A NEW DESCENT ALGORITHM FOR SOLVING QUADRATIC BILEVEL PROGRAMMING PROBLEMS*

HAN JIYE (韩继业)

(*Institute of Applied Mathematics, the Academy of Mathematics and Systems Sciences, the Chinese Academy of Sciences, Beijing* 100080, *China*)

LIU GUOSHAN (刘国山)

(*School of Business Administration, Renmin University of China, Beijing* 100872, *China*)

WANG SHOUYANG (汪寿阳)

(*Institute of Systems Science, the Academy of Mathematics and Systems Science, the Chinese Academy of Sciences, Beijing* 100080, *China*)

Abstract

In this paper, we give a descent algorithm for solving quadratic bilevel programming problems. It is proved that the descent algorithm finds a locally optimal solution to a quadratic bilevel programming problem in a finite number of iterations. Two numerical examples are given to illustrate this algorithm.

Key words. Bilevel programming, quadratic programming, descent algorithm, finite convergence

1. Introduction

A bilevel programming problem (BLPP) involves two sequential optimization problems where the constraint region of the upper one is implicitly determined by the solution of the lower. It is proved in [1] that even to find an approximate solution of a linear BLPP is strongly NP-hard. A number of algorithms have been proposed to solve BLPPs. Among them, the descent algorithms constitute an important class of algorithms for nonlinear BLPPs. However, it is assumed for almost all those descent algorithms that the solution set of the lower level problem is a singleton for any given value of the upper level variables. Under this assumption, a BLPP can be transformed into a single level optimization problem where the lower level variables are taken as a function of the upper level variables. Those

Received July 26, 1998.

* This work is supported by the National Natural Science Foundation of China (No.19731001) and the Croucher Foundation of Hong Kong.

descent algorithms heavily depend on the information about this implicit function. On the basis of the gradient information generated from the lower level optimization problem, Kolstad and Lasdon[2] proposed a heuristic descent algorithm for BLPPs. Friesz et al.[3] analyzed some heuristic algorithms based on sensitivity analysis and suggested some rules to determine the step length along the descent directions. The computational results showed that the heuristic algorithms are quite efficient in computing an approximate solution of BLPPs, especially for nonlinear BLPPs. Unfortunately these algorithms do not necessarily converge, even to a locally optimal solution of a BLPP. Under some other assumptions, Dempe[4], Outrata, Zowe[5], and Falk, Liu[6] proposed three different methods to compute a subgradient of the implicit function determined by the lower level optimization problem respectively. Hence, the bundle method can be applied to compute a locally optimal solution of a BLPP. Pang et al.[7] proposed a method with a linear search scheme to solve nonsmooth unconstrained optimization problems and also applied it to solve BLPPs. It was mentioned in [8] that this method requires that the objective function of the problem is pseudo-regular. However, the objective function of the nonsmooth optimization problem converted from a nonlinear BLPP is generally not pseudo-regular. A few descent methods have been recently developed for some particular cases of BLPPs. Chen[9] designed a descent method to solve a BLPP appearing in transportation networks. Vicente et al.[10] presented a descent method for solving quadratic BLPPs. The descent direction is computed by solving a sequence of linear complementary problems, and an interesting technique was also introduced for determining the exact search step size along the descent direction in the induced region. Unfortunately, this method also lacks any convergence.

In this paper, we propose a new descent method for quadratic BLPPs. This method is finitely convergent and can be also applied to solve a BLPP where the upper level constraints is linear and the lower level programming problem is quadratic. The paper is organized as follows. Some preliminaries are introduced in Section 2. In Section 3, we discuss convergence of the steepest descent algorithm for quadratic BLPPs and explain why most of the descent algorithms appearing in the literature can not be applied to solve a quadratic BLPP, even for finding a locally optimal solution. A new descent algorithm for solving quadratic BLPPs is presented in Section 4. The convergence of this algorithm is analyzed in Sections 5. Finally, we give a few conclusions in Section 6.

2. Preliminaries

In this paper, we consider the following quadratic bilevel programming problem (QBP):

$$\begin{aligned} \min_{x\geq 0}\ & F(x,y)=\frac{1}{2}\begin{bmatrix}x\\ y\end{bmatrix}^T\begin{bmatrix}C_1, & C_3\\ C_3^T, & C_2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}+\begin{bmatrix}c_1\\ c_2\end{bmatrix}^T\begin{bmatrix}x\\ y\end{bmatrix}\\ \text{s.t.}\ & A_1x+B_1y\leq b_1,\end{aligned}$$

where y is a solution of

$$\begin{aligned} \min_{y\geq 0\|x}\ & f(x,y)=\tfrac{1}{2}y^TQy+y^TSx+d^Ty\\ \text{s.t.}\ & A_2x+B_2y\leq b_2,\end{aligned}$$

where $c_1\in R^n$, $c_2,d\in R^m$, $b_1\in R^p$, $b_2\in R^q$, $C_1\in R^{n\times n}$, $Q,C_2\in R^{m\times m}$, $S,C_3^T\in R^{m\times n}$, $A_1\in R^{p\times n}$, $A_2\in R^{q\times n}$, $B_1\in R^{p\times m}$, $B_2\in R^{q\times m}$. "$\|x$" stands for "when x is fixed".

The above (QBP) can be reformulated as the following single level programming problem (NP):

$$\begin{array}{ll} \min\limits_{x,y} & F(x,y) \\ \text{s.t.} & A_1x + B_1y \leq b_1, \\ & y \in R(x), \quad x \geq 0, \end{array}$$

where

$$R(x) = \arg\min_y \{f(x,y) \mid A_2x + B_2y \leq b_2, \ y \geq 0\}.$$

Denote

$$\Psi = \{(x,y) \mid A_1x + B_1y \leq b_1, \ y \in R(x), \ x \geq 0\}.$$

On the basis of this reformulation, we can introduce the concept of an optimal solution to (BLPP).

Definition 2.1. Ψ is called *the induced region* of (BLPP). $(\bar{x},\bar{y}) \in \Psi$ is called *an optimal solution* to (BLPP) if $F(\bar{x},\bar{y}) \leq F(x,y)$ for any $(x,y) \in \Psi$. $(\bar{x},\bar{y}) \in \Psi$ is called *a locally optimal solution* to (BLPP) if there is a neighborhood $N(\bar{x},\bar{y})$ of $(\bar{x},\bar{y})$ such that $F(\bar{x},\bar{y}) \leq F(x,y)$ for any $(x,y) \in \Psi \cap N(\bar{x},\bar{y})$.

Denote

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_2 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1, & C_3 \\ C_3^T, & C_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

In the sequel, we assume that both C and Q are symmetric and positive semi-definite. For simplicity, we also assume that the set

$$\{(x,y) \mid Ax + By \leq b, \ x \geq 0, \ y \geq 0\}$$

is compact. Replacing the lower level optimization problem in (QBP) by the K-K-T optimality condition, we have the following equivalent single level programming problem (SLP):

$$\begin{array}{ll} \min\limits_{x,y,u} & F(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}^T C \begin{bmatrix} x \\ y \end{bmatrix} + c^T \begin{bmatrix} x \\ y \end{bmatrix} \\ \text{s.t.} & Qy + Sx + d + B_2^T u_1 - u_2 = 0, \\ & Ax + By \leq b, \qquad x \geq 0, \ y \geq 0, \\ & u_1^T(b_2 - A_2x - B_2y) = 0, \qquad u_1 \geq 0, \\ & u_2^T y = 0, \qquad u_2 \geq 0, \end{array}$$

where $u^T = (u_1^T, u_2^T)$. Denote by D the feasible set of (SLP). An interesting property of D can be given if we denote the relaxed feasible set of (SLP) by $\overline{D}$, i.e.,

$$\overline{D} = \{(x,y,u) \mid Qy + Sx + d + B_2^T I_1 u - I_2 u = 0, \ Ax + By \leq b, \ x \geq 0, \ y \geq 0, \ u \geq 0\},$$

where $I_1 = (I_1', 0)$, $I_2 = (0, I_2')$, I_1' and I_2' being the $q \times q$ unit matrix and the $m \times m$ unit matrix respectively.

Theorem 2.1. The feasible set D of (SLP) is the union of some faces of $\overline{D}$.

Proof. It is obvious that $D \subset \overline{D}$ and $\overline{D}$ is a polyhedral set in R^{n+2m+q}. For any fixed $(\bar{x},\bar{y},\bar{u}) \in D$, there is a smallest face F of $\overline{D}$ satisfying $(\bar{x},\bar{y},\bar{u}) \in F$. We are only required to show that $F \subset D$. Let the set of all the vertices of F be $\{(x^1,y^1,u^1), \cdots, (x^k,y^k,u^k)\}$, where $k \geq 1$. There are $\lambda_i > 0$, $i = 1, \cdots, k$ satisfying $\sum\limits_{i=1}^{k} \lambda_i = 1$ such that

$$(\bar{x},\bar{y},\bar{u}) = \sum_{i=1}^{k} \lambda_i (x^i, y^i, u^i).$$

Because $(\bar{x}, \bar{y}, \bar{u}) \in D$, we have

$$\sum_{i=1}^{k} \lambda_i (u_1^i)^T \Big(b_2 - A_2 \sum_{i=1}^{k} \lambda_i x^i - B_2 \sum_{i=1}^{k} \lambda_i y^i\Big) = \bar{u}_1^T (b_2 - A_2 \bar{x} - B_2 \bar{y}) = 0.$$

Since $(x^i, y^i, u^i) \in \overline{D}$, for $i = 1, \cdots, k$, by the definition of $\overline{D}$,

$$(u_1^j)^T (b_2 - A_2 x^i - B_2 y^i) \geq 0, \qquad j = 1, \cdots, k, \quad i = 1, \cdots, k.$$

Hence,

$$\sum_{i=1}^{k} \lambda_i (u_1^i)^T \Big(b_2 - A_2 \sum_{i=1}^{k} \lambda_i x^i - B_2 \sum_{i=1}^{k} \lambda_i y^i\Big)$$
$$= \sum_{i=1}^{k} \sum_{j=1}^{k} \lambda_i \lambda_j (u_1^i)^T (b_2 - A_2 x^j - B_2 y^j) \geq \sum_{i=1}^{k} \lambda_i^2 (u_1^i)^T (b_2 - A_2 x^i - B_2 y^i) \geq 0.$$

Therefore,

$$(u_1^i)^T (b_2 - A_2 x^i - B_2 y^i) = 0, \qquad \text{for } i = 1, \cdots, k.$$

By a similar analysis, we can prove

$$(u_2^i)^T y^i = 0, \qquad \text{for } i = 1, \cdots, k.$$

So

$$(x^i, y^i, u^i) \in D, \qquad \text{for } i = 1, \cdots, k.$$

This implies that $F \subset D$. The proof is completed.

Before ending this section, we introduce the definition of two adjacent vertices of a polyhedral set.

Definition 2.2. Assume z_1 and z_2 are vertices of a polyhedral set F. If the convex hull $\mathrm{conv}\{z_1, z_2\}$ of the set $\{z_1, z_2\}$ is a face of F, then we call z_1 and z_2 the two adjacent vertices of F.

3. An Observation

In this section, we give a numerical example to illustrate that the steepest descent method is not necessarily convergent when it is applied to find a local minimum of a quadratic function on the union of some polyhedral sets. This observation can explain why some descent algorithms for BLPPs may not converge, even to a local minimum of a bilevel programming problem.

Consider applying the steepest-descent method to solve the following unconstrained quadratic programming problem (QP):

$$\min \ x^2 + ay^2,$$

where $x \in R^1$, $y \in R^1$ and a is a parameter satisfying $a > 1$. First we prove a lemma as follows.

Lemma 3.1. For any given initial point (x_0, y_0) satisfying $x_0 \neq 0$ and $y_0 \neq 0$, the steepest descent method with the exact linear search generates a sequence of infinite iteration points $\{(x_k, y_k)\}$ satisfying

$$x_{k+1} x_k > 0, \qquad y_{k+1} \neq 0, \qquad \text{for any } k \geq 0,$$

if it is applied to solve (QP).

Proof. We show this lemma by the induction. Without loss of generality, we suppose that $x_0 > 0$ and $y_0 \neq 0$. Assume that $x_j > 0$ and $y_j \neq 0$ for all $j = 0, 1, \cdots, k$, where k is a nonnegative integer. It suffices to verify that $x_{k+1} > 0$ and $y_{k+1} \neq 0$. The steepest descent direction of the function $x^2 + ay^2$ at (x_k, y_k) is $(-2x_k, -2ay_k)$. The linear search generates the next iteration point $(x_k(1-2t), y_k(1-2at))$, where

$$t = \frac{x_k^2 + a^2 y_k^2}{2(x_k^2 + a^3 y_k^2)}.$$

Because $a > 1$ and $y_k \neq 0$, we get that $1 - 2t > 0$. Hence, $x_{k+1} > 0$.

It is obvious that $2at \neq 1$. So $y_{k+1} \neq 0$. The proof is completed.

Example 3.1. Consider the following quadratic bilevel programming problem:

$$\begin{aligned} \min_{x} \quad & x_1^2 + 4x_2^2 + (y-1)^2 \\ \text{s.t.} \quad & -2 \leq x_1 \leq 2, \quad -1 \leq x_2 \leq 1, \end{aligned}$$

where y is a solution of

$$\begin{aligned} \min_{y \,\|\, x} \quad & y^2 \\ \text{s.t.} \quad & y - x_1 \geq 0, \qquad y \leq 2. \end{aligned}$$

It is not hard to show that the induced region of the above bilevel programming problem is the union of the following two polyhedral sets:

$$\Omega_1 = \{(x_1, x_2, y) \,|\, -2 \leq x_1 \leq 0, \ -1 \leq x_2 \leq 1, \ y = 0\}$$

and

$$\Omega_2 = \{(x_1, x_2, y) \,|\, 0 \leq x_1 \leq 2, \ -1 \leq x_2 \leq 1, \ y = x_1\}.$$

Let the initial iteration point be $(-2, -1, 0)$. The first iteration point generated by the steepest descent algorithm in [10] is $(-\frac{24}{17}, \frac{3}{17}, 0)$. The objective function value at this iteration point is $(\frac{24}{17})^2 + 4 \times (\frac{3}{17})^2$. By Lemma 3.1, this steepest descent algorithm generates a sequence of infinite iteration points in the relatively interior set $\mathrm{ri}\,\Omega_1$ of Ω_1. It is not difficult to show that this sequence converges to $(0,0,0)$. However, we can easily show that $(0,0,0) \in \Omega_2$ is not any local minimum of the bilevel programming problem.

This example reveals that any method based on the technique of the steepest descent algorithm is not convergent when it is applied to solve quadratic bilevel programming problems.

4. A New Descent Algorithm

As mentioned in [10], at least one local minimum of (QBP) is an extreme point of the induced region if the objective function F of the upper level is concave and thus their descent algorithm terminates finitely at a local minimum of (QBP) under a nondegeneracy assumption about the induced region. Unfortunately, the conclusion is not true when F is not concave because the iteration points generated by the descent algorithms for (QBP) are often not any extreme points of the induced region. In this section, we propose a new descent algorithm to solve (QBP). Our new algorithm utilizes a particular property of an extreme point related to the iteration point at each iteration.

Denote

$$\begin{aligned} &\Pi(x, y, u) = (x, y), \qquad \text{for any } (x, y, u) \in \overline{D}, \\ &\Pi(D) = \big\{(x, y) \,|\, (x, y, u) \in D\big\}, \\ &\Lambda(x, y) = \{u \,|\, (x, y, u) \in D\}, \qquad \text{for any } (x, y) \in \Pi(D). \end{aligned}$$

Let H be a polyhedral set in R^{n+2m+q} and z be a vertex of $\overline{D}$. We denote by $E(H)$ the set of all the vertices of H, and by $A(z)$ the set of all the vertices adjacent to z in $\overline{D}$. If $z \in \overline{D}$ and G is a nonempty convex subset of $\overline{D}$, we denote the smallest face of $\overline{D}$ including z and G respectively by $S(z)$ and $S(G)$.

Denote by $h'(z;d)$ the directional derivative of h at z in the direction d. Now we can state our new descent algorithm as follows.

Algorithm 4.1.
Step 0. Find an initial feasible solution (x_0, y_0) of (QBP) and let $k=0$.
Step 1. Set $V_k = \phi$.
Step 2. If $E(\Lambda(x_k, y_k)) = V_k$ then stop, and (x_k, y_k) is a locally optimal solution to (QBP); otherwise, proceed.
Step 3. Take a $\bar{u} \in E(\Lambda(x_k, y_k)) \setminus V_k$.
Step 4. Take a $z_k \in E(S(x_k, y_k, \bar{u}))$ and let $A_k = \phi$.
Step 5. If $A(z_k) \cap D = A_k$, then set $V_k = V_k \cup \{\bar{u}\}$ and go to Step 2; otherwise, proceed.
Step 6. Take a $\bar{z} \in A(z_k) \cap D \setminus A_k$.
Step 7. If $F(\bar{z}) < F(x_k, y_k)$, then set $(x_{k+1}, y_{k+1}) = \Pi(\bar{z})$, $k = k+1$ and go to Step 1; otherwise, proceed.
Step 8. If $F'\big((x_k, y_k); \Pi(\bar{z}) - (x_k, y_k)\big) \geq 0$, then set $A_k = A_k \cup \{\bar{z}\}$ and go to Step 5; otherwise, proceed.
Step 9. If $\frac{1}{2}\big((x_k, y_k, \bar{u}) + \bar{z}\big) \notin D$, then set $A_k = A_k \cup \{\bar{z}\}$ and go to Step 5; otherwise, proceed.
Step 10. Choose a maximal face T of $\overline{D}$ satisfying $\{(x_k, y_k, \bar{u}), \bar{z}\} \subset T \subset D$ and solve the following quadratic programming problem $(\mathrm{QP})_T$

$$\begin{aligned} \min_{x,y,u} \quad & F(x,y) \\ \text{s.t.} \quad & (x,y,u) \in T. \end{aligned}$$

Let $(x_{k+1}, y_{k+1}) \in \Pi\big(\arg\min\{F(x,y) \mid (x,y,u) \in T\}\big)$, $k = k+1$ and go to Step 1.

Before discussing the finite convergence of the algorithm, we illustrate this algorithm via a numerical example.

Example 4.1. Consider the following quadratic bilevel programming problem:

$$\begin{aligned} \min_{x \geq 0} \quad & \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}\Big(x_2 - \frac{2}{5}\Big)^2 + \frac{1}{2}\Big(y - \frac{4}{5}\Big)^2 \\ \text{s.t.} \quad & y \in \arg\min_{y \,\|\, x} \Big\{\frac{1}{2}y^2 - y - x_1 y + 3x_2 y \,\Big|\, 0 \leq y \leq 1\Big\}, \\ & x_1 \leq 1, \ \ x_2 \leq 1. \end{aligned}$$

Take $(1,0,1)$ as the initial iteration point. It is easy to verify that $\Lambda(1,0,1) = \{(0,1)\}$ and that $(1,0,1,0,1)$ is a vertex of $\overline{D}$. In the first iteration, we get a descent vertex $\big(1, \frac{1}{3}, 1, 0, 0\big)$ of $\overline{D}$ and then the first iteration point $\big(1, \frac{1}{3}, 1\big)$ with $\Lambda\big(1, \frac{1}{3}, 1\big) = \{0,0\}$. In the second iteration, the vertex $\big(1, \frac{2}{3}, 0, 0, 0\big)$ of $\overline{D}$ is found to get a descent direction. We choose $T = \{(x,y,u) \mid y - x_1 + 3x_2 = 1, \ u_1 = 0, \ u_2 = 0, \ 0 \leq x_1 \leq 1, \ 0 \leq x_2 \leq 1, \ 0 \leq y \leq 1\}$. T is a maximal face of $\overline{D}$ satisfying $\big\{\big(1, \frac{2}{3}, 0, 0, 0\big), \ \big(1, \frac{1}{3}, 1, 0, 0\big)\big\} \subset T \subset D$. Solving the corresponding quadratic programming problem $(\mathrm{QP})_T$:

$$\begin{aligned} \min_{(x,y,u) \geq 0} \quad & \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}\Big(x_2 - \frac{2}{5}\Big)^2 + \frac{1}{2}\Big(y - \frac{4}{5}\Big)^2 \\ \text{s.t.} \quad & y - x_1 + 3x_2 = 1, \quad u_1 = 0, \quad u_2 = 0, \\ & 0 \leq x_1 \leq 1, \ \ 0 \leq x_2 \leq 1, \ \ 0 \leq y \leq 1, \end{aligned}$$

we get the second iteration point $\left(1,\frac{2}{5},\frac{4}{5}\right)$. It is easy to show that $\Lambda\left(1,\frac{2}{5},\frac{4}{5}\right) = \{(0,0)\}$. In the third iteration, no extreme descent direction can be found and the algorithm terminates at the local minimum $\left(1,\frac{2}{5},\frac{4}{5}\right)$. In fact, $\left(1,\frac{2}{5},\frac{4}{5}\right)$ is also a global minimum of the problem discussed in this example.

At the k-th iteration of the above descent algorithm, we implicitly enumerate the vertices of $\Lambda(x^k, y^k)$. For each vertex u of $\Lambda(x^k, y^k)$, we check whether or not (x^k, y^k, u) is a local minimum of (SLP). If not, a new face of $\overline{D}$ is found to generate a new iteration point. The objective value of the upper level decreases as the algorithm is iterated. The crucial step in the algorithm is to choose a vertex z^k of the smallest face T of $\overline{D}$ satisfying $(x^k, y^k, u) \in T$ and to check all the vertices in D which are adjacent to z^k. If a feasible descent vertex or a feasible descent vertex direction is found, then the required face can be easily constructed. Otherwise, we can conclude that (x^k, y^k, u) is a local minimum of (SLP). If (x^k, y^k, u) is a local minimum of (SLP) for each vertex u of $\Lambda(x^k, y^k)$, then (x^k, y^k) is a local minimum of (QBP). When a descent direction is found, we solve a convex quadratic programming problem instead of a linear search. This can avoid searching repeatedly in one face of $\overline{D}$ and guarantee the finite convergence of the algorithm.

Now we apply the new descent algorithm to solve the quadratic bilevel programming problem given in Example 3.1. We denote by $\overline{D}_E$ the relaxed feasible solution set of the single level programming problem converted by replacing the lower level programming problem with the K-K-T optimality condition, i.e.,

$$\overline{D}_E = \{(x_1, x_2, y, u_1, u_2) \mid -2 \le x_1 \le 2,\ -1 \le x_2 \le 1,\ y - x_1 \ge 0,\ y \le 2,\ u_1 \ge 0,\ u_2 \ge 0\}.$$

Taking $(-2,-1,0)$ as the initial point, we have that $\Lambda(x,y) = \{(0,0)\}$. It is easy to verify that $(-2,-1,0,0,0)$ is a vertex of $\overline{D}_E$. At the first iteration, $(0,-1,0,0,0)$ is found which is a feasible descent vertex of $\overline{D}_E$ and is also a point in D adjacent to $(-2,-1,0,0,0)$. Thus, we get the first iteration point $(0,-1,0)$. At the second iteration, the vertex $(0,1,0,0,0)$ provides a feasible descent direction. We construct a maximal face T_2 of $\overline{D}_E$ satisfying $\{(0,-1,0,0,0),\ (0,1,0,0,0)\} \subset T_2 \subset D$ as follows:

$$T_2 = \{(x_1, x_2, y, u_1, u_2) \mid -2 \le x_1 \le 0,\ -1 \le x_2 \le 1,\ y = 0,\ u_1 = 0,\ u_2 = 0\}.$$

Solving the corresponding subproblem $(\mathrm{QP})_T$:

$$\begin{array}{ll} \min & F(x,y) = x_1^2 + 4x_2^2 + (y-1)^2 \\ \text{s.t.} & (x_1, x_2, y, u_1, u_2) \in T_2, \end{array}$$

we get an optimal solution $(0,0,0,0,0)$. Thus, the second iteration point is $(0,0,0)$. It is obvious that

$$S(0,0,0,0,0) = \{(x_1, x_2, y, u_1, u_2) \mid -1 \le x_2 \le 1,\ x_1 = 0,\ y = 0,\ u_1 = 0,\ u_2 = 0\}.$$

The point $(0,1,0,0,0)$ is a vertex of $S(0,0,0,0,0)$ and the vertex $(2,1,2,4,0)$ of $\overline{D}_E$ provides a feasible descent direction. The maximal face T_3 of $\overline{D}_E$ satisfying $\{(0,0,0,0,0),\ (2,1,2,4,0)\} \subset T_3 \subset D$ can be constructed as follows:

$$T_3 = \{(x_1, x_2, y, u_1, u_2) \mid 0 \le x_1 \le 2,\ -1 \le x_2 \le 1,\ y = x_1,\ u_1 = 2y,\ u_2 = 0\}.$$

Solving the following quadratic programming problem $(\mathrm{QP})_T$:

$$\begin{array}{ll} \min & F(x,y) = x_1^2 + 4x_2^2 + (y-1)^2 \\ \text{s.t.} & (x_1, x_2, y, u_1, u_2) \in T_3, \end{array}$$

we get an optimal solution $\left(\frac{1}{2},0,\frac{1}{2},1,0\right)$. Hence, the third iteration point is $\left(\frac{1}{2},0,\frac{1}{2}\right)$. It is not hard to show that among all the faces of $\overline{D}_E$, T_3 is also the smallest face of $\overline{D}_E$ which includes $\left(\frac{1}{2},0,\frac{1}{2},1,0\right)$. We can not find any feasible descent direction at this iteration point. Because $\Lambda\left(\frac{1}{2},0,\frac{1}{2}\right)=\{(1,0)\}$, we can conclude that $\left(\frac{1}{2},0,\frac{1}{2}\right)$ is a locally optimal solution of the quadratic programming problem in Example 3.1. In fact, it is a globally optimal solution to this bilevel programming problem.

5. Finite Convergence

First we prove three lemmas.

Lemma 5.1. Let $(\bar{x},\bar{y})$ be a feasible solution of (QBP). If $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP) for any vertex $\bar{u}$ of $\Lambda(\bar{x},\bar{y})$, then $(\bar{x},\bar{y})$ is a locally optimal solution of (QBP).

Proof. Suppose that $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP) for any vertex $\bar{u}$ of $\Lambda(\bar{x},\bar{y})$. If $(\bar{x},\bar{y})$ was not a locally optimal solution of (QBP), there would be a sequence of feasible points $\{(x_i,y_i)\}$ of (QBP) converging to $(\bar{x},\bar{y})$ and satisfying

$$F(x_i,y_i)<F(\bar{x},\bar{y}),\qquad \text{for } i=1,2,\cdots.$$

Choose a vertex u_i of $\Lambda(x_i,y_i)$ for $i=1,2,\cdots$. There must be a subsequence $\{u_{i_j}\}$ of $\{u_i\}$ which converges to $\bar{u}$. It can be verified that $\bar{u}$ is a vertex of $\Lambda(\bar{x},\bar{y})$. This contradicts the assumption that $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP). Hence, $(\bar{x},\bar{y})$ is a locally optimal solution of (QBP). The proof is completed.

Lemma 5.2. Let $(\bar{x},\bar{y},\bar{u})\in D$. If $F(\bar{x},\bar{y},\bar{u})\le F(z)$ for any face T of $\overline{D}$ satisfying $(\bar{x},\bar{y},\bar{u})\in T\subset D$ and any $z\in T$, then $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP).

Proof. Suppose that $F(\bar{x},\bar{y},\bar{u})\le F(z)$ for any face T of $\overline{D}$ satisfying $(\bar{x},\bar{y},\bar{u})\in T\subset D$ and any $z\in T$. If $(\bar{x},\bar{y},\bar{u})$ was not a locally optimal solution of (SLP), there would be a sequence of points $\{(x_i,y_i,u_i)\}$ in D converging to $(\bar{x},\bar{y},\bar{u})$ and satisfying

$$F(x_i,y_i,u_i)<F(\bar{x},\bar{y},\bar{u}),\qquad \text{for } i=1,2,\cdots.$$

Because there are only finite faces of $\overline{D}$, there are a subsequence $\{(x_{i_k},y_{i_k},u_{i_k})\}$ of $\{(x_i,y_i,u_i)\}$ and a face $\overline{T}$ of $\overline{D}$ such that

$$S(x_{i_k},y_{i_k},u_{i_k})=\overline{T},\qquad \text{for } k=1,2,\cdots.$$

Therefore, $(\bar{x},\bar{y},\bar{u})\in\overline{T}$ and $S(\bar{x},\bar{y},\bar{u})\subset\overline{T}$. From Theorem 2.1, we have $\overline{T}\in D$. This contradicts the assumption that $F(\bar{x},\bar{y},\bar{u})\le F(z)$ for any face T of $\overline{D}$ satisfying $(\bar{x},\bar{y},\bar{u})\in T\subset D$ and any $z\in T$. So $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP).

Lemma 5.3. Let $(\bar{x},\bar{y},\bar{u})\in D$ and $\bar{z}\in E\big(S(\bar{x},\bar{y},\bar{u})\big)$ with $F(\bar{z})\ge F(\bar{x},\bar{y},\bar{u})$. If $F'\big((\bar{x},\bar{y},\bar{u});\,z-(\bar{x},\bar{y},\bar{u})\big)\ge 0$ for any $z\in A(\bar{z})\cap D$ satisfying $\frac{1}{2}(z+\bar{z})\in D$, then $(\bar{x},\bar{y},\bar{u})$ is a locally optimal solution of (SLP).

Proof. It suffices to verify that the condition of Lemma 5.2 is satisfied if $F'\big((\bar{x},\bar{y},\bar{u});\,z-(\bar{x},\bar{y},\bar{u})\big)\ge 0$ for any $z\in A(\bar{z})\cap D$ satisfying $\frac{1}{2}(z+\bar{z})\in D$. Let T be a face of $\overline{D}$ satisfying $(\bar{x},\bar{y},\bar{u})\in T\subset D$ and z be a point of T. Thus there are a positive integer t, t numbers $\lambda_i\ge 0$ and t points $z_i\in A(\bar{z})\cap T$, $i=1,\dots,t$ such that

$$z=\bar{z}+\sum_{i=1}^{t}\lambda_i(z_i-\bar{z}).\tag{1}$$

If $\sum\limits_{i=1}^{t}\lambda_i\le 1$, let $\lambda_0=1-\sum\limits_{i=1}^{t}\lambda_i$ and $z_0=\bar{z}$. Note that

$$F'\big((\bar{x},\bar{y},\bar{u});\ z-(\bar{x},\bar{y},\bar{u})\big)=\sum_{i=0}^{t}\lambda_i F'\big((\bar{x},\bar{y},\bar{u});\ z_i-(\bar{x},\bar{y},\bar{u})\big)\geq 0.$$

By the convexity of F, we have

$$F(z)\geq F(\bar{x},\bar{y},\bar{u}).$$

Now we suppose $\sum_{i=1}^{t}\lambda_i>1$. We have

$$(\bar{x},\bar{y},\bar{u})=\mu_0\bar{z}+\sum_{i=1}^{t_1}\mu_i z_i+\sum_{i=t+1}^{t_2}\mu_i z_i, \tag{2}$$

where $\mu_0>0$, $\mu_i\geq 0$ and $z_i\in E\big(S(\bar{x},\bar{y},\bar{u})\big)\cap A(\bar{z})$ for $i=1,2,\cdots,t_1$, $\mu_i\geq 0$ and $z_i\in E\big(S(\bar{x},\bar{y},\bar{u})\big)\setminus A(\bar{z})$ for $i=t+1,\cdots,t_2$ and $\sum_{i=0}^{t_1}\mu_i+\sum_{i=t+1}^{t_2}\mu_i=1$. Without loss of generality, we can assume $t_1\leq t$. Let $M=\sum_{i=1}^{t}\lambda_i$, $\mu_i=0$ for $i=t_1+1,\cdots,t$ and $\lambda_i=0$ for $i=t+1,\cdots,t_2$. By (2), we can replace $\bar{z}$ in (1) by $(\bar{x},\bar{y},\bar{u})$. Therefore,

$$z=\frac{1}{\mu_0}\sum_{i=0}^{t_2}\big(\mu_0\lambda_i+(M-1)\mu_i\big)\big(z_i-(\bar{x},\bar{y},\bar{u})\big)+(\bar{x},\bar{y},\bar{u}).$$

Similarly, we can also get

$$F(z)\geq F(\bar{x},\bar{y},\bar{z}).$$

This completes the proof.

Now we can state our main result as follows.

Theorem 5.1. Algorithm 4.1 finds a locally optimal solution of (QBP) in a finite number of iterations.

Proof. The algorithm generates a new iteration point which is a globally optimal solution of the objective function $F(x,y)$ of the upper level on a face of $\overline{D}$ at each iteration. Because $\overline{D}$ has only a finite faces and any face of $\overline{D}$ is at most checked once, the algorithm is determinate after a finite number of iterations. By Lemmas 5.1, 5.2 and 5.3, the final iteration point generated by the algorithm is a locally optimal solution of (QBP).

6. Conclusions

We have presented a new descent algorithm for solving quadratic bilevel programming problems. The algorithm finds a locally optimal solution to a quadratic bilevel programming problem in a finite number of iterations. No assumption of any strict convexity is made to the lower level optimization problem when applying the new descent algorithm to solve QBPPs. But such an assumption is necessary for the other algorithms applied to solve quadratic bilevel programming problems. From the proofs of Theorem 5.1 and Lemmas 5.1, 5.2 and 5.3, we observe that Algorithm 4.1 can also be applied to solve the bilevel programming problems in which the upper level objective function is convex but not necessarily quadratic and that the algorithm terminates at a locally optimal solution of the problem in a finite number of iterations in this situation. When the upper level objective function is not convex, the algorithm can still be applied without any modification and the algorithm terminates in a finite number of iterations as well. In this case, however, the final iteration point is not necessarily a locally optimal solution of the bilevel programming problem.

The main feature of the new descent algorithm is its finite convergence. The maximal number of iterations is not more than the number of the faces H of $\overline{D}$ satisfying $H\subset D$.

If the set of all the optimal solutions of the lower level optimization problem in a general quadratic bilevel programming problem is not a singleton, most of the methods based on sensitivity analysis can not be applied to solve the bilevel problem because no descent direction can be found at an iteration point in this situation. In Algorithm 4.1, the extreme point algorithm for linear-quadratic bilevel programming problems proposed in [10] is applied, with some modifications, to find a descent direction for a convex quadratic bilevel programming problem. Hence, the algorithm proposed in this paper possesses the main advantage of the extreme point algorithm even when the iteration point is not an extreme point of the induced region.

Finally, we would like to mention that computational experiments are required to be made to test this new descent algorithm. We will report the numerical results in a separate paper.

References

1 X. Deng, Q. Wang, S. Wang. On Complexity of Linear Bilevel Programming. Operations Research and It's Applications, edited by D.Z. Du, X.S. Zhang and K. Chuan, World Publishing Corporation, Singapore, 205–212, 1995

2 C. Kolstad, L. Lasdon. Derivative Evaluation and Computational Experience with Large Bilevel Mathematical Programs. *Journal of Optimization Theory and Applications*, 1990, 65: 485–499

3 T. Friesz, R. Tobin, H. Cho, N. Mehta. Sensitivity Analysis Based Heuristic Algorithms for Mathematical Programs with Variational Inequality Constraints. *Mathematical Programming*, 1990, 48: 265–284

4 S. Dempe. On Generalized Differentiability of Optimal Solutions and Its Application to an Algorithm for Solving Bilevel Optimization Problems. Recent Advances in Nonsmooth Optimization, edited by D.Z. Du, L. Qi and R.S. Womersley, World Scientific Publishing Co., Singapore, 36–56, 1995

5 J. Outrata, J. Zowe. A Numerical Approach to Optimization Problems with Variational Inequality Constraints. *Mathematical Programming*, 1995, 68: 105–130

6 J.E. Falk, J. Liu. On Bilevel Programming, Part I: General Cases. *Mathematical Programming*, 1995, 70: 47–72

7 J.S. Pang, S.H. Han, N. Rangaraj. Minimization of Locally Lipschitzian Functions. *SIAM Journal on Optimization*, 1991, 1: 57-82

8 R. Poliquin, L. Qi. Iteration Function for Some Nonsmooth Optimization Algorithms. *Mathematics of Operations Research*, 1995, 20: 479–496

9 Y. Chen. Bilevel Programming Problems: Analysis, Algorithms and Applications. Universite'de Montré, Montré, Ph.D. Thesis, 1993

10 L. Vicente, G. Savard, J. Judice. Descent Approaches for Quadratic Bilevel Programming. *Journal of Optimization Theory and Applications*, 1994, 81: 379–399

11 A. Frangioni. A New Class of Bilevel Programming Problems and Its Use for Reformulating Mixed Inter Problems. *European Journal of Operational Research*, 1995, 82: 615–648

12 P. Hansen, B. Jaumard, G. Savard. New Branch and Bound Rules for Linear Bilevel Programming. *SIAM Journal of Scientific Statistics and Computations*, 1992, 13: 1194–1217

13 J.J. Judice, A. Faustino. The Linear-quadratic Bilevel Programming Problems. *INFOR*, 1994, 32: 87–98

14 A. Migdalas, P.M. Pardalos. Nonlinear Bilevel Problems with Convex Second Level Problem—Heuristics and Descent Methods. Operations Research and It's Applications, edited by D.Z. Du, X.S. Zhang and K. Chuan, World Publishing Corporation, Singapore, 194–204, 1995

15 D. Ralph, S. Dempe. Directional Derivatives of the Solution of a Parametric Nonlinear Program. *Mathematical Programming*, 1995, 70: 159–172