lower level problem have an optimal solution for each $x \in K$. Then, by linear optimization theory, there is an optimal basic solution:

$$y = (y_B, y_N) : y_B = B^{-1}(d - Ax) \geq 0, y_N = 0 \text{ for some basic matrix } B.$$

Assume without loss of generality that the rank $r(D)$ of the matrix D is equal to l . Then, B is a square matrix with l rows and $r(B) = l$. B is a submatrix composed by columns of D . The matrix B is a basic matrix, the number of basic matrices corresponding to feasible basic solutions is not less than q . A basic matrix corresponds to an optimal solution of the lower level problem if $B^{-1}(d - Ax) \geq 0$, the basic matrix is then (primally) feasible. Assume without loss of generality that there are exactly q different basic matrices

$$B^1, B^2, \dots, B^q.$$

Remark 2.2. A square matrix B with l rows and $r(B) = l$ but $B^{-1}(d - Ax) \geq 0$ for no $x \in \mathbb{R}^n$ does not need to be considered in what follows. For each set X^i as defined above there exists one basic matrix. But this basic matrix does not need to be unique. In that case, take one of them.

Let y_{B^i} be the subvector of y with components corresponding to the rows of D belonging to the basic matrix B^i , y_{N^i} is the subvector of the other variables: $y = (y_{B^i}, y_{N^i})$. The vector $c = (c_{B^i}, c_{N^i})$ is partitioned analogously. Then, $y_{B^i} = (B^i)^{-1}(d - Ax) \geq 0$, $y_{N^i} = 0$ and (y_{B^i}, y_{N^i}) is an optimal solution of the lower level problem for all

$$x \in X^i := \{x \in \mathbb{R}^n : (B^i)^{-1}(d - Ax) \geq 0\}$$

provided $c_{B^i}^\top (B^i)^{-1} D - c^\top \leq 0$. The set X^i is then called *region of stability*.

Within X^i , $\varphi(x) = c^\top y = c_{B^i}^\top (B^i)^{-1}(d - Ax)$ is an affine function of the upper level variable x . Thus, problem (2.4) reads as

$$\begin{aligned} a^\top x + s^\top (d - Ax) + \gamma c_{B^i}^\top (B^i)^{-1}(d - Ax) &\rightarrow \min_{x, s, \gamma, i} \\ \text{s.t. } x &\in K, \\ D^\top s + \gamma c &\geq b, \\ (B^i)^{-1}(d - Ax) &\geq 0. \end{aligned} \tag{2.5}$$

Theorem 2.3. Let $i \in \{1, \dots, q\}$ be fixed. The optimal solution of the problem (2.1) is constant for $\bar{x} \in X^i \cap K$.

Proof. A vertex $(\bar{s}, \bar{\gamma})$ of the set $\{(s, \gamma) : D^\top s + \gamma c \geq b\}$ is an optimal solution of the problem (2.1) if there exists z such that

$$Dz = d - A\bar{x}, \quad c^\top z = \varphi(\bar{x}), \quad z \geq 0, \quad (\bar{s}^\top D + \bar{\gamma} c^\top - b^\top)z = 0$$

by linear optimization, cf. Theorem 2.1. These conditions are satisfied for an optimal solution $\bar{y}(\bar{x})$ of problem (1.4), hence, $z = \bar{y}(\bar{x})$ is valid. The complementarity condition is satisfied as long as the set of basic variables does not change and is independent of \bar{x} , i.e. the optimal solution of the problem (2.1) remains constant for $\bar{x} \in X^i$. \square

Hence, problem (2.5) reduces to the problem (2.6) used in the algorithm below.

Now, if we compute all dual feasible basic matrices B^i from D , $i = 1, \dots, q$, that is, all basic matrices B^i satisfying $c_{B^i}^\top (B^i)^{-1} D - c^\top \leq 0$, then we can solve the pessimistic linear bilevel optimization problem globally.

Algorithm 2.1. For all basic matrices B^i , $i = 1, \dots, q$ of the matrix D :

If $c_{B^i}^\top (B^i)^{-1} D - c^\top \leq 0$, then

1. select $x^i \in X^i = \{x \in \mathbb{R}^n : (B^i)^{-1}(d - Ax) \geq 0\}$ with $x^i \in K$,
2. compute $y^i = (y_{B^i}^i, y_{N^i}^i)$ with $y_{B^i}^i = (B^i)^{-1}(d - Ax^i)$, $y_{N^i}^i = 0$, $\varphi(x^i) = c_{B^i}^\top (B^i)^{-1}(d - Ax^i)$,
3. compute an optimal solution (s^i, γ^i) of problem (2.1) with $x = x^i$,
4. if $(y^i)^\top (D^\top s^i + \gamma^i c - b) = 0$ then let \bar{x}^i be an optimal solution of the problem

$$\begin{aligned} a^\top x + (s^i)^\top (d - Ax) + (\gamma^i) c_{B^i}^\top (B^i)^{-1}(d - Ax) &\rightarrow \min \\ x &\in K \cap X^i \end{aligned} \tag{2.6}$$

end if

end if

Let x^* be the best of the computed solutions \bar{x}^i :

$$a^\top x^* + \varphi_p(x^*) = \min\{a^\top \bar{x}^i + \varphi_p(\bar{x}^i) : i \in \{1, \dots, q\}\}.$$

Remark 2.4. 1. If B^i is a basic matrix with $c_{B^i}^\top (B^i)^{-1} D - c^\top \leq 0$ only two possibilities exist in Step 1: Either, the system of linear inequalities $(B^i)^{-1}(d - Ax) \geq 0$ has no solution or there is a solution x^i of this system. In the first case, we take the next basic matrix. In the second one, we can take an arbitrary point $x^i \in X^i$. Then, $y^i = (y_{B^i}^i, y_{N^i}^i)$ with $y_{B^i}^i = (B^i)^{-1}(d - Ax^i)$, $y_{N^i}^i = 0$ is an optimal solution of the lower level problem (1.3).

2. If y^i is not an optimal solution of the problem

$$\max\{b^\top y : y \in \Psi(x^i)\}, \quad (2.7)$$

its dual problem has no optimal solution satisfying the complementarity slackness conditions. This is checked in Steps 3 and 4. If y^i is not an optimal solution of the problem (1.4), the next basic matrix is taken.

3. If y^i is an optimal solution of the problem (2.7), an optimal solution (s^i, γ^i) remains optimal for problem (2.1) as long as $y = y(x) = ((B^i)^{-1}(d - Ax), 0)$ is feasible for the lower level problem. This is guaranteed if $x \in X^i$, see Theorem 2.3.

4. Taking the preceding remarks into account, problem (2.3) reduces to (2.6).

Candler *et al.* [4] and others proposed algorithms for smart enumeration of all different basic matrices B^i .

Theorem 2.5. x^* , as computed by Algorithm 2.1, is a global optimal solution of the pessimistic linear bilevel optimization problem (1.5).

Proof. Assuming that the assertion is not correct, let $\tilde{x} \in K$ be an optimal solution of the pessimistic linear bilevel optimization problem (1.5). Then, there is an optimal basic solution \tilde{y} of the problem (2.7) with

$$\Psi(\tilde{x}) = \{y : Dy = d - A\tilde{x}, y \geq 0, c^\top y = \varphi(\tilde{x})\}.$$

By linear optimization, $\{y : Dy = d - A\tilde{x}, y \geq 0, c^\top y = \varphi(\tilde{x})\}$ is a face of the set $\{y : Dy = d - A\tilde{x}, y \geq 0\}$, \tilde{y} is a vertex of the second set or a basic optimal solution of the lower level problem

$$\min_y \{c^\top y : Dy = d - A\tilde{x}, y \geq 0\}.$$

Hence there exists a basic matrix B such that $\tilde{y} = (\tilde{y}_B, \tilde{y}_N)$ with $\tilde{y}_B = B^{-1}(d - A\tilde{x})$, $\tilde{y}_N = 0$. By construction, $B \in \{B^i : i = 1, \dots, q\}$ considered in the algorithm. The matrix B can be taken such that $c_B^\top (B)^{-1} D - c^\top \leq 0$.

Since \tilde{y} is an optimal solution of the linear optimization problem and using linear optimization duality, there exists an optimal solution $(\tilde{s}, \tilde{\gamma})$ of the dual problem (2.1) with $x = \tilde{x}$ and we have

$$\tilde{y}^\top (D^\top \tilde{s} + \tilde{\gamma}c - b) = 0.$$

Hence, \tilde{x} is an optimal solution of one of the problems (2.6) which leads to the desired contradiction. The theorem is correct. \square

3 A Descent Algorithm

Algorithm 2.1 is an algorithm computing a global optimum of the pessimistic bilevel optimization problem (1.1). Such algorithms can have large numerical (non-polynomial) effort. Hence, we will suggest an algorithm computing a local optimal solution in what follows. This will be a descent algorithm investigating

one after the other certain regions of stability. Since subsequent regions of stability have nonempty intersections it seems to be obvious that the basic matrices describing them are neighboring matrices. This is the content of the next two theorems.

Problem (2.6) is a linear optimization problem. Optimal solutions of such problems can be found at vertices of the feasible set

$$K \cap X^i = \{x \in K : (B^i)^{-1}(d - Ax) \geq 0\}$$

with $y_{B^i} = (B^i)^{-1}(d - Ax)$, $y_{N^i} = 0$.

Theorem 3.1. *Let $\bar{x} \in K \cap X^i$ be an optimal solution of problem (2.6). If $(B^i)^{-1}(d - A\bar{x})_j > 0$ for $j = 1, \dots, l$, then (\bar{x}, \bar{y}) with $\bar{y}_{B^i} = (B^i)^{-1}(d - A\bar{x})$ and $\bar{y}_{N^i} = 0$ is a local optimal solution of problem (1.1).*

Proof. By construction, the point (\bar{x}, \bar{y}) is feasible for (1.1). If it is not locally optimal for this problem, there exists a sequence $\{x^k, y^k\}_{k=1}^\infty$ converging to (\bar{x}, \bar{y}) with

$$a^\top x^k + \max_{y^k \in \Psi(x^k)} b^\top y^k < a^\top \bar{x} + \max_{y \in \Psi(\bar{x})} b^\top y \quad (3.1)$$

and $x^k \in K$, $y^k \in \Psi(x^k)$, $b^\top y^k \geq b^\top y \forall y \in \Psi(x^k)$ for all k . Since $(B^i)^{-1}(d - A\bar{x})_j > 0$ for $j = 1, \dots, l$ we have $(B^i)^{-1}(d - Ax^k)_j > 0$ for $j = 1, \dots, l$ and sufficiently large k . Hence, $x^k \in K \cap X^i$ and (x^k, y^k) is feasible for the problem (1.1) for sufficiently large k . This contradicts (3.1). \square

Result of the transformations in the previous section is to transform the pessimistic bilevel optimization problem into

$$\begin{aligned} a^\top x + (s^i)^\top (d - Ax) + (\gamma^i) c_{B^i}^\top (B^i)^{-1} (d - Ax) &\rightarrow \min_{x, i} \\ x &\in K \cap X^i. \end{aligned} \quad (3.2)$$

Local and global optimal solutions of this problem can be found at vertices of the sets $K \cap X^i$. If $\bar{x} \in K \cap X^i$ is a (local or global) optimal solution, the basic matrix B^i used for describing X^i can be used to compute the point $\bar{y} = (\bar{y}_{B^i}, \bar{y}_{N^i})$ with $\bar{y}_{B^i} = (B^i)^{-1}(d - A\bar{x})$, $\bar{y}_{N^i} = 0$ such that (\bar{x}, \bar{y}) is a (local or global) optimal solution of the pessimistic bilevel optimization problem (1.1).

If the point (\bar{x}, \bar{y}) with $\bar{x} \in K$ being an optimal solution of the problem (2.6) is not a local optimal solution of problem (1.1), then there exists a sequence $\{(x^k, y^k)\}_{k=1}^\infty$ converging to (\bar{x}, \bar{y}) , $x^k \in K$ with (3.1). Without loss of generality we can assume that $y^k \in \Psi(x^k)$ is a basic solution of the lower level problem (1.2). Since there exists only a finite number of basic matrices B^i we have

$$y_{B^k}^k = (B^k)^{-1}(d - Ax^k) \geq 0, \quad y_{N^k}^k = 0,$$

where the basic matrix for y^k is denoted by B^k . Hence, there is an infinite subsequence of $\{y^k\}_{k=1}^\infty$ (denoted again by $\{y^k\}_{k=1}^\infty$) for which the basis matrix is constant. Let B^j be this basic matrix. Then, $x^k \in K \cap X^j$ for all k and, since the regions of stability are closed, $\bar{x} \in K \cap X^j$.

Theorem 3.2. *Let $\bar{x} \in K \cap X^i$ be an optimal solution of the problem (2.6), $\bar{y} \in \Psi(\bar{x})$ with basic matrix B^i . If (\bar{x}, \bar{y}) is not locally optimal for (1.1), then there exists a region of stability X^j corresponding to a basic matrix B^j such that: B^j is obtained from B^i by replacing columns D^p in B^i with columns D^q of the matrix D not being columns in B^i .*

1. the rank $r(B^j) = l$,
2. $(\bar{y}_{B^i})_p = ((B^i)^{-1}(d - A\bar{x}))_p = 0$,
3. Let $\Delta := c_{B^i}^\top (B^i)^{-1} D - c^\top \leq 0$ be the reduced costs for basic matrix B^i . Then,

$$\frac{\Delta_q}{D_{pq}} \geq \frac{\Delta_j}{D_{pj}} \forall j : D_{pj} > 0$$

4. problem (2.1) has an optimal solution $(\bar{s}, \bar{\gamma})$ and we have

$$\bar{y}^\top (D^\top \bar{s} + \bar{\gamma}c - b) = 0, (B^j)^\top \bar{s} + \bar{\gamma}c_{B^j} - b_{B^j} = 0.$$

Here, b_{B^j}, c_{B^j} are the subvectors of b, c corresponding to basic variables for the basic matrix B^j .

Proof. If (\bar{x}, \bar{y}) is not a local optimal solution of problem (1.1), then there exists a sequence $\{(x^k, y^k)\}_{k=1}^\infty$ converging to (\bar{x}, \bar{y}) with

$$a^\top x^k + \max_{y \in \Psi(x^k)} b^\top y < a^\top \bar{x} + \max_{y \in \Psi(\bar{x})} b^\top y, \quad c^\top y^k = \max_{y \in \Psi(x^k)} b^\top y, \quad y^k \in \Psi(x^k) \forall k.$$

Since the number of different basic matrices is finite, this implies that there exists without loss of generality a basic matrix B^j such that

$$r(B^j) = l, \quad \lim_{k \rightarrow \infty} y_{B^j}^k = \lim_{k \rightarrow \infty} (B^j)^{-1}(d - Ax^k) = \bar{y}_{B^j}$$

and \bar{y}_p is not a basic variable in \bar{y}_{B^j} . This implies the first two assertions.

Since $y^k \in \Psi(x^k)$ for all k we have $x^k \in X^j$ for all k and, since X^j is closed, $\bar{x} \in X^j$. This implies that

$$c_{B^j}^\top (B^j)^{-1} D - c^\top \leq 0$$

and B^j is a result of replacing columns of B^i with columns of D which do not belong to B^i . Executing such replacement, dual feasibility is secured which, by the rules of the simplex algorithm, implies that the third assertion holds.

Since y^k is an optimal solution of problem $\max\{b^\top y : y \in \Psi(x^k)\}$, problem (2.1) has an optimal solution $(\hat{s}, \hat{\gamma})$ which is independent of k and we have

$$(y^k)^\top (D^\top \hat{s} + \hat{\gamma}c - b) = 0 \quad \forall k.$$

This implies the last assertion by linear optimization duality and $\lim_{k \rightarrow \infty} y^k = \bar{y}$. □

Using this theorem, an algorithm can be suggested computing a sequence of basic matrices B^k , corresponding regions of stability X^k and solutions x^k of the problems (2.6) such that

$$a^\top x^{k+1} + \varphi_p(x^{k+1}) < a^\top x^k + \varphi_p(x^k), \quad k = 1, 2, \dots, t-1$$

and x^t is a local optimal solution of the problem (1.5).

Algorithm 3.1. Select $x^0 \in K$, compute y^0 as an optimal basic solution of problem (1.4) and the corresponding basic matrix B^0 , $k := 0$.

1. Set $X^k := \{x : (B^k)^{-1}(d - Ax) \geq 0\}$ and compute (s^k, γ^k) as optimal solution of the problem (2.1) with $x = x^k$.
2. Compute an optimal solution x^{k+1} of the problem (2.6) with $B^i = B^k$.
3. Use Theorem 3.2 to find a new basic matrix B^{k+1} , set $k := k + 1$, go to Step 1.

There is some vagueness in Step 3 of this algorithm. Namely, Theorem 3.2 does not imply that the new basic matrix is uniquely determined. Aim of this step is to determine a region of stability X^{k+1} with $x^{k+1} \in X^k \cap X^{k+1}$ such that we can find $x^{k+2} \in X^{k+1}$ with

$$a^\top x^{k+2} + \varphi_p(x^{k+2}) < a^\top x^{k+1} + \varphi_p(x^{k+1}) \tag{3.3}$$

in the next iteration. If there are more such regions of stability X^i , it is possibly necessary to investigate all these regions until (3.3) is satisfied. If this search is not successful, the algorithm stops with a local optimal solution.

Theorem 3.3. Let x^* be the last solution computed with Algorithm 3.1. The point x^* is a local optimal solution of (1.5).

The proof of this theorem is straightforward.

4 An Example

In the previous sections, we have shown that the pessimistic bilevel linear optimization can be converted to a single-level optimization problem. Aim of this section is to illustrate the behavior of the algorithm for computing a local optimal solution.

Example 4.1 The data of the bilevel optimization problem are as follows:

$$\begin{aligned}
 & 4y_1 - 40y_2 - 4y_3 - 8x_1 - 4x_2 \rightarrow \min_x \\
 & \text{s.t. } x \geq 0, \\
 & \text{where } y = (y_1, y_2, \dots, y_6)^\top \text{ solves} \\
 & \quad -y_1 - y_2 - y_3 \rightarrow \min_y \\
 & \quad \text{s.t. } \begin{array}{ccccccccc} y_1 & +y_2 & +y_3 & +y_4 & & & & & = 1 \\ -y_1 & +2y_2 & -0.5y_3 & & +y_5 & & +2x_1 & & = 1 \\ 2y_1 & -y_2 & -0.5y_3 & & & +y_6 & & +2x_2 & = 1 \\ y & \geq & 0 & & & & & & \end{array}
 \end{aligned} \tag{4.1}$$

Let $x^0 = (0.5, 0.5)^\top$. Then, $y^0 = (0.2, 0, 0.8, 0, 1, 0)^\top$ and $(0, 0, 1, 0, 0.5, 0.5)^\top$ are optimal basic solutions for the lower level problem. The optimal pessimistic solution (*i.e.* the optimal solution of problem (1.4)) is y^0 . The optimal basic matrix and its inverse are:

$$B^1 = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -0.5 & 1 \\ 2 & -0.5 & 0 \end{pmatrix} \quad (B^1)^{-1} = \begin{pmatrix} \frac{1}{5} & 0 & \frac{2}{5} \\ \frac{4}{5} & 0 & -\frac{2}{5} \\ \frac{3}{5} & 1 & \frac{1}{5} \end{pmatrix}.$$

Furthermore,

$$\begin{aligned}
 X^1 &= \left\{ x \geq 0 : \begin{pmatrix} \frac{1}{5} & 0 & \frac{2}{5} \\ \frac{4}{5} & 0 & -\frac{2}{5} \\ \frac{3}{5} & 1 & \frac{1}{5} \end{pmatrix} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \geq 0 \right\} \\
 &= \{x : -0.5 \leq x_2 \leq 0.75, x_1 \leq -0.2x_2 + 0.9\}
 \end{aligned}$$

is the resulting region of stability. The basic variables are y_1, y_3, y_5 . Problem (2.1) reduces to

$$\begin{array}{ccccccc}
 s_1 & & & -\gamma & \rightarrow & \min \\
 s_1 & -s_2 & +2s_3 & -\gamma & \geq & 4 \\
 s_1 & +2s_2 & -s_3 & -\gamma & \geq & -40 \\
 s_1 & -0.5s_2 & -0.5s_3 & -\gamma & \geq & -4 \\
 s_i & & & & \geq & 0, \quad i = 1, 2, 3
 \end{array}$$

with optimal solution $(s^1, \gamma^1) = ((0, 0, 2), 0)$.

Then, problem (2.6):

$$-8x_1 - 4x_2 + 2(1 - 2x_2) \rightarrow \min$$

subject to

$$0 \leq x_2 \leq 0.75, \quad 0 \leq x_1 \leq -0.2x_2 + 0.9$$

gives the point $x^1 = (0.75, 0.75)^\top$ as next iterate. One, not unique lower level optimal solution is $y(x^1) = (0, 0, 1, 0, 0, 0)^\top$ which is the solution of problem (1.4) at $x = x^1$. Here, a new basic matrix and a new region of stability are selected and the computation proceeds similarly.

One neighboring basic matrix is:

$$B^1 = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -0.5 & 1 \\ -1 & -0.5 & 0 \end{pmatrix} \quad (B^1)^{-1} = \begin{pmatrix} -1 & 0 & -2 \\ 2 & 0 & 2 \\ 3 & 1 & 5 \end{pmatrix}$$

with the region of stability

$$X^2 = \{x \in \mathbb{R}^2 : 0.75 \leq x_2 \leq 1, 4.5 - 5x_2 \geq x_1\}.$$

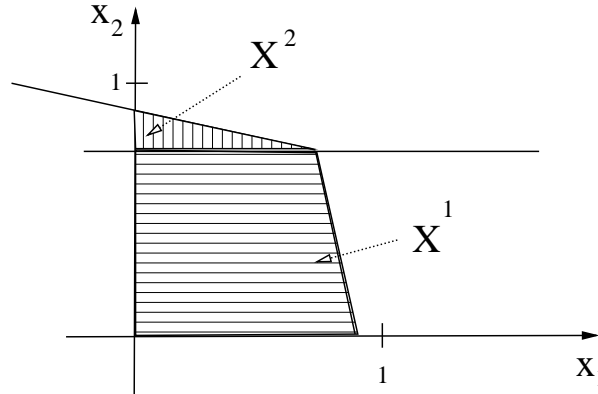


Figure 1: Regions of stability in Example 4.1 as subsets of \mathbb{R}_+^2

The regions of stability are depicted in Fig. 1.

At $x = x^1$ an optimal solution of problem (1.4) is $(\bar{s}, \bar{\gamma}) = ((0, 0, 3.2), 2.4)$. Using the region of stability X^2 at $x = x^1$, we obtain again x^1 as optimal solution.

□

5 Conclusion

Regarding pessimistic bilevel linear optimization problems (PBLOP), it is harder to capture its essence than in the optimistic case while it is very important in reality. By means of duality theory and using the lower level optimal value function, we converted the PBLOP to a single-level optimization problem with nonconvexity. Then we proposed algorithms for computing a global and a local optimal solution respectively. Applications and solution methods of nonlinear pessimistic bilevel optimization problem will be studied in the future.

Acknowledgement: The first and the third author have been supported by the Deutsche Forschungsgemeinschaft, grant DE-650/7-1. The second author has been supported by the Humanities and Social Sciences of Ministry of Education Planning Fund, grant 15YJA79004.

References

- [1] Aboussoror A, Mansouri A (2005) Weak linear bilevel programming problems: existence of solutions via penalty method. *J Math Anal Appl*, **304**: 399-408.
- [2] Bank B, Guddat J, Klatte D, Kummer B, Tammer K (1983) Non-linear parametric optimization. Birkhäuser Verlag Basel - Boston - Stuttgart.
- [3] Beer K (1977) Lösung großer linearer Optimierungsaufgaben. Deutscher Verlag der Wissenschaften, Berlin.
- [4] Candler W, Townsley R (1982) A linear two-level programming problem, *Comput Oper Res*, **9**(1): 59-76.
- [5] Cecchini M, Ecker J, Kupferschmid M, Leitch R (2013) Solving nonlinear principal agent problems using bilevel programming. *Eur J Oper Res*, **230**: 364-373.
- [6] Chiou SW (2009) A bi-level programming for logistics network design with system-optimized flows. *Inform Sciences*, **179**: 2434-2441.

- [7] Chiou SW (2009) Optimization of limited network capacity with toll settings. *Inform Sciences*, **179**: 109-119.
- [8] Dempe S (2002) *Foundations of Bilevel Programming*. Kluwer Academic Publishers Dordrecht.
- [9] Dempe S (2003) Annotated Bibliography on Bilevel Programming and Mathematical Programs with Equilibrium Constraints. *Optimization*, **52**: 333-359.
- [10] Dempe S, Richter K (2000) Bilevel programming with knapsack constraints. *Central Eur J Oper Res*, **8**: 93-107.
- [11] Dempe S, Dutta J, Mordukhovich BS (2007) New necessary optimality conditions in optimistic bilevel programming. *Optimization*, **56**: 577-604.
- [12] Dempe S, Dutta J (2012) Is bilevel programming a special case of a mathematical program with complementarity constraints? *Math Program*, **131**: 37-48.
- [13] Dempe S, Zemkoho AB (2012) On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem. *Nonlinear Anal-Theor*, **75**: 1202-1218.
- [14] Dempe S, Mordukhovich BS, Zemkoho AB (2014) Necessary optimality conditions in pessimistic bilevel programming. *Optimization*, **63**(4): 505-533.
- [15] Deng X (1998) Complexity issues in bilevel linear programming. In: Migdalas A, Pardalos PM, Värbrand P, Eds. Multilevel Optimization: Algorithms and Applications. *Kluwer Academic Publishers Dordrecht*, pp. 149-164.
- [16] Hoffman AJ (1952) On approximate solutions of systems of linear inequalities. *J Res Nat Bur Standards*, **49**: 263-265.
- [17] Jeroslow RG (1985) The polynomial hierarchy and a simple model for competitive analysis. *Math Program* **32**: 146-164.
- [18] Klatte D, Kummer B (1984) Stability properties of infima and optimal solutions of parametric optimization problems. In: Demjanov VF, Ed. Nondifferentiable Optimization: Motivations and Applications. Proceedings of the IIASA Workshop Sopron, Springer Berlin, *Lect. Notes Econ. Math. Syst.*, **255**: 215-229.
- [19] Luderer B (1983) Über der Äquivalenz nichtlinearer Optimierungsaufgaben. *Wissenschaftliche Zeitschrift der TH Karl-Marx-Stadt*, **26**: 257-258.
- [20] Stackelberg Hv (1952) Marktform und Gleichgewicht. Springer-Verlag, Berlin. 1934. (engl. transl.: *The Theory of the Market Economy*. Oxford University Press.)
- [21] Tsoukalas A, Wiesemann W, Rustem B (2009) Global optimisation of pessimistic bi-level problems. In: Pardalos PM, Coleman TF, Eds. *Lectures on Global Optimization*, Fields Inst. Commun. 55. American Mathematical Society. Providence, pp. 215-243.
- [22] Unger T, Dempe S (2010) Lineare Optimierung: Modell, Lösung, Anwendung. *Vieweg+Teubner Verlag Wiesbaden*.
- [23] Wiesemann W, Tsoukalas A, Kleniati PM, Rustem B (2013) Pessimistic bi-level optimization. *SIAM J Optimiz*, **23**(1): 353-380.
- [24] Zheng Y, Wan Z, Jia S, Wang G (2015) A new method for strong-weak linear bilevel programming problem. *Journal of Industrial and Management Optimization*, **2**(11): 529-547.