Journal of Global Optimization **22:** 285–317, 2002.

Optimization over the efficient set: overview

YOSHITSUGU YAMAMOTO
Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba 305-8573, Japan
(E-mail: yamamoto@shako.sk.tsukuba.ac.jp)

Abstract. Over the past several decades, the optimization over the efficient set has seen a substantial development. The aim of this paper is to provide a state-of-the-art survey of the development. Given p linear criteria $c^1x, \dots, c^px$ and a feasible region X of R^n, the linear multicriteria problem is to find a point x of X such that no point x' of X satisfies $(c^1x', \dots, c^px') \geqslant (c^1x, \dots, c^px)$ and $(c^1x', \dots, c^px') \neq (c^1x, \dots, c^px)$. Such a point is called an *efficient point*. The *optimization over the efficient set* is the maximization of a given function ϕ over the set of efficient points. The difficulty of this problem is mainly due to the nonconvexity of this set. The existing algorithms for solving this problem could be classified into several groups such as adjacent vertex search algorithm, nonadjacent vertex search algorithm, branch-and-bound based algorithm, Lagrangian relaxation based algorithm, dual approach and bisection algorithm. In this paper we review a typical algorithm from each group and compare them from the computational point of view.

Key words: Multicriteria programming, Efficient set, Global optimization, Minimum maximal flow

1. Introduction

The problem we consider in this paper is the optimization over the set of efficient points of the linear multiple criteria program

$$(MC) \qquad \left| \begin{array}{ll} \text{vector max} & Cx \\ \text{s.t.} & x \in X, \end{array} \right.$$

where C is a $p \times n$ matrix with rows c^i's, and X is a polyhedral set of R^n defined as $X = \{x \mid x \in R^n; Ax = b; x \geqslant 0\}$. To avoid the technicality we assume throughout the paper that X is nonempty and bounded. Let X_E denote the set of efficient points, whose definition will be given in the next section. Then the problem is formulated as

$$(P_E) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X_E, \end{array} \right.$$

where $\phi\colon R^n \to R$ is a continuous function to be maximized.

The main difficulty of the problem arises from the nonconvexity of the efficient set X_E, which is the union of several faces of X. This problem was first considered by Philip (1972), in which an algorithm based on moving to adjacent efficient vertices is outlined when ϕ is a linear funciton, and lots of papers followed his work.

The purpose of this paper is to survey the existing algorithms for Problem (P_E) as well as some variations. We will not discuss the merits and demerits of the algorithms because we have not yet had enough computational experience to evaluate them. Theoretically interesting algorithms do not always work efficiently, on the contrary, naive methods can surpass sophisticated algotrithms in computation time. We should be careful not to nip the promising algorithms in the bud.

After reviewing the well-known facts concerning problems (MC) and (P_E) in Section 2, the adjacent vertex search algorithms and the nonadjacent vertex search algorithms will be explained in Section 3 and Section 4. In Section 5 we introduce the face search algorithm, which is based on the enumeration of faces that constitute the efficient set. Section 6 is devoted to the branch-and-bound method based on the conical partition, and Section 7 to the Lagrangian relaxation methods. The dual approach and the bisection algorithm will be explained in Section 8 and Section 9. Some other methods are briefly outlined in Section 10. Some conclusion will be given in the last section.

2. Basic results of linear multicriteria program

Throughout this paper we use the following notations: R^k denotes the set of k-dimensional real column vectors, $R^k_+ = \{ x \mid x \in R^k; x \geqslant 0 \}$ and $R^k_{++} = \{ x \mid x \in R^k; x > 0 \}$. R_k is the set of k-dimensional real row vectors, and R_{k+} and R_{k++} are defined in the similar way. We use e and 1 to denote a row vector and a column vector of ones of an appropriate dimension. X_V denotes the set of vertices (extreme points) of X.

DEFINITION 2.1. A point $x \in R^n$ is said to be an *efficient point* of Problem (MC) if $x \in X$ and there is no point $x' \in X$ such that $Cx \leqslant Cx'$ and $Cx \neq Cx'$. We denote the set of efficient points of (MC) by X_E. A point $x \in R^n$ is said to be a *weakly efficient point* of Problem (MC) if $x \in X$ and there is no point $x' \in X$ such that $Cx < Cx'$. We denote the set of weakly efficient points of (MC) by X_W.

DEFINITION 2.2. The set $Y = CX = \{ y \mid y \in R^p; y = Cx \text{ for some } x \in X \}$ is called the *outcome set*. The set $Y^{\leqslant} = Y + R^p_- = \{ y \mid y \in R^p; y \leqslant Cx \text{ for some } x \in X \}$ is called the *lower outcome set*, and $Y^< = Y + R^p_{--} = \{ y \mid y \in R^p; y < Cx \text{ for some } x \in X \}$ is called the *strictly lower outcome set*.

DEFINITION 2.3. A point $y \in Y$ is said to be an *efficient outcome* if there is no point $y' \in Y$ such that $y \leqslant y'$ and $y \neq y'$, in other words, $Y \cap (y + R^p_+) = \{y\}$. We denote the set of efficient outcomes by Y_E. A point $y \in Y$ is said to be a *weakly efficient outcome* if there is no point $y' \in Y$ such that $y < y'$, in other words, $Y \cap (y + R^p_{++}) = \emptyset$. We denote the set of weakly efficient outcomes by Y_W.

The following lemma is a restatement of these definitions.

LEMMA 2.4. *(i)* $X_E = \{ x \mid x \in X; Cx \in Y_E \}$. *(ii)* $X_W = \{ x \mid x \in X; Cx \in Y_W \}$.

DEFINITION 2.5. For $\lambda \in R_{p++}$ and $x \in X$ let

$$g_\lambda(x) = \max \{ \lambda C x' \mid x' \in X; Cx' \geqslant Cx \} - \lambda C x, \tag{1}$$

which is called the *gap function*. When $\lambda = e$, we will omit the subscript λ and denote g_λ simply by g.

As can be seen readily, $x \in X$ is in X_E if and only if $g_\lambda(x) = 0$, and a point x' which solves $\max \{ \lambda C x' \mid x' \in X; Cx' \geqslant Cx \}$ is in X_E. The theory of parametric linear program shows that g_λ is a piecewise linear concave function. Furthermore the following property is readily seen.

LEMMA 2.6. *If* $Cx = Cx'$ *then* $g_\lambda(x) = g_\lambda(x')$.

We first introduce several well-known results, whose proof can be found in, for example Benson (1995), Sawaragi et al. (1985), Steuer (1985), and White (1982).

THEOREM 2.7.

$$\begin{aligned}
X_E &= \left\{ x \;\middle|\; \begin{array}{l} x \in X; \exists \lambda \in R_{p++} \textit{ such that} \\ \lambda C x \geqslant \lambda C x' \textit{ for } \forall x' \in X \end{array} \right\} && (2)\\
&= \left\{ x \;\middle|\; \begin{array}{l} x \in X; \nexists x' \in R^n \textit{ such that} \\ Cx' \geqslant 0; Cx' \neq 0; Ax' = 0; \\ x'_i \geqslant 0 \textit{ for } i \textit{ with } x_i = 0 \end{array} \right\} && (3)\\
&= \left\{ x \;\middle|\; \begin{array}{l} x \in X; \exists (\lambda, \mu, \nu) \in R_{p++} \times R_m \times R_{n+} \textit{ such that} \\ \lambda C - \mu A + \nu = 0; \nu x = 0 \end{array} \right\} && (4)\\
&= \left\{ x \;\middle|\; \begin{array}{l} x \in X; \exists (\lambda, \mu) \in R_{p++} \times R_m \textit{ such that} \\ \lambda C - \mu A \leqslant 0; \lambda C x - \mu b = 0 \end{array} \right\} && (5)\\
&= \{ x \mid x \in X; g_\lambda(x) = 0 \}. && (6)
\end{aligned}$$

Furthermore, there is an $M > 0$ *such that* R_{p++} *above can be replaced by the* $(p-1)$*-dimensional simplex defined by*

$$\Lambda = \{ \lambda \mid \lambda \in R_{p+}; \lambda \geqslant e; \lambda 1 = M \}. \tag{7}$$

Proof. The equivalence among (2), (3), (4) and (5) follows from the duality theorem of linear program. We will prove only that Λ defined by (7) can replace R_{p++} in (2), (4) and (5). By (2) X_E is the union of finitely many faces, say $F^1, \ldots, F^L$ of X such that F^ℓ is the optimal set of maximizing $\lambda^\ell C x$ over X for some $\lambda^\ell \in R_{p++}$. Let $\alpha^\ell = 1/(\min_{i=1,\ldots,p} \lambda_i^\ell)$ and $M = \max_{\ell=1\ldots,L} \alpha^\ell(\lambda^\ell 1)$, where 1 is the p-dimensional column vector of ones. Then for $\ell = 1, \ldots, L$ $(M/\lambda^\ell 1)\lambda^\ell$ lies in Λ defined by (7) and F^ℓ remains the optimal set of maximizing $(M/\lambda^\ell 1)\lambda^\ell C x$ over X. □

Denote

$$(SC(\lambda)) \quad \left| \begin{array}{ll} \max & \lambda Cx \\ \text{s.t.} & x \in X, \end{array} \right.$$

then (2) means that every efficient point is an optimal solution of the single criterion problem ($SC(\lambda)$) defined for some $\lambda \in \Lambda$. The condition (4) remains identical as long as the set of binding constraints at x does not change. Therefore, if points x and x' lie in the relative interior of the same face of X, we see that $x \in X_E$ if and only if $x' \in X_E$.

THEOREM 2.8. *The set X_E is connected. Any two vertices in X_E are connected by a path of efficient edges, where an efficient edge is an edge of X contained in X_E.*

For the proof of Theorem 2.8 see Theorem 9.19 and Theorem 9.23 in Steuer (1985), Theorem 3.31 in Sawaragi et al. (1985) and Naccache (1978).

Let $x = (x_B, x_N) = (B^{-1}b, 0)$ be a basic feasible solution of X and let $A = [B, N]$ and $C = [C_B, C_N]$ be the partitions of A and C corresponding to the basic and nonbasic parts of x, respectively.

LEMMA 2.9.
- *(i) Let $x = (x_B, x_N) = (B^{-1}b, 0)$ be a basic feasible solution of X. Then $x \in X_E$ if and only if $\lambda(C_N - C_B B^{-1}N) - \nu_B B^{-1}N \leqslant 0$ for some $\lambda \in \Lambda$ and $\nu_B \in R_{m+}$ such that $\nu_B x_B = 0$.*
- *(ii) If $x = (x_B, x_N) = (B^{-1}b, 0)$ is a nondegenerate basic solution, the above condition is reduced to $\lambda(C_N - C_B B^{-1}N) \leqslant 0$ for some $\lambda \in \Lambda$.*
- *(iii) Let c^j and a^j be the columns of C_N and N, respectively, corresponding to a nonbasic variable x_j. If $\lambda(C_N - C_B B^{-1}N) \leqslant 0$ and $\lambda(c^j - C_B B^{-1}a^j) = 0$ for some $\lambda \in \Lambda$, then the edge obtained by increasing x_j is an efficient edge.*

Note that the condition of Lemma 2.9 for an efficient basic solution $x = (B^{-1}b, 0)$ and a nonbasic varaible x_j holds if and only if

$$\max\{\lambda(c^j - C_B B^{-1}a^j) \mid \lambda \in \Lambda; \lambda(C_N - C_B B^{-1}N) \leqslant 0\} = 0. \tag{8}$$

The problems we consider in this paper are the following optimization over the efficient set X_E and the weakly efficient set X_W:

$$(P_E) \quad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X_E \end{array} \right.$$

and

$$(P_W) \quad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X_W. \end{array} \right.$$

For these problems we write $\phi(P_E)$ and $\phi(P_W)$ to denote their optimal values, respectively.

Theorem 2.7 will provide several equivalent formulations of Problem (P_E). By (2) we have a infinitely constrained equivalence

$$\left|\begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X; \lambda \in \Lambda \\ & \lambda C x \geqslant \lambda C x' \text{ for all } x' \in X. \end{array}\right.$$

By (4) and (5) we have

$$\left|\begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X; \lambda \in \Lambda; \mu \in R_m; \nu \in R_{n+} \\ & \lambda C - \mu A + \nu = 0; \nu x = 0, \end{array}\right.$$

and

$$\left|\begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X; \lambda \in \Lambda; \mu \in R_m \\ & \lambda C - \mu A \leqslant 0; \lambda C x - \mu b = 0. \end{array}\right.$$

Note that even if ϕ is linear, these problems contain a nonlinear equality constraint. Using the gap function we obtain another equivalent form

$$\left|\begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X; g_\lambda(x) = 0, \end{array}\right.$$

where λ is arbitrarily chosen from Λ and fixed. Since $g_\lambda(x) \geqslant 0$ for all $x \in X$, the last equality constraint $g_\lambda(x) = 0$ can be replaced by $g_\lambda(x) \leqslant 0$, which yields

$$\left|\begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X; g_\lambda(x) \leqslant 0. \end{array}\right.$$

Since g_λ is a concave function, the constraint $g_\lambda(x) \leqslant 0$ is a reverse convex constraint. See Tuy (1998) and Horst and Tuy (1996) for the reverse convex constrained optimization problems.

3. Adjacent vertex search algorithms

The algorithms proposed in Philip (1972), Ecker and Song (1994) and Fülöp (1994) for a linear function ϕ, and in Bolintineanu (1993) for a quasi-convex function ϕ are mainly based on the two techniques: moving from an efficent vertex to an efficient neighbor with a larger objective function value via an efficient edge, and cutting off the portion of X where ϕ takes a smaller value than the incumbent objective function value. We assume for the time being the quasi-convex function ϕ and will follow the line of Bolintineanu (1993).

For $x, x' \in X_V$ let $[x, x']$ denote the edge connecting x and x'. For $x \in X_V \cap X_E$ let

$$N_E(x) = \{ x' \mid x' \in X_V \cap X_E; [x, x'] \subseteq X_E \}, \tag{9}$$

i.e., the set of efficient vertices linked to x by an efficient edge. Then by the quasi-convexity of ϕ we have the lemma.

LEMMA 3.1. *Let $x \in X_V \cap X_E$ and suppose $\{ x' \mid x' \in N_E(x); \phi(x') > \phi(x) \} = \emptyset$. Then x is a local maximum point for (P_E).*

The algorithm is outlined as follows. Here we denote the two halfspaces determined by a hyperplane $H = \{ x \mid x \in R^n; ax = \alpha \}$ by $H_+ = \{ x \mid x \in R^n; ax \geqslant \alpha \}$ and $H_- = \{ x \mid x \in R^n; ax \leqslant \alpha \}$, and their interiors by H_{++} and H_{--}, respectively. (See Fig. 1).

$\langle 0 \rangle$ (Initialization)
Set $p = k = 0$, $X^0 = X$ and find $x^0 \in X_V \cap X_E$. If $N_E(x^0) = \emptyset$ then x^0 is the optimal solution of (P_E). Otherwise, go to the major cycle $\langle p \rangle$.

$\langle p \rangle$ (Major cycle)
- $\langle p1 \rangle$ If $\{ x \mid x \in N_E(x^p); \phi(x) > \phi(x^p) \} \neq \emptyset$, choose x^{p+1} from this set, $p = p + 1$ and go to $\langle p \rangle$.
- $\langle p2 \rangle$ Otherwise, let $L^p = \{ x \mid \phi(x) \leqslant \phi(x^p) \}$ and go to the minor cycle $\langle k \rangle$.

$\langle k \rangle$ (Minor cycle)
- $\langle k1 \rangle$ Find $v^k \in \mathrm{argmax}\{ \phi(x) \mid x \in X^k \}$. If $\phi(x^p) \geqslant \phi(v^k) - \epsilon$ for some tolerance $\epsilon > 0$, then stop with x^p as an ϵ-approximate optimal solution. Otherwise, go to $\langle k2 \rangle$.
- $\langle k2 \rangle$ Find a supporting hyperplane H^k of L^p such that $L^p \subseteq H^k_+$ and $v^k \in H^k_{--}$.
- $\langle k3 \rangle$ If there is an efficient edge $[u', u'']$ such that $[u', u''] \cap H^k \neq \emptyset$ and $\max\{\phi(u'), \phi(u'')\} > \phi(x^p)$, then set x^{p+1} be one of u' and u'' with a larger objective function value. Set $p = p + 1$ and go to the major cycle $\langle p \rangle$. Otherwise, go to $\langle k4 \rangle$.
- $\langle k4 \rangle$ Set $X^{k+1} = X^k \cap H^k_+$, $k = k + 1$ and go to the minor cycle $\langle k \rangle$.

The alogrithm generates a sequences of efficient vertices $x^0, x^1, \ldots$ and polytopes $X^0, X^1, \ldots$ such that $\phi(x^0) < \phi(x^1) < \cdots$ and $X = X^0 \supseteq X^1 \supseteq \cdots$. Let u^k denote the point at which H^k supports L^p. It can be seen that if the angle between $v^k - u^k$ and the normal vector of H^k pointing toward v^k is less than some constant δ, then $\lim_{k \to \infty} \phi(v^k) = \phi(x^p)$. Then we see that for a given positive ϵ the minor cycle does not repeat infinitely.

LEMMA 3.2. *If the above condition on the angle between $v^k - u^k$ and the normal vector of H^k is satisfied, the minor cycle terminates after a finite number of iterations for each p.*

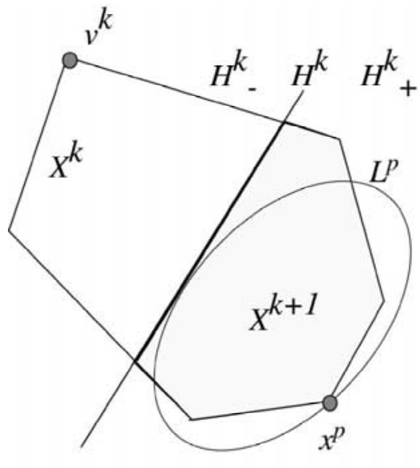

Figure 1. Level set L^p and cutting plane H^k.

Proof. The condition implies $\lim_{k\to\infty}\phi(v^k) = \phi(x^p)$, hence the stopping criterion $\phi(x^p) \geqslant \phi(v^k) - \epsilon$ will be satisfied within a finite number of iterations. □

The most costly and crucial step would be $\langle k3\rangle$ as well as $\langle k1\rangle$, in which a quasi-convex maximization problem is to be solved. We will not go into detail of how to solve the quasi-convex maximization problem. See, for example, Horst and Tuy (1996).

Step $\langle k3\rangle$ is based on the following observation.

LEMMA 3.3. *Let $F^k = X^k \cap H^k$ and F_E^k be the set of efficient points of vector max $\{Cx \mid x \in F^k\}$. Then $X_E \cap F^k \subseteq F_E^k$.*

Proof. If $x \in X_E \cap F^k$, there is no point $x' \in X$ such that $Cx' \geqslant Cx$ and $Cx' \neq Cx$. Then clearly no points in F^k meet this condition, which means $x \in F_E^k$. □

This lemma shows that if we enumerate all the efficient vertices of F_E^k, we can see if there is the edge desired in step $\langle k3\rangle$. Namely, step $\langle k3\rangle$ is carried out by generating the efficient vertices of F^k by a standard algorithm for linear multicriteria optimization such as ADBASE by Steuer (1995) till one of them turns out to be in X_E, and then for such a point, checking if it lies on an efficient edge of X^k with endpoints u' and u'' such that $\max\{\phi(u'), \phi(u'')\} > \phi(x^p)$.

LEMMA 3.4. *$X_E \subseteq X^k$ for $k = 0, 1 \ldots$.*

Proof. Since $X_E \subseteq X^0 = X$, suppose $X_E \subseteq X^k$ as the inductive hypothesis. If $X_E \nsubseteq X^{k+1}$, there is $x' \in X_V \cap X_E$ such that $x' \notin H_+^k$. By the construction of H^k we see $\phi(x') > \phi(x^p)$. Then by Theorem 2.8 there is an efficient edge $[u', u'']$ with $[u', u''] \cap H^k \neq \emptyset$ and $\max\{\phi(u'), \phi(u'')\} > \phi(x^p)$. This is contrary to the fact that X^{k+1} was generated. □

LEMMA 3.5. *When the algorithm terminates with x^p and v^k satisfying $\phi(x^p) \geqslant \phi(v^k) - \epsilon$, x^p is an ϵ-approximate optimal solution of (P_E).*

Proof. By Lemma 3.4 we obtain

$$\begin{aligned}\phi(P_E) &= \max\{\phi(x) \mid x \in X_E\} \leqslant \max\{\phi(x) \mid x \in X^k\} \\ &= \phi(v^k) \leqslant \phi(x^p) + \epsilon \leqslant \phi(P_E) + \epsilon.\end{aligned}$$

□

THEOREM 3.6. *The algorithm provides an ϵ-approximate optimal solution of Problem (P_E) after a finite number of iterations.*

Proof. The minor cycle terminates within finitely many iterations for each p as shown in Lemma 3.2, and points x^p's are efficient vertices of X satisfying $\phi(x^0) < \phi(x^1) < \cdots$, and hence distinct. Therefore the finiteness of $X_V \cap X_E$ and Lemma 3.5 imply the theorem. □

A preliminary computational experiment for small problems up to $n = 7, m = 7$, $p = 4$ with a convex quadratic or linear objective function is reported in Bolintineanu (1993), where it is observed that the vertices, including those on the cutting planes, generated by the algorithm are fewer than the efficient vertices of X.

When ϕ is a linear function dx for $d \in R_n$, the algorithm is substantially simplified. Suppose we have obtained $x^p \in X_V \cap X_E$ with $\{x \mid x \in N_E(x^p); dx > dx^p\} = \emptyset$ after several repetitions of the major cycle. Then the lower level set is the half space $L^p = \{x \mid dx \leqslant dx^p\}$ and the supporting hyperplane of this set is uniquely determined by $H^p = \{x \mid dx = dx^p\}$. Then the efficient vertices of $F^k = X \cap H^k$ are enumerated to check if H^k intersects an efficient edge $[u', u'']$ of X such that $\max\{du', du''\} > dx^p$. When no such edge exists, we conclude from the connectedness of X_E that

$$X_E \subseteq \{x \mid dx \leqslant dx^p\} \tag{10}$$

and hence x^p is an optimal solution of (P_E). Thus, k is never incremented through the algorithm.

In the enumeration of efficient vertices of $F^k = X \cap H^k$ Fülöp (1994) proposed a cutting plane algorithm based on convexity and disjunctive cuts. Assume we have a vertex $\bar{x} \in F^k$ which is not efficient, i.e., $g_\lambda(\bar{x}) > 0$, where g_λ is the gap fucntion defined in Definition 2.5. The portion of F^k with $g_\lambda(x) > 0$, which is a convex set, should be cut off and eliminated for further enumeration. Fülöp proposed to introduce a convexity cut $tx \geqslant 1$, where $t \in R_n$, and reduce the set F^k to $F^k \cap \{x \mid tx \geqslant 1\}$. Suppose the nondeneracy at $\bar{x}$, and for each nonbasic variable x_j let z^j be the direction of edge of F^k adjacent to $\bar{x}$ obtained by increasing x_j. Note that z^j is easily obtained from the dictionary corresponding to $\bar{x}$. Let

$$s_j = \sup\{s \mid s \in R; C(\bar{x} + sz^j) \leqslant Cx; x \in F^k\}, \tag{11}$$

then we have the convexity cut as follows. Note that the constraint $C(\bar{x}+sz^j) \leqslant Cx$ together with $x \in F^k$ means that $C(\bar{x}+sz^j)$ be in the lower outcome set of CF^k.

LEMMA 3.7. *Suppose $s_j > 0$ for every nonbasic variable x_j of $\bar{x}$. Let $t \in R_n$ be defined by*

$$t_j = \begin{cases} 1/s_j & \text{if } x_j \text{ is a nonbasic variable and } s_j < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

Then $t\bar{x} < 1$, and $tx \geqslant 1$ for all efficient points x in F^k.

See Horst and Tuy (1996) for further detail of convexity cut. Everytime a nonefficient vertex is found, F^k is reduced by the convexity cut, which might lighten the computational burden. No computational experiment is reported in Fülöp (1994).

Ecker and Song (1994) proposed to solve $\max\{c^i x \mid x \in X \cap H^k_+\}$ for $i = 1, \ldots, p$ to find the next iterate x^{p+1} before resorting to the vertex enumeration of F^k.

4. Nonadjacent vertex search algorithm

The algorithms which trace the adjacent vertices need a step of enumerating all efficient vertices of a polyhedral set with a lower dimension. This section explains a nonadjacent vertex search algorithm proposed by Benson (1992), which dispenses with the vertex enumeration.

We assume the linear objective function $\phi(x) = dx$. Suppose we have $k+1$ efficient points $x^0, x^1, \ldots, x^k \in X_E$ and let $\alpha^k = \max\{dx^j \mid j = 0, 1, \ldots, k\}$ and (P^k) be the problem, which plays a central role in the algorithm, of finding a point $(x, \lambda) \in R^n \times R_p$ satisfying

$$(P^k) \qquad \left| \begin{array}{l} \lambda Cx \geqslant \lambda Cx^j \text{ for } j = 0, 1, \ldots, k \\ x \in X \\ \lambda \in \Lambda \\ dx > \alpha^k. \end{array} \right.$$

REMARK 4.1. *If $(\bar{x}, \bar{\lambda}) \in X \times \Lambda$ satisfies the constraints*

$$\lambda Cx \geqslant \lambda Cx^j \text{ for } j = 0, 1, \ldots, k$$

of (P^k), we see that $\bar{x}$ is an efficient point of the convex hull of $x^0, \ldots, x^k$ and $\bar{x}$ itself. In this sense Problem (P^k) is an inner approximation of Problem (P_E).

We start with the case where Problem (P^k) has no solution.

LEMMA 4.2. *Suppose $x^0, x^1, \ldots, x^k \in X_E$ and Problem (P^k) has no solution. Then $x^* \in \operatorname{argmax}\{dx^j \mid j = 0, 1, \ldots, k\}$ is an optimal solution of (P_E).*

Proof. Since (P^k) has no solution, if $x \in X$ together with some $\lambda \in \Lambda$ satisfies

$$\lambda Cx \geqslant \lambda Cx' \text{ for all } x' \in X \tag{12}$$

then $dx \leqslant \alpha^k$, i.e., $x \in X_E$ implies $dx \leqslant \alpha^k$. This and $x^* \in X_E$ yield the lemma. □

Leaving the method of solving (P^k) till later on, we give the algorithm first.

$\langle 0 \rangle$ (Initialization) Find an efficient vertex x^0, set $k = 0$ and go to $\langle k \rangle$.

$\langle k \rangle$ (Iteration k)

$\langle k1 \rangle$ Find a solution $(x, \lambda) \in R^n \times R_p$ of (P^k). If no solution exists, $x^* \in \operatorname{argmax}\{ dx^j \mid j = 0, \ldots, k \}$ is an optimal solution of (P_E). Otherwise, set $(\bar{x}^{k+1}, \bar{\lambda}^{k+1})$ be the solution found.

$\langle k2 \rangle$ Solve the linear program

$$(Test^k) \quad \left| \begin{array}{ll} \max & eCx \\ \text{s.t.} & Cx \geqslant C\bar{x}^{k+1} \\ & x \in X \end{array} \right.$$

for a solution $\hat{x}$. If $eC\hat{x} = eC\bar{x}^{k+1}$, go to $\langle k3 \rangle$. Otherwise, go to $\langle k5 \rangle$.

$\langle k3 \rangle$ If $\bar{x}^{k+1}$ is a vertex of X, then set $x^{k+1} = \bar{x}^{k+1}$, $k = k + 1$ and go to $\langle k \rangle$. Otherwise, go to $\langle k4 \rangle$.

$\langle k4 \rangle$ Let F be a face of X whose relative interior contains $\bar{x}^{k+1}$, and solve the linear program

$$(Face^k) \quad \left| \begin{array}{ll} \max & dx \\ \text{s.t.} & x \in F \end{array} \right.$$

for an extreme point x^{k+1}. Set $k = k + 1$ and go to $\langle k \rangle$.

$\langle k5 \rangle$ Solve $(SC(\bar{\lambda}^{k+1}))$ for a solution x^{k+1}, set $k = k + 1$ and go to $\langle k \rangle$.

Note that whether $\bar{x}^{k+1}$ is a vertex of X can be seen by checking the linear independence of colunms of A corresponding to positive components of $\bar{x}^{k+1}$.

There may be various ways of determining the face F of $\langle k4 \rangle$. One possible way is

$$F = \{ x \mid x \in X; x_j = 0 \text{ for } j \text{ with } \bar{x}_j^{k+1} = 0 \}. \tag{13}$$

Benson proposes to define it by

$$F = \{ x \mid x \in X; (e + u)Cx = v \}, \tag{14}$$

where u is an optimal dual variable vector corresponding to the constraint $Cx \geqslant C\bar{x}^{k+1}$ of $(Test^k)$ and $v = \max\{ (e + u)Cx \mid x \in X \}$.

The following lemma shows that x^k's are efficient vertices of X.

LEMMA 4.3. *$x^k \in X_V \cap X_E$ for $k = 0, 1, \ldots$.*

Proof. Since it is clear that $x^j \in X_V$, we only show that $x^j \in X_E$. When x^{k+1} is computed in either $\langle k3\rangle$ or $\langle k5\rangle$, it is an optimal solution of either $(Test^k)$ or $(SC(\bar{\lambda}^{k+1}))$. Then clearly $x^{k+1} \in X_E$. When x^{k+1} is generated in $\langle k4\rangle$, it lies in the face whose relative interior contains the efficient point $\bar{x}^{k+1}$. Then by Theorem 2.7 we see $x^{k+1} \in X_E$. □

Now we show that the algorithm always generates a sequence of distinct vertices of X_E.

LEMMA 4.4. $x^{k+1} \notin \{ x^j \mid j = 0, 1, \ldots, k \}$.

Proof. Three cases should be considered. In $\langle k3\rangle$ x^{k+1} is given by $x^{k+1} = \bar{x}^{k+1}$, which satisfies $d\bar{x}^{k+1} > \max\{ dx^j \mid j = 0, \ldots, k \}$, and hence x^{k+1} differs from any point of $x^0, \ldots, x^k$. By construction $dx^{k+1} \geqslant d\bar{x}^{k+1}$ in $\langle k4\rangle$ and the same argument applies. Now suppose x^{k+1} is generated in $\langle k5\rangle$. Then $\bar{x}^{k+1} \notin X_E$, i.e., there is a point, say $\tilde{x} \in X$ with $C\tilde{x} \geqslant C\bar{x}^{k+1}$ and $C\tilde{x} \neq C\bar{x}^{k+1}$. Since $\bar{\lambda}^{k+1} > 0$ we see $\bar{\lambda}^{k+1}C\tilde{x} > \bar{\lambda}^{k+1}C\bar{x}^{k+1}$. Since x^{k+1} solves $(SC(\bar{\lambda}^{k+1}))$, we also see $\bar{\lambda}^{k+1}Cx^{k+1} \geqslant \bar{\lambda}^{k+1}C\tilde{x}$. Then for $j = 0, \ldots, k$

$$\bar{\lambda}^{k+1}Cx^{k+1} \geqslant \bar{\lambda}^{k+1}C\tilde{x} > \bar{\lambda}^{k+1}C\bar{x}^{k+1} \geqslant \bar{\lambda}^{k+1}Cx^j \tag{15}$$

holds. This means that $x^{k+1} \notin \{ x^j \mid j = 0, \ldots, k \}$. □

Note that in either case of $\langle k3\rangle$ and $\langle k4\rangle$ $dx^{k+1} > \max\{ dx^j \mid j = 0, \ldots, k \}$, i.e., monotone increasing of the objective function value, but in case $\langle k5\rangle$ it may decrease. Combining the above lemmas we have the following theorem.

THEOREM 4.5. *Suppose Problem (P^k) is solved within a finite number of iterations. Then the algorithm provides an optimal solution x^* of Problem (P_E) after a finite number of iterations.*

Now we go back to Problem (P^k) and explain the algorithm proposed by Benson (1991). For a solution of (P^k) it suffices to solve

$$(\overline{P^k}) \qquad \left| \begin{array}{ll} \max & dx \\ \text{s.t.} & \lambda Cx \geqslant \lambda Cx^j \text{ for } j = 0, 1, \ldots, k \\ & x \in X \\ & \lambda \in \Lambda. \end{array} \right.$$

Let $\overline{Y} = \{ y \mid y \in R^p; \min\{-c^i x \mid x \in X\} \leqslant y_i \leqslant \max\{-c^i x \mid x \in X\} \text{ for } i = 1, \ldots, p \}$ and $\overline{\Lambda}$ be a p-dimensional hypercube containing Λ, for example $\overline{\Lambda} = \{ \lambda \mid \lambda \in R_p; e \leqslant \lambda \leqslant (M + p - 1)e \}$. Then $(\overline{P^k})$ is equivalent to

$$\left|\begin{array}{ll} \max & dx \\ \text{s.t.} & \lambda y + \lambda C x^j \leqslant 0 \text{ for } j = 0, 1, \ldots, k \\ & y + Cx = 0 \\ & x \in X \\ & y \in \overline{Y} \\ & \lambda \in \overline{\Lambda}. \end{array}\right.$$

The constraint $\lambda 1 = M$ could be added, but is not necessary. The bilinear term λy makes the problem difficult to solve and hence should be relaxed. The algorithm in Benson (1991) is based on the successive partition of the hypercube $\overline{Y} \times \overline{\Lambda}$ into smaller hypercubes and the relaxation of the problem restricted to the smaller hypercubes to a linear program. Let $\overline{Y}' \times \overline{\Lambda}' = \prod_{i=1}^{p}[\underline{\alpha}_i, \overline{\alpha}_i] \times \prod_{i=1}^{p}[\underline{\beta}_i, \overline{\beta}_i]$ be a smaller hypercube contained in $\overline{Y} \times \overline{\Lambda}$. Note that $\overline{Y}' \times \overline{\Lambda}' = \prod_{i=1}^{p}([\underline{\alpha}_i, \overline{\alpha}_i] \times [\underline{\beta}_i, \overline{\beta}_i])$ by rearranging the coordinates and λy is the sum of bilinear terms $\lambda_i y_i$ defined on $[\underline{\alpha}_i, \overline{\alpha}_i] \times [\underline{\beta}_i, \overline{\beta}_i]$. Al-Khayyal and Falk (1983) show that the *convex envelope* of $\lambda_i y_i$ on the two-dimensional cube $[\underline{\alpha}_i, \overline{\alpha}_i] \times [\underline{\beta}_i, \overline{\beta}_i]$, the pointwise supremum of all convex functions underestimating $\lambda_i y_i$ on $[\underline{\alpha}_i, \overline{\alpha}_i] \times [\underline{\beta}_i, \overline{\beta}_i]$, is given by the piecewise linear convex function $\max\{\underline{\beta}_i y_i + \underline{\alpha}_i \lambda_i - \underline{\beta}_i \underline{\alpha}_i, \overline{\beta}_i y_i + \overline{\alpha}_i \lambda_i - \overline{\beta}_i \overline{\alpha}_i\}$. Then the convex envelope of λy is given by $\sum_{i=1}^{p} \max\{\underline{\beta}_i y_i + \underline{\alpha}_i \lambda_i - \underline{\beta}_i \underline{\alpha}_i, \overline{\beta}_i y_i + \overline{\alpha}_i \lambda_i - \overline{\beta}_i \overline{\alpha}_i\}$ and the constraint $\lambda y + \lambda C x^j \leqslant 0$ is relaxed to

$$\sum_{i=1}^{p} \max\{\underline{\beta}_i y_i + \underline{\alpha}_i \lambda_i - \underline{\beta}_i \underline{\alpha}_i, \overline{\beta}_i y_i + \overline{\alpha}_i \lambda_i - \overline{\beta}_i \overline{\alpha}_i\} + \lambda C x^j \leqslant 0. \tag{16}$$

This constraint is, by introducing variables w_i's, rewritten as

$$\underline{\beta}_i y_i + \underline{\alpha}_i \lambda_i - \underline{\beta}_i \underline{\alpha}_i \leqslant w_i \text{ for } i = 1, \ldots, p \tag{17}$$

$$\overline{\beta}_i y_i + \overline{\alpha}_i \lambda_i - \overline{\beta}_i \overline{\alpha}_i \leqslant w_i \text{ for } i = 1, \ldots, p \tag{18}$$

$$\sum_{i=1}^{p} w_i + \lambda C x^j \leqslant 0. \tag{19}$$

Thus we yield a linear programming relaxation of $(\overline{P^k})$ restricted to a smaller hypercube $\overline{Y}' \times \overline{\Lambda}'$ contained in $\overline{Y} \times \overline{\Lambda}$. In Benson (1991) (16) is further relaxed to a single linear inequality.

It would be a routine to construct a branch-and-bound algorithm based on this relaxation. If we employ the bisection procudure to divide a hypercube, i.e., to divide it into two hypercubes with equal volumes such that the midpoint of one of the longest edges is a vertex of both new hypercubes, we will see the following theorem.

THEOREM 4.6. *If the branch-and-bound procedure does not terminate after a finite number of iterations, any accumulation point of the sequence* $(x^\nu, y^\nu, \lambda^\nu, w^\nu)$ *generated by the procedure is an optimal solution of* $(\overline{P^k})$.

See, for example, Section 4 of Chapter VII in Horst and Tuy (1996) for the convergence proof.

5. Face search algorithm

In this section we introduce the algorithm for Problem (P_E) proposed by Sayin (2000), which is based on the enumeration method of efficient faces in Sayin (1996).

For a point $x \in X$ let $I(x)$ be the index set of zero components of x, i.e., $I(x) = \{ i \mid i \in \{1, \ldots, n\}; x_i = 0 \}$. For $I \subseteq \{1, \ldots, n\}$ let

$$F(I) = \{ x \mid x \in X; x_i = 0 \text{ for } i \in I \}, \tag{20}$$

which is a, possibly vacant, face of X. Then the efficient set X_E is decomposed as

$$X_E = \bigcup_{I \subseteq \{1,\ldots,n\}} (X_E \cap F(I)). \tag{21}$$

Therefore Problem (P_E) reduces to the family of following problems

$$(P_E(I)) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X_E \cap F(I), \end{array} \right.$$

each of which is corresponding to $I \subseteq \{1, \ldots, n\}$. For a mutually disjoint decomposition of X_E see Corollary 3.3 in Benson (1995). Since $X_E \cap F(I) \subseteq X \cap F(I) = F(I)$,

$$(\overline{P}_E(I)) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in F(I). \end{array} \right.$$

is a relaxation problem of $(P_E(I))$. Note that this is a linear program when ϕ is a linear function.

Suppose we have at hand an incumbent, i.e., a point $x^* \in X_E$, and the list of problems $(P_E(I))$ to solve. At the beginning the list consists of the single problem $(P_E(\emptyset))$, which is identical to (P_E) since $F(\emptyset) = X$. Choosing a problem $(P_E(I))$ on the list and solving its relaxation $(\overline{P}_E(I))$, we have the following cases.

1. $(\overline{P}_E(I))$ is infeasible: Problem $(P_E(I))$ is fathomed and deleted from the list.
2. $(\overline{P}_E(I))$ has an optimal solution x.
   (a) $\phi(x) < \phi(x^*)$: Problem $(P_E(I))$ is fathomed and deleted from the list.
   (b) $\phi(x) > \phi(x^*)$:

(i) $x \in X_E$: The incumbent is updated as $x^* = x$, and Problem $(P_E(I))$ is fathomed and deleted from the list.

(ii) $x \notin X_E$: Problem $(P_E(I))$ is fathomed and deleted from the list, and for each index $k \in \{1, \ldots, n\} \setminus I$ Problem $(P_E(I \cup \{k\}))$ is added to the list.

The last case where $x \notin X_E$ may need an explanation. We see from Theorem 2.7 that no point in the relative interior of $F(I)$ is efficient in this case. Since any point in the relative boundary of $F(I)$ belongs to $F(I \cup \{k\})$ for some $k \in \{1, \ldots, n\} \setminus I$, Problem $(P_E(I))$ is fathomed and can be deleted from the list.

In the case of $x \notin X_E$, if $x_k = 0$, it remains optimal to Problem $(\overline{P}_E(I \cup \{k\}))$, which therefore needs not be solved. Even if this is not the case, Probelm $(\overline{P}_E(I \cup \{k\}))$ differs slightly from $(\overline{P}_E(I))$.

The key issue of implementation would be the list-management as it is always the case in the branch-and-bound method. Especially, a subset $I = \{i_1, \ldots, i_\ell\}$ of $\{1, \ldots, n\}$ would be generated from each of $\{i_2, \ldots, i_\ell\}$, $\{i_1, i_3, \ldots, i_\ell\}, \ldots$, and $\{i_1, \ldots, i_{\ell-1}\}$. The redundacy can be avoided by a simple technique. Even incorporating the technique, the list grows very rapidly and becomes too larg to keep in the memory. Due to the rapid growth of problem list, the computational experiment reported in Sayin (2000) is restricted in problem size.

6. Branch-and-bound algorithm

This section is devoted to introducing the branch-and-bound algorithm for Problem (P_W) with a concave function ϕ proposed by Horst and Thoai (1999) and Thoai (2001).

First they observe the following characterization of the weakly efficient outcome set Y_W.

LEMMA 6.1. *Let $\partial Y^{\leqslant}$ denote the boundary of $Y^{\leqslant}$. Then $Y_W = Y \cap \partial Y^{\leqslant}$.*

Proof. This lemma follows the equivalence $Y \cap \operatorname{int} Y^{\leqslant} = Y \setminus Y_W$. If $y \in Y \cap \operatorname{int} Y^{\leqslant}$, $y < y'$ for some $y' \in Y^{\leqslant}$, for which there is $y'' \in Y$ such that $y' \leqslant y''$. Therefore $y \notin Y_W$. If $y \in Y \setminus Y_W$, there is $y' \in Y$ with $y < y'$, and hence its neighbor $\{ z \mid z \in R_p;\ y - (y' - y) \leqslant z \leqslant y' \}$ is contained in $Y^{\leqslant}$. This implies $y \in \operatorname{int} Y^{\leqslant}$. □

Then Problem (P_W) is rewritten as

$$\max \{ \phi(x) \mid x \in X;\ Cx \in \partial Y^{\leqslant} \}. \tag{22}$$

Introducing additional variables $y \in R^p$ and $t \in R$, it is cast into the following problem called *Master Problem*

$$(MP) \quad \left| \begin{array}{ll} \max & t \\ \text{s.t.} & t \leqslant \phi(x) \\ & x \in X \\ & y = Cx \\ & y \in \partial Y^{\leqslant}, \end{array} \right.$$

for which the following theorem holds.

THEOREM 6.2. *If x^* is an optimal solution of (P_W), then (x^*, y^*, t^*) with $y^* = Cx^*$, $t^* = \phi(x^*)$ is an optimal solution of (MP). If (x^*, y^*, t^*) is an optimal solution of (MP), then x^* is an optimal solution of (P_W) with $\phi(x^*) = t^*$.*

Since we assume that the feasible region X is bounded, there is a point $y^0 \in R^p$ whose ith component y_i^0 satisfies

$$y_i^0 \leqslant \min\{\, y_i \mid y \in Y \,\} = \min\{\, c^i x \mid x \in X \,\}. \tag{23}$$

Then

$$Y_W \subseteq (y^0 + R_+^p) \cap \partial Y^{\leqslant} \subseteq (y^0 + R_+^p) \cap Y^{\leqslant}. \tag{24}$$

The key idea of the algorithm is to decompose the truncated lower outcome set $(y^0 + R_+^p) \cap Y^{\leqslant}$ into cones K with vertex at y^0 and consider the subproblem $(MP(K))$ with variable y restricted to $\partial Y^{\leqslant} \cap K$

$$(MP(K)) \quad \left| \begin{array}{ll} \max & t \\ \text{s.t.} & t \leqslant \phi(x) \\ & x \in X \\ & y = Cx \\ & y \in \partial Y^{\leqslant} \cap K. \end{array} \right.$$

There are two things to have done: to replace the concave function ϕ by a function easier to handle, and to construct a polyhedral set containing $Y \cap \partial Y^{\leqslant} \cap K$. They propose a piecewise linear concave function Φ to replace ϕ. Suppose we have a finite number of points $x^1, \ldots, x^k$ in the domain of ϕ and a subgradient $s^i \in R_n$ of ϕ at x^i. Then

$$\Phi(x) = \min\{\, \phi(x^i) + s^i(x - x^i) \mid i = 1, \ldots, k \,\} \tag{25}$$

is a piecewise linear concave function which overestimates ϕ, i,e., $\Phi(x) \geqslant \phi(x)$ at any point x. Furthermore note that the constraint $t \leqslant \phi(x)$ with ϕ replaced by Φ is equivalent to the k linear inequality constraints

$$t \leqslant \phi(x^i) + s^i(x - x^i) \text{ for } i = 1, \ldots, k. \tag{26}$$

Let $r^1, \ldots, r^p \in R^p$ be p extreme rays generating the cone $K - y^0$ and for each $i = 1, \ldots, p$ let y^i be the intersection point of the ray $\{ y \mid y = y^0 + \alpha r^i; \alpha \geqslant 0 \}$ and $\partial Y^{\leqslant}$. The intersection point y^i is found by solving the linear program

$$\max \{ \alpha \mid y^0 + \alpha r^i \leqslant Cx; x \in X; \alpha \geqslant 0 \}. \tag{27}$$

Once we have these points $y^1, \ldots, y^p$ and the hyperplane, say H, passing through them, we see the following lemma. See Fig. 2.

LEMMA 6.3. *Let H^+ be the half space defined by H that does not contain y^0. Then*

$$Y \cap \partial Y^{\leqslant} \cap K \subseteq Y \cap Y^{\leqslant} \cap K \cap H_+ = Y \cap K \cap H_+. \tag{28}$$

Therefore as a relaxation problem of $(MP(K))$ we obtain

$$(\overline{MP}(K)) \quad \left| \begin{array}{ll} \max & t \\ \text{s.t.} & t \leqslant \phi(x^i) + s^i(x - x^i) \text{ for } i = 1, \ldots, k. \\ & x \in X \\ & y = Cx \\ & y \in K \cap H_+. \end{array} \right.$$

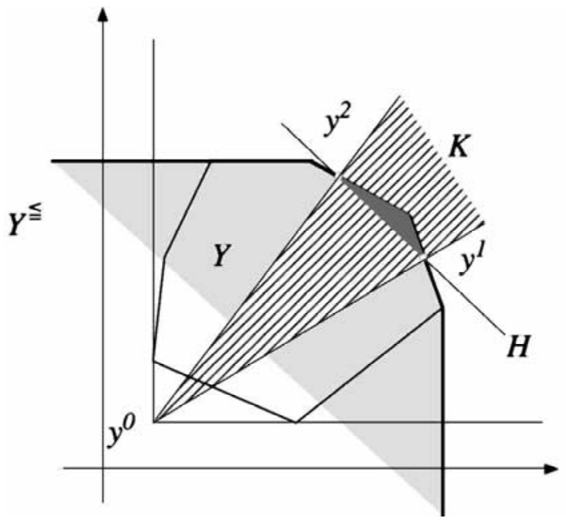

Figure 2. Problem $(\overline{MP}(K))$.

Let V be the $p \times p$ matrix consisting of columns $y^1 - y^0, \ldots, y^p - y^0$, then the last two constraints are equivalent to

$$Cx = V\mu + y^0; \quad e\mu \geqslant 1; \quad \mu \in R^p_+. \tag{29}$$

Clearly the optimal value of $(\overline{MP}(K))$ provides an upper bound of the optimal value of $(MP(K))$.

Once the relaxation problem is so constructed, it will be a routine to make a branch-and-bound algorithm and we omit the description. To guarantee the convergence

1. the piecewise linear approximation Φ of ϕ should become better, and
2. the conical partition should become finer

as the process proceeds. Everytime an optimal solution $(x(K), y(K), t(K))$ of Problem $(\overline{MP}(K))$ is obtained, the set of points $x^1, \ldots, x^k$ is incremented by $x(K)$, which improves the approximation accuracy of Φ. Concerning the conical partition, the desired property is referred to as *exhaustiveness* and defined as

DEFINITION 6.4. The partition procedure is said to be *exhaustive* when $\bigcap_k K^k$ is a ray for any nested sequence $\{K^k\}_{k=1,\ldots}$ of cones generated by the procedure.

See Horst and Tuy (1996) for a full detail of exhaustiveness.

THEOREM 6.5. *Assume that the conical partition procedure is exhaustive. Then every cluster point (x^*, y^*, t^*) of the sequence of points (x^ν, y^ν, t^ν) generated by the branch-and-bound algorithm is a solution of Master Problem (MP). Hence x^* is a solution of (P_W).*

Preliminary computational results are reported in Thoai (2001) for linear case. He ran the algorithm on randomly generated test problems with $p = 2$ to 4, $m = 10$ to 50 and $n = 35$ to 250, and reported the average number of iterations, the maximal number of cones stored at an iteration and the average CPU time.

7. Lagrangian relaxation methods

White (1996) considered Problem (P_E) with linear function $\phi(x) = dx$ and presented several equivalent formulations. Dauer and Fosnaugh (1995) considered the problem with quasi-convex function ϕ and showed a way of converting it to a bicriteria problem, which could be viewed as a Lagrangian relaxation of Problem (P_E). An, Tao and Muu An, Tao and Muu (1996) showed that there is no duality gap for a sufficiently large Lagrangian multiplier. We will explain the common idea in terms of the Lagrangian relaxation method. The central role will be played by the gap function $g\colon X \to R$ defined by

$$g(x) = \max\{\, eCx' \mid x' \in X; Cx' \geqslant Cx \,\} - eCx. \tag{30}$$

We call a point x' that attains the maximum above a *projected point* of x. It is easily seen from the theory of parmetric linear program that g is a piecewise linear concave function on X. As stated in Theorem 2.7 $g(x) \geqslant 0$ for $x \in X$, and $x \in X_E$ if and only if $g(x) = 0$ for $x \in X$. See Theorem 4.1 of Benson (1995). Thus Problem (P_E) is reformulated as follows:

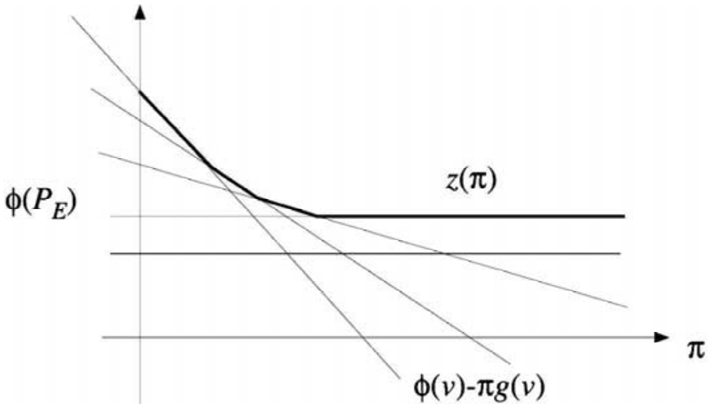

Figure 3. $z(\pi)$ and $\phi(v) - \pi g(v)$.

$$(P_E) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X \\ & g(x) \leqslant 0. \end{array} \right.$$

Note that the last constraint $g(x) \leqslant 0$ is a reverse convex constraint, which has been attracting attention. See, for example, Horst and Tuy (1996) and Tuy (1998). To solve Problem (P_E) we combine the objective function $\phi(x)$ with the constraint $g(x) \leqslant 0$ multiplied by a Lagrangian multiplier $\pi \geqslant 0$ to have the *Lagrangian relaxation problem*

$$(Q(\pi)) \qquad \left| \begin{array}{lll} z(\pi) = & \max & \phi(x) - \pi g(x) \\ & \text{s.t.} & x \in X. \end{array} \right.$$

In the sequel $x(\pi)$ denotes an optimal solution of $(Q(\pi))$ and $x'(\pi)$ denotes its projected point. Note that $(Q(\pi))$ is a quasi-convex maximization and that the optimality is always attained at a vertex of X. For we assume that X is a polytope, we reformulate Problem $(Q(\pi))$ in terms of the vertices of X and obtain

$$z(\pi) = \max\{\phi(v) - \pi g(v) \mid v \in X_V\}. \tag{31}$$

Note that for each vertex $v \in X_V$ the function $\phi(v) - \pi g(v)$ is a linear function with nonpositive slope in variable π. In Figure 3 are shown these linear functions as well as $z(\pi)$ depicted by a bold piecewise linear line. Notice that horizontal lines, meaning $g(v) = 0$, correspond to vertices in X_E.

Though the following lemmas are straightforward from this observation, brief proofs will be given.

LEMMA 7.1. *If $g(x(\pi)) = 0$ for some $\pi \geqslant 0$, then $x(\pi)$ is an optimal solution of (P_E).*

Proof. For any x in X_E, we readily see $\phi(x(\pi)) = \phi(x(\pi)) - \pi g(x(\pi)) \geqslant \phi(x) - \pi g(x) = \phi(x)$. □

Concerning $z(\pi)$ we have the following property.

LEMMA 7.2. *Let* $0 \leqslant \pi \leqslant \pi'$ *and let* $x'(\pi)$ *be a projected point of* $x(\pi)$. *Then*

$$\phi(x'(\pi)) \leqslant \phi(P_E) \leqslant z(\pi') \leqslant z(\pi). \tag{32}$$

Proof. Since the projected point lies in X_E, the first inequality is trivial. By the definition of $z(\pi')$, it holds that $\phi(x) - \pi' g(x) \leqslant z(\pi')$ for any $x \in X$ and also for any $x \in X_E$. Then we see $\phi(x) \leqslant z(\pi')$ for any $x \in X_E$, which implies the second inequality. The last inequality is derived from $z(\pi') = \phi(x(\pi')) - \pi' g(x(\pi')) \leqslant \phi(x(\pi')) - \pi g(x(\pi')) \leqslant \phi(x(\pi)) - \pi g(x(\pi)) = z(\pi)$. □

This lemma means that $z(\pi)$ gives an upper bound of $\phi(P_E)$ and also $x'(\pi)$, the projected point of $x(\pi)$, gives a lower bound. Above two lemmas suggest that solution $x(\pi)$ of $(Q(\pi))$ for a sufficiently large $\pi > 0$ solves Problem (P_E). In fact, because of the finiteness of X_V we readily see the following theorem. See Figure 3.

THEOREM 7.3. *There is a* $\pi^* > 0$ *such that for any* $\pi > \pi^*$ $x(\pi)$ *is an optimal solution of* (P_E).

An, Tao and Muu showed the same result for a convex funtion ϕ in Lemma 4 of An et al. (1996). Dauer and Fosnaugh showed in (1995) that $z(\pi)$ converges to $\phi(P_E)$ as π goes to infinity for a more general setting.

Muu (2000) reduces the variables of the gap function by using Lemma 2.6. Let r be the rank of C and without loss of generality we assume that the first r rows $c^1, \ldots, c^r$ are linearly independent. Let L be the range space of matrix $C^\top$ and $L^\perp$ be its orthogonal complement in R^n and suppose we have a basis $b^{r+1}, \ldots, b^n$ of $L^\perp$. Then any $x \in R^n$ is uniquely represented as $x = \overline{C}^\top \alpha + B\beta$ for $\alpha \in R^r$ and $\beta \in R^{n-r}$, where $\overline{C}$ is the matrix of rows $c^1, \ldots, c^r$ and B is the matrix of columns $b^{r+1}, \ldots, b^n$. Then Problem $(Q(\pi))$ is rewritten as

$$\left|\begin{array}{lll} z(\pi) = & \max & \phi(\overline{C}^\top \alpha + B\beta) - \pi g(\overline{C}^\top \alpha + B\beta) \\ & \text{s.t.} & \overline{C}^\top \alpha + B\beta \in X. \end{array}\right.$$

We see, however, by Lemma 2.6 that

$$g(\overline{C}^\top \alpha + B\beta) = g(\overline{C}^\top \alpha), \tag{33}$$

then Problem $(Q(\pi))$ is

$$\left| \begin{array}{lll} z(\pi) = & \max & \phi(\overline{C}^\top \alpha + B\beta) - \pi g(\overline{C}^\top \alpha) \\ & \text{s.t.} & A\overline{C}^\top \alpha + AB\beta = b \\ & & \overline{C}^\top \alpha + B\beta \geqslant 0. \end{array} \right.$$

The rank r of C is no more than p, which is usually much smaller than n. When ϕ is a linear function, the above problem contains a small number of nonconvex variables.

Dauer and Fosnaugh (1995) also showed that when ϕ is a linear function dx and d is a linear combination of rows c^i's of C, i.e., $d = \gamma C$ for some $\gamma \in R_p$, the π^* in Theorem 7.3 is given by $\|\gamma\|_\infty$. Notice that this value is 1 if $d = \pm c^i$ for some $i = 1, \ldots, p$. Muu (2000) generalized this result to the nonlinear case where $\phi(x)$ is given by $\varphi(Cx)$ for some function φ.

The transformation of Problem (P_E) by White (1996) is based on Theorem 2.7. Note that Problem (P_E) is equivalent to

$$\max \left\{ \phi(x) \;\middle|\; \begin{array}{l} x \in X; \lambda \in \Lambda; \mu \in R_m; \\ \mu A - \lambda C \geqslant 0; \lambda C x - \mu b = 0 \end{array} \right\}. \tag{34}$$

By multiplying the bilinear constraint $\lambda Cx - \mu b = 0$ by π we have its Lagrangian relaxation

$$\max \left\{ \phi(x) + \pi(\lambda C x - \mu b) \;\middle|\; \begin{array}{l} x \in X; \lambda \in \Lambda; \mu \in R_m; \\ \mu A - \lambda C \geqslant 0 \end{array} \right\} \tag{35}$$

which is to maximize a bilinear objective function under linear inequality constraints. Several properties of this relaxation are discussed in White (1996).

8. Dual approach

Nonconvex duality is one of the most promising subject in the global optimization. We will not go into details of the duality theory in this paper. The readers who are interested in it should refer Atteia and El Qortobi (1981) and Thach (1991, 1993, 1994). In this section we will briefly explain the dual approach of Thach et al. (1996).

Let

$$C^{\leqslant} = \{\, y \mid y \in R^n; Cy \leqslant 0; c^i y < 0 \text{ for some } i = 1, \ldots, p \,\}. \tag{36}$$

Then the efficient set X_E is written as the difference of two convex sets. See Figure 4.

LEMMA 8.1.

$$X_E = X \setminus (X + C^{\leqslant}). \tag{37}$$

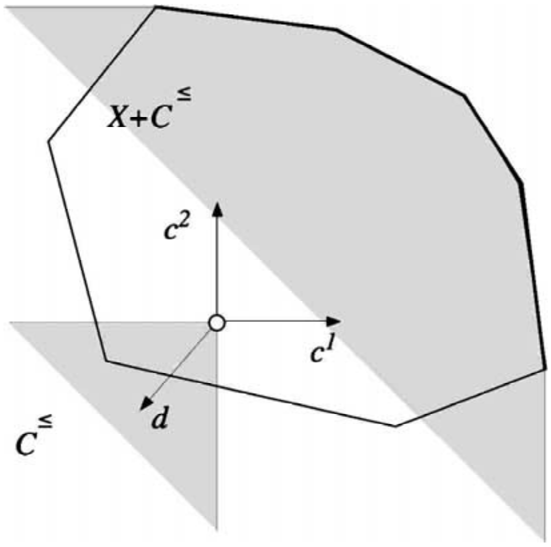

Figure 4. $X_E = X \setminus (X + C^{\leqslant})$.

Proof.

$$\begin{aligned} X_E &= \{ x \mid x \in X; \nexists x' \text{ such that } Cx' \geqslant Cx, c^i x' > c^i x \text{ for some } i \} \\ &= X \setminus \{ x \mid \exists x' \text{ such that } Cx' \geqslant Cx, c^i x' > c^i x \text{ for some } i \} \\ &= X \setminus \{ x \mid \exists x' \text{ such that } C(x - x') \leqslant 0, c^i (x - x') < 0 \text{ for some } i \} \\ &= X \setminus \{ x + y \mid x \in X; Cy \leqslant 0; c^i y < 0 \text{ for some } i \} \\ &= X \setminus (X + C^{\leqslant}). \end{aligned}$$

□

Then Problem (P_E) is written as

$$(P_E) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X \setminus (X + C^{\leqslant}). \end{array} \right.$$

Since X is now assumed to be a polytope, we show that the set $X + C^{\leqslant}$ can be replaced by the interior of a closed convex set. Let E be the $p \times p$ matrix all of whose elements are unity, and for a positive parameter s define a $p \times p$ matrix C_s, sets $C_s^{\leqslant}$ and X_s by

$$C_s = (I + sE)C \tag{38}$$

$$C_s^{\leqslant} = \{ y \mid C_s y \leqslant 0 \} \tag{39}$$

$$X_s = X \setminus \operatorname{int}(X + C_s^{\leqslant}), \tag{40}$$

where I is the $p \times p$ indentity matrix. Note that X_s is also the difference of two convex sets.

LEMMA 8.2.

$$X_s = \left\{ x \;\middle|\; \begin{array}{l} x \in X; \exists \lambda \in R_{p+} \setminus \{0\} \textit{ such that} \\ \lambda C_s x \geqslant \lambda C_s x' \textit{ for all } x' \in X \end{array} \right\}. \tag{41}$$

Proof. Let x be a point in X_s. By the separation theorem, there is a $v \neq 0$ satisfying $vx \geqslant vz$ for all $z \in X + C_s^{\leqslant}$. Hence $vx \geqslant v(x + y)$ holds for all y such that $C_s y \leqslant 0$. Applying Farkas' alternative theorem, we have $v = \lambda C_s$ for some $\lambda \in R_{p+} \setminus \{0\}$, and hence $\lambda C_s x \geqslant \lambda C_s z$ holds for all $z \in X + C_s^{\leqslant}$. Noting that $0 \in C_s^{\leqslant}$ we see that $\lambda C_s x \geqslant \lambda C_s x'$ for all $x' \in X$, and hence x is contained in the set on the right side.

Suppose x maximizes $\lambda C_s x$ over X for some $\lambda \in R_{p+} \setminus \{0\}$. Then clearly it also maximizes $\lambda C_s x$ over $X + C_s^{\leqslant}$ and does not lie in the interior of $X + C_s^{\leqslant}$. □

By this lemma we see that X_s coincides with X_E when s is sufficiently small.

LEMMA 8.3. *There is an* $\hat{s} > 0$ *such that* $X_s = X_E$ *if* $0 < s < \hat{s}$.

Proof. To show that $X_s \subseteq X_E$, choose arbitrarily $x \in X_s$. Then by the above lemma, there is a $\lambda \in R_{p+} \setminus \{0\}$ such that x maximizes $\lambda C_s x$ over X. Here we assume that $\lambda 1 = 1$ without loss of generality. Substituting the definition for C_s, we see $\lambda C_s = (\lambda + se)C$. This and the equality (2)

$$X_E = \{ x \mid x \in X; \exists \lambda \in R_{p++} \text{ such that } \lambda C x \geqslant \lambda C x' \text{ for } \forall x' \in X \}$$

of Theorem 2.7 imply that $x \in X_E$.

By Theorem 2.7 X_E is the union of finitely many faces $F^1, \ldots, F^L$ of X such that F^ℓ is the optimal set of maximizing $\lambda^\ell C x$ over X for some $\lambda^\ell \in R_{p++}$ such that $\lambda^\ell 1 = 1$. Let $\hat{s} = \min\{ \lambda_i^\ell/(1 - p\lambda_i^\ell) \mid \lambda_i^\ell < 1/p \}$ and choose s such that $0 < s < \hat{s}$. Then $s/(1 + sp) < \lambda_i^\ell$ for all $\ell = 1, \ldots, L$ and $i = 1, \ldots, p$. Let $\theta_i^\ell = \lambda_i^\ell - \frac{s}{1+sp}$ for $\ell = 1, \ldots, L$, $i = 1, \ldots, p$. Then we readily see that $\theta_i^\ell > 0$ and

$$\lambda^\ell C = \theta^\ell C_s. \tag{42}$$

This means that $F^\ell \subseteq X_s$ by Lemma 8.2, and hence $X_E \subseteq X_s$. □

We assume hereafter that $0 < s < \hat{s}$. Then Problem (P_E) is equivalently rewritten as

$$(P_E) \qquad \left| \begin{array}{ll} \max & \phi(x) \\ \text{s.t.} & x \in X_s = X \setminus \operatorname{int}(X + C_s^{\leqslant}). \end{array} \right.$$

For $v \in R_n$ let

$$\xi(v) = \sup\{ \phi(x) \mid x \in X; vx \geqslant 1 \}, \tag{43}$$

where $\xi(v) = -\infty$ when $\{ x \mid x \in X; vx \geqslant 1 \} = \emptyset$.

DEFINITION 8.4. For $Z \subseteq R^n$ the set $\{ v \mid v \in R_n; vx \leqslant 1$ for all $x \in Z \}$ is called the *polar set* of Z and denoted by Z°.

See for example Section 2.14 of Stoer and Witzgall (1970), and Section E of Chapter 11 in Rockafellar and Wets (1998) for the properties of polar set. We here assume that $0 \in \mathrm{int}X$, $\mathrm{int}C_s^{\leqslant} \neq \emptyset$ and ϕ is a concave function. Then by the nonconvex duality theory of Thach (1991) we obtain the following duality theorem between Problem (P_E) and its dual problem

$$(D_s) \qquad \left| \begin{array}{ll} \max & \xi(v) \\ \text{s.t.} & v \in (X + C_s^{\leqslant})^\circ. \end{array} \right.$$

THEOREM 8.5. *Let $\xi(D_s)$ denote the optimal value of (D_s), then*

$$\phi(P_E) = \xi(D_s).$$

Proof. See Thach (1991) and Chapter 4 of Konno, Thach and Tuy (1997). □

Since $0 \in \mathrm{int}X$, $(X + C_s^{\leqslant})^\circ \subseteq (C_s^{\leqslant})^\circ$, which is identical to $\{ \gamma C_s \mid \gamma \in R_{p+} \}$. Therefore $v \in (X + C_s^{\leqslant})^\circ$ if and only if $v = \gamma C_s$ for some $\gamma \in R_{p+}$ and $\sup\{ v(x + y) \mid x \in X; y \in C_s^{\leqslant} \} \leqslant 1$. The latter condition can be replaced by $\sup \{ vx \mid x \in X \} \leqslant 1$ from the definition of $C_s^{\leqslant}$ and $v = \gamma C_s$. Letting

$$\Gamma = \{ \gamma \mid \gamma \in R_{p+}; \sup_{x \in X} \gamma C_s x \leqslant 1 \}, \tag{44}$$

we have

$$(X + C_s^{\leqslant})^\circ = \{ \gamma C_s \mid \gamma \in \Gamma \}. \tag{45}$$

Now let

$$\Xi(\gamma) = \sup \{ \phi(x) \mid x \in X; \gamma C_s x \geqslant 1 \}. \tag{46}$$

The above argument yields an equivalent form of (D_s) in variable $\gamma \in R_p$.

THEOREM 8.6. *Problem (D_s) is equivalent to*

$$\left| \begin{array}{ll} \max & \Xi(\gamma) \\ \textit{s.t.} & \gamma \in \Gamma. \end{array} \right.$$

We will see that this problem is a quasi-convex maximization over a convex polyhedral set.

LEMMA 8.7.
(i) Γ *is a convex polyhedral subset of* R_p.
(ii) Ξ *is a quasi-convex function.*

Proof. The first assertion can be seen from the finitely constrained representation

$$\Gamma = \{ \gamma \mid \gamma \in R_{p+}; \gamma C_s x \leqslant 1 \text{ for } x \in X_V \}.$$

To show the second assertion let γ^1, γ^2 be two point of the level set $\{ \gamma \mid \Xi(\gamma) \leqslant \alpha \}$, meaning $\sup \{ \phi(x) \mid x \in X; \gamma^k C_s x \geqslant 1 \} \leqslant \alpha$ for $k = 1, 2$, and suppose $\sup\{ \phi(x) \mid x \in X; (\beta\gamma^1 + (1-\beta)\gamma^2)C_s x \geqslant 1 \} > \alpha$ for some $\beta \in (0, 1)$. Then there is $\tilde{x} \in X$ such that $(\beta\gamma^1 + (1-\beta)\gamma^2)C_s\tilde{x} \geqslant 1$ and $\phi(\tilde{x}) > \alpha$. For $\tilde{x}$ either $\gamma^1 C_s \tilde{x} \geqslant 1$ or $\gamma^2 C_s \tilde{x} \geqslant 1$ holds. Hence we obtain either $\sup\{ \phi(x) \mid x \in X; \gamma^1 C_s x \geqslant 1 \} \geqslant \phi(\tilde{x}) > \alpha$ or $\sup\{ \phi(x) \mid x \in X; \gamma^2 C_s x \geqslant 1 \} \geqslant \phi(\tilde{x}) > \alpha$, which is a contradiction. $\square$

They exploited the outer approximation method to solve the dual problem in Theorem 8.6 and proposed the following algorithm. See Figure 5.

$\langle 0 \rangle$ (Initialization) Construct a polyhedral set Γ^0 such that $\Gamma \subseteq \Gamma^0$ and the vertex set of Γ^0 is easily enumerated. Set $k = 0$ and go to $\langle k \rangle$.

$\langle k \rangle$ (Iteration k)

$\langle k1 \rangle$ Solve the relaxation problem

$$\left| \begin{array}{ll} \max & \Xi(\gamma) \\ \text{s.t.} & \gamma \in \Gamma^k \end{array} \right.$$

to obtain a solution γ^k.

$\langle k2 \rangle$ Solve the linear program

$$\left| \begin{array}{ll} \max & \gamma^k C_s x \\ \text{s.t.} & x \in X \end{array} \right.$$

to obtain a vertex solution x^k and the optimal value $\sigma^k = \gamma^k C_s x^k$.

$\langle k3 \rangle$ If $\sigma^k \leqslant 1$, meaning that γ^k is in Γ and hence solves $\max\{ \Xi(\gamma) \mid \gamma \in \Gamma \}$, then solve $\max \{ \phi(x) \mid x \in X; \gamma^k C_s x \geqslant 1 \}$ and obtain a solution x^*. Stop with x^* as an optimal solution of (P_E).

$\langle k4 \rangle$ If $\sigma^k > 1$, meaning $\gamma^k \notin \Gamma$, reduce Γ^k to $\Gamma^{k+1} = \Gamma^k \cap \{ \gamma \mid \gamma C_s x^k \leqslant 1 \}$. Set $k = k + 1$ and go to $\langle k \rangle$.

THEOREM 8.8. *The algorithm terminates after a finite number of iterations and provides an optimal solution of* (P_E).

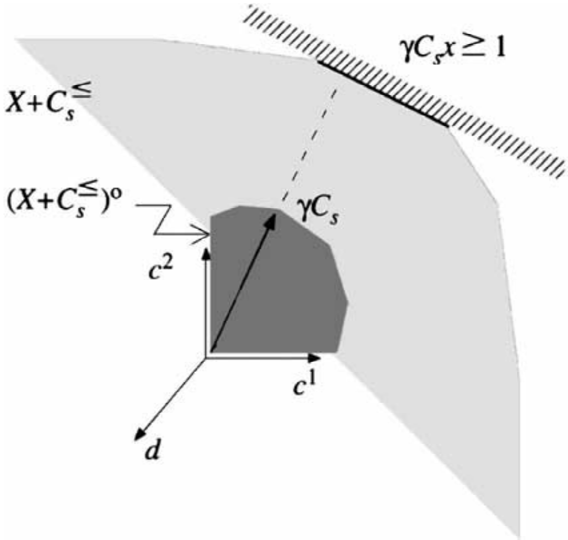

Figure 5. $X + C_s^{\leqslant}$ and its polar.

Proof. The theorem is readily seen from the fact that Γ is a polyhedral set defined by a finite number of constraints each of which corresponds to a vertex of X and that $\{x^k\}_{k=0,1,\ldots}$ generated by the algorithm is a sequence of distinct vertices of X. □

The most costly step of the algorithm is $\langle k1 \rangle$ of maximizing $\Xi(\gamma)$ over Γ^k. Thach, Konno and Yokota (1996) proposed to enumerate the vertex set of Γ^{k+1} from that of Γ^k in this step. Numerical results are reported in Thach et al. (1996) with two different objective functions: absolute deviation $\phi(x) = -\sum_{i=1}^n w_i|x_i - \bar{x}_i|$ and linear function $\phi(x) = -\sum_{i=1}^n w_i x_i$. They used the enumeration method by Thieu et al. (1983) in $\langle k1 \rangle$. They fixed $m = 20$ and varied $p = 2$–$5, n = 20$–100, and concluded that the number of vertices generated through the computation does not grow very rapidly as long as p is kept small, and also most of the computation time was spent in solving linear programs.

Based on the same duality concept Yamada et al. (2000) proposed an algorithm to the problem (P_W) with the concave objective function ϕ and the closed convex feasible region X satisfying Slater's constraint qualification.

9. Bisection search algorithm

This section is devoted to the explanation of the algorithm proposed by Phong and Tuyen (2000) for Problem (P_E) with linear objective function $\phi(x) = dx$. The main idea is the bisection method for locating $\phi(P_E)$. Namely, they start with an interval $[\ell_0, u_0]$ which is known to contain $\phi(P_E)$, solve for $\alpha = (\ell_k + u_k)/2$

$$(P_\alpha) \qquad \Big|\ \text{Find } x \in X_E \ \text{ such that } dx \geqslant \alpha$$

and then reduce the interval $[\ell_k, u_k]$ to either $[\alpha, u_k]$ when (P_α) has a solution or $[\ell_k, \alpha]$ when (P_α) has no solution. Thus after a finitely many iterations they obtain an ϵ-approximate solution.

For $\lambda \in \Lambda$ let $\sigma(\lambda)$ denote the optimal value of Problem $(SC(\lambda))$, i.e.,

$$\sigma(\lambda) = \max\{\, \lambda C x \mid x \in X \,\}, \tag{47}$$

and

$$\tau_\alpha(\lambda) = \max\{\, \lambda C x \mid x \in X;\, dx \geqslant \alpha \,\}. \tag{48}$$

Since X is the convex hull of its vertex set X_V and an efficient vertex solves Problem $(SC(\lambda))$ for $\lambda \in \Lambda$, we see

LEMMA 9.1.
(i) $\sigma(\lambda) = \max\{\, \lambda C v \mid v \in X_E \cap X_V \,\}$ *for* $\lambda \in \Lambda$.
(ii) $\sigma(\cdot)$ *is a piecewise linear convex function on* Λ.

Proof. From (i) σ is the maximum of finitely many linear functions $\lambda C v$ each of which corresponds to a vertex v of $X_E \cap X_V$. Thus it is piecewise linear convex. □

In the same way we obtain

LEMMA 9.2.
(i) $\tau_\alpha(\lambda) = \max\{\, \lambda C v \mid v$ *is an efficient vertex of* $X \cap \{\, x \mid dx \geqslant \alpha \,\}\}$.
(ii) $\tau_\alpha(\lambda) \leqslant \sigma(\lambda)$ *for any* $\lambda \in R_p$.
(iii) $\tau_\alpha(\cdot)$ *is a piecewise linear convex function on* Λ.
(iv) $\tau_\alpha(\lambda)$ *is a nonincreasing function in* $\alpha \in R$.

Let us denote the epigraph of σ by $\operatorname{epi}\sigma$, i.e.,

$$\operatorname{epi}\sigma = \{\, (\lambda, \mu) \mid (\lambda, \mu) \in \Lambda \times R;\, \sigma(\lambda) \leqslant \mu \,\}. \tag{49}$$

For the existence of a solution of (P_α) we have the following theorem.

THEOREM 9.3.
(i) $X_E \cap \{\, x \mid dx \geqslant \alpha \,\} \neq \emptyset$ *if and only if* $\sigma(\lambda) = \tau_\alpha(\lambda)$ *for some* $\lambda \in \Lambda$.
(ii) $\sigma(\lambda) = \tau_\alpha(\lambda)$ *for some* $\lambda \in \Lambda$ *if and only if there is a vertex* $(\bar\lambda, \bar\mu)$ *of* $\operatorname{epi}\sigma$ *such that* $\bar\mu = \tau_\alpha(\bar\lambda)$.

Proof. We show only the first assertion because the second assertion is clear from the piecewise linearity of σ and the fact that $\tau_\alpha \leqslant \sigma$.

Suppose $x \in X_E \cap \{\, x \mid dx \geqslant \alpha \,\}$, then $\sigma(\lambda) = \lambda C x$ for some $\lambda \in \Lambda$. Since $dx \geqslant \alpha$, $\lambda C x \leqslant \tau_\alpha(\lambda) \leqslant \sigma(\lambda)$. Therefore $\sigma(\lambda) = \tau_\alpha(\lambda)$.

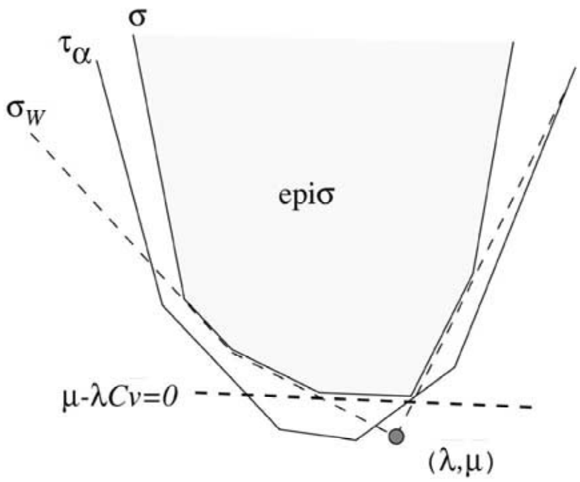

Figure 6. σ and τ_α.

Suppose $\sigma(\lambda) = \tau_\alpha(\lambda)$ at $\lambda \in \Lambda$ and let x be a point that attains $\max\{\lambda Cx \mid x \in X; dx \geqslant \alpha\} = \tau_\alpha(\lambda)$. Then, since $\sigma(\lambda) = \tau_\alpha(\lambda)$, x maximizes λCx over X, meaning $x \in X_E$. □

Now let W be a nonempty subset of $X_E \cap X_V$ and let

$$\sigma_W(\lambda) = \max\{\lambda Cv \mid v \in W\}. \tag{50}$$

Then for any $\lambda \in \Lambda$

$$\sigma_W(\lambda) \leqslant \sigma(\lambda) \tag{51}$$

and we have the following corollary from Theorem 9.3 and the piecewise linearity of $\sigma_W(\lambda)$.

COROLLARY 9.4.
(i) $\tau_\alpha(\lambda) < \sigma_W(\lambda)$ *for any* $\lambda \in \Lambda$*, then* $X_E \cap \{x \mid dx \geqslant \alpha\} = \emptyset$.
(ii) $\tau_\alpha(\lambda) \geqslant \sigma_W(\lambda)$ *for some* $\lambda \in \Lambda$ *if and only if there is a vertex* $(\bar\lambda, \bar\mu)$ *of epi* σ_W *such that* $\bar\mu \leqslant \tau_\alpha(\bar\lambda)$.

This corollary means that we can check whether $\tau_\alpha(\lambda) = \sigma_W(\lambda)$ at some $\lambda \in \Lambda$ by evaluating $\tau_\alpha(\bar\lambda)$ at vertices $(\bar\lambda, \bar\mu)$ of epi σ_W. If $\tau_\alpha(\bar\lambda) < \bar\mu$ for every vertex $(\bar\lambda, \bar\mu)$, we conclude that $\tau_\alpha < \sigma$, and hence $X_E \cap \{x \mid dx \geqslant \alpha\} = \emptyset$ by (i) of Theorem 9.3. Otherwise, i.e., we have found a vertex $(\bar\lambda, \bar\mu)$ with $\tau_\alpha(\bar\lambda) \geqslant \bar\mu$. Two possible cases occur. If $\sigma(\bar\lambda) \leqslant \bar\mu$, implying $\sigma(\bar\lambda) = \bar\mu = \tau_\alpha(\bar\lambda)$, we see that $X_E \cap \{x \mid dx \geqslant \alpha\} \neq \emptyset$ by Theorem 9.3. If $\sigma(\bar\lambda) > \bar\mu$, a vertex $\bar v$ of X that attains $\max\{\bar\lambda Cx \mid x \in X\}$ is not in W. Then W is incremented by this vertex $\bar v$ to make a better underestimation $\sigma_{W\cup\{\bar v\}}$ of σ. (See Fig. 6).

LEMMA 9.5. *The above procedure terminates after a finite number of incrementation of W and shows whether $X_E \cap \{ x \mid dx \geqslant \alpha \}$ is empty or not.*

Proof. Clear from the finiteness of the vertices of X. □

The main technique used in the procedure is generating the vertex set of epi $\sigma_{W\cup\{\bar{v}\}}$ from that of epi σ_W. Note first that epi σ_W is represented by finitely many linear inequalities each of which corresponds to a vertex of W:

$$\operatorname{epi}\sigma_W = \{ (\lambda, \mu) \mid (\lambda, \mu) \in \Lambda \times R;\ \mu - \lambda C v \geqslant 0 \text{ for } v \in W \}. \tag{52}$$

Suppose that we have known the vertex set of epi σ_W, the second case above occurs and we find a vertex $\bar{v}$ of X by maximizing $\bar{\lambda} C x$ over X. This vertex will add an inequality $\mu - \lambda C\bar{v} \geqslant 0$, which cuts off the vertex $(\bar{\lambda}, \bar{\mu})$ of epi σ_W. To generate the vertex set of epi $\sigma_{W\cup\{\bar{v}\}}$ we have only to generate the vertex set of (epi σ_W) $\cap$ $\{ (\lambda, \mu) \mid \mu - \lambda C\bar{v} = 0 \}$. There have been proposed a lot of algorithms for this purpose, e.g., Horst et al. (1988), Chen et al. (1991) and Thieu et al. (1983). See also Section 4.2, Chapter II of Horst and Tuy (1996).

For a given tolerance $\epsilon > 0$ after finitely many bisections we obtain an interval $[\ell_k, u_k]$ such that (P_{u_k}) has no solution while (P_{ℓ_k}) has a solution together with $\bar{\lambda} \in \Lambda$ at which σ coincides with τ_{ℓ_k}. Then solve $\max\{ \bar{\lambda} C x \mid x \in X;\ dx \geqslant \ell_k \}$ to obtain x^*. This is an ϵ-approximate solution of Problem (P_E), i.e., $x^* \in X_E$ and $dx^* \geqslant dx - \epsilon$ for any $x \in X_E$.

Phong and Tuyen (2000) report that an illustrative example of $p = 2, n = 3, m = 4$ required 11 iterations for $\epsilon = 0.1$.

10. Other methods

Benson and Sayin (1994) consider four special cases of linear (P_E), and propose simple linear programming procedures. Benson and Lee (1996) consider (MC) with two criteria and propose an algorithm for maximizing an upper semicontinuous function ϕ. In this case the outcome set Y is of dimension at most two, and Y_E is of dimension at most one, i.e., Y_E consists of edges and vertices.

Thoai (2000) considers the case where $\phi(x) = \varphi(Cx)$ and propose an outer approximation algorithm. He assumes that φ is a quasi-convex function and *nondecreasing* in the sense that $y' \geqslant y$ implies $\varphi(y') \geqslant \varphi(y)$. It is seen that

$$\max\{ \varphi(Cx) \mid x \in X_E \} = \max\{ \varphi(Cx) \mid x \in X \}. \tag{53}$$

His algorithm makes a sequence of polyhedral sets Y^k shrinking to the lower outcome set $Y^{\leqslant}$, solves the relaxation problem $\max\{ \varphi(y) \mid y \in Y_E^k \}$ to find a solution y^k, where Y_E^k is the set of efficient points of Y^k. If $y^k \in Y^{\leqslant}$, any point $x \in X$ with $Cx = y^k$ is an optimal solution of (P_E). Otherwise, it generates a cutting plane defined by the linear equation $\ell^k(y) = 0$ to cut y^k off the set Y^k and reduces Y^k to $Y^{k+1} \cap \{ y \mid y \in R_p;\ \ell^k(y) \leqslant 0 \}$. Since φ is quasi-convex, a vertex of Y^k attains $\max\{\varphi(y) \mid y \in Y_E^k\}$. Thus for solving the relaxation problem he proposes to compute all the vertices of Y^{k+1} from the vertex set of Y^k. The key of the algorithm is the step of checking whether y^k lies in $Y^{\leqslant}$ and generating the cutting plane. Note that $X = \{x \mid x \in R_+^n; Ax = b\}$, then $y^k \in Y^{\leqslant}$ if and only if the system

$$y^k \leqslant Cx; \quad Ax = b; \quad x \geqslant 0 \tag{54}$$

has a solution x. By the linear programming duality theorem this is equivalent to

$$\max\{-\lambda y^k + \mu b \mid -\lambda C + \mu A \leqslant 0; \lambda \geqslant 0\} = 0. \tag{55}$$

When this problem has a positive optimal value, $y^k \notin Y^{\leqslant}$ and further $\ell^k(y) = -\lambda^k y + \mu^k b = 0$ is the desired cutting plane, where (λ^k, μ^k) is an optimal solution of this problem. In Theorem 4.1 of Thoai (2000) the procedure is shown to be finite. Thoai also considers the nonlinear case, namely $\phi(x) = \varphi(c^1(x), \ldots, c^p(x))$, $c^i(x)$'s are concave functions, and also X is a closed convex set defined by nonlinear inequalities. A preliminary experiment for the quadratically constrained problems with quadratic c^i's shows that the most expensive step of the algorithm is the enumeration of vertices, whose number grows rapidly as the number p of criteria increases.

One of the often occured objective functions ϕ is $\phi(x) = -c^i x$, i.e., (P_E) is to minimize the ith objective function $c^i x$ of Cx. To estimate the optimal value of this problem, the process of using the *payoff table* was proposed by several authors. See for example Section 9.13 of Steuer (1985). Consider the linear program

$$\max\{c^j x \mid x \in X\} \tag{56}$$

and let x^j be its optimal solution for $j = 1, \ldots, p$. Then the payoff table is the matrix whose (i, j)-element is $c^i x^j$ (Table 1). The popular way of estimating $\min\{c^i x \mid x \in X_E\}$ is to scan the table and determine the minimum of each column. Notice that this column-wise minimum value gives neither an upper bound nor an lower bound of $\min\{c^i x \mid x \in X_E\}$ because x^j might not be efficient. In order to ensure that x^j is efficient, lexicographical maximization could be employed, i.e., to find x^1 first maximize $c^1 x$ over X and obtain the optimal value z^1, maximize $c^2 x$ over $X \cap \{x \mid c^1 x = z^1\}$, and maximize $c^3 x$ over $X \cap \{x \mid c^1 x = z^1; c^2 x = z^2\}$

Table 1. Payoff table.

	1	2	$\cdots$	p
1	$c^1 x^1$	$c^1 x^2$	$\cdots$	$c^1 x^p$
		$\cdots$		
p	$c^p x^1$	$c^p x^2$	$\cdots$	$c^p x^p$

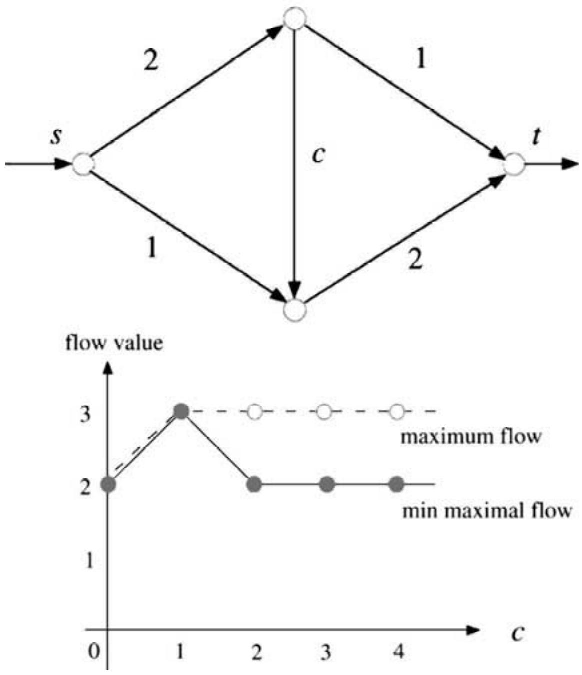

Figure 7. Maximum flow versus minimum maximal flow.

and so on. Then each column-wise minimum of the payoff table thus obtained gives an upper bound of $\min\{c^i x \mid x \in X_E\}$. In Isermann and Steuer (1987) and Reeves and Reid (1988) it is reported how a good approximation is obtained from the payoff table based on the computational experience of randomly generated problems.

11. Conclusion

Most of the algorithms reviewed in this paper anticipate a small number of criteria of Problem (MC) and convert Problem (P_E) to a global optimization problem in p or so variables. However, there are interesting and important problems that do not enjoy the low dimensionality of p. An example is the minimum maximal flow problem that has a close relation with the uncontrolable flow problem raised by Iri (1994, 1996). Let $(V, s, t, E, \partial^+, \partial^-, c)$ denote a network with node set V, arc set E, source node s, sink node t, incidence functions ∂^+ and ∂^-, and a nonnegative capacity c_e for each arc e. A vector $x \in R^{|E|}$ is said to be a *feasible flow* if it satisfies the conservation equations and capacity constraints:

$$\sum_{\partial^+ e = v} x_e = \sum_{\partial^- e = v} x_e \text{ for all } v \in V \tag{57}$$

$$0 \leqslant x_e \leqslant c_e \text{ for all } e \in E. \tag{58}$$

A feasible flow x is said to be a *maximal feasible flow* if there is no feasible flow x' such that $x' \geqslant x$ and $x' \neq x$. The flow value, denoted by $v(x)$, of flow x is

$$v(x) = \sum_{\partial^+ e = s} x_e - \sum_{\partial^- e = s} x_e. \tag{59}$$

The problem is to find a maximal flow with the minimum flow value. Note that this problem embraces the minimum maximal matching problem, which is known to be NP-hard, (e.g. Garay and Johnson, 1979). An example of Iri (1996) is shown in Figure 7. The maximum flow value grows as the arc capacity c increases, while the minimum maximal flow value does not.

Let (MC) be defined for $C = I$, the indentity matrix of dimension $|E|$, and the set of feasible flows X, and let $\phi(x) = -v(x)$. Then the minimum maximal flow problem reduces to Problem (P_E). Problem (MC) has the objective functions as many as the variables, the algorithms that exploit the low dimensionality of p would not work efficiently. The algorithm based on the outer approximation in Shi and Yamamoto (1997) is not satisfactory. Further research is needed.

Even when p is small, few algorithms are yet tried and tested, and we hardly derive any conclusion about the efficiency of the algorithms. Organized computational experiment should be carried out.

Acknowledgements

This research is supported in part by the Grant-in-Aid C10680419 for Scientific Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Al-Khayyal, F.A. and Falk, J.E. (1983), Jointly constrained biconvex programming. *Mathematics of Operations Research* 8, 273–286.

An, L.T.H., Tao, P.D. and Muu, L.D. (1996), Numerical solution for optimization over the efficient set by d.c. optimization algorithms. *Operations Research Letters* 19, 117–128.

Atteia, M. and El Qortobi, Quasi-convex duality. In *Optimization and Optimal Control*, Lecture Notes in Control and Information Science 35, Springer, Berlin, 1981, pp. 3–8.

Benson, H.P. (1991), An all-linear programming relaxation algorithm for optimizing over the efficient set. *Journal of Global Optimization* 1, 83–104.

Benson, H.P. (1992), A finite nonadjacent extreme-point search algorithm for optimization over the efficient set. *Journal of Optimization Theory and Applications* 73, 47–64.

Benson, H.P. (1995), A geometric analysis of the efficient outcome set in multiple objective convex program with linear criteria functions. *Journal of Global Optimization* 6, 213–251.

Benson, H.P. and Lee, D. (1996), Outcome-based algorithm for optimizing over the efficient set of a bicriteria linear programming. *Journal of Optimization Theory and Applications* 88, 77–105.

Benson, H.P. and Sayin, S. (1994), Optimization over the efficient set: four special case. *Journal of Optimization Theory and Applications* 80, 3–18.

Bolintineanu, S. (1993), Minimization of a quasi-concave function over an efficient set. *Mathematical Programming* 61, 89–110.

Chen, P.C., Hansen, P. and Jaumard, B. (1991), On-line and off-line vertex enumeration by adjacency lists. *Operations Research Letters* 10, 403–409.

Dauer, J.P. and Fosnaugh, T.A. (1995), Optimization over the efficient set. *Journal of Global Optimization* 7, 261–277.

Ecker, J.G. and Song, J.H. (1994), Optimizing a linear function over an efficient set. *Journal of Optimization Theory and Applications* 83, 541–563.

Fülöp, J. (1994), A cutting plane algorithm for linear optimization over the efficient set. In S. Komlösi, T. Rapcsàk and S. Shaible (eds.), *Generalized Convexity*, Lecture Notes in Economics and Mathematical Systems 405, Springer, Berlin, pp. 374–385.

Garay, M.R. and Johnson, D.S. (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco.

Horst, R., de Vries, J. and Thoai, N.V. (1988), On finding new vertices and redundant constraints in cutting plane algorithms for global optimization. *Operations Research Letters* 7, 85–90.

Horst, R. and Thoai, N.V. (1999), Maximizing a concave function over the efficient or weakly-efficient set. *European Journal of Operational Research* 117, 239–252.

Horst, R. and Tuy, H. (1996), *Global Optimization: Deterministic Approach*. Springer-Verlag, Berlin.

Iri, M. (1994), An essay in the theory of uncontrollable flows and congestion. Technical Report, Department of Information and System Engineering, Faculty of Science and Engineering, Chuo University, TRISE 94-03.

Iri, M. (1996), Network flow-theory and applications with practical impact. In J. Doležal and J. Fidler (eds.), *System Modelling and Optimization*, Chapman & Hall, London, 1996, pp. 24-36.

Isermann, H. and Steuer, E. (1987), Computational experience concerning payoff tables and minimum criterion values over the efficient set. *European Journal of Operational Research* 33, 91–97.

Konno, H., Thach, P.T. and Tuy, H. (1997), *Optimization on Low Rank Nonconvex Structures*. Kluwer, Dordrecht.

Muu, L.D. (2000), A convex-concave programming method for optimizing over the efficient set. *Acta Mathematica Vietnamica* 25(1), 67–85.

Naccache, P.H. (1978), Connectedness of the set of nondominated outcomes in multicriteria optimization. *Journal of Optimization Theory and Applications* 25, 459–467.

Philip, J. (1972), Algorithms for the vector maximization problem. *Mathematical Programming* 2, 207–229.

Phong, T.Q. and Tuyen, J.Q. (2000), Bisection search algorithm for optimizing over the efficient set. to appear in *Vietman Journal of Mathematics*.

Reeves, G.R. and Reid, C. (1988), Minimum values over the efficient set in multiple objective decision making. *European Journal of Operational Research* 36, 334–338.

Rockafellar, R.T. and Wets, R.J-B. (1998), *Variational Analysis*. Springer, Berlin.

Sayin, S. (1996), An algorithm based on facial decomposition for finding the efficient set in multiple objective linear programming. *Operations Research Letters* 19, 87–92.

Sayin, S. (2000), Optimizing over the efficient set using a top-down search of faces. *Operations Research* 48, 65–72.

Sawaragi, Y., Nakayama, H. and Tanino, T. (1985), *Theory of Multiobjective Optimization*. Academic Press, Orland.

Shi, J.M. and Yamamoto, Y. (1997), A global optimization method for minimum maximal flow problem. *Acta Mathematica Vietnamica* 22, 271-287.

Steuer, R.E. (1985), *Multiple Criteria Optimization: Theory, Computation and Application*. Wiley, New York.

Steuer, R.E. (1995), Manual for the ADBASE multiple objective linear programming package. Department of Management Science and Information Technology, University of Georgia.

Stoer, J. and Witzgall, C. (1970), *Convexity and Optimization in Finite Dimensions I*. Springer-Verlag, Berlin.

Thach, P.T. (1991), Quasiconjugates of functions, duality relationship between quasiconvex minimization under a reverse convex constraints and quasiconvex maximization under a cnvex constraint, and applications. *Journal of Mathematical Analysis and Applications* 159, 299–322.

Thach, P.T., (1993), Global optimality criterion and a duality with a zero gap in nonconvex optimization. *SIAM Journal on Mathematical Analysis* 24, 1537–1556.

Thach, P.T. (1994), A nonconvex duality with zero gap and applications. *SIAM Journal on Optimization* 4, 44–64.

Thach, P.T., Konno, H. and Yokota, D. (1996), Dual approach to nminimization on the set of pareto-optimal solutions. *Journal of Optimization Theory and Applications* 88, 689–707.

Thieu, T.V., Tam, B.T. and Ban, V.T. (1983), An outer-approximation method for globally minimizing a concave function over a compact convex set. *Acta Mathematica Vietnamica* 8, 21–40.

Thoai, N.V. (2000), A class of optimization problems over the efficient set of a multiple criteria nonlinear programming problem. *European Journal of Operational Research* 122, 58–68.

Thoai, N.V. (2001), Conical algorithm in global optimization for optimizing over efficient sets *Journal of Global Optimization*, 18, 321–336.

Tuy, H. (1998), *Convex Analysis and Global Optimization*. Kluwer, Dordrecht.

White, D.J. (1982), *Optimality and Efficiency*. John Wiley & Sons, Chichester.

White, D.J. (1996), The maximization of a function over the efficient set via a penalty function approach, *European Journal of Operational Research* 94, 143–153.

Yamada, S., Tanino, T. and Inuiguchi, M. (2000), An inner approximation method for optimization over the weakly efficient set. *Journal of Global Optimization* 16, 197–217.