Vietnam Journal of Mathematics **35**:4(2007) 523–539

Vietnam Journal of MATHEMATICS

On Branch-and-Bound Algorithms for Global Optimal Solutions to Mathematical Programs with Affine Equilibrium Constraints*

Le Dung Muu[1] **and Nguyen Van Quy**[2]

[1]*Hanoi Institute of Mathematics,*
18 Hoang Quoc Viet Road, Hanoi, Vietnam
[2]*Hanoi Academy of Finance, Vietnam*

Dedicated to Professor Hoang Tuy on the occasion of his 80th-birthday

Received November 9, 2007

Abstract. Mathematical programming problems with affine equilibrium constraints, shortly AMPEC, have many applications in different fields of engineering and economics. AMPEC contains several classes of optimization problems such as bilevel convex quadratic programming, optimization over the efficient set as special cases. AMPEC is known to be very difficult to solve globally due to its nested structure. We propose a relaxation algorithm for globally solving AMPEC by using a branch-and-bound strategy. The proposed algorithm uses a binary tree enumeration search for bounding and branching. We also discuss bounding operations by linear programming relaxation and the convex envelope. Numerical experiences and results for the proposed algorithm are discussed and reported.

1991 Mathematics subject classification: 90 C29.

Keywords: Optimization with equilibrium constraints, branch-and-bound, relaxation, bilevel convex quadratic program, optimization over the efficient set, enumeration binary tree.

*This work was supported in part by the National Basic Program in Natural Science, Vietnam, C12.

1. Introduction

A mathematical programming problem with equilibrium constraints, shortly MPEC, is an optimization problem over two-types of variables in which some or all of its constraints are defined by a parametric variational inequality. This class of optimization problems is known to be very difficult to solve even for linear case due to its nonconvexity and nondifferentiability. However, such problems arise frequently in applications, for example, in shape optimization, the design transportation network, economic modeling and data mining. A natural way to handle this two-level problem is to reduce it into an one-level optimization problem by using the Kuhn-Tucker theorem for the lower variational inequality. The reduced problem obtained in this way, in general, is nondifferentiable. Moreover the complementarity condition appeared in the reduced problem is extremely difficult to handle.

In [18] Outrata and in [19] Outrata and Zowe converted a MPEC problem into an unconstrained nonsmooth Lipschitz optimization problem. Then they used nonsmooth numerical methods for solving the reduced problem. In [6] Facchinei et al reformulated a MPEC problem as an one-level nonsmoothly constrained optimization problem and solved the latter by a sequence of smooth problems. These methods converge to a Kuhn-Tucker point. Recently in [22] Quy and Muu used the Kuhn-Tucker theorem to convert a certain MPEC problem into a mathematical program with complementarity constraints and proposed branch-and-bound algorithms using a binary tree search and conical subdivision for finding a global optimal solution to the latter. The readers are referred to the monograph [14] which gives an extensive study of MPEC and presents first and second-order optimality conditions for this problem; some iterative algorithms for computing stationary points are also described therein. A multiobjective optimization approach to equilibrium problems with equilibrium constraints, which contain MPEC as a special case, is studied in [16].

An important class of MPEC is linear programming with affine equilibrium constraints, shortly AMPEC, where the lower problem is defined by a parametric affine variational inequality. AMPEC contains some important classes of optimization problems such as bilevel convex quadratic programming, linear optimization over the efficient set and linear optimization over the Nash-Cournot equilibrium sets in noncooperative game. AMPEC is known to be extremely difficult to solve globally due to its nested structure which causes that there many local optimal solutions may not be global ones. In this paper by employing the linearity structure of the lower variational inequality we convert an AMPEC problem into a linear program with complementarity constraints, shortly called LPCC. The latter problem has attracted much interest in recent years and some solution approaches has been developed for approximating a stationary point or a local solution of the problem (see e.g. [3, 8, 9, 12, 15] and the references therein). While there are a lot efficient and robust algorithms based upon different approaches such as interior point method, sequential quadratic programming, local d.c. optimization for locally solving LPCC problems, very few algorithms exist for globally solving these problems [20, 24]. In [20] a value-at-risk model is formulated as a LPCC and a branch-and-bound algorithm strongly employing the special structure of the model is proposed for approximating a global solution of the resulting problem. In [24] algorithms using exhaustive simplicial and rectangular subdivisions for branching and linear programming relaxations for bounding are proposed for globally solving LPCC problems.

It can be observed that LPCC problems difficulty is relatively low when the number of the complementarity constraint pairs is small, but it becomes high when this number is fairly large. This fact suggests us to use the enumeration binary tree technique to solve the resulting LPCC problem. In fact, instead of handling all pair complementarity constraints altogether, we relax the problem by first choosing some of them and solve the relaxed problem. If the obtained solution of the relaxed problem satisfies all of complementarity constraints, then it is also an optimal solution to the original problem. Otherwise we choose one or several violated constraints and continue the process using the path associated with the current solution as a starting leaf of the tree. In this way we hope to avoid handling all complementarity constraints. To solve a relaxed problem we propose a branch-and-bound procedure using a binary tree search built up by the help of the dual variables. This paper can be regarded as a continuation of our work in [22]. In fact the binary tree search algorithm proposed in this paper is an improvement of that in [22] using a more reflexive adaptive branching operation for problems where the feasible domain of the follower variational inequality may not be constant set as in [22]. Furthermore, we point out in this paper that the well-known lower boundings by linearization and convex envelope are trivial and therefore not suitable for several special cases of AMPEC such as optimization over the efficient set of a multivalue linear objective program.

The remainder of this paper is organized as follows. In the next section we state the problem to be considered and present AMPEC formulations of several important classes of optimization problems. In Sec. 3 we formulate the problem as a linear program with additional complementarity constraints and its penalized problem. Section 4 is devoted to describe a relaxation scheme and a binary tree branch-and-bound algorithm. At the end of this section we discuss linear programming relaxation and convex envelope lower bounding and compare them with the enumeration binary tree technique.

2. The Problem Statement and Its Special Cases

Consider the following mathematical programming problem with affine variational inequality constraints

$$\min f(x, y) \tag{P}$$

subject to

$$x \in X,\ y \in Y,\ (x, y) \in Z, \tag{1}$$

$$x \in C(y),\ \langle Ax + By + a, v - x\rangle \geq 0,\ \forall v \in C(y), \tag{2}$$

where $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$, $Z \subseteq \mathbb{R}^{n+m}$, $C(y) \subset \mathbb{R}^n$ are nonempty polyhedral convex sets for each $y \in Y$, $f : \mathbb{R}^{n+m} \to \mathbb{R}$ is a convex function, and A, B

are appropriate real matrices. In Problem (P), as usual, we shall refer to x as *primary variable* or *decision variable* and to y as the *parameter*. A pair (x, y) is *feasible* for Problem (P) if $(x, y) \in Z, x \in X, y \in Y$ and x solves variational inequality (2). Problem (P) has some applications in machine learning, data mining, measure value-at-risk and other fields (see e.g. [15, 20]). Problem (P) contains many important problems as special cases. Here we list some of them.

2.1. Mathematical Programs with Complementarity Constraint

It is easy to see that when $X \equiv \mathbb{R}^n$, $Y \equiv \mathbb{R}^m$ and

$$C = \{x \in R^n : x \geq 0\},\ Z := \{(x, y) \in R^{n+m} : Ax + By + a \geq 0\},$$

Problem (P) becomes a mathematical program with linear complementarity constraint that is given by

$$\min f(x, y) \tag{CP}$$

subject to

$$x \geq 0, Ax + By + a \geq 0,\ \langle x, Ax + By + a\rangle = 0. \tag{3}$$

2.2. Bilevel Convex Quadratic Programming

Note that when A is a symmetric positive semidefinite matrix, the variational inequality (2) is equivalent to the parametric convex quadratic problem

$$\min_{x \in C}\{\varphi(x, y) := \frac{1}{2}x^T Ax + (By + a)^T x\}.$$

In this case Problem (P) becomes the bilevel convex program

$$\min\{f(x, y) :\ x \in X, y \in Y, (x, y) \in Z\} \tag{BP}$$

where x solves the convex quadratic program

$$\min_{x \in C}\{\varphi(x, y) := \frac{1}{2}x^T Ax + (By + a)^T x\}. \tag{4}$$

In general case, when A is indefinite, or nonsymmetric, the variational inequality (2) is not necessarily equivalent to the programming problem (4). So Problem (P), in general, cannot be formulated as a bilevel problem of form (BP).

2.3. Optimization over The Efficient Set

Let $\emptyset \neq C \subset \mathbb{R}^n$ be a bounded polyhedral convex set and F be a $(p \times n)$-real matrix. Consider the vector optimization problem

$$\text{vmin}\{Fx : x \in C\}. \tag{VP}$$

We recall that a point x^* is said to be an *efficient solution* or a *Pareto solution* to (VP) if whenever $x \in C, Fx \leq Fx^*$ then $Fx = Fx^*$. Let $E(F, X)$ denote the set of all efficient solutions to (VP). Consider the optimization problem over the efficient set

$$\min\{f(x):\ x\in E(F,C)\},$$

where f is a real valued function on $\mathbb{R}^n$. This problem has some applications in decision making and recently has been studied in a lot of articles (see e.g. [4, 11, 13, 21] and the references therein). Note that, since the efficient set is rarely convex, this is a nonconvex optimization problem.

It has been shown in [21] that one can find a simplex Y in $\mathbb{R}^p$ such that a point x^* is efficient for (VP) if and only if there exists $y\in Y$ such that

$$\langle F^T y, x-x^*\rangle \geq 0\ \forall\ x\in C.$$

Thus the above optimization problem over the efficient set can be formulated as the following mathematical program with linear equilibrium constraint

$$\min\{f(x):(x,y)\in C\times Y, \langle F^T y, v-x\rangle\geq 0\ \ \forall v\in C\}.$$

2.4. Stackelberg-Nash-Cournot Equilibrium Market Models

In this model, there are one distinguished individual (called the leader) and n firms (followers). Every follower supplies a homogeneous product. Suppose that the leader has an advantage over the producers. Specially, the leader is able to commit his action before the others are allowed to follow suit. Suppose further that each policy x_0 from the strategy set X_0 of the leader determine a strategy set $X_j(x_0)$ of firm j $(j=1,...,n)$.

Let p denote the price of the product that is assumed to be depend on total producing quantity and is given by

$$p(\sum_{j=1}^n x_j):=\alpha-\beta\sum_{j=1}^n x_j$$

where x_j is the quantity (to be determined) of the goods supplied by firm j and $\alpha\geq 0$, $\beta>0$ are given coefficients. Let $h_j(x_j)$ denote the cost of firm j when its product is x_j. Let u_j $(j=1,...,n)$ be the benefit function of firm j. Then the benefit function of firm j is given as

$$u_j(x_j):=x_j p(\sum_{j=1}^n x_j)-h_j(x_j).$$

Naturally, the leader and each firm j seeks a strategy from $X_j(x_0)$ such that their benefits are maximal. However, such a production strategy in general, does not exist. To overcome this difficulty, the Nash equilibrium notion can be used. By definition, a strategy

$$x^*=(x_1^*,...,x_n^*)^T\in X(x_0)\equiv X_1(x_0)\times ...\times X_n(x_0)$$

is called a Nash equilibrium point of $u:=(u_1,...,u_n)$ with respect to x_0 if

$$u_j(x_0,x^*)\geq u_j(x_0,x^*[x_j])\ \forall x_j\in X_j(x_0)\ \forall j=1,...,n,$$

where $x^*[x_j]$ stands for the vector obtained from x^* by replaced j-th component by x_j. Let $E(x_0)$ denote the set of all Nash equilibrium points with respect to x_0. Then the problem to be considered is

$$\max\{u_0(x_0,x_1,...,x_n)\ :x_0\in X_0, x\in E(x_0)\}, \tag{SC}$$

where u_0 denotes the benefit function of the leader.

Suppose that each cost function h_j is linear of the form

$$h_j(x_j) = c_j x_j + \delta_j.$$

Define the function $F : X(x_0) \to \mathbb{R}^n$ by taking for each $x^T = (x_1, ..., x_n) \in X(x_0)$,

$$F(x_1, ..., x_n) := c - \nabla\Big(x_j p(\sum_{j=1}^{n} x_j)\Big)\ j = 1, ..., n,$$

where $c^T = (c_1, ..., c_n)$.

It has been shown (see e.g., [10]) that if $X(x_0)$ is closed convex then $E(x_0)$ is just the solution-set of the parametric variational inequality

$$\text{find } x^* \in X(x_0) : \langle F(x_0, x^*), x - x^* \rangle \geq 0\ \forall x \in X(x_0).$$

Thus Problem (SC) can be written equivalently as

$$\max u_0(x_0, x)$$

$$\text{s. t. } x_0 \in X_0, x \in X(x_0), \langle F(x^*), x - x^* \rangle \geq 0\ \forall x \in X(x_0)$$

which is a MPEC.

3. Equivalent Formulations

As we have mentioned in the introduction part, a main difficulty of Problem (P) is that its feasible domain is not given in an explicit form as a system of equalities and/or inequalities. A commonly used approach is to formulate (P) as an standard mathematical programming problem. Such smoothing formulations are proposed in [6, 9, 18, 19, 22] for some special cases of Problem (P). Smoothing formulations are mainly used for local optimization. Unlike the smoothing formulations in the mentioned papers, here we apply the Lagrange duality for the parametric variational inequality constraint to derive equivalent formulations that allows globally solving Problem (P) by using the branch-and-bound strategy. To be specific, in what follows we suppose that the polyhedral convex set C is given as

$$C(y) := \{x \in \mathbb{R}^n : Px + Qy + b \leq 0\}, \tag{5}$$

where P and Q are $\nu \times n$ and $\nu \times m$ matrices respectively, and $b \in \mathbb{R}^\nu$.

Note that x solves variational inequality (2) if and only if x is an optimal solution to the linear program

$$\min_v \{\langle Ax + By + a, v \rangle : Pv + Qy + b \leq 0\}.$$

Applying the Lagrange duality theorem to the latter linear programming problem we obtain the following lemma.

Lemma 3.1. *Suppose that X is given by* (5). *Then Problem* (P) *can be formulated equivalently as*

$$\beta(I_\nu) := \min_{x,y,\lambda} f(x,y) \tag{P1}$$

subject to

$$x \in X, y \in Y, (x,y) \in Z, \tag{6}$$
$$Ax + By + a + \lambda^T P = 0, \tag{7}$$
$$\lambda_i \lambda_i^* = 0 \ \forall i \in I_\nu \equiv \{1, ..., \nu\} \ \text{ with } \lambda_i^* := (Px + Qy + b)_i, \tag{8}$$
$$\lambda \geq 0, Px + Qy + b \leq 0. \tag{9}$$

When f is linear and X, Y, Z are polyhedra given in explicit forms, Problem (P1) is a linear program with additional linear complementarity constraints that makes problem difficult. Especially when the number of complementarity constraints ν is large. Theoretically, one may have that $\lambda = 0$. In this case in Problem (P1) the complementarity constraints disappear, and the problem then becomes a standard convex program when f is convex. Thus in the sequel, we always suppose that $\lambda \neq 0$.

Note that the complementarity constraints can be cast into a single reverse convex constraint by using, for example, the function

$$p(x,y,\lambda) := \sum_{i=1}^{p} \min\big\{\lambda_i, (-Px - Qy - b)_i\big\}.$$

It is clear, that p is a concave function and that under the condition $\lambda \geq 0$, $Px+b \leq 0$, the inequality $p(x,y,\lambda) \leq 0$ holds true if and only if $\lambda^T(Px+Qy+b) = 0$. Thus Problem (P1) can be rewritten as the following linear program with an additional reverse convex constraint

$$\min_{x,y,\lambda} f(x,y) \tag{P2}$$

subject to

$$x \in X, y \in Y, (x,y) \in Z, \tag{10}$$
$$Ax + By + a + \lambda^T P = 0, \tag{11}$$
$$p(x,y,\lambda) \leq 0, \tag{12}$$
$$\lambda \geq 0, Px + Qy + b \leq 0. \tag{13}$$

For a fixed $t > 0$ we define the penalized problem

$$\min_{x,y,\lambda}\{f_t(x,y,\lambda) \equiv f(x,y) + tp(x,y,\lambda)\} \tag{P$_t$}$$

subject to

$$x \in X, y \in Y, (x,y) \in Z, \tag{14}$$
$$Ax + By + a + \lambda^T P = 0, \tag{15}$$
$$\lambda \geq 0, Px + Qy + b \leq 0. \tag{16}$$

Suppose that $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$ are nonempty bounded polyhedra and f is linear, then Problem (P_t) admits a global optimal solution. Thus there exists $t_* \geq 0$ such that for every $t > t_*$ the sets of global optimal solutions of (P2) and (P_t) coincide [2]. Note that the objective function of (P_t) is the sum of the linear function f and the concave function tp. Hence (P_t) is a concave minimization problem.

4. A Binary Tree Relaxation Branch-and-Bound Algorithm

In this section we consider Problem (P1) where, as usual, we suppose that X, Y, Z are polyhedral convex sets given in explicit forms as systems of linear equalities and/or inequalities, and that f is a linear function given by $f(x, y) = c^T x + d^T y$. In this case Problem (P1) becomes a linear program with additional complementarity constraints that make the problem nonconvex. Motivated by the fact that difficulty of this problem increases as the number of the complementarity constraints gets larger, in order to avoid handling all complementarity constraints, we relax the problem by taking a subset of indices $I_0 \subset I_\nu$ and solve the corresponding relaxed problem (P1I_0) obtained from (P1) by ignoring the complementarity constraints $\lambda_i(Px + Qy + b)_i$ for all $i \in I_\nu \setminus I_0$. If it happens that the obtained solution of this relaxed problem satisfies all complementarity constraints of the original problem (P1), we are done. Otherwise, the optimal solution of the relaxed problem yields a lower bound for the optimal value of the original problem. In the case the difference between upper and lower bounds is less or equal to a given tolerance, we can terminate the procedure to obtain an approximate solution. In the other case, we use the obtained solution of the relaxed problem to define a new relaxed problem to improve the lower and upper bounds.

4.1. Bounding, Branching and Cutting

The algorithm we are going to describe is a branch-and-bound procedure using finite rooted binary trees. We recall (see. e.g., [5]) that in a rooted binary tree each node which is not a leaf has two children: left and right children. We will call a left (resp. right) child a left (resp. right) node. For a binary rooted tree τ, we adopt the following notations:

- $\mathcal{N}(\tau)$: the set of the nodes of τ;
- $\mathcal{L}(\tau)$: the set of the leaves of τ;
- $\mathcal{P}(N)$: the path from the root to node N.

For simplicity of notation for each j we take

$$\lambda_j^* := (Px + Qy + b)_j, \ \delta_j := \lambda_j \lambda_j^* \ (j = 1, ..., \nu),$$

and we call λ_j^* the dual value of λ_j and δ_j the jth complementarity value.

We define a rooted binary tree by the help of complementarity constraints

$$\delta_j := \lambda_j \lambda_j^* = 0 \ \ (j = 1, ..., \nu).$$

Namely, the left child of a node is associated with $\lambda_i = 0$, and the right child with $\lambda_i^* := 0$ for some $i \in I_\nu$. Every pair $(\lambda_i = 0, \lambda_i^* = 0)$ is the left and right children of one and only node, and we say that the index i is associated with these two nodes. Since the number of complementarity constraints is ν, if we use all complementarity constraints $\lambda_j \lambda_j^* = 0$ $(j = 1, ..., \nu)$ to define a tree in this way, we obtain a rooted binary that we will denote by τ_ν. Clearly, this is a finite tree having 2^ν leaves. Let $I(N) \subset \{1, ..., \nu\}$ be the set of the indices associated with the nodes (children) belonging to the path $\mathcal{P}(N)$ from the root to a node N and let

$$I_0(N) = \{i \in I(N) : \lambda_i = 0\},\ I_*(N) := \{i \in I(N) : \lambda_i^* = 0\}.$$

We assign to each leaf L of $\tau(I_\nu)$ a real number that is the optimal value of the problem defined as

$$\beta(L) := \min f(x, y) \tag{PL}$$

subject to

$$\begin{gathered} x \in X, y \in Y, (x, y) \in Z, \\ Ax + By + \lambda^T P = 0, \\ \lambda_i = 0 \quad \forall i \in I_0(L), \quad \lambda_i \geq 0 \quad \forall i \notin I_0(L), \\ (Px + Qy + b)_i = 0 \quad \forall i \in I_*(L),\ (Px + Qy + b)_i \leq 0 \quad \forall i \notin I_*(L). \end{gathered}$$

Let $L_* \in \mathcal{L}$ be such that

$$\beta(L_*) = \min\{\beta(L) : L \in \mathcal{L}(\tau_\nu)\}.$$

Take $\beta(\tau_\nu) = \beta(L_*)$. Clearly, $\beta(\tau_\nu)$ is the optimal value β_* of Problem $(P1)$. Thus for solving $(P1)$ we can use a complex enumeration-search throughout all paths from the root to every leaf of T_ν. However, computing $\beta(L)$ for all leaves is very cost when ν is high. In order to avoid the complex enumeration we will use a subset $I_0 \subset I_\nu$ and start the procedure with the binary tree τ_{I_0} built up from I_0 in the same way as presented above. To each leaf L of τ_{I_0} we also assign the number $\beta(L)$ defined as above for τ_ν. Since $I_0 \subset I_\nu$, the number $\beta(\tau_{I_0})$ is a lower bound for the optimal value β_* of (P1). As computing lower bounds we may obtain feasible solutions for (P1), thereby we may reduce the best upper bound. If the difference between upper and lower bounds is less than a given tolerance, we may terminate the procedure. Otherwise, we expand τ_{I_0} to improve lower and upper bounds, and so on.

At each step of the algorithm we will expand the currently considered tree by using the obtained optimal solution to the relaxed problem. Specially, let τ be the tree under consideration at the current step: $\tau \equiv \tau(I_0)$ or τ is a descendant of $\tau(I_0)$. Clearly, $\beta(\tau)$ is a lower bound for β_*. Let $(x^\tau, y^\tau, \lambda^\tau)$ be an optimal solution to Problem $P(L_*)$, Hence $\beta(\tau) = f(x^\tau, y^\tau)$.

Suppose

$$I(x^\tau, y^\tau, \lambda^\tau) := \{i \in I_\nu : \lambda_i^\tau (Px^\tau + Qy^\tau + b)_i = 0\}.$$

Since τ is a descendant of $\tau(I_0)$, we have $I_0 \subseteq I(L_*) \subseteq I(x^\tau, y^\tau, \lambda^\tau)$. Now take one or more indices from $I(x^\tau, y^\tau, \lambda^\tau) \setminus I(L_*)$ and expand τ by adding new nodes as descendants of L_* that are associated with each newly taking index in the way described above. Then for each leaf L of the newly obtained tree we solve

Problem (PL) in order to improve the currently best known lower and bounds until their difference is small as desired. Note that if $\beta(L) \geq \alpha_c$ with α_c being the currently best known upper bound, then L can be deleted; L is a dead leaf.

As usual, with a given tolerance $\epsilon \geq 0$, we call a point $(x^\epsilon, y^\epsilon, \lambda^\epsilon)$, an *$\epsilon$-global optimal solution* to Problem (P1) if it is feasible for (P1) and $f(x^\epsilon, y^\epsilon) - \beta(\tau) \leq \epsilon(|f(x^\epsilon, y^\epsilon)|+1)$. Now we can describe a binary-tree branch-and-bound algorithm for Problem (P1) in detail as follows.

ALGORITHM (Binary-tree branch-and-bound)

To start the procedure, if no priori informations are available, we may take a subtree as the rooted full tree defined by some subset of I_ν.

At the beginning of each iteration we have at the hand a rooted binary tree τ defined by a subset of I_ν.

Step 1. (Computing lower bound) For each $L \in \mathcal{L}(\tau)$ solve the linear program

$$\beta(L) := \min f(x, y) \tag{PL}$$

subject to

$$\begin{gathered} x \in X, y \in Y, (x, y) \in Z, \\ Ax + By + \lambda^T P = 0, \\ \lambda_i = 0 \quad \forall i \in I_0(L), \quad \lambda_i \geq 0 \quad \forall i \notin I_0(L), \\ (Px + Qy + b)_i = 0 \quad \forall i \in I_*(L),\ (Px + Qy + b)_i \leq 0 \quad \forall i \notin I_*(L). \end{gathered}$$

Let

$$\beta(\tau) := \min\{\beta(L) : L \in \mathcal{L}(\tau)\} = \beta(L_*)$$

and $I(x^\tau, y^\tau, \lambda^\tau)$ be an optimal solution to Problem (PL_*) such that $\beta(\tau) = f(x^\tau, y^\tau)$. Take

$$I(x^\tau, y^\tau, \lambda^\tau) := \{i \in I_\nu : \lambda_i^\tau (Px^\tau + Q, y^\tau + b)_i = 0\}.$$

Step 2. (Updating upper bound) Let α_c be the smallest upper bound currently known for the optimal value of Problem $(P1)$ and (x^c, y^c, λ^c) be the corresponding feasible solution such that $f(x^c, y^c) = \alpha_c$ (feasible solutions of $(P1)$ can be obtained as solving Problems (PL) with $L \in \mathcal{L}(\tau)$).

Step 3. (Checking optimality) If $\alpha_c - \beta(\tau) \leq \epsilon(|\alpha_c| + 1)$, then terminate: (x^c, y^c, λ^c) is an ϵ-optimal solution.

Otherwise go to Step 4.

Step 4. (Cutting) Delete all edges of τ belonging to the leaf L if

$$\alpha_c - \beta(L) \leq \epsilon(|\alpha_c| + 1)$$

(we call such a leaf L a dead leaf).

Step 5. (Branching) Take

$$\emptyset \neq I_a \subseteq I(x^\tau, y^\tau, \lambda^\tau) \setminus I(L_*).$$

Expand the current tree τ by adding new nodes as descendants of L_* that corresponding to each index of I_a. Then go back to Step 1 with the new expanded tree.

Remark 1. Step 5 in the above algorithm can be replaced by the following, so called mostly violated constraint-branching. Take an index $j \in I_\nu \setminus I(x^\tau, y^\tau, \lambda^\tau)$ such that the complementarity constraint $\lambda_j(Px^\tau, y^\tau, \lambda^\tau)_j = 0$ is mostly violated, that is

$$\lambda_j(Px^\tau, y^\tau, \lambda^\tau)_j = \max\{\lambda_k(Px^\tau, y^\tau, \lambda^\tau)_k : k \in I_\nu \setminus I(x^\tau, y^\tau, \lambda^T)\}.$$

Then we expand τ by adding to the leaf L_* two children: the left child is associated to $\lambda_j = 0$ and the right one to $(Px + Qy + b)_j = 0$.

Remark 2. At Step 2, having an optimal solution $(\lambda^\tau, x^\tau, y^\tau)$ of Problem (PL$_*$) we may update the current upper bound by solving the following linear program

$$\min_{x,y,}\{c^T x + d^T y\}$$

subject to

$$x \in X, y \in Y, (x, y) \in Z, c^T x + d^T y \leq \alpha_c \tag{17}$$

$$Ax + By + a + (\lambda^\tau)^T P = 0, (Px + Qy + b)_i \leq 0 \; i \in I(x^\tau, y^\tau, \lambda^\tau), \tag{18}$$

$$(Px + Qy + b)_i = 0 \quad \text{if} \quad i \in I_\nu \setminus I(x^I, y^I, \lambda^I). \tag{19}$$

Clearly, an optimal solution (x, y) to this linear program, if any, is feasible for (P1) and $f(x, y) \leq \alpha_c$.

Illustrative example In this example $n = 2, m = 1, \nu = 3$ and

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, a = \begin{pmatrix} -8 \\ -12 \end{pmatrix},$$

$$P = \begin{pmatrix} 1 & 2 \\ 2 & -1 \\ 1 & 3 \end{pmatrix}, Q = \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}, b = \begin{pmatrix} -6 \\ -7 \\ -14 \end{pmatrix}.$$

$$f(x, y) = 2x_1 - x_2 + y_1, \;\; X = \mathbb{R}^2_+, Y = \mathbb{R}_+, Z = \mathbb{R}^2.$$

Thus Problem (P1) to be solved is

$$\min\{2x_1 - x_2 + 3y_1\}$$

$$\text{s. t.} \begin{cases} x_1 + 3x_2 + 3y_1 + \lambda_1 + 2\lambda_2 + \lambda_3 = 8 \\ 4x_1 + 2x_2 + y_1 + 2\lambda_1 - \lambda_2 + 3\lambda_3 = 12 \\ x_1 + 2x_2 + y_1 \leq 6 \\ 2x_1 - x_2 + y_1 \leq 7 \\ x_1 + 3x_2 - 3y_1 \leq 14 \\ \lambda_1(6 - (x_1 + 2x_2 + y_1)) = 0 \\ \lambda_2(7 - (2x_1 - x_2 + y_1)) = 0 \\ \lambda_3(14 - (x_1 + 3x_2 - 3y_1)) = 0 \\ x_1 \geq 0, x_2 \geq 0, y_1 \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0. \end{cases} \tag{P1}$$

Let

$$\delta_1 := \lambda_1\lambda_1^* \text{ with } \lambda_1^* := 6 - x_1 - 2x_2 - y_1,$$
$$\delta_2 := \lambda_2\lambda_2^* \text{ with } \lambda_2^* := 7 - 2x_1 + x_2 - y_1,$$
$$\delta_3 := \lambda_3\lambda_3^* \text{ with } \lambda_3^* := 14 - x_1 - 3x_2 + 3y_1.$$

We take $I_0 = \{3\}$ and start with the tree τ_0 having two leaves associated with $\lambda_3 = 0$ and $\lambda_3^* := 14 - (x_1 + 3x_2 - 3y_1) = 0$ (see Fig. 1).

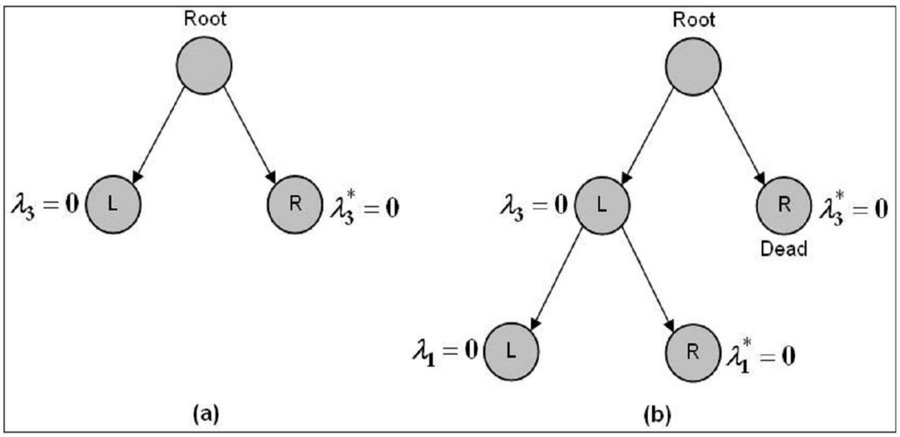

Fig. 1. (a) Tree τ_0, (b) Tree τ_1

For the left leaf L associated with $\lambda_3 = 0$, we solve Problem (PL). The optimal value and optimal solution obtained are

$$\beta(L) = -1,\ x^* = (0, 1),\ y^* = (0),\ \lambda^* = (5, 0, 0).$$

For the right leaf R associated with $\lambda_3^* = 0$, Problem (PR) has nonfeasible solution. Hence $\beta(R) = +\infty$. The leaf R is dead (it can be deleted).

Thus $\beta(L) = -1$ is a lower bound for the optimal value of Problem $(P1)$ and the complementarity values associated with L are

$\delta_1 = 20,\ \delta_2 = 0,\ \delta_3 = 0$ (the first complementarity constraint is violated)

We expand τ by adding to L two children: the leaf child associated with $\lambda_1 = 0$ and the second one with $\lambda_1^* = 0$. We obtain a new tree that have two leaves (the third leaf was dead) (see Fig. 1b). For the left leaf (left child) L associated with $\lambda_1 = 0$ we solve Problem (PL) and obtain the optimal value and solution

$$\beta(L) = 2,\ x^* = (2, 2),\ y^* = (0),\ \lambda^* = (0, 0, 0).$$

For the right leaf (right child) R associated with $\lambda_1^* = 0$ the optimal value and optimal solution to (PR) are

$$\beta(R) = 2,\ x^* = (2, 2),\ y^* = (0),\ \lambda^* = (0, 0, 0).$$

These two problems have the same optimal value. The new lower bound thus is 2. Since the optimal solution to the latter problem satisfies all complementarity constraints, it is feasible for the original problem (P1). Thus $(x^*, y^*) = (2, 2, 0)$ is a global optimal solution to (P1) and its associated Lagrange multiplier is $\lambda^* = (0, 0, 0)$.

4.2. On Lower Bounding by Linear Relaxation and the Convex Envelope

In this subsection we discuss lower bounding by linear programming relaxation and the convex envelope. For easy exposition, let us introduce the slack variable w by setting $-w = Px + Qy + b$. Take, for example,

$$W = \{w \geq 0 : \exists x \in X, y \in Y : -w = Px + Qy + b\}. \tag{20}$$

Since $-W = P(X) + Q(Y) + b$ and X, Y are polyhedral convex sets, W is a polyhedron, and if, for example, the set $\{(x, y) \in X \times Y : \; Px + Qy + b \leq 0\}$ is bounded, then W is bounded too. Using the slack variable w we can rewrite Problem $(P1)$ in the following form

$$\beta_* := \min_{x,y,\lambda,w} f(x, y) \tag{P1w}$$

subject to

$$x \in X, y \in Y, (x, y) \in Z, \tag{21}$$
$$Ax + By + a + \lambda^T P = 0, \tag{22}$$
$$\lambda \geq 0, w \in W \subset \mathbb{R}^n_+, \; \lambda^T w = 0, \tag{23}$$
$$Px + Qy + b + w = 0. \tag{24}$$

Suppose that W is bounded and let $\{v^1, ..., v^q\}$ be the vertices of W. Then every $w \in W$ can be represented as

$$w = \sum_{i=1}^{q} t_i v^i, \;\; t_i \geq 0 \;\forall i, \;\; \sum_{i=1}^{q} t_i = 1.$$

Then we have

$$0 = w^T \lambda = \sum_{i=1}^{q} \langle t_i v^i, \lambda\rangle = \sum_{i=1}^{q} \langle t_i \lambda, v^i\rangle = \sum_{i=1}^{q} \langle \lambda^i, v^i\rangle,$$

where $\lambda^i := t_i \lambda \;\; (i = 1, ..., q)$. Since $\lambda \geq 0, \; t_i \geq 0$, we have $\lambda^i \geq 0$ for every i and $\sum_{i=1}^{q} \lambda^i = \lambda$.

Substituting into (P1w) we obtain the following relaxed problem

$$\min_{x,y,\lambda^i,t_i} f(x, y) \tag{LP1}$$

subject to

$$x \in X, y \in Y, (x, y) \in Z, \tag{25}$$

$$Ax + By + a + \sum_{j=1}^{q} (\lambda^j)^T P = 0, \tag{26}$$

$$\lambda^i \geq 0 \ (i = 1, ..., q), \ \sum_{i=1}^{q} \langle \lambda^i, v^i \rangle = 0, \tag{27}$$

$$Px + Qy + b + \sum_{i=1}^{q} t_i v^i = 0, \ \sum_{i=1}^{q} t_i = 1, \ t_i \geq 0 \ \forall i = 1, ..., q. \tag{28}$$

This problem of the variables x, y, λ^i and t_i is a linear relaxation of (P1) (here we disregard the constraint $\lambda^i = t_i \lambda$ for all i). In general, finding all vertices of W is very costly. Thus this linear programming relaxation is recommended to use when the number of vertices of W is relatively small.

To analyze this linear programming relaxation, let us consider, for example, the optimization problem over the efficient set of a multicriteria linear program described in Sec. 2. As we have mentioned in the Sec. 2, this problem can be formulated as

$$\min\{f(x) : (x, y) \in C \times Y, \langle F^T y, v - x \rangle \geq 0 \ \ \forall v \in C\}.$$

This is a special case of Problem (P) with $X \equiv C$, $Z \equiv \mathbb{R}^m \times \mathbb{R}^p$ and $A \equiv 0$, $a \equiv 0$, $B \equiv F^T$. Suppose that $C := \{x : \ Px + b \leq 0\}$. Then the problem under consideration can take the form

$$\min f(x)$$

subject to

$$Px + b + w = 0, \ y \in Y, \ F^T y + \lambda^T P = 0,$$
$$\lambda^T w = 0, \ \lambda \geq 0, \ w \geq 0,$$

where, by Philip [21], Y can be given as

$$Y = \{y^T = (y_1, ..., y_p) \geq \delta : \ \sum_{i=1}^{p} y_i \leq 1\}$$

with $\delta > 0$ sufficiently small. In this case

$$W = \{w : \ \exists x \in C, -w = Px + b\} = \{w \geq 0 : \ -w = Px + b\}.$$

As usual, we assume that C is a bounded polyhedral convex set. Then so is W. As before let $v^1, ..., v^q$ be the vertices of W. Then, for this case, the linear programming relaxed problem $(LP1)$ takes the form

$$\min f(x)$$

subject to

$$Px + b + \sum_{i=1}^{q} t_i v^i = 0, \sum_{i=1}^{q} t_i = 1, t_i \geq 0 \ \forall i, \tag{29}$$

$$y \in Y, F^T y + (\sum_{i=1}^{q} \lambda^i)^T P = 0, \lambda^i \geq 0, \sum_{i=1}^{q} \langle \lambda^i, v^i \rangle = 0. \tag{30}$$

Since the constraint (30) is independent of the objective function $c^T x$, this problem actually is

$$\min f(x)$$

subject to

$$Px + b + \sum_{i=1}^{q} t_i v^i = 0, \sum_{i=1}^{q} t_i = 1, t_i \geq 0 \ \forall i, \tag{31}$$

which in turn is

$$\min\{f(x) : Px + b + w = 0, w \geq 0\} \equiv \min\{f(x) : Px + b \leq 0\}.$$

Hence, for this example, the linear programming relaxation problem $(LP1)$ is just the linear program obtained by simply ignoring all complementarity constraints. Lower bounds obtained by this linear relaxation thus is not hope to use effectively.

Now let us consider lower bounding by using the convex envelope of each bilinear term $w_i\lambda_i$. Observe that, since $w_i \geq 0$, $\lambda_i \geq 0$, the constraint $\lambda_i w_i = 0$ can be equivalently written as $\min\{\lambda_i, w_i\} \leq 0$. Thus Problem $(P1w)$ can be rewritten as

$$\min f_{x,y,\lambda,w}(x, y)$$

subject to

$$\begin{gathered} x \in X, y \in Y, (x, y) \in Z, \\ Ax + By + a + \lambda^T P = 0, \\ Px + Qy + b + w = 0, \\ \lambda \geq 0, w \geq 0, \sum_{i=1}^{\nu} \min\{w_i, \lambda_i\} \leq 0. \end{gathered}$$

Assume that the polyhedron defined by the constraint $Px + Qy + b \leq 0$ is bounded and that the set of Lagrange multipliers λ is bounded uniformly with respect to y. Then we can write

$$0 \leq \underline{w_i} \leq w_i \leq \overline{w_i} \ \ \forall i, \tag{32}$$

$$0 \leq \underline{\lambda_i} \leq \lambda_i \leq \overline{\lambda_i} \ \ \forall i. \tag{33}$$

We observe that if both $\underline{\lambda_i} > 0, \underline{w_i} > 0$, then, under the conditions (32) and (33), the constraint $\lambda_i w_i = 0$ does not fulfill. Moreover $\underline{\lambda_i} > 0$ and $w_i\lambda_i = 0$ implies $w_i = 0$. Similarly, $\underline{w_i} > 0$ and $w_i\lambda_i = 0$ implies $\overline{\lambda_i} = 0$. Thus we may assume that both $\underline{\lambda_i} = 0$, $\underline{w_i} = 0$.

Now let $l_i(w_i, \overline{\lambda_i})$ be the convex envelope of the two-variables concave function $\min\{w_i, \lambda_i\}$ over the simplex $S_i = \text{cov}\{v^0, v^1, v^2\} \subset \mathbb{R}^2$, where $\text{cov}\{v^0, v^1, v^2\}$ denotes the convex hull of points $v^0 = (0, 0)$, $v^1 = (0, \underline{\lambda_i})$, $v^2 = (\underline{w_i}, 0)$. The convex envelope of the function $\min\{\lambda_i, w_i\}$ over the simplex $S_i = \text{cov}\{v^0, v^1, v^2\}$ is a linear function that can be given explicitly (see e.g. [7]). However, since the values of the function $\min\{\lambda_i, w_i\}$ and its convex envelope coincide at the vertices v^0, v^1, v^2, the convex envelope $l_i(w_i, \lambda_i)$ must vanish on the whole simplex $S_i = cov\{v^0, v^1, v^2\}$. Note that (see [1]) the convex envelope $l(w, \lambda)$ of the function $\sum_{i=1}^{\nu} \min\{w_i, \lambda_i\}$ is $\sum_{i=1}^{\nu} l_i(w_i, \lambda_i)$. Thus the convex envelope $\sum_{i=1}^{\nu} l_i(w_i, \lambda_i)$ also vanishes on the set $S \equiv S_1 \times S_2 \times ... \times S_\nu$. Then the linear relaxation of Problem $(P1w)$ by using the convex envelope of function defining the complementarity constraint $\lambda^T w = 0$ actually takes the form

$$\min f_{x,y,\lambda,w}(x, y)$$

subject to

$$x \in X, y \in Y, (x, y) \in Z,$$
$$Ax + By + a + \lambda^T P = 0,$$
$$Px + Qy + b + w = 0, \lambda \geq 0, w \geq 0$$

which shows that the lower bound obtained by this way is obvious, since, as before, it just simply ignored the hard complementarity constraints.

References

1. F. A. Al-Khayyal and J. E. Falk, Jointly constrained biconvex programming, *Math. Operations Research* **8** (1983) 273–286.
2. L. T. H. An, Dinh T. Pham, and L. D. Muu, Exact penalty in d.c. programming, *Vietnam J. Math.* **27** (1999) 169–178.
3. H. Benson, A. Sen, D. Shanno, and R. Vanderbei, Interior-point algorithms, penalty methods and equilibrium problems, *Computational Optimization and Applications* **34** (2005) 155–182.
4. H. Benson, Optimization over the efficient set, *J. Math. Anal. Appl.* **98** (1984) 562–580.
5. R. Diestel, *Graph Theory.* Springer-Verlag, NewYork, 2000.
6. F. Facchinei, H. Jiang, and L. Qi, A smoothing method for mathematical programs with equilibrium constraints, *Mathematical Programming Series* **A** **85** (1999) 107–134.
7. R. Horst and H. Tuy, *Global Optimization: Deterministic Approach*, Springer-Verlag, Berlin, 1992.
8. X. X. Huang, X. Q. Yang, and K. L. Teo, Partial augmented Lagrangian method and mathematical programs with complementarity constraints, *J. Global optimization* **35** (2006) 235–254.
9. H. Jiang and R. Ralph, Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, *SIAM J. Optimization* **10** (2000) 779–807.
10. I. Konnov, *Combined Relaxation Methods for Variational Inequalities*, Springer, Berlin, 2001.
11. N. T. B. Kim and L. D. Muu, On the projection of the efficient set and potential applications, *Optimization* **51** (2002) 451–421.
12. X. Liu, G. Perakis, and J. Sun, A robust SQP method for mathematical programs with linear complementarity constraints, *Computational Optimization and Applications* **34** (2006) 5–33.

13. L. T. Luc and L. D. Muu, *Global Optimization Approach to optimizing over the efficient set*, in Recent Advances in Optimization, Edited by P. Gritzmann, R. Horst, E. Sachs, and R. Tichatschke, Springer Verlag, Berlin, Germany, 1997, pp. 183–195.
14. J. Q. Luo, J. S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, 1996.
15. O. L. Mangasarian and J. S. Pang, Exact penalty functions for linear programs with linear complementarity constraints, *Optimization* **42** (1997) 1–8.
16. B. S. Mordukhovich, Equilibrium problems with equilibrium constraints via multiobjective optimization, *Optimization Methods and Software* **19** (2004) 479–492.
17. L. D. Muu and W. Oettli, Optimization over equilibrium sets, *Optimization* **49** (2000) 179–189.
18. J. V. Outrata, On optimization problems with variational inequality constraints, *SIAM J. Optimization* **4** (1994) 340–357.
19. J. V. Outrata and J. Zowe, A numerical approach to optimization problems with variational inequality constraints, *Mathematical Programming* **68** (1995) 105–130.
20. J. S. Pang and S. Leyffer, On the global optimization of the value-at-risk, *Optimization Methods and Software* **19** (2004) 611–631.
21. J. Philip, Algorithms for the Vector Maximization Problems, *Mathematical Programming* **2** (1972) 207–229.
22. N. V. Quy and L. D. Muu, *Methods for finding global optimal solution to linear programs with equilibrium constraints*, Hanoi Institute of Mathematics, preprint 2000/38.
23. R. T. Rockafellar, *Convex Analysis*, Princeton University, Princeton 1970.
24. N. V. Thoai, Y. Yamamoto, and A. Yoshise, Global optimization method for solving mathematical programs with linear complementarity constraints, *J. of Optimization Theory and Applications* **124** (2005) 467–490.
25. H. Tuy, *Convex Analysis and Global Optimization*, Kluwer Academic Publishers, 1997.
26. J. J. Ye, Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints, *SIAM J. Optimization* **10** (2000) 943–962.