Stephan Dempe, Rizo Saboiev

Linear Bilevel Optimization with Parameters in the Objective Function of the Lower Level Problem
Linear Bilevel Optimization with Parameters in the Objective Function of the Lower Level Problem
LINEAR BILEVEL OPTIMIZATION WITH PARAMETERS IN THE OBJECTIVE FUNCTION OF THE LOWER LEVEL PROBLEM

S. DEMPE, R. SABOIEV

Abstract. In the paper has been proposed bilevel programming for dealing with decision processes including two decision makers with hierarchical structure. A solution algorithm for linear bilevel programming is presented, for computing a local optimal solution. Some convergence results are also shown.

1. Introduction

Many problems in the real world represent a hierarchical relationship between two decision makers, for instance the management of environmental regulation, planning the economical electric power pricing and oil production, optimal design in engineering and so on. In such a structure, the upper level represents decision-makers whose decision bring to some reaction with a particular market or social entity, which corresponds to the lower level of the problem [1]. We can also refer to the hierarchical structure the planning of transportation network, where on the upper level operator dealing with improvement of the network performance, but in the lower level users of a such network making their travel choices. Such investigation were considered in [14], for example network design, signal setting and origin/destination matrix adjustment problems. Pricing problems, where the company assigns the price for a set of products so that to maximize its revenue, anticipating the reaction of potential clients are considered in [13].

2. Bilevel programming problem

The bilevel programming problem is an hierarchical optimization problem where the feasible set of the upper level problem (leader’s level) is restricted by the solution set mapping of the lower level (follower’s level) problem. Such problems were introduced by Bracken and McGill [4] as mathematical programs with optimization problems in the constraints, but the term ”bilevel” were introduced later by Candler and Norton [5]. The research work [4] has close relation with the economic model of Stackelberg game [16]. In the Stackelberg problem, the hierarchical model of the market situation is described in which different decision-makers try to optimize their decisions depending on their own different objectives subject to a certain hierarchy.

From the terminology of bilevel programming follows the sequentiality: the follower chooses his/her optimal solution when the leader’s choice is already known.

Key words and phrases. Bilevel programming; parametric linear optimization; region of stability; shortest path problem.
and the leader optimizes his/her choice knowing that the follower always reacts optimally to it. The follower’s problem may have multiple solutions for certain values of the upper level variables. Depending on the follower’s behavior two modeling approaches can be suggested. The first is the optimistic case, meaning cooperation between the two players, when the leader assumes the follower’s choice is always one most favorable to him/her. And the pessimistic solution when the leader anticipate the worst reaction of the follower [7]. An annotated bibliography of bilevel optimization can be found in [8].

Denote by $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ decision vectors of the leader and the follower, resp., $F, f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ objective functions, $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ be closed sets. The solution set mapping of the lower level problem is $\Psi : \mathbb{R}^n \Rightarrow \mathbb{R}^m$ given by

$$\Psi(x) := \text{Argmin}_{y} \{ f(x, y) : g(x, y) \leq 0 \},$$

where $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$. Considering the graph of the solution set mapping $\text{gph}\Psi := \{(x, y) : y \in \Psi(x)\}$ the bilevel programming problem can be stated as

$$\min_{x, y} \{ F(x, y) : G(x) \leq 0, (x, y) \in \text{gph}\Psi, x \in X \},$$

where $G : \mathbb{R}^n \to \mathbb{R}^q$.

We consider the optimistic formulation of the bilevel programming problem and therefore we can minimize the objective function in problem (2.2) with respect to $x$ and $y$. So that we can avoid the uncertainty in the case of the existence of nonunique optimal solutions in problem (2.1). The mapping $y \mapsto F(x, y)$ and the functions $f, G_j (j = 1, ..., p), g_i (i = 1, ..., q)$ are linear, the set $X \subseteq \mathbb{R}^n$ is a polyhedron. The bilevel problem that we investigate is

$$F(x, y) \to \min_{x, y}$$

s.t. $Ax \leq b$

and

$$y \in \Psi(x) = \text{Argmin}_{y} \{ x^\top y : By \leq d \}$$

where $x, y \in \mathbb{R}^n, b \in \mathbb{R}^p, d \in \mathbb{R}^q$ and $A : m \times n, B : p \times n$ are given matrices.

There exist different solution algorithms to solve bilevel linear problems, for instance the $k$-best algorithm, where the vertices of a certain polyhedron are enumerated according to non-decreasing values of the upper level objective functions. The algorithm stops when the first feasible solution of the problem is found [3]. In [12] is proposed a method based on genetic algorithm approach for solving a bilevel problem. An algorithm solving the nonsmooth reformulation of problem (2.3), (2.4) using the optimal value function of the problem (2.4) can be found in [10]. Here, a sequence of optimization problems with quadratic constraint needs to be solved. For problems, where the lower level problem (2.1) is a linear optimization problem with right-hand side perturbations, a solution algorithm solving a sequence of linear optimization problems can be found in [6].

3. Reformulation of bilevel programming problem

In order to solve the problem (2.3), (2.4) we need to transform it into a one-level problem. For that, there exist different approaches [7]:
(1) The problem (2.4) can be replaced using the optimal value function
\[ \varphi(x) = \min_y \{ x^\top y : By \leq d \} \]

Then, (2.3), (2.4) is replaced with
\[(3.1) \min_{x,y} \{ F(x,y) : x \in X, f(x,y) \leq \varphi(x), g(x,y) \leq 0 \}, \]
where
\[ X = \{ x : Ax \leq b \}, \]
\[ g(x,y) = By - d, \]
\[ f(x,y) = x^\top y. \]

By Ye and Zhu [17] problem (3.1) is fully equivalent to (2.3), (2.4) without any assumptions. But the optimal value function is nonsmooth, even if the functions involved are affine linear one.

**Theorem 3.1** (Beer [2]). The function \( \varphi(\cdot) \) is a piecewise linear concave function over \( Q := \{ x : |\varphi(x)| < \infty \} \). Moreover, \( Q \) is a convex polyhedron.

This result has been used in [10] to approximate problem (3.1) by
\[(3.2) \min_{x,y} \{ F(x,y) : x \in X, f(x,y) \leq \xi_K(x), g(x,y) \leq 0 \}, \]
where
\[ \xi_K(x) = \min \{ x^\top y^t : y^t \in K \} \]
and \( K = \{ y^1, \ldots, y^p \} \) is a finite set of feasible solutions to the lower level problem. Since
\[ \varphi(x) \leq \xi_K(x) \forall x \]
for all (finite) sets \( K \subseteq \{ y : By \leq d \} \),
\[ \min_{x,y} \{ F(x,y) : x \in X, f(x,y) \leq \xi_K(x), g(x,y) \leq 0 \} \]
\[ \leq \min_{x,y} \{ F(x,y) : x \in X, f(x,y) \leq \varphi(x), g(x,y) \leq 0 \}. \]

This is used in [10] to suggest an algorithm for globally solving (2.4), (2.3) by solving problems (3.2) for increasing sets \( K \).

(2) The lower level problem can be replaced using the Karush-Kuhn-Tucker (KKT) necessary optimality conditions. Since the lower level problem (2.4) is a linear parametric optimization problem, these conditions are necessary and sufficient optimality conditions for it:
\[ \Psi(x) = \{ y : \exists \lambda \text{ such that } By \leq d, x + B^\top \lambda = 0, \lambda \geq 0, \lambda^\top (By - d) = 0 \}. \]

Hence, problem (2.3), (2.4) is replaced by
\[(3.3) F(x,y) \rightarrow \min_{x,y,\lambda} \]
\[ \text{s.t. } Ax \leq b, \]
\[ By \leq d, \]
\[ x + B^\top \lambda = 0, \]
\[ \lambda \geq 0, \]
\[ \lambda^\top (By - d) = 0. \]
The main complexity of the KKT reformulation is that we introduce new variables implying that the problems (2.3), (2.4) and (3.3) are equivalent only if we consider global optimal solutions of both problems.

**Theorem 3.2** (Dempe and Dutta [9]). A feasible point \((\bar{x}, \bar{y})\) is a local optimal solution of problem (2.3), (2.4) if and only if \((\bar{x}, \bar{y}, \lambda)\) is a local optimal solution of problem (3.3) for all

\[
\lambda \in \Lambda(\bar{x}, \bar{y}) := \{u : x + B^\top u = 0, u \geq 0, u^\top (By - d) = 0\}.
\]

Moreover, problem (3.3) is a nonconvex optimization problem for which the Mangasarian-Fromovitz constraint qualification is violated at every feasible point [15].

Let \((\hat{x}, \hat{y}, \hat{\lambda})\) be a feasible solution of problem (3.3) and use a partitioning of \(\{1, \ldots, q\}\) as follows:

\[
(y, \lambda) : \lambda^\top (By - d) = 0, \lambda \geq 0, By \leq d, B^\top \lambda = -x = \bigcup_{I} \{(y, \lambda) : B^\top \lambda = -x, By \leq d, \lambda \geq 0, \lambda_i = 0, \forall i \in I; (By - d)_i = 0, \forall i \notin I\},
\]

where the union is taken over all sets \(I\) satisfying

\[
I_0 := \{i : (B\hat{y} - d)_i < 0\} \subseteq I \subseteq I_0 \cup \{i : (B\hat{y} - d)_i = 0, \hat{\lambda}_i = 0\}.
\]

Denote by \(I\) the family of all sets \(I\) satisfying the inclusions (3.5).

Thus, problem (3.3) is equivalent to

\[
F(x, y) \to \min_{x, y, \lambda, I} \quad \text{s.t.} \quad Ax \leq b, \\
x + B^\top \lambda = 0, \\
\lambda_i = 0, \forall i \in I, \\
\lambda_i \geq 0, \forall i \notin I, \\
(By - d)_i = 0, \forall i \notin I, \\
(By - d)_i \leq 0, \forall i \in I, \\
I \in I.
\]

4. Solution algorithm and convergence results

To solve problem (3.6) we proceed as follows starting with a feasible solution \((\bar{x}, \bar{y}, \bar{\lambda})\) of problem (3.3):

**Algorithm:**

**Step 1:** Fix \(x\) to \(\bar{x}\) and minimize the objective function of problem (3.3) over \(\Psi(\bar{x})\), i.e. solve

\[
(4.1) \quad \min_y \{F(\bar{x}, y) : y \in \Psi(\bar{x})\}.
\]

**Step 2:** Fix \(y\) to an optimal solution \(\bar{y}\) of this problem and minimize the objective function of problem (3.3) with respect to \(x\) and \(\lambda\), and obtain the point \(\hat{x}\), i.e.
\[
\min_{x, \lambda, I} F(x, \bar{y}), \quad Ax \leq b,
\]
\[
\text{s.t.} \quad x + B^T \lambda = 0,
\]
\[
\lambda_i = 0, \forall i \in I,
\]
\[
\lambda_i \geq 0, \forall i \notin I,
\]
\[
(By - d)_i = 0, \forall i \notin I,
\]
\[
(By - d)_i \leq 0, \forall i \in I,
\]
\[
I \in \mathcal{I}
\]

**Step 3:** For the obtained optimal point \( \hat{x} \) solve the lower level problem with respect to \( y \) in order to compute the optimal \( \hat{y} \), i.e.
\[
\hat{y} \in \arg\min_y \{ F(\hat{x}, y) : y \in \Psi(\hat{x}) \}
\]

Go to step 2.

\[\square\]

**Remark 4.1.** The minimum of the objective function in the problem (4.2) is realized for the smallest set \( I \) (with respect to inclusion). This smallest set \( I = I_0 \) corresponds to a vertex \( \bar{y} \) of \( \{ y : By \leq d \} \), since convex combination \( \bar{y} = \sum_{i=1}^{q} \lambda_i y^i \) of the vertices of this set \( (\lambda_i \geq 0, \sum_{i=1}^{q} \lambda_i = 1) \) leads to larger set \( \{ i : (By - d)_i < 0, i = 1, ..., q \} \). Hence, we can restrict ourself to vertices of \( \{ y : By \leq d \} \) in steps one and three.

This algorithm is illustrated in fig. 1.

**Figure 1.** Illustration of the algorithm: *The algorithm starts in the point \( (x, y_1) \). In Step 2 the new parameter \( x_1 \) is computed. Then, fixing \( x \) to \( x_1 \) the point \( y_2 \) is computed in Step 1. Here, the upper level objective is independent on \( x \) and the problem in Step 2 reduces e.g. to taking a vertex of the feasible set.*
Note that in the problem in fig. 1 each vertex in the graph of the solution set mapping of the lower level (the dashed lines) is a local optimal solution for suitable chosen \( x \), see also [11]. But this is not the general case.

In order to compute a first value for variable \( x \) in the above algorithm we drop simply the objective function of the lower level problem (2.4) and solve the following problem

\[
F(x, y) \rightarrow \min_{x,y} \\
\text{s.t. } Ax \leq b, By \leq d
\]

Then we get, for instance the point \((\hat{x}, \hat{y})\). Fix the parameter \( x \) to \( \hat{x} \) solve the problem (4.1) w.r.t variable \( y \) where \( y \in \Psi(\hat{x}) \) and get the optimal solution \( \hat{y} \). The obtained point \((\hat{x}, \hat{y})\) is the initial point of the problem (3.6).

With the object of solving the problem (4.3) in step 3 we can use the KKT necessary and sufficient condition to solve the lower level problem \( y \in \Psi(\hat{x}) : \)

\[
\Psi(\hat{x}) = \{ y : \exists \lambda \text{ such that } By \leq d, \hat{x} + B^\top \lambda = 0, \lambda \geq 0, \lambda^\top (By - d) = 0 \}.
\]

**Definition 4.1.** Let \( y \) be a vertex of the set \( \{ y : By \leq d \} \). The region of stability \( \mathcal{R}(y) \) is the set of parameters \( x \) such that \( y \in \Psi(x) : \)

\[
\mathcal{R}(y) = \{ x : \exists \lambda \geq 0, \lambda^\top (By - d) = 0, -x = B^\top \lambda \} = \\
\bigcup_{i \in I} \{ x : \exists \lambda \geq 0, \lambda_i = 0, \forall i \in I, x + B^\top \lambda = 0 \}.
\]

The problem to be solved in Step 2 of the above algorithm reduces to

\[
(4.5) \begin{align*}
\min_x & \{ F(x, \hat{y}) : \hat{y} \in \Psi(x), Ax \leq b \} \\
= & \min_x \{ F(x, \hat{y}) : x \in \mathcal{R}(\hat{y}), Ax \leq b \}
\end{align*}
\]

\[
= \min_{x, \lambda, j} \{ F(x, \hat{y}) : \lambda_i = 0, \forall i \in I, j : (By - d)_i < 0, \lambda_i \geq 0, \forall i \notin I, x + B^\top \lambda = 0, Ax \leq b \}
\]

\[
= \min_{x, \lambda} \{ F(x, \hat{y}) : \lambda_i = 0, \forall i \in I_0 = \{ j : (By - d)_i < 0, \lambda_i \geq 0, \forall i \notin I_0, x + B^\top \lambda = 0, Ax \leq b \}
\]

If \((\pi, \lambda, T)\) is an optimal solution of problem (4.5), then we can select an arbitrary set \( T \in \mathcal{T} \) such that

\[
I_0 \subseteq T \subseteq \{ i : (B\hat{y} - d)_i = 0, \lambda_i = 0 \} \cup I_0.
\]

Then, \((\pi, \lambda, T)\) is an optimal solution of problem (4.5). Hence,

\[
\Psi(\pi) = \{ y : (By - d)_i = 0, \forall i \notin T, (By - d)_i \leq 0, \forall i \in T \}.
\]

Note that, for \( x = \pi \)

\[
\lambda \in \Lambda(\pi, y) = \{ \lambda : \lambda \geq 0, \lambda^\top (B\pi - d) = 0 \}.
\]

Problem (4.5) is a linear optimization problem, as optimal solution we can take a vertex \((\pi, \lambda)\) of the feasible set. In that case, \( \lambda \) is a vertex of the set \( \Lambda(\pi, y) \) and the set

\[
\{ y : (By - d)_i = 0, \forall i \notin T, (By - d)_i \leq 0, \forall i \in T \}
\]

is a (in general not unique due to nonuniquely determined set \( T \)) face of the set \( \{ y : By \leq d \} \). Hence, the set \( \text{gph} \Psi \) equals the union of faces of the set \( \{ y : By \leq d \} \), as it can be seen in fig. 1.

Summing up, the following convergence results can be shown:
Remark 4.2. In step 2 of the above algorithm we compute the global optimal solution for problem (4.2).

**Theorem 4.1.** Assume that the set \( Y := \{ y : By \leq d \} \) is bounded. Let the sequence \( \{x^k, y^k\}_{k=1}^P \) be computed by the above algorithm. Then,

\[
F(x^{k+1}, y^{k+1}) \leq F(x^k, y^k) \quad \text{for all } k.
\]

The algorithm stops after a finite number of iterations (i.e. \( P < \infty \)), provided that strong inequalities hold in (4.7) for all \( k \).

**Proof.** The inequality (4.7) is a consequence of the formulation of the algorithm. Due to these inequalities, no calculated solution can be repeated, if strong inequalities hold. Due to linear optimization, the set \( \Psi(\overline{\pi}) \) is a face of the set \( Y := \{ y : By \leq d \} \) which is equal to the convex hull of certain vertices of the set \( Y \). Hence, \( y^P \) is a solution of a linear optimization problem (4.1) being attained at a vertex of the set \( Y \). Due to strict inequalities in (4.7), this point can also not be repeated in the algorithm. Since the number of vertices of the set \( Y \) is finite, the theorem is correct. \( \square \)

**Theorem 4.2.** Consider the above algorithm and assume that \( (x^P, y^P) \) is the last computed point. Assume that strong inequalities hold in (4.7) for all \( k \) and that \( Y \) is bounded. Then, \( (x^P, y^P) \) is a local optimal solution of problem (2.3), (2.4).

**Proof.** Assume that the assertion of the theorem is not correct, i.e. there is a sequence \( \{(u^t, v^t)\}_{t=1}^\infty \) of feasible points of (2.3), (2.4) converging to \( (x^P, y^P) \) with

\[
F(u^t, v^t) < F(x^P, y^P) \quad \text{for all } t.
\]

Using the same arguments as in the proof of Theorem 4.1, there is an infinite subsequence of \( \{(u^t, v^t)\}_{t=1}^\infty \) such that \( v^{t_j} = \overline{\pi} \) for all \( j \) and \( \overline{\pi} \) is a vertex of \( Y \). Due to finiteness of the number of vertices and convergence of \( \{(u^t, v^t)\}_{t=1}^\infty \) we have \( v^{t_j} = \overline{\pi} = y^P \) for sufficiently large \( t \).

If \( Y \) is bounded, the region of stability \( R(y^P) \) is equal to the convex hull of normal directions of the faces of the set \( Y \) containing \( y^P \), where the normal directions belong to the set \( X \). Since the number of faces is finite, the set \( R(y^P) \) is a bounded convex polyhedron.

Let the point \( v^{t_j} \) is the optimal solution of the problem (4.1) or (4.3). Then \( R(v^{t_j}) \) is constant. Hence, also \( x^{t_j} = x^P \) for all \( j \), because \( v^{t_j} \) is a global solution of problem (4.2) (without lost of generality, if the set of global solution does not reduce to a singleton). \( \square \)

**Theorem 4.3.** Let the point \( (x^P, y^P) \) is the last computed point and the strong inequalities do not hold in (4.7) for all \( k \), and that \( Y \) is bounded. If we not get the better solution then terminate above algorithm and the point \( (x^P, y^P) \) is a local optimal solution.

**Proof.** Consider the graph of the solution set of bilevel problem as

\[
gph \Psi \subseteq \{x : Ax \leq b\} \times \bigcup_I \{y : By \leq d, (By - d)_i = 0, \forall i \in I\},
\]

where

\[
\bar{Y} = \bigcup_I \{y : By \leq d, (By - d)_i = 0, \forall i \in I\}
\]

is a bounded set.
Consider the set \( \{x^p, y^p\} \) with respect to \( \text{gph} \, \Psi \):

\[
T_{\text{gph}}(x, y) = \left\{ \left. \left( \frac{dx}{dy} \right) \in \mathbb{R} : \exists \{x^p, y^p\} \in \text{gph} \, \Psi, \exists \{t_k\}_{k=1}^{\infty} \subseteq \mathbb{R}_+ \setminus \{0\}, \text{ and } \lim_{k \to \infty} t_k = 0 \right\}
\]

\[
\{(x^p, y^p)\} \to (x, y), \quad \lim_{k \to \infty} t_k = 0, \quad \lim_{k \to \infty} \frac{1}{t_k}((x^p, y^p) - (x, y)) = 0
\]

Define the contingent (or Bouligand) cone at a point \((x, y)\) with respect to \(\text{gph} \, \Psi\):

\[
\lim_{p \to \infty} \frac{o\left(\left\|\left(\frac{x^p}{y^p}\right) - \left(\frac{x}{y}\right)\right\|\right)}{o\left(\left\|\left(\frac{x^p}{y^p}\right) - \left(\frac{x}{y}\right)\right\|\right)} = 0
\]

Define \(d = \left(\frac{dx}{dy}\right)\). Obviously, we get

\[
\left(\frac{x^p}{y^p}\right) - \left(\frac{x}{y}\right) \to \left(\frac{dx}{dy}\right) \in T_{\text{gph}}(x, y).
\]

Since we are consider only the vertices of the set \(\tilde{Y}\), we have \(\hat{y} = \bar{y}\) and \(y^p = \bar{y}\), \(\forall p\) is global solution with respect to \(x\) in iteration \(p\). Hence, we obtain \(F(\hat{x}, \bar{y}) < F(x, y)\), i.e. \(F(x, \bar{y}) < F(\hat{x}, y)\), \(\forall p\). But this inequality contradicts (4.2).

\[\square\]

**Theorem 4.4.** Consider the set \(\tilde{Y}\) of vertices of \(\{y : By \leq d\}\) and a mapping \(y \to \text{Argmin}_{y}\{F(x, y) : x \in \mathcal{R}(y), Ax \leq b\} : \tilde{Y} \to \mathbb{R}^n\).

Then, a global optimal solution of (2.3),(2.4) without loss of generality are obtained in

\[
\bar{W} = \{(x, y) : y \in \tilde{Y}, x \in \text{Argmin}_{x}\{F(x, y) : x \in \mathcal{R}(y), Ax \leq b\}\}
\]

**Proof.** Assume that \((\bar{x}, \bar{y})\) is a global optimal solution of (2.3),(2.4) and \((\bar{x}, \bar{y}) \notin \bar{W}\). Then

\[
F(\hat{x}, \bar{y}) < F(\bar{x}, \bar{y}), \forall (\bar{x}, \bar{y}) \in \bar{W},
\]

By the steps 1 and 3 of the above algorithm, we have:

\[
F(\hat{x}, \bar{y}) \geq F(\bar{x}, \bar{y})
\]

for some \(\bar{y} \in \tilde{Y}\), \((\bar{y})\) is the solution of a linear optimization problem (4.1) or (4.3), and

\[
F(\hat{x}, \bar{y}) \geq F(\bar{x}, \bar{y})
\]
with \( \tilde{x} \) solves (4.2) for \( \tilde{y} = \bar{y} \). This contradicts (4.10), where the point \((\tilde{x}, \bar{y})\) minimize the objective function \( F(x, y) \) from \( \bar{W} \). \( \square \)

**Corollary 4.1.** If the polyhedron \( \{ y : By \leq d \} \) is compact, only its vertices need to be considered for the computation of optimal solution of the lower level problem (2.4), where the upper variable is fixed. This means that algorithm stops after a finite number of iterations.

5. Example to the algorithm

In this section we would like to illustrate the procedure of the algorithm by an example, where \( n = 2 \).

**Example 5.1.** Consider the following bilevel problem

\[
\begin{align*}
\min_{x,y} F(x, y) &= x_1y_1 + y_2 \\
\text{s.t.} &\quad |x_1| \leq 1, \ x_2 = -1 \\
y &\in \Psi(x) := \operatorname{Argmin}_y \{ x^T y : -2y_1 + y_2 \leq 0, y_1 \leq 2, 0 \leq y_2 \leq 2 \}
\end{align*}
\]

In order to find the initial solution of the above bilevel problem, simply drop the objective function of the lower level and solve the problem:

\[
\begin{align*}
\min_{x,y} F(x, y) &= x_1y_1 + y_2 \\
\text{s.t.} &\quad |x_1| \leq 1, \ x_2 = -1 \\
-2y_1 + y_2 &\leq 0, y_1 \leq 2, 0 \leq y_2 \leq 2
\end{align*}
\]

Fix the variable \( \tilde{x}_1 = 1 \) and solve the problem with respect to \( y \):

\[
\begin{align*}
\min_{y} F(x, y) &= y_1 + y_2 \\
\text{s.t.} &\quad -2y_1 + y_2 \leq 0, y_1 \leq 2, 0 \leq y_2 \leq 2
\end{align*}
\]

We get the initial solution \( \tilde{y} = (2, 2) \).

Now fix the lower variable \( y \) to \( \bar{y} \) and minimize the following problem w.r.t. \( x \) :

\[
\begin{align*}
\min_{\tilde{x}} F(x, y) &= 2x_1 + 2 \\
\text{s.t.} &\quad |x_1| \leq 1, \ x_2 = -1
\end{align*}
\]

Solution of this problem is \( \tilde{x}_1 = -1 \). As in the previous step fix it and solve the problem with respect to lower level variable \( y \) :

\[
\begin{align*}
\min_{y} F(x, y) &= -y_1 + y_2 \\
\text{s.t.} &\quad -2y_1 + y_2 \leq 0, y_1 \leq 2, 0 \leq y_2 \leq 2
\end{align*}
\]

Now for lower level variable we get the optimal solution \( \hat{y} = (1, 2) \).

The global optimal solution of this problem is \( \tilde{x} = -1 \). Now the lower level solution with \( x = (-1, -1) \) is equal to the one found in the preceding step. The algorithm stops with the optimal solution \( \tilde{x} = (-1, 1), \tilde{y} = (1, 2) \) and \( F(\tilde{x}, \tilde{y}) = 1 \) of the bilevel programming problem.
6. Conclusion

In this paper we have analyzed the bilevel problems with linear parametric lower level problem. In general our problem is non-convex optimization problem. To derive the optimality condition first we reformulated this problem to one-level with aid of KKT-necessary optimality conditions and then used the region of stability approach to shown that for the upper level variable, the region of stability remains optimal. The algorithm is also presented, for computing the local optimal solution and some results convergence are showed.

References