A Quasiconcave Minimization Method for Solving Linear Two-Level Programs

HOANG TUY, ATHANASIOS MIGDALAS, and PETER VÄRBRAND
Department of Mathematics, Linköping Institute of Technology, S-581 83 Linköping, Sweden

(Received: 4 March 1992; accepted: 29 May 1993)

Abstract. In this paper the linear two-level problem is considered. The problem is reformulated to an equivalent quasiconcave minimization problem, via a reverse convex transformation. A branch and bound algorithm is developed which takes the specific structure into account and combines an outer approximation technique with a subdivision procedure.

Key words. Linear two-level program, global optimization, Stackelberg game, quasiconcave minimization, branch and bound, outer approximation, subdivision procedure.

1. Introduction

Linear two-level programming, a special case of multi-level programming, deals with optimization problems in which the constraint region is implicitly determined by another optimization problem.

The model can be considered as a two-person game where one of the players, the leader, knows the cost function mapping of the second player, the follower, who may or may not know the cost function of the leader. The follower knows however the strategy of the leader and takes this into account when computing his own strategy. The leader can foresee the reactions of the follower and can therefore optimize his choice of strategy. The problems associated with leader and the follower are often referred to as the outer problem and the inner problem respectively.

A number of algorithms have been proposed to solve the linear two-level problem since the first solution technique proposed by Falk [13]. Falk studied the general max-min problem – a special case of the linear two-level problem, where the inner objective function is the negative of the outer objective function. His algorithm is based on branch and bound and linear programming techniques.

Bialas and Karwan [9] have shown that the solution of the problem must occur at an extreme point of the feasible set (obtained by all linear constraints to both the outer and the inner problem). Based on this observation they proposed the "k^{th} best algorithm" that finds the optimal solution by an explicit enumeration scheme. The same idea has been utilized in the enumeration method by Candler and Townsley [12] and the B&B-algorithm by Moore and Bard [7]. Other B&B-algorithms have also been developed by Judice and Faustino [19], Bard and Moore [6] and Hansen *et al.* [15].

Another common solution technique is to replace the inner problem by its

Journal of Global Optimization **4**: 243–263, 1994.

Karush–Kuhn–Tucker conditions and hence obtain an ordinary mathematical programming problem with a single objective function. Here, the difficulty occurs in the set of constraints – the complementary slackness conditions. Bard and Falk [4] used this idea and proposed (for the nonlinear two-level problem) a non-convex programming algorithm based on branch and bound. The method can fail to find the global optimal solution. Fortuny and McCarl [14] introduced 0/1-variables in order to take care of the complementarity slackness conditions and instead solve a large mixed integer programming problem. The Karush–Kuhn–Tucker approach has also been used by Bialas and Karwan [9] and Bialas, Karwan and Shaw [10] in what they call "the parametric complementarity pivot algorithm".

Narula and Nwosu [23] and [24] proposed a solution procedure that solves the problem in three phases. In phase I it finds the reaction set (the set of possible optimal solutions to the inner problem) of the inner problem with respect to the outer problem and, if necessary, partitions the reaction set into non overlapping linear sets. In phase II the "subproblems" are solved while the global optimum is found in phase III. However, it is shown in Ben-Ayed [8] that this approach may fail to find the global optimum.

Another class of solution methods tries to solve the linear two-level programming problem via multiple objective linear programming [5], [31]. Here the two objective functions are weighted together to give a standard linear programming problem. However, Wen and Hsu [32] have shown that, in general, there is no such relationship between bilevel and bicriteria programming problems.

Tuy *et al.* [30] restated the linear two-level problem as a global optimization problem and a new method based on this approach have been developed. The most important feature of this new solution method is that it attempts to maximally exploit the structure of the constraints based on recent global optimization techniques. Recently, it has also been discovered that multilevel problems are actually complementarity convex optimization problems with a very specific structure. This seems also to be a very promising approach to this important class of problems.

For the nonlinear two-level problem, Bard and Falk [4] and Bard [5] developed a one-dimensional search algorithm that yields a locally optimal solution. Local optimization methods based on nonlinear programming such as penalty or barrier function methods and direct gradient methods have also been used to approach the optimal solution smoothly, see e.g. Loridan and Morgan [22] and Kolstad and Lasdon [21]. Aiyoshi and Shimisu [1] made a reformulation of the problem into a one level optimization problem and applied a penalty function method to solve it. Anandalingam and White [2] presented a solution procedure for the linear two-level program where the duality gap of the lower problem is appended to the objective of the upper problem with a penalty. Judice and Faustino [20] proposed a method for the linear two-level program which consists of solving a sequence of linear complementarity problems by using a hybrid enumeration method.

The computational complexity of the linear two-level problem is treated in Blair [11]. The paper by Ben-Ayed [8] gives a survey on the linear two-level problem. The recent volume on hierarchical optimization, edited by Anandalingam and Friesz [3] contains papers on the whole spectrum of bilevel programs: linear and nonlinear, theory and algorithms as well as applications.

The linear two-level programming problem can be formulated as

$$
\begin{aligned}
\textbf{[P]}\quad & \min_{x\geq 0} \quad c_1^T x + d_1^T y \\
& \text{s.t.} \quad A_1 x + B_1 y \leq g_1 \\
& \qquad \text{and } y \text{ solves } R(x), \text{ where } R(x) \text{ is the problem}
\end{aligned}
$$

$$
\begin{aligned}
\textbf{[R(x)]}\quad & \min_{y\geq 0} \quad d_2^T y \\
& \text{s.t.} \quad A_2 x + B_2 y \leq g_2
\end{aligned}
$$

$$(x \in R^p,\ y \in R^q,\ g_1 \in R^{m_1},\ g_2 \in R^{m_2}).$$

In an earlier paper [30] we have shown that this problem (**P**) is equivalent to the following reverse convex program

$$
\begin{aligned}
\textbf{[Q]}\quad & \min \quad c_1^T x + d_1^T y \\
& \text{s.t.} \quad A_1 x + B_1 y \leq g_1 \\
& \qquad A_2 x + B_2 y \leq g_2 \\
& \qquad\quad x, y \geq 0 \\
& \qquad\quad \varphi(x) \geq d_2^T y \\
& \qquad \text{where } \varphi(x) \text{ is the optimal value of } \mathbf{R(x)}\ (\varphi(x) \text{ is convex}).
\end{aligned}
$$

By exploiting the specific structure of the reverse convex constraint $\varphi(x) \geq d_2^T y$, a polyhedral annexation method was proposed in [30] which works basically in a space of much smaller dimension than $p + q$.

In the present paper we will discuss a branch and bound algorithm for solving (**P**) based on the reduction of this problem to a quasiconcave minimization problem in a space of dimension $1 + \mathit{rank}\, A_2$.

2. The Equivalent Quasiconcave Minimization Problem

Setting

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, W = \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix}, c = \begin{pmatrix} c_1 \\ d_1 \end{pmatrix}, g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \Phi(u) = \varphi(x), h(u) = d_2^T y\ ,$$

we can rewrite (**Q**) in the form

$$\min\{c^T u \mid Wu \leq g, \Phi(u) - h(u) \geq 0, u \geq 0\}\ .$$

Let u^0 be a basic optimal solution of the linear program

$$\min\{c^T u \mid Wu \leqslant g, u \geqslant 0\} .$$

We can assume that $\Phi(u^0) - h(u^0) < 0$, otherwise u^0 would solve (**P**). Let

$$D = \{u \mid Wu \leqslant g, u \leqslant 0\} , \quad C = \{u \mid \Phi(u) - h(u) \leqslant 0\} ,$$
$$\hat{D} = D - u^0 , \quad \hat{C} = C - u^0 ,$$

where $\hat{D}$ is a polyhedron, assumed, for simplicity, to be nonempty and bounded. Furthermore, the origin 0 is a vertex of $\hat{D}$ and $\hat{D} \subset \hat{C}$.

The function $\varphi(x) = \min\{d_2^T y : B_2 y \leqslant g_2 - A_2 x,\ y \geqslant 0\}$ is a convex function and we shall assume that $\varphi(x) < +\infty\ \forall x$. Consequently, $\varphi(x)$ is continuous everywhere, hence $\Phi(u)$ is continuous everywhere, C is closed and

$$int\, C = \{u : \Phi(u) - h(u) < 0\} .$$

The general case when $\varphi(x)$ may be $+\infty$ for some x (i.e. **R(x)** may be infeasible), can be reduced to the previous one, see Appendix A.

We can then reformulate the problem as

$$[\hat{\mathbf{Q}}] \quad \min\{c^T u : u \in \hat{D} \backslash int\, \hat{C}\} .$$

Since u^0 is a vertex of D and $\Phi(u^0) - h(u^0) < 0$, i.e. $u^0 \in int\, C$, it follows that 0 is a vertex of $\hat{D}$ and

$$0 \in \hat{D} \cap int\, \hat{C} . \tag{1}$$

An important structural feature of the problem that should be exploited is the following property that has been established in ([30]).

PROPOSITION 1. *The convex set $\hat{C}$ contains the cone*

$$K = \{u = (x, y) \mid A_2 x \leqslant 0, d_2^T y \geqslant 0\} .$$

Denote by a^i, $i = 1, \ldots, m_2$ the vector of R^{p+q} whose first p components form the i-th row of A_2 and whose last q components are all zero. Also, let $a^0 = (0, -d_2) \in R^{p+q}$. Without loss of generality it can be assumed that $a^0, a^1, \ldots, a^{m_2}$ are linearly independent (this amounts to assuming that $d_2 \neq 0$ and the rows of A_2 are linearly independent). It follows from Proposition 1 that

$$\hat{C}^* \subset K^* = cone(a^0, a^1, \ldots, a^{m_2}) \tag{2}$$

where * denotes the polar. This suggests considering the dual problem of ($\hat{\mathbf{Q}}$) in the sense of [26]. As shown in the latter paper this dual problem will be a quasiconcave minimization problem.

Define a function $f: K^* \to (-\infty, +\infty]$ as follows:

$$f(v) = \inf\{c^T u : u \in \hat{D},\ \langle v, u\rangle \geq 1\} \tag{3}$$

(with the usual convention $\inf \emptyset = +\infty$).

PROPOSITION 2. *The function $f(v)$ is quasiconcave on K^*.*

Proof. Let $[v', v'']$ be any line segment in K^*. For any $v \in [v', v'']$, we obviously have $\langle v, u\rangle \leq \max\{\langle v', u\rangle, \langle v'', u\rangle\}$, for all u, hence $f(v) \geq \min\{f(v'), f(v'')\}$. □

Actually, the formula (3) defines a quasiconcave function over the whole space generated by K^*.

THEOREM 1. *Problem* $(\hat{\mathbf{Q}})$ *is equivalent to the quasiconcave minimization problem*

$$[(\mathbf{QCM_1})] \quad \min f(v) \quad s.t.\ v \in \hat{C}^*$$

in the following sense:

The optimal value in the two problems are equal and if $\bar{v}$ solves $(\mathbf{QCM_1})$ *then the vector $\bar{u} \in \operatorname{argmin}\{c^T u : u \in \hat{D},\ \langle \bar{v}, u\rangle \geq 1\}$ solves* $(\hat{\mathbf{Q}})$.

Proof. First observe that since $\hat{C}$ is closed, convex and contains 0 it is known and can be checked easily that $int\hat{C} = \{u : vu < 1,\ \forall v \in \hat{C}^*\}$. Now if u is feasible to $(\hat{\mathbf{Q}})$ then $u \in \hat{D} \backslash int\hat{C}$, hence $vu > 1$ for at least one $v \in \hat{C}^*$; so u must be feasible to the problem defining $f(v)$, and therefore $c^T u \geq f(v)$. This shows that the optimal value of $(\hat{\mathbf{Q}})$ is never less than that of $(\mathbf{QCM_1})$. On the other hand, if $v \in \hat{C}^*$ then any u feasible to the problem defining v must satisfy $u \in \hat{D}$, $vu \geq 1$, hence $u \not\in int\hat{C}$, i.e. u must be feasible to $(\hat{\mathbf{Q}})$. This shows that the optimal value of $(\mathbf{QCM_1})$ is never less than that of $(\hat{\mathbf{Q}})$. Therefore, the optimal values in the two problems are equal. The conclusion then follows easily. □

Thus, to solve $(\mathbf{P})$ it suffices to solve $(\mathbf{QCM_1})$.

3. Basic Properties of $(\mathbf{QCM_1})$

The objective function $f(v)$ and the feasible set $\hat{C}^*$ in the problem $(\mathbf{QCM_1})$ have some nice properties that should be taken into account when designing an efficient solution method. In this section we discuss these properties.

PROPOSITION 3. *The function $f(v)$ is lower semicontinuous.*

Proof. Let $v^k \to v^0$ and $f(v^k) \leq \alpha$. Then for each k there exist a $u^k \in \hat{D}$ such that $\langle v^k, u^k\rangle \geq 1$ and $f(u^k) = c^T u^k \leq \alpha$. By taking a subsequence $u^k \to u^0 \in \hat{D}$ such that $\langle v^0, u^0\rangle \geq 1$ and since $\hat{D}$ is bounded by hypothesis, it then follows that $f(v^0) \leq c^T u^0 \leq \alpha$.

PROPOSITION 4. *$f(v)=+\infty$ if and only if $\theta v \in \partial \hat{D}^*$ for some $\theta > 1$.*

Proof. The problem in (3) has empty feasible set if and only if $v^T u < 1$ for all $y \in \hat{D}$, i.e. since $\hat{D}$ is compact, $\max\{v^T u : u \in \hat{D}\} < 1$. Hence the conclusion. □

PROPOSITION 5. *The function $F_v(\lambda) = f(\frac{v}{1-\lambda})$ is convex polyhedral. If*

$$\xi := \min\{c^T u : u \in \hat{D}\} < f(v) < +\infty$$

then $f(\theta v) < f(v)$ for all $\theta > 1$.

Proof. Setting $\theta = \frac{1}{1-\lambda}$ we can write

$$\begin{aligned} F_v(\lambda) = f(\theta v) &= -c^T u^0 + \min\{c^T u : u \in D, v^T(u-u^0) \geq 1-\lambda\} \\ &= -c^T u^0 + \min\{c^T u : Wu \leq g, v^T u \geq v^T u^0 + 1 - \lambda, u \geq 0\} . \end{aligned} \tag{4}$$

If $f(v) < +\infty$ then, since $f(0) = +\infty$, it follows from the quasiconcavity of f that $f(\theta v) < +\infty \, \forall \theta > 1$ and by the duality theorem of linear programming,

$$\begin{aligned} f(\theta v) = &-c^T u^0 \\ &+ \sup\{-g^T t + (v^T u^0 + 1 - \lambda) t_0 : -W^T t + t_0 v \leq c, (t, t_0) \geq 0\} . \end{aligned} \tag{5}$$

This shows in particular that $F_v(\lambda) = f(\theta v)$ is a convex polyhedral function of λ. Furthermore, from linear parametric programming theory it is also known that this parametric program in λ has a basic optimal solution $(\bar{t}, \bar{t}_0)$ which is optimal for all λ in some interval $[0, \lambda_1) \subset [0, 1)$ (note that for $\lambda = 0$ this program has a finite optimal value which is $F_v(0) = f(v)$). Hence,

$$f(\theta v) = -c^T u^0 - g^T \bar{t} + (v^T u^0 + 1 - \lambda) \bar{t}_0 \; \forall \lambda \in [0, \lambda_1) . \tag{6}$$

Here it is easy to see that $\bar{t}_0 > 0$. Indeed, if $\bar{t}_0 = 0$ then

$$\begin{aligned} -g^T \bar{t} &= \sup\{-g^T t : -W^T t \leq c, t \geq 0\} = \min\{c^T u : Wu \leq g, u \geq 0\} \\ &= c^T u^0 + \xi , \end{aligned}$$

hence for $\lambda = 0$ in (6) we get $f(v) = -c^T u^0 - g^T \bar{t} = \xi$, a contradiction. □

For every $v \in K^*$ let us define

$$\sigma(v) = \max\{v^T u : u \in \hat{D}\} , \quad \hat{v} = \frac{v}{\sigma(v)} \tag{7}$$

Since $\hat{D}$ is compact and $0 \in \hat{D}$ we always have $0 \leq \sigma(v) < +\infty$. If $\sigma(v) = 0$, $\hat{v}$ will be understood as a point at infinity in the direction v.

COROLLARY 1. *For every v with $\sigma(v) > 0$ the point $\hat{v} = \frac{v}{\sigma(v)}$ belongs to the boundary of the polytope $\hat{D}^*$. The function $f(\theta v)$ (for fixed v) is equal to $+\infty$ in the* *interval* $0 \leq \theta \leq \frac{1}{\sigma(v)}$ *and monotonically decreases from* $f(\hat{v})$ *to* ξ *in the interval* $\frac{1}{\sigma(v)} \leq \theta < +\infty$. *If* $\sigma(v) = 0$ *then* $\theta v \in \hat{D}^*$ *and* $f(\theta v) = +\infty \, \forall \theta \geq 0$.

Proof. Clearly, $\max\{v^T u : u \in \hat{D}\} = 1$ while $\max\{\theta v^T u : u \in \hat{D}\} > 1$, $\forall \theta > 1$. That is, $v \in \hat{D}^*$ but $\theta v \notin \hat{D}^*$, for any $\theta > 1$. Hence, $v \in \partial \hat{D}^*$. The rest follows from Proposition 3, 4 and 5. □

REMARK. Since $c^T u^0$ is inferior to the optimal value of (**Q**), it is easily seen that the optimal value of $\hat{\mathbf{Q}}$ must be positive. Hence, $f(v) > 0$, $\forall v \in \hat{C}^*$.

Before proving the next property it is convenient to mention the following simple but useful fact.

LEMMA 1. *Let* $\eta(u)$ *be a convex function defined as*

$$\eta(u) = \sup\{\langle Hu + s, t\rangle : t \in \Delta\} ,$$

where H is a matrix, s a vector and Δ *an arbitrary set of values t. For any* $\tilde{u}$ *where* $\eta(\tilde{u})$ *is finite, denote by* $\tilde{t}$ *the value such that* $\eta(\tilde{u}) = \langle H\tilde{u} + s, \tilde{t}\rangle$, *i.e.* $\tilde{t}$ *achieves the supremum in the problem defining* $\eta(\tilde{u})$. *Then*

$$H^T \tilde{t} \in \partial \eta(\tilde{u}) .$$

Proof. It is easily seen that $\eta(u) - \eta(\tilde{u}) \geq \langle Hu + s, \tilde{t}\rangle - \langle H\tilde{u} + s, \tilde{t}\rangle = \langle H(u - \tilde{u}), \tilde{t}\rangle = \langle H^T \tilde{t}, u - \tilde{u}\rangle$, hence the conclusion. □

Now consider any feasible solution $\tilde{u} = (\tilde{x}, \tilde{y})$ of ($\hat{\mathbf{Q}}$). Then $\tilde{u} \in \hat{D} \cap \partial \hat{C}$. The next property deals with the question of how to derive, from $\tilde{u}$, a point $\tilde{v} \in \hat{C}^*$ with $f(\tilde{v}) \leq c^T \tilde{u}$. Let $\bar{u} = (\bar{x}, \bar{y})$, $\bar{x} = \tilde{x} + x^0$ and $\bar{y} = \tilde{y} + y^0$. Since

$$\varphi(\bar{x}) = \min\{d_2^T y : -B_2 y \geq A_2 \bar{x} - g_2, \, y \geq 0\} = d_2^T \bar{y}$$

it follows that

$$\varphi(\bar{x}) = \max\{\langle A_2 \bar{x} - g_2, t\rangle : -B_2^T t \leq d_2, t \geq 0\} . \tag{8}$$

PROPOSITION 6. *Let* $\tilde{z} = (A_2^T \bar{t}, -d_2) \in R^{p+q}$, *where* $\bar{t}$ *is an optimal solution of the linear program in* (8). *Then for the vector*

$$\tilde{v} = \frac{\tilde{z}}{\langle \tilde{z}, \tilde{u}\rangle}$$

we have $f(\tilde{v}) \leq c^T \tilde{u}$.

Proof. By Lemma 1, $A_2^T \bar{t} \in \partial \varphi(\bar{x})$, hence $\langle \tilde{z}, u - \tilde{u}\rangle = \langle A_2^T, x - \tilde{x}\rangle - \langle d_2, y - \tilde{y}\rangle \leq (\varphi(x) - d_2^T y) - (\varphi(\bar{x}) - d_2^T \bar{y}) = \varphi(\bar{x}) - d_2^T \bar{y} \leq 0$ whenever $\varphi(x) - d_2^T y \leq 0$. That is, $\langle \tilde{z}, u\rangle \leq \langle \tilde{z}, \tilde{u}\rangle$, $\forall u \in \hat{C}$. Here $\tilde{z} \neq 0$ because $d_2 \neq 0$ and since $0 \in int\hat{C}$, it follows that $\langle \tilde{z}, \tilde{u} \rangle > 0$. Thus, $\tilde{v}$ is well defined and $\langle \tilde{v}, u \rangle \leqslant 1$, $\forall u \in \hat{C}$, while $\langle \tilde{v}, \tilde{u} \rangle = 1$. But then from (3), $f(\tilde{v}) \leqslant c^T \tilde{u}$. □

We close this section by the following property of C.

PROPOSITION 7. *The set C is a polyhedron which is the projection on the (x, y) space of the polyhedron defined by the inequalities*

$$A_2 x + B_2 z \leqslant g_2 , \quad d_2^T z - d_2^T y \leqslant 0 , \quad z \geqslant 0 . \tag{9}$$

Proof. If $(x, y) \in C$, i.e. $\varphi(x) - d_2^T y \leqslant 0$, then $\varphi(x) < +\infty$ and for any optimal solution z of the subproblem $\mathbf{R(x)}$ we have (9). Conversely, if for (x, y) there exists z satisfying (9) then z is a feasible solution to $\mathbf{R(x)}$, hence $\varphi(x) \leqslant d_2^T z \leqslant d_2^T y$, i.e. $(x, y) \in C$. Thus, C is actually the projection of the polyhedron (9) on the (x, y) space. It is known then that C itself is a polyhedron. □

COROLLARY 2. *Let $v \in K^*$, i.e. $v = \Sigma_{i=0}^{m_2} t_i a^i$ with $t_i \geqslant 0$, and denote by $\rho(v)$ the optimal value of the linear program*

$$\begin{aligned}
(\mathbf{LP(v)}) \quad & \max \sum_{i=1}^{m_2} t_i \langle a^i, x - x^0 \rangle - t_0 \langle d_2, y - y^0 \rangle \\
& s.t. \quad A_2 x + B_2 z \leqslant g_2 \\
& \qquad\quad d_2^T z - d_2^T y \leqslant 0 \\
& \qquad\qquad\qquad z \geqslant 0
\end{aligned} \tag{10}$$

Then $v \in \hat{C}^*$ if and only if $\rho(v) \leqslant 1$.

Proof. This follows, since by definition $v \in \hat{C}^*$ if and only if $v^T u \leqslant 1$, $\forall u \in \hat{C}$. □

4. Outline of the Solution Method

It follows from Proposition 7 that $(\mathbf{QCM_1})$ is actually a quasiconcave minimization problem with linear constraints. However, since these linear constraints are not known explicitly, we propose to solve $(\mathbf{QCM_1})$ by a method which combines a branch and bound technique with outer approximation to generate these constraints as needed in the solving process.

We start with a simplex $S \supset \hat{C}^*$ considered as a rough outer approximation of $\hat{C}^*$ and usually with a value α which is the value of f at the best feasible solution of $(\mathbf{QCM_1})$ available (briefly, the best feasible value of f). Then the first approximating problem for $(\mathbf{QCM_1})$ is

$$\min\{f(v) : v \in S\} . \tag{11}$$

To solve this approximating problem we subdivide the space into cones and for each cone M we estimate a lower bound $\beta(M) \leq \min f(S \cap M)$. Cones M with $\beta(M) \geq \alpha$ are deleted and the cone $\bar{M}$ with smallest lower bound is considered the most promising. To $\bar{M}$ a point $\bar{\omega} \in S \cap \bar{M}$ is associated such that if $\bar{\omega} \in \hat{C}^*$ then the outer approximation S can be considered tight enough and we continue the branch and bound process for solving the approximating problem (11). But if $\bar{\omega} \notin \hat{C}^*$, then the outer approximation must be refined. We do this by adding to S a linear constraint that cuts off $\bar{\omega}$ without cutting off any point of $\hat{C}^*$. With the new polytope S', replacing the old one, we then continue the branch and bound process.

Thus, the method is a branch and bound procedure for solving the approximating problem (11) in which the approximating polytope S is made tighter and tighter as the algorithm proceeds. At the beginning we have a value α which is the best feasible value available. Since at each iteration new feasible solutions may be generated, we update the current best feasible value α whenever possible. The algorithm terminates when the current best value α is sufficiently near to the smallest lower bound $\beta(\bar{M})$ or there is evidence that the problem is infeasible.

The four basic operations in the just outlined method are

- Branching
- Bounding
- Tightening the approximating polytope S
- Updating the incumbent

These are discussed in more detail below.

4.1. BRANCHING

To simplify the notation we will drop the subscript 2 in m_2 and write m instead of m_2. Recall that we assume that the vectors a^i $(i = 0, 1, \ldots, m)$ are linearly independent. Then the cone K^*, which contains $\hat{C}^*$, has exactly $m + 1$ edges, each passing through one a^i. Branching will be performed by subdividing K^* into subcones, where by cone we will always understand a cone with vertex at 0 and exactly $m + 1$ edges. The subdivision proceeds as follows.

Let $M_0 := K^*$ and $T_0 = [a^0, a^1, \ldots, a^m]$. For each $v \in k^*$ denote by $\pi(v)$ the intersection of T_0 with the ray through v. Given a cone $M = cone(v^0, v^1, \ldots, v^m) \subset M_0$ with $v^i \in T_0$, a subdivision of M is defined by the selection of a point $w \in M$ not lying on any edge of M (so that $\pi(w) \neq v^i$, $\forall i$). This point w can be represented in a unique way in the form $w = \Sigma\, \lambda_i v^i$ with $\lambda_i \geq 0$ $(i = 0, \ldots, m)$. Let $J = \{j : \lambda_j > 0\}$, $|J| \geq 2$. Then the partition of M in this subdivision consists of all the cones M_j, $j \in J$, such that $M_j = cone(v^0, \ldots, v^{j-1}, \pi(w), v^{j+1}, \ldots, v^m)$ (so $T_j = \pi(M_j)$ is the simplex whose vertex set is obtained from that of M by replacing v^j with $\pi(w)$). We say that the cone M is split upon w.

When w is the midpoint of a longest edge $T=[v^0, v^1, \dots, v^m]$ the subdivision is called *bisection*. A subdivision rule (process) is called exhaustive if for any infinite nested sequence of cones $M_k = cone(v^{k0}, v^{k1}, \dots, v^{km})$ $(v^{ki} \in T_0)$ such that M_{k+1} is obtained from M_k via a subdivision obeying this rule, we always have

$$diam[v^{k0}, v^{k1}, \dots, v^{km}] = \max_{i<j} |v^{ki} - v^{kj}| \to 0 (k \to +\infty),$$

i.e. the cone M_k shrinks to a ray as $k \to +\infty$. The simplest exhaustive rule is the bisection rule. For the convergence of the procedure, however, the bisection rule is rather slow, therefore more efficient exhaustive subdivision rules have been proposed, see [16] and [28] for more details.

4.2. BOUNDING

At a given iteration we have an incumbent $\bar{v}$ with $f(\bar{v}) = \alpha$ and, in addition, a polytope $S \supset \hat{C}^*$ which is the current approximation of $\hat{C}^*$. Let $M = cone(v^0, v^1, \dots, v^m)(v^i \in T_0)$ be any newly generated cone. Using definition (7) we can compute the points $\hat{v}^i$ $(i = 0, 1, \dots, m)$. Let $\hat{v}^i = v^i/\sigma_i$, where $\sigma_i = \sigma(v_i)$ and let the polytope S be defined by the system $Ev \leqslant r$.

Consider the linear program

$$\begin{aligned} \textbf{(LP(S, M))} \quad & \max \sum_{i=0}^{m} \lambda_i \\ & \text{s.t.} \quad \sum_{i=0}^{m} \lambda_i E\hat{v}^i \leqslant r \\ & \qquad \lambda \geqslant 0 . \end{aligned} \tag{12}$$

Let $\mu(M)$ be the optimal value of this program. Define

$$\hat{\hat{v}}^i = \mu(M)\hat{v}^i, \quad \nu(M) = \min\{f(\hat{\hat{v}}^i) : i = 0, 1, \dots, m\}. \tag{13}$$

PROPOSITION 8. *$S \cap M$ is entirely contained in the simplex $[\hat{\hat{v}}^0, \hat{\hat{v}}^1, \dots, \hat{\hat{v}}^m]$. If $\mu(M) < 1$ then $f(v) = +\infty$, $\forall v \in S \cap M$, while for $\mu(M) \geqslant 1$ we have*

$$\nu(M) \leqslant \min\{f(v) : v \in S \cap M\} \leqslant \min\{f(v) : v \in \hat{C}^* \cap M\}.$$

Proof. Since any point $v \in M$ is of the form $v = \Sigma \lambda_i \hat{v}^i$ with $\lambda_i \geqslant 0$, and since $\mu(M) = \max\{\Sigma \lambda_i : \Sigma \lambda_i \hat{v}^i \in S \cap M\}$, it is easily seen that $S \cap M \subset \{v = \sum \lambda_i \hat{v}^i : \lambda \geqslant 0,\ \Sigma \lambda_i \leqslant \mu(M)\} = [\hat{\hat{v}}^0, \hat{\hat{v}}^1, \dots, \hat{\hat{v}}^m]$. The second assertion follows from the quasiconcavity of f, which implies that the minimum of f over a simplex is achieved at a vertex, and the fact that if $\mu(M) < 1$ then $f(v^i) = +\infty$, $\forall i$.

Thus, if $\mu(M) < 1$ then the cone M can be discarded from further consideration. If $\mu(M) \geq 1$ we define

$$\beta(M_0) = \nu(M_0)\,, \quad \beta(M) = \max\{\beta(M_{fa}), \nu(M)\}\,, \quad M \neq M_0 \tag{14}$$

where M_{fa} denotes the father cone of M, then $\beta(M)$ yields a lower bound for $\min\{f(v) : v \in S \cap M\}$ which has the monotonicity property, $\beta(M) \leq \beta(M')$ if $M' \subset M$.

REMARK. Based on Proposition 5 and Corollary it is easy to determine, for each i, the value θ_i such that $f(\theta_i v^i) = \alpha$. Clearly, $\theta_i \geq v^i/\sigma_i$ we take $\hat{v}^i = \theta_i v^i$. It should be, however, that the determination of the value θ_i may be more expensive than that of σ_i.

4.3. TIGHTENING THE APPROXIMATING POLYTOPE S

Let $\bar{M} \in argmin\{\beta(M) : M \in \mathcal{R}\}$, where $\mathcal{R}$ is the set of cones that remain for further exploration at a given iteration. Let $\bar{\omega} = \Sigma\, \bar{\lambda}_i v^i/\sigma_i$, where $(\bar{\lambda}_0, \ldots, \bar{\lambda}_m)$ is a basic optimal solution of **LP(S, $\bar{\text{M}}$)**. (We will refer to $\bar{\omega}$ as a basic optimal solution of **LP(S, $\bar{\text{M}}$)**.)

By solving **LP($\bar{\omega}$)** (see Proposition 7 and Corollary 2) we obtain $\rho = \rho(\bar{\omega})$. If $\rho \leq 1$ then, by Corollary 2, $\bar{\omega} \in \hat{C}^*$ and we let the set S be unchanged. Otherwise, if $\bar{\omega} \not\in \hat{C}^*$, let $(\bar{x}, \bar{y}, \bar{z})$ be a basic optimal solution of **LP($\bar{\omega}$)**, so that the vector $\bar{u} = (\bar{x}, \bar{y})$ satisfies $\langle \bar{\omega}, \bar{u} \rangle = \rho > 1$. Then the linear inequality

$$\langle \bar{u}, v \rangle \leq 1 \tag{15}$$

will be violated by $\bar{\omega}$ but will still be satisfied by all $v \in \hat{C}^*$. Indeed, since $(\bar{x}, \bar{y}, \bar{z})$ is feasible to (10) it follows that $\bar{u} = (\bar{x}, \bar{y}) \in C$, hence $\langle \bar{u}, v \rangle \leq 1$, $\forall v \in \hat{C}^*$. Therefore, by adding (15) to S we define a polytope smaller than S but still containing $\hat{C}^*$.

Since $\langle \bar{u}, \bar{\omega}/\rho \rangle = 1$ and $\langle \bar{u}, v \rangle \leq 1$, $\forall v \in \hat{C}^*$, we see that (15) defines a supporting hyperplane to $\hat{C}^*$ at the point $\bar{\omega}/\rho$. It can be proved, see, e.g. Tuy [27], that (15) is even one of the defining constraints for the polyhedron $\hat{C}^*$ (but we shall not need this fact for the justification of the algorithm).

4.4. UPDATING THE INCUMBENT

At each iteration, for every newly generated point v, we compute $\rho(v)$ by solving **LP(v)**, see (10). Then $\xi(v) = v/\rho(v)$ is the intersection of the ray through v with $\partial \hat{C}^*$, since $\rho(v) = \max\{\langle v, u \rangle : u \in \hat{C}\}$ implies $\max\{\langle \xi(v), u \rangle : u \in \hat{C}\} = 1$. Therefore, the incumbent can be updated by considering $f(\xi(v))$ for the newly generated points v. Note that always $\rho(v) > 0$ because $v \neq 0$ and $0 \in int\hat{C}$.

5. The Algorithm

If a feasible solution $\tilde{u}$ to $^*(\hat{\mathbf{Q}})$ is available we can compute the corresponding feasible solution $\tilde{v}$ to (QCM_1), see Proposition 6, and set $\alpha = f(\tilde{v})$. Otherwise, let $\alpha = +\infty(\tilde{v} = \emptyset)$. Take a simplex $S \supset \hat{C}^*$ (or any polytope S which is known to contain an optimal solution of $(\mathbf{QCM}_1)$). Set $M_0 = cone(a^0, a^1, \ldots, a^m)$, $\mathcal{M} = \{M_0\}$. Select a subdivision rule which is exhaustive.

Step 1: For every $M = cone(v^0, v^1, \ldots, v^m) \in \mathcal{M}$ with $v^i \in T_0 = [a^0, a^1, \ldots, a^m]$ compute $\hat{v}_i = v^i/\sigma_i$, $i = 0, 1, \ldots, m$ and solve $\mathbf{LP(S, M)}$ to obtain its optimal value $\mu(M)$ and a basic optimal solution $\omega(M)$. If $\mu(M) \geqslant 1$ then also compute $\beta(M)$ by formulas (13) (14).

Step 2: Delete any cone M with $\mu(M) < 1$ or $\beta(M) \geqslant \alpha$. Let $\mathcal{R}$ be the set of remaining cones. If $\mathcal{R} = \emptyset$ then terminate. If $\tilde{v}$ exists it solves $(\mathbf{QCM}_1)$, otherwise, if $\tilde{v} = \emptyset$ the problem is infeasible.

Step 3: Select $\bar{M} \in argmin\{\beta(M) : M \in \mathcal{R}\}$. Let $\bar{\omega} = \omega(\bar{M})$. Compute $\rho(\bar{\omega})$, by solving $\mathbf{LP}(\bar{\omega})$. Update $\tilde{v}$ and α if the point $\xi(\bar{\omega}) = \bar{\omega}/\rho(\bar{\omega})$ is better than $\tilde{v}$. If $\rho(\bar{\omega}) > 1$, which means that $\bar{\omega} \notin \hat{C}^*$, then let $(\bar{x}, \bar{y}, \bar{z})$ be a basic optimal solution of $\mathbf{LP}(\bar{\omega})$ and reset $S \leftarrow S \cap \{v : \langle \bar{u}, v \rangle \leqslant 1\}$, where $\bar{u} = (\bar{x}, \bar{y})$.

Step 4: Split $\bar{M}$ according to the chosen subdivision rule. Let $\mathcal{P}$ be the partition of $\bar{M}$. Reset $\mathcal{M} \leftarrow \mathcal{R} \backslash \{\bar{M}\}) \cap \mathcal{P}$ and go back to *Step 1*.

REMARKS. (i) The initial simplex S can be constructed as in the method proposed in [30], where a simplex P inscribed in $\hat{C}$ is taken and $S = P^*$. If a polytope is available that is known to contain an optimal solution, then it can be taken as S.

(ii) In the implementation it is convenient to take $a^0, a^1, \ldots, a^m$ as a basis of the space (of dimension $m + 1$) spanned by K^*. Then a point $v \in K^*$ is given by its coordinates $t_0, t_1, \ldots, t_m$ with respect to this basis, and a constraint like (15) is written as

$$\sum_{i=0}^{m} t_i \langle a^i, \bar{u} \rangle \leqslant 1 .$$

Also note that the linear programs for computing $f(v)$ at different v differ only by the last constraint, while for given S and α the programs $\mathbf{LP(S, M)}$, for two "brother" cones M and M', differ only by one v^i (i.e. their duals differ only by one constraint). Moreover, all the programs $\mathbf{LP(v)}$ have the same constraint set. These properties should be exploited for saving computational efforts in solving all these problems.

(iii) To avoid having to deal with point at infinity, like v^i/α_i when $\sigma_i = 0$, the following "bounded version" of the algorithm can be used:

Take a very large positive number L and set $\hat{v} = L(v/|v|)$ for $\hat{v} = v/\sigma$ with

$\sigma = 0$. This amounts to replacing $\hat{D}^*$, which may be unbounded, by the convex compact set

$$G = \hat{D}^* \cap \{v : |v| \leq L\}$$

It is easily seen that the bounded version is nothing but the above algorithm applied to problem $(\mathbf{QCM_1})'$ that is obtained from $(\mathbf{QCM_1})$ by redefining $f(v) = -\infty$ at points v outside the ball of radius L around 0. Obviously, these two problems are equivalent.

(iv) Exhaustive subdivision rules that are more efficient than the bisection rule are generally complicated, see [16]. Computational experiments suggest that in practice it will suffice to use the following rule:

Split $\bar{M}$ (in step 4) upon the point $\bar{\omega} = \omega(\bar{M})$, as long as the algorithm runs normally. Use the bisection rule when the algorithm seems to slow down.

6. Convergence

Denote by S_k, M_k, $\omega^k, \ldots$ the approximating polytope S, the cone $\bar{M}$ and the point $\bar{\omega} = \omega(\bar{M})$ at iteration k. Also let $\bar{v}^k$ and α_k be the incumbent and its value when entering iteration k.

PROPOSITION 9. *For sufficiently large k we have $\omega^k \in \partial \hat{C}^*$ (so that $S_{k'} = S_k$, $\forall k' > k$).*

Proof. Each new constraint of the form (15) added to the current approximating polytope defines a supporting hyperplane to the polytope $\hat{C}^*$ at some boundary point of $\hat{C}^*$ (see Section 3, III). Furthermore, this supporting hyperplane cuts off the current point ω^k which is feasible to all the previous constraints. Therefore, the supporting hyperplanes corresponding to different constraints are all distinct and touch $\hat{C}^*$ at all distinct points. Since $\hat{C}^*$ is a polytope, (Proposition 7) the number of these supporting hyperplanes must be finite. Hence, for large enough k we must have $\omega^k \in \hat{C}^*$, and consequently, $\omega^k \in \partial \hat{C}^*$ since always $\omega^k \in \partial S_k$ and $S_k \supset \hat{C}^*$. □

THEOREM 2. *If the algorithm is infinite, then it generates at least one infinite nested sequence of cones M_{k_r} such that M_{k_r+1} is a son of M_{k_r}. For any such sequence of cones the corresponding sequence of points ω^{k_r} will converge to a global optimal solution of (QCM_1).*

Proof. If the algorithm is infinite then the tree describing the branching process has at least one infinite path, which corresponds to an infinite nested sequence M_{k_r} such that M_{k_r+1} is a son of M_{k_r}. Consider any such sequence. To simplify the notation, we will drop the subscript r and write k instead of k_r.

Let $M_k = cone(v^{k0}, v^{k1}, \ldots, v^{km})$ with $v^{ki} = [a^0, a^1, \ldots, a^m]$. By exhaustiveness of the subdivision process the simplex $[v^{k0}, v^{k1}, \dots, v^{km}]$ will shrink to a point v^* as $k \to +\infty$, i.e.

$$v^{ki} \to v^*(k \to +\infty)\,, \quad i = 0, 1, \dots, m\,. \tag{17}$$

By Proposition 9 we may assume $\omega^k \in \partial \hat{C}^*$, $\forall k$. Since, C^* is compact we may also assume that

$$\omega^k \to \omega^* \in \partial \hat{C}^*\,. \tag{18}$$

Now denote by s^k and $\hat{\omega}^k$ the points where the ray through ω^k meets the simplex $[v^{k0}, v^{k1}, \dots, v^{km}]$ and the boundary $\partial \hat{D}^*$ and $\hat{D}^*$, respectively (the term "simplex" should be understood as "generalized simplex" if certain vertices are at infinity). Since $\mu(M_k) \geqslant 1$, s^k always exists and

$$s^k \in [0; \omega^k] \subset \hat{C}^*\,. \tag{19}$$

We contend that the ray through v^* meets the boundary $\partial \hat{D}^*$ of $\hat{D}^*$ at a unique point. Indeed, if this ray is entirely contained in $\hat{D}^*$ then $|v^{ki}| \to +\infty$, $\forall i$ and hence (since s^k belongs to the simplex $[v^{k0}, v^{k1}, \dots, v^{km}]$), $|s^k| \to +\infty$, contradicting (19) in view of the boundedness of $\hat{C}^*$. Therefore, this ray meets $\partial \hat{D}^*$. Since, on the other hand, for any $v \in K^* = M_0$, $\theta v \in \hat{D}^*$ for all small enough $\theta > 0$ (i.e. 0 is an interior point of $\hat{D}$ in the relative topology of K^*; see Corollary 1) it follows from the convexity of $\hat{D}^*$ that the intersection is a unique point.

Thus, the ray through v^* meets $\partial \hat{D}^*$ at a unique point $\hat{v}^*$. But then from (17) it follows that

$$\hat{v}^{ki} \to \hat{v}^*(i = 0, 1, \dots, m)\,, \quad \hat{\omega}^k \to \hat{v}^*\,, \quad s^k \to \hat{v}^*\,.$$

Furthermore, since $\omega^k \in [s^k, \hat{\omega}^k]$ while s^k and $\hat{\omega}^k$ both tend to $\hat{v}^*$, we must have

$$\hat{v}^* = \omega^*\,. \tag{20}$$

Also clearly,

$$\mu(M_k) = 1 + \frac{|\omega^k - s^k|}{|s^k|} \to 1\,,$$

hence

$$\hat{\hat{v}}^{ki} = \mu(M_k)\hat{v}^{ki} \to \omega^*(i = 0, 1, \dots, m)\,.$$

Now, from the definition of $\beta(M_k)$ and assuming without loss of generality that $\min\{f(\hat{\hat{v}}^{ki}) : i = 0, 1, \dots, m\} = f(\hat{\hat{v}}^{k0})$, we have $\beta(M_k) \geqslant f(\hat{\hat{v}}^{k0})$, hence, by the lower semicontinuity of f (Proposition 3), as $\hat{\hat{v}}^{k0} \to \omega^*$:

$$\underline{\lim}\, \beta(M_k) \geq \underline{\lim}\, f(\hat{v}^{k0}) \geq f(\omega^*) .$$

But $\beta(M_k) = \min\{\beta(M) : M \in \mathscr{R}_k\}$, so that $\beta(M_k) \leq \min\{f(v) : v \in C^*\}$. Therefore, $f(\omega^*) \leq \min\{f(v) : v \in \hat{C}^*\}$ and since $\omega^* \in \hat{C}^*$ this implies that ω^* is a global optimal solution of (**QCM**$_1$). □

Note that even though $\omega^k \in \partial \hat{C}^*$ it may not belong to $\partial \hat{D}^*$ and so $f(\omega^k)$ may be equal to $+\infty$.

THEOREM 3. *If the algorithm is infinite, then any cluster point of the sequence $\tilde{v}^k$ yields a global optimal solution of* (**QCM**$_1$).

Proof. Let $\tilde{v}^* = \lim\{\tilde{v}^k\}$ for some subsequence $\{\hat{v}^k, k \in \Delta\}$. Among the cones that are sons of M_0 at least one must contain infinitely many points $\tilde{v}^k$, $k \in \Delta$. This cone must be split at some iteration (otherwise it could not contain infinitely many $\hat{v}^k$), i.e. it must be M_{k_1} for some k_1. Analogously, among the cones sons of M_{k_1} at least one, say M_{k_2}, contains infinitely many $\tilde{v}^k$, $k \in \Delta$. Continuing this way we find that there exists an infinitely nested sequence $\{M_{k_r}\}$, containing each infinitely many $\tilde{v}^k$, $k \in \Delta$. To simplify the notation, let us henceforth write k instead of k_r. As in the previous proof, all the points $\hat{v}^{ki}$, ω^k will tend to a common limit $\hat{v}^* = \omega^*$ which is the unique intersection point of the ray though v^* with $\partial \hat{D}^*$. Since, however, $f(\tilde{v}^k) < +\infty$, we have $\tilde{v}^k \in \partial \hat{D}^*$. Hence,

$$\tilde{v}^k \to \hat{v}^* = \omega^*, \text{i.e. } \tilde{v}^* = \omega^* ,$$

and by the previous Theorem, $\tilde{v}^*$ is a global optimal solution of (**QCM**$_1$). □

Clearly, by an argument analogous to the one used at the beginning of the proof of Theorem 3, one can prove that any cluster point of the sequence $\{\omega^k\}$ is the limit of a subsequence corresponding to an infinite nested sequence of cones $\{M_{k_r}\}$. Therefore, any cluster point of the sequence $\{\omega^k\}$ also yields a global optimal solution.

Thus, either the algorithm terminates after finitely many steps with a global optimal solution or by an evidence that the problem is infeasible, or it generates infinite sequences of points $\{\tilde{v}^k\}$, $\{\omega^k\}$ and any cluster point of these sequences provides a global optimal solution.

7. Illustrative Example

To illustrate how the algorithm works, we consider the following small example:

$$\begin{array}{lll} (\mathbf{P}) & \min & 3x_1 + 2x_2 + y_1 + y_2 \\ & \text{s.t.} & x_1 + x_2 + y_1 + y_2 \leq 4 \\ & & x_1, x_2 \geq 0 \end{array}$$

where y solves

$$\begin{array}{lll}(\mathbf{R(x)}) & \min & 4y_1 + y_2 \\ & \text{s.t.} & 3x_1 + 5x_2 + 6y_1 + 2y_2 \geq 15 \\ & & y_1, y_2 \geq 0 .\end{array}$$

Preliminary transformations

D is a polyhedron

$$D = \{x_1, x_2, y_1, y_2 \geq 0 : x_1 + x_2 + y_1 + y_2 \leq 4, 3x_1 + 5x_2 + 6y_1 + 2y_2 \geq 15\}$$

while C is the projection on the (x, y) space of the polyhedron (Proposition 7):

$$\{x_1, x_2, y_1, y_2, z_1, z_2 : 3x_1 + 5x_2 + 6z_1 + 2z_2 \geq 15, 4z_1 + z_2 \leq 4y_1 + y_2, z_1, z_2 \geq 0\} .$$

By eliminating z_1, z_2 above we get

$$C = \{x_1, x_2, y_1, y_2 : 3x_1 + 5x_2 + 6y_1 + 2y_2 \geq 15, 4y_1 + y_2 \geq 0\} \tag{21}$$

(in the general case it is not necessary to have the explicit inequalities defining C). Let $u = (x_1, x_2, y_1, y_2)$ and $c = (3, 2, 1, 1)$. A basic optimal solution of the linear program associated with (**P**), i.e. of the program

$$\min\{c^T u : u \in D\}$$

is $u^0 = (0, 0, \frac{5}{2}, 0)$ and the sets $\hat{C} = (C - u^0)$ and $\hat{D} = (D - u^0)$ can be defined. We are now able to transform the original problem into ($\hat{\mathbf{Q}}$) ($\min\{c^T u : u \in \hat{D} \backslash int\, \hat{C}\}$). The final transformation leading to the quasiconcave minimization problem (**QCM$_1$**) ($\min f(v)$ s.t. $v \in \hat{C}^*$), is carried out below. The set $\hat{C}^* = (C - u^0)^*$ is contained in the cone K^* generated by the two vectors

$$a^0 = (0, 0, -4, -1), \quad a^1 = (-3, -5, 0, 0) . \tag{22}$$

Taking a^0, a^1 as a basis of the space spanned by K^*, we represent each point $v \in K^*$ by a vector $t(v) = (t_0, t_1) \in R^2_+$ such that $v = t_0 a^0 + t_1 a^1$.

Since $c^T u^0 = \frac{5}{2}$ and for $t(v) = (t_0, t_1)$ we have $v^T u = -(4y_1 + y_2)t_0 - (3x_1 + 5x_2)t_1$ while $v^T u^0 = -10t_0$, we can define

$$\begin{aligned} f(v) &= -c^T u^0 + \inf\{c^T u : u \in D, \langle v, u - u^0 \rangle \geq 1\} \\ &= -\tfrac{5}{2} + \inf\{3x_1 + 2x_2 + y_1 + y_2 : \text{s.t. (20) and} \\ &\quad -(4y_1 + y_2 - 10)t_0 - (3x_1 + 5x_2)t_1 \geq 1\} . \end{aligned} \tag{23}$$

This completes the reformulation into (**QCM$_1$**). Also,

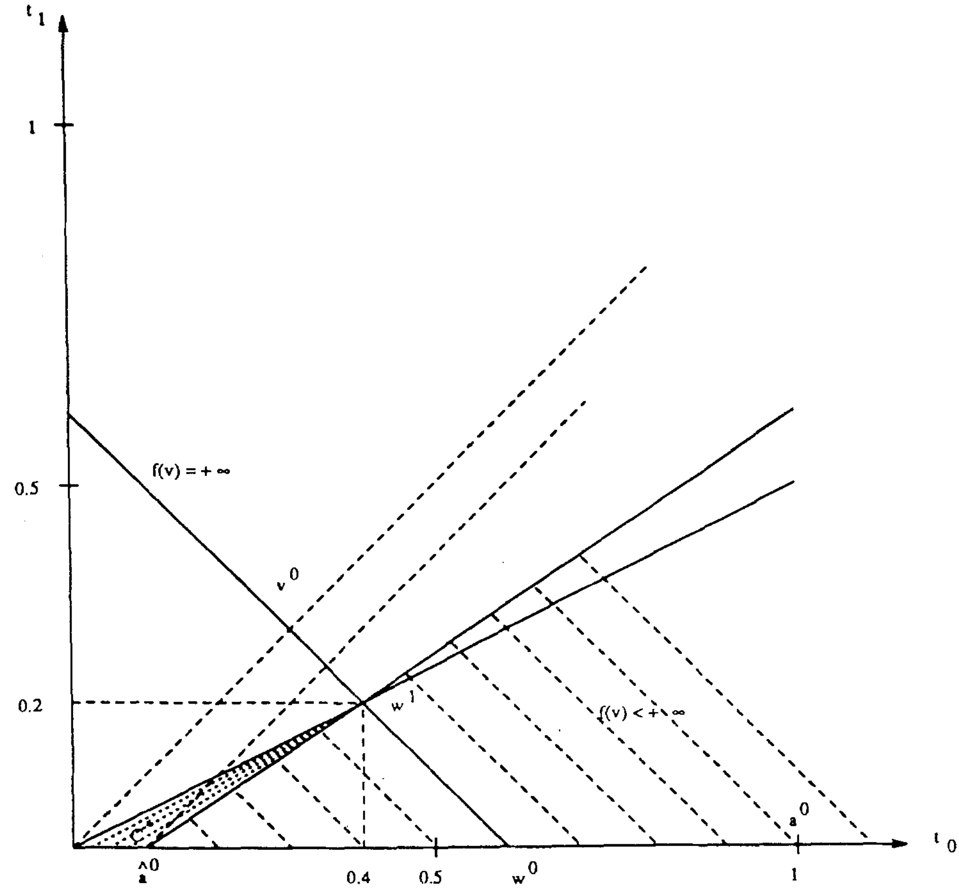

Fig. 1. Problem (**QCM**$_1$) in the example.

$$\begin{aligned}\sigma(v) &= -v^T u^0 + \max\{v^T u : u \in D\} \\ &= 10t_0 + \max\{-(4y_1 + y_2)t_0 - (3x_1 + 5x_2)t_1 : \text{s.t. } (20)\}\end{aligned} \tag{24}$$

$$\begin{aligned}\rho(v) &= -v^T u^0 + \max\{v^T u : u \in C\} \\ &= 10t_0 + \max\{-(4y_1 + y_2)t_0 - (3x_1 + 5x_2)t_1 : \\ &\qquad 3x_1 + 5x_2 + 6y_1 + 2y_2 \geqslant 15, 4y_1 + y_2 \geqslant 0\} .\end{aligned} \tag{25}$$

Initialization

To construct the initial simplex S approximating $\hat{C}^*$ from outside, we observe that the point $s = -0.6a^0 - 0.6a^1 = (1.8, 3, 2.4, 0.6)$ belongs to $\hat{C}$ (see (21)), therefore the cone $P = s + K$ is contained in $\hat{C}$ and its polar $S_0 = P^*$ contains $\hat{C}^*$. We thus take $S_0 = P^*$, i.e.

$$S_0 = \{(t_0, t_1) : t_0 + t_1 \leqslant 0.6, t_0 \geqslant 0, t_1 \geqslant 0\} . \tag{26}$$

Let $M_0 = K^* = \{t : t_0 \geq 0,\ t_1 \geq 0\}$. Also let $v^0 = \emptyset$, $\alpha = +\infty$.

Iteration 1

Step 1: We have from (24) $\sigma(a^0) = 10$, $\sigma(a^1) = 0$. Since a^0 and a^1 are vectors of the basis of the space spanned by K^* and by (7) $\hat{a}^0 = \frac{a^0}{\sigma(a^0)} = \frac{a^0}{10}$, $\hat{a}^1 = \frac{a^1}{\sigma(a^1)} = \frac{a^1}{0}$ it follows that $t(\hat{a}^0) = (0.1,\ 0)$, $t(\hat{a}^1) = (0, \infty)$ i.e. point at infinity in the direction $(0, 1)$). Solving

$$(\mathbf{LP}(S_0, M_0)) \quad \max \lambda_0 + \lambda_1$$
$$\text{s.t. } \lambda \cdot t(\hat{a}^0) + \lambda_1 \cdot t(\hat{a}^1) \leq 0.6$$
$$\lambda_0, \lambda_1 \geq 0$$

yields $\lambda^* = (6, 0)$ i.e. $\mu(M_0) = 6$ and $\omega(M_0) = \frac{6a^0}{10}$, i.e. $t(\omega^0) = (0.6,\ 0)$. We do not need to compute $\beta(M_0)$ because it is the only cone at this stage.

Step 2: $\mathcal{R}_0 = \{M_0\}$.

Step 3: We have from (25) $\rho(\omega^0) = 6$, hence $\xi(\omega^0) = \frac{\omega^0}{6} = \hat{a}^0 \in \hat{C}^*$. Also from (23) $f(\hat{a}^0) = 3.5$. Therefore, the incumbent is $\bar{v}^1 = \hat{a}^0 = \frac{a^0}{10}$ with $\alpha_1 = f(\bar{v}^1) = 3.5$. Since $\rho(\omega^0) > 1$ and the vector $\bar{u}^0$ such that $\bar{u}^0 + u^0$ solves (25) for $v = \omega^0$ is $(0, 3, -2.5, 0)$ we define the new S_1 by adding the following constraint to S_0 (note that $\langle a^0, \bar{u}^0 \rangle = 10$, $\langle a^1, \bar{u}^0 \rangle = -15$):

$$10t_0 - 15t_1 \leq 1\,.$$

Step 4: Split M_0 upon $\omega^0 = \frac{1}{2}(a^0 + a^1)$ with $t(\omega^0) = (0.5, 0.5)$. $\mathcal{M}_1 = \{M_{00}, M_{01}\}$, where M_{0i} contains the vector a^i, $i = 0, 1$.

Iteration 2

Step 1: We have from (24) $\sigma(\omega^0) = \frac{3}{8}$, hence $\hat{\omega}^0 = \frac{8}{3}\omega^0$. Solving $\mathbf{LP}(\mathbf{S}_1, \mathbf{M}_{0i})$ for $i = 0, 1$ yields $\mu(M_{01}) < 1$, $\mu(M_{00}) = 2.15 > 1$ and $\omega(M_{00}) = 0.4a^0 + 0.2a^1$, i.e. $t(\omega(M_{00})) = (0.4,\ 0.2)$. We do not need to compute $\beta(M_{00})$.

Step 2: $\mathcal{R}_1 = \{M_{00}\}$.

Step 3: $M_1 = M_{00}$ is chosen. We have $\omega^1 = \omega(M_{00})$, so $t(\omega^1) = (0.4,\ 0.2)$. Since from (25) $\rho(\omega^1) = 1$, i.e. $\omega^1 \in \hat{C}^*$, we let $S_2 = S_1$. Since $f(\omega^1) = 3.5$ the incumbent is unchanged and $\bar{v}^2 = \bar{v}^1 = \hat{a}^0$. $\alpha_2 = \alpha_1 = 3.5$.

Step 4: Split M_1 upon ω^1. $\mathcal{M}_2 = \{M_{10}, M_{11}\}$, where M_{1i} is the cone that contains a^0.

Iteration 3

Step 1: We have from (24) $\sigma(\omega^1) = 1$, i.e. $\hat{\omega}^1 = \omega^1$. Solving $\mathbf{LP}(\mathbf{S}_2, \mathbf{M}_{1i})$ for $i = 0, 1$ yields $\mu(M_{10}) = \mu(M_{11}) = 1$. Computing $\beta(M_{1i})$ by formula (14) we get $\beta(M_{10}) = \beta(M_{11}) = f(\omega^1) = 3.5$.

Step 2: Since $\beta(M_{10}) = \beta(M_{11}) = \alpha_2$, $\mathcal{R}_2 = \emptyset$. Hence the optimal value of $(\mathbf{QCM}_1)$, i.e. $(\hat{\mathbf{Q}})$, is 3.5. Noting that the vector $\bar{u}$ that solves (25) for $\bar{v} = \hat{a}^0$ (or

$\tilde{v}=\omega^1$)) is $\tilde{u}=(0, 3, 0, 0)$ we thus conclude that an (exact) global optimal solution of (**P**) is

$$x_1=0, x_2=3, y_1=0, y_2=0$$

with objective function value $3.5+2.5=6$.

This optimal solution has already been obtained in the second iteration, but one more iteration was needed to recognize it as such.

Computational results with this algorithm as well as comparisons with the algorithm presented in [30], will be discussed in a subsequent paper.

Appendix A

In this appendix we will show that the assumption $\varphi(x)$ *is continuous* does not restrict the generality.

Consider problem **R(x)**, $\min\{d_2^T y : -B_2 y \geqslant A_2 x - g_2,\ y \geqslant 0\}$, where $\varphi(x)$ is the optimal value of **R(x)**. The dual of **R(x)** is

$$[\mathbf{RD}]\quad \max\{\langle A_2 x - g_2, t\rangle : -B_2^T t \leqslant d_2, t \geqslant 0\}\ .$$

If **R(x)** is infeasible for some x, $\varphi(x)=+\infty$, hence $\varphi(x)$ is not continuous and the dual (**RD**) will be unbounded.

Let M_0 be any number such that $\forall M \geqslant M_0$, the set

$$\{t : 0 \leqslant t \leqslant M\} \tag{1}$$

contains all the vertices of

$$\{-B_2^T t \leqslant d_2, t \geqslant 0\}\ . \tag{2}$$

Such an M can always be found, see Papadimitriou and Steiglitz [25]. If the set of constraints (1) is added to the dual, the objective function value of (**RD**) is bounded and the corresponding primal problem is modified to

$$\varphi_M(x) = \min\{d_2^T y + Mz : B_2 y - z \leqslant g_2 - A_2 x,\ y, z \geqslant 0\}\ ,$$

where z is the multipliers corresponding to the constraints (1). If $\varphi(x) < +\infty$ the optimal value of the dual must be achieved at a vertex of (2), hence $\varphi(x)=\varphi_M(x)$. Therefore, the problem will not change if we replace $\varphi(x)$ by $\varphi_M(x)$ and this conclusion ends this appendix.

Acknowledgement

Research supported in part by the Swedish Transport Research Board (TFB) and Marianne and Marcus Wallenberg foundation.

References

1. E. Aiyoshi and K. Shimisu (1981), Hierarchical Decentralized System and Its New Solution by a Barrier Method, *IEEE Transactions on Systems, Man and Cybernetics SMC-11*, pp. 444–449.
2. G. Anandalingam and D. J. White (1990). A Solution Method for the Linear Stackelberg Problem, *IEEE Trans. Auto. Contr. AC-35*, pp. 1170–1175.
3. G. Anandalingam and T. L. Friesz (eds.) (1992), Hierarchical Optimization, *Annals of Operations Research* **34**. J.C. Baltzer AG, Basel.
4. J. F. Bard and J. E. Falk (1982), An Explicit Solution to the Multi-level Programming Problem, *Computers and Oper. Res.* **9**(1), 77–100.
5. J. F. Bard (1983), An Algorithm for Solving the General Bi-level Programming Problem, *Math. of Ops. Res.* **8**(2), 260–272.
6. J. F. Bard and J. T. Moore (1990), A Branch and Bound Algorithm for the Bilevel Programming Problem, *SIAM J. Sci. Statist. Comput.* **11**, 281–292.
7. J. T. Moore and J. F. Bard (1990), The Mixed Integer Linear Bilevel Programming Problem, *Ops. Res.* **38**(5), 911–921.
8. O. Ben-Ayed (1993), Bilevel Linear Programming, *Computers and Oper. Res.* **20**(5), 485–501.
9. W. F. Bialas and M. H. Karwan (1984), Two-level Linear Programming, *Management Science* **30**, 1004–1020.
10. W. F. Bialas, M. H. Karwan, and J. P. Shaw (1980), A Parametric Complementarity Pivot Approach for Two-level Linear Programming, Research Report No. 80-2, Operations Research Program, Dept. of Industrial Eng., State University of New York at Buffalo.
11. C. Blair (1992), The Computational Complexity of Multi-level Linear Programs, *Annals of Operations Research* **34**, 13–19.
12. W. Candler and R. Townsley (1982), A Linear Two-level Programming Problem, *Comput. & Ops. Res.* 9(1), 59–76.
13. J. E. Falk (1973), A Linear Max-Min Problem, *Mathematical Programming* **5**, 169–188.
14. J. Fortuny-Amat and B. McCarl (1981), A Representation and Economic Interpretation of a Two-level Programming Problem, *J. Ops. Res. Soc.* **32**, 783–792.
15. P. Hansen, B. Jaumard, and G. Savard (1990), New Branching and Bounding Rules for Linear Bilevel Programming, To appear in *SIAM Journal on Scientific and Statistical Computing*.
16. R. Horst and H. Tuy (1990), *Global Optimization: deterministic Approaches*, Springer Verlag, Berlin.
17. R. Horst, N. V. Thoai, and J. de Vries (1988), On Finding New Vertices and Redundant Constraints in Cutting Plane Algorithms for Global Optimization, *Operations Research Letters* **7**, 85–90.
18. R. Horst and N. V. Thoai (1989), Modification, Implementation and Comparison of Three Algorithms for Globally Solving Linearly Constrained Concave Minimization Problems, *Computing* **42**, 271–289.
19. J. J. Judice and A. M. Faustino (1988), The Solution of the Linear Bilevel Programming Problem by Using the Linear Complementarity Problem, *Invest. Opnl.* **8**, 75–95.
20. J. J. Judice and A. M. Faustino (1992), A Sequential LCP Method for Bilevel Linear Programming, *Annals of Operations Research* **34**, 89–106.
21. C. D. Kolstad and L. S. Lasdon (1986), Derivative Evaluation and Computational Experience with Large Bi-level Mathematical Programs, Faculty Working Paper No. 1266, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign.
22. P. Loridan and J. Morgan (1989), New Results of Approximate Solutions in Two-level Optimization, *Optimization* **20**(6), 819–836.
23. S. C. Narula and A. D. Nwosu (1983), Two-level Hierarchical programming problem, *Multiple Criteria Decision Making – Theory and Application*, edited by P. Hansen, New York: Springer Verlag, pp. 290–299.
24. S. C. Narula and A. D. Nwosu (1985), An Algorithm to Solve a Two-level Resource Control Preemptive hierarchical Programming Problem, In *Mathematics of Multi-Objective Optimization* (Edited by P. Serafini), pp. 353–373, Springer, New York.

25. C. H. Papadimitriou and K. Steiglitz (1982), *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Inc., New Jersey (1982).
26. P. T. Thach and H. Tuy, Dual Outer Approximation Methods for Concave Programs and Reverse Convex Programs, Submitted.
27. H. Tuy (1985), Concave Minimization Under Linear Constraints with Special structure, Optimization 16, pp. 335–352.
28. H. Tuy (1991), Effect of Subdivision Strategy on Convergence and Efficiency of Some Global Optimization Algorithms, *Journal of Global Optimization* 1(1), 23–36.
29. H. Tuy, The Normal Conical Algorithm for Concave Minimization over Polytopes, To appear in *Mathematical Programming*.
30. H. Tuy, S. Migdalas, and P. Värbrand (1993), A Global Optimization Approach for the Linear Two-level Program, *Journal of Global Optimization* **3**, 1–23.
31. G. Ünlü (1987), A Linear Bi-level Programming Algorithm Based on Bicriteria Programming, *Comput. Ops. Res.* **14**(2), 173–179.
32. U.-P. Wen and S.-T. Hsu (1989). A Note on a Linear Bilevel Programming Algorithm Based on Bicriteria Programming, *Comput. & Ops. Res.* **16**(1), 79–83.