The Complementary Convex Structure in Global Optimization

HOANG TUY
Institute of Mathematics, Hanoi.

(Received: 5 July 1991; accepted: 8 October 1991)

Abstract. We show the importance of exploiting the complementary convex structure for efficiently solving a wide class of specially structured nonconvex global optimization problems. Roughly speaking, a specific feature of these problems is that their nonconvex nucleus can be transformed into a complementary convex structure which can then be shifted to a subspace of much lower dimension than the original underlying space. This approach leads to quite efficient algorithms for many problems of practical interest, including linear and convex multiplicative programming problems, concave minimization problems with few nonlinear variables, bilevel linear optimization problems, etc. . .

Key words. Complementary convex structure, Generalized Rank k Property, global optimization.

1. Introduction

In the deterministic approach to global optimization a question of fundamental importance is to exploit the mathematical structure of the problem under study. Therefore, it seems natural to ask: which kind of structure is most favourable?

It is common knowledge that convex analysis has been one of the cornerstones of (local) optimization theory in the past four decades. In global optimization, although the convex structure will certainly still play a major role (at least because in most nonconvex problems convexity is present in a certain limited sense), there are sound reasons to believe that the complementary convex structure will be predominant.

In fact, the space R^n in which optimization problems are formulated is such that any point of it has a base of convex neighbourhoods. On the basis of this local convex structure of the space it can be shown that every closed set in R^n is the projection of a set in R^{n+1} which is the difference of two convex sets, i.e., a complementary convex set (see [1] and appendix). Therefore, any continuous optimization problem in finite-dimensional space can be restated in the following general form

$$\text{minimize } l(x) \quad \text{subject to} \quad x \in D \setminus C\,, \tag{1}$$

where $l(x)$ is a linear function and C, D are convex sets.

On the other hand, a conspicuous limitation of conventional local optimization methods is their ability of being trapped at a local minimum (or even a stationary point). Therefore, the core of a global optimization method is to deal with the

Journal of Global Optimization **2**: 21–40, 1992.

question of how to *transcend stationarity*, i.e., how to recognize that a given stationary point is actually a global minimizer and if it is not, how to proceed to a better feasible point. In a recent paper by Tuy and Horst [2] it has been shown that for a large variety of nonconvex global optimization problems the question of transcending stationarity can always be reduced to solving a particular problem of the following type:

> Given two convex sets C and D in R^n, find a point $x \in D \setminus C$ or else establish that $D \subset C$ (i.e., $D \setminus C$ is empty). (*Geometric Complementarity Problem (GCP)*.)

Thus, a complementary convex structure underlies every global optimization problem of a very wide class. Perhaps this structure is not always apparent and it is the hard work of the problem solver to disclose this structure and to bring it into a form amenable to computational analysis. In many cases of interest, however, this work turns out to be far rewarding:

EXAMPLE 1. The problem of minimizing the product $(c^1x + d_1)(c^2x + d_2)$ of two affine functions over a polytope $M \subset \{x: c^ix + d_i > 0 \ (i = 1 = 1, 2)\}$ appears in certain applications in microeconomics, VLSI chip design, bond portfolio optimization, etc. . . [3–6]. Various solution methods have been proposed in the literature for this problem [7–10]. It can be shown that the underlying GCP here involves a convex set C whose recession cone contains the cone $\{x: c^ix \geqslant 0 \ (i = 1, 2)\}$. Based on this specific property, an algorithm (even simpler than that of Konno–Kuno in [10]) has been devised in [11] that merely consists of solving the linear parametric program in usual form:

$$\text{minimize}\langle c^1 + \alpha(c^2 - c^1), x\rangle \quad \text{s.t.} \quad x \in M \quad (\alpha \in [0, 1]) \tag{2}$$

to obtain the sequence of breaking values $\alpha_0 = 0 < \alpha_1 < \cdots < \alpha_N = 1$ along with, for each interval $[\alpha_{j-1}, \alpha_j](j = 1, \ldots, N-1)$, a basic solution x^j optimal for all $\alpha \in [\alpha_{j-1}, \alpha_j]$. Then a global optimal solution of the problem is x^{j^*}, where

$$j^* \in \text{argmin}\{(c^1x^j + d_1)(c^2x^j + d_2): j = 1, \ldots, N-1\}. \tag{3}$$

EXAMPLE 2. A class of location problems [12] can be formulated as maximizing a function $\Sigma_{j=1}^p q_j(h_j(x))$ over $x \in R^n$, where each $q_j: R^n_+ \to R_+$ is a convex decreasing function such that $\lim q_j(t) = 0 \ (t \to \infty)$ and each $h_j: R^n \to R_+$ is a convex function such that $\lim h_j(x) = +\infty \ (|x| \to \infty)$. Since the objective function is neither convex nor concave, a method was developed in [12] that only computes a local maximizer. However, here again the underlying GCP has a special structure, so that, as shown in [13], the problem can be solved through a sequence of unconstrained *convex minimization* subproblems of the form

$$\text{minimize} \sum_{i=1}^{p} s_j h_j(x) \quad \text{over} \quad x \in R^n \quad (s_j \geqslant 0, j = 1, \ldots, p). \tag{4}$$

EXAMPLE 3. Certain hierarchical decision-making situations can be modelled by the following two level optimization problem (Stackelberg game):

$$\underset{x}{\text{minimize}}\, c^1x + d^1 \quad \text{s.t.} \quad A_1x + B_1y \leq g_1\,, \quad x \in R^p_+$$

$$\text{where } y \text{ solves } \min\{d^2y\colon A_2x + B_2y \leq g_2,\ y \in R^q_+\}\,. \tag{5}$$

Although the model involves only linear functions, this is a difficult nonconvex problem, fraught with pitfalls (see [14]), for which so far developed algorithms have not always been successful. However, it appears that the complementary convex structure of this problem can be disclosed and analyzed. As it turns out [15] the recession cone of the set C in the corresponding GCP contains the cone $\{(x, y)\colon A_2x \leq 0, d^2y \geq 0\}$. Due to this property, if $\text{rank}(A_2) = k$ ($k \leq p$) then the original problem in R^{p+q} can be converted into a problem in R^{k+1}, i.e., usually of much smaller dimension. In [15] a new solution method along this line has been developed that is better founded and more efficient than existing ones, at least when k is small compared to $p + q$.

These examples demonstrate the importance of exploiting the complementary convex structure in deterministic global optimization.

A common feature of all these examples is the possibility of associating with the given problem a GCP having some special structure that allows the problem to be reduced to a form more amenable to efficient solution methods.

The purpose of the present paper is to generalize the situation encountered in the above examples and to develop a framework for exploiting the complementary convex structure in a wide class of global optimization problems.

The paper consists of 5 sections. In Section 2 we will introduce a general property (called "Generalized Rank k Property") that allows a given GCP to be converted into a problem of smaller dimension. In Section 3 we will present a general method for handling GCP's with the Generalized Rank k Property. In Section 4 we will discuss the convergence of the method and some other issues. Finally, Section 5 is devoted to some applications.

2. Generalized Rank *k* Property

One of the most successful ideas in convex optimization is the dualization of various concepts (polar sets, conjugate functions, dual programs, Lagrange multipliers, etc. . .). By dualization we can sometimes simplify or clarify a given situation, either by replacing a constrained problem with an unconstrained one, or by reducing the dimension of the problem, or by incorporating certain constraints into the objective to make the problem easier, and so on.

In combinatorial optimization as well as in other fields of mathematics, the idea of dualization has proven very fruitful, too. Therefore, no wonder that this idea could play a major role in providing insight into the complementary convex structure in global optimization. This seems the more natural because in the complementary convex structure convexity is present, either partially or in the other (opposite) way.

A number of research works have already been done in this direction, though sometimes incidentally or indirectly and dealing only with some particular situations [11, 14–23]. In the sequel we shall discuss a general transformation scheme based on dualization which may lead to quite efficient solution methods for a wide class of originally difficult problems.

Consider the global optimization problem:

$$(P) \qquad \text{minimize } f(x) \quad \text{subject to} \quad x \in G,$$

where G is a closed subset of R^n and $f(x)$ is a continuous function defined on some open neighbourhood Ω of G. Let $\bar{x}$ be the best feasible solution known so far ($\bar{x}$ may or may not be a stationary point obtained via a local optimization procedure). Denote $F = \{x \in \Omega : f(x) \geq f(\bar{x})\}$. Then in its primitive form the question of transcending the incumbent $\bar{x}$ amounts to solving the following subproblem

$$\text{Find a point } x \in G \backslash F \text{ (or else establish that } G \subset F) \tag{6}$$

In [2] Tuy and Horst have shown that the latter subproblem can always be reduced to a GCP as formulated in the Introduction. We are now interested in which conditions to impose on the structure of the sets F, G in order that the subproblem (6) can be transformed into a GCP of much smaller dimension than the original problem, and thereby efficiently solvable by currently available methods.

For convenience we shall refer to a problem like (6) as a *(G, F)-problem.*

DEFINITION 1. We shall say that the (G, F)-problem (6), where F is closed, has the *Generalized Rank k Property* if the following conditions hold:

(i) There exist a mapping $\varphi : R^n \to R^p$ $(p \leq n)$ together with a closed convex set $\tilde{C}$ in R^p such that $x \in F$ if and only if $\varphi(x) \in \tilde{C}$;

(ii) $\tilde{C}$ has the Rank k Property $(k \leq p)$ in the sense that its recession cone contains a polyhedral cone

$$K = \{t \in R^p : At \leq 0\} \tag{7}$$

where A is a $q \times p$ matrix with rank $A = k$;

(iii) For every $u \in R^q_+$ the problem

$$\text{maximize } \langle A^T u, \varphi(x) \rangle \quad \text{subject to} \quad x \in G \tag{8}$$

can be solved by some known efficient algorithm.

REMARK 1. When $p = n$ and φ is the identity mapping the above property reduces to the Rank k Property as was introduced in [11] and earlier studied in [16] (note that rank $A = \dim K^*$, where K^* is the polar of K).

REMARK 2. In each example given in the Introduction the problem of transcending an incumbent $\bar{x}$ as formulated above has the Generalized Rank k Property. Indeed, in Example 1, $k = 2$ (actually the Rank 2 Property holds), $G = M$, $\varphi =$ identity mapping, $\tilde{C} = F$, $K = \{t \in R^n : c^i t \geq 0 \ (i = 1, 2)\}$.

In Example 2, $k = p$, $\varphi(x) = (h_1(x), \ldots, h_p(x))$, $\tilde{C} = \{t \in R^p_+ : \sum_{j=1}^p q_j(t_j) \geq \sum_{j=1}^p q_j(h_j(\bar{x}))\}$, $K = R^p_+$ (i.e., $A = -I$, $I =$ identity matrix). Problem (8) here has the form (4). Example 3 will be discussed later (see Section 5).

Let us mention two more examples among a host of others that can easily be constructed.

EXAMPLE 4. Consider the problem

$$\text{minimize } f(x) := \prod_{j=1}^{p} f_j(x) \quad \text{subject to} \quad x \in M \tag{9}$$

where $f_j : \Omega \rightarrow R_+$ ($j = 1, \ldots, p$) are positive-valued convex functions on an open convex set Ω, while M is a convex subset of Ω. Clearly the function $f(x)$ is neither convex nor concave. For $p = 2$ this problem, which appears in several applications (see [3–6]), has been investigated by Kuno and Konno [24].

As above, let $F = \{x \in \Omega : f(x) \geq f(\bar{x})\}$, $G = M$ (i.e., consider the problem of transcending an incumbent $\bar{x}$). If we take $\varphi(x) = (f_1(x), \ldots, f_p(x))$ and

$$\tilde{C} = \{t \in R^p_+ : \prod_{j=1}^{p} t_j \geq f(\bar{x})\}, \tag{10}$$

then since the function $t \rightarrow \prod_{j=1}^p t_j$ is quasiconcave on R^p_+, $\tilde{C}$ is a closed convex set and $x \in F$ if and only if $\varphi(x) \in \tilde{C}$. Furthermore, from (10) we have

$$R^p_+ = \{t \in R^p : t \geq 0\} \subset \tilde{C},$$

so that $\tilde{C}$ has the Rank p Property with $K = R^p_+$ ($A = -I$, $q = p$). Finally, for any $u \in R^p_+$ the problem (8) which here reads as

$$\text{maximize} - \sum_{j=1}^{p} u_j f_j(x) \quad \text{subject to} \quad x \in M \tag{11}$$

is an ordinary convex programming problem, hence can be solved by efficient standard algorithms. Therefore, the Generalized Rank p Property holds.

EXAMPLE 5. Consider the problem

$$\text{minimize } f(x) := \sum_{j=1}^{p} \log f_j(x) + \min\{f_1(x), \ldots, f_p(x)\} \quad \text{s.t.} \quad x \in M, \tag{12}$$

where f_i and M are as in Example 4.

If we take $\varphi(x) = (f_1(x), \ldots, f_p(x))$, and

$$\tilde{C} = \left\{ t \in R^p_+ : \sum_{j=1}^{p} \log t_j + \min\{t_1, \ldots, t_p\} \geq f(\bar{x}) \right\},$$

then, since the function $t \to \Sigma_{j=1}^p \log t_j + \min\{t_1, \ldots, t_p\}$ is concave on R^p_+, the set $\tilde{C}$ is closed and convex and we have $f(x) \geq f(\bar{x})$ if and only if $\varphi(x) \in \tilde{C}$. Again $\tilde{C}$ has the Rank p Property (with $K = R^p_+$) and for any $u \in R^p_+$ the problem (8) has the same form as (11). Therefore, the Generalized Rank p Property holds for transcending an incumbent $\bar{x}$ in this problem.

Thus, for many interesting problems arising from applications, the basic question of transcending stationarity, when formulated as a (G, F)-problem, has the Generalized Rank k Property. In the next section we proceed to show how, due to this property, each of these problems can be transformed into an easier problem in dimension k which can then be solved through a sequence of subproblems (8).

3. How to Solve Problems with Generalized Rank k Property

Let $\bar{x} \in G$ be the best feasible solution so far known for problem (P) and let $F = \{x \in \Omega : f(x) \geq f(\bar{x})\}$ (recall that Ω is an open neighbourhood of G over which $f(x)$ is defined). Let us examine how to solve the corresponding (G, F)-problem, assuming the Generalized Rank k Property stated above.

Define the set

$$\tilde{D} = \varphi(G) = \{t \in R^p : t = \varphi(x) \text{ for some } x \in G\} . \tag{13}$$

PROPOSITION 1. *We have $G \subset F$ if and only if $\tilde{D} \subset \tilde{C}$.*

Proof. If $G \subset F$ then for any $t \in \tilde{D}$, since $t = \varphi(x)$ for some $x \in G \subset F$, it follows that $t \in \tilde{C}$, i.e., $\tilde{D} \subset \tilde{C}$. Conversely, if $\tilde{D} = \tilde{C}$ then for any $x \in G$, since $t \in \varphi(x) \in \tilde{D}$, it follows that $t \in \tilde{C}$, hence $x \in F$, i.e., $G \subset F$.

Clearly, whenever $t \in \tilde{D} \backslash \tilde{C}$ then $t = \varphi(x)$ with $x \in G \backslash F$. Therefore, by Proposition 1 the (G, F)-problem is equivalent to the following problem in R^p:

$(\tilde{D}, \tilde{C})$ Find a point $t \in \tilde{D} \backslash \tilde{C}$ (or else establish that $\tilde{D} \subset \tilde{C}$) .

This is almost a GCP except that the set $\tilde{D}$ needs not be convex. However, thanks to condition (iii) in the Generalized Rank k Property (namely, every problem of the form (8) can be efficiently solved), we shall see that this $(\tilde{D}, \tilde{C})$-problem can be solved by the polyhedral annexation method earlier developed in [16] (see also [23] and [19]).

Assume that we know a point $t^0 \in \tilde{D} \cap \operatorname{int} \tilde{C}$, e.g., $t^0 = \varphi(x^0)$ where $x^0 \in G \cap \operatorname{int} F$, i.e., x^0 is a feasible point with $f(x^0) > f(\bar{x})$ (we shall later comment on this assumption). Now define

$$C = \tilde{C} - t^0 , \quad D = \tilde{D} - t^0 . \tag{14}$$

Then

$$0 \in D \cap \operatorname{int} C . \tag{15}$$

Let C^*, D^* denote the polars of C, D respectively, i.e.,

$$C^* = \{v \in R^p : \langle v, t\rangle \leq 1 \quad \forall t \in C\} ,$$
$$D^* = \{v \in R^p : \langle v, t\rangle \leq 1 \quad \forall t \in D\} . \tag{16}$$

Observe that if $C^* \subset D^*$ then $(C^*)^* \supset (D^*)^*$, hence $D \subset C$ because always $(D^*)^* \supset D$, while $(C^*)^* = C$ by the convexity and closedness of C. Furthermore, if a point $v \in C^* \backslash D^*$, then $\langle v, s\rangle > 1$ for at least some $s \in D$, while $\langle v, t\rangle \leq 1$ for all $t \in C$, hence $s \in D \backslash C$. Therefore, instead of solving the $(\tilde{D}, \tilde{C})$-problem we can solve the "dual" (C^*, D^*)-problem

(C^*, D^*) Find a point $v \in C^* \backslash D^*$ or else establish that $C^* \subset D^*$.

From condition (ii) in the Generalized Rank k Property we have $t^0 + K \subset \tilde{C}$, i.e., $K \subset C$, hence

$$C^* \subset K^* = \{v : v = A^T u \text{ for some } u \in R^q_+\} \tag{17}$$

and since rank $A = k$ it follows that dim $K^* = k$ (see, e.g., [26]). Specifically, if $a^1, \ldots, a^k$ are k linearly independent rows of A, then K^*, and hence also C^*, is contained in the subspace L of R^p spanned by $a^1, \ldots, a^k$. We see that (C^*, D^*) is actually a problem in L with dim $L = k$.

Now define for each $v \in K^*$:

$$g(v) = \sup\{\langle v, t\rangle : t \in D\} . \tag{18}$$

Obviously $g(v)$ is a convex function on K^*. Since $v = A^T u$ for some $u \in R^q_+$ we have:

$$\begin{aligned} g(v) &= \sup\{\langle v, t\rangle : t \in \tilde{D} - t^0\} = \sup\{\langle v, t\rangle : t = \varphi(x) - \varphi(x^0) , \quad x \in G\} \\ &= \sup\{\langle A^T u, t\rangle : t = \varphi(x) - \varphi(x^0) , \quad x \in G\} \\ &= \sup\{\langle A^T u, \varphi(x)\rangle : x \in G\} - \langle A^T u, \varphi(x^0)\rangle , \end{aligned} \tag{19}$$

so that computing the value of $g(v)$ amounts to solving a problem of the form (8) (for which efficient algorithm is assumed to exist). On the other hand, since $D^* = \{v : g(v) \leq 1\}$, the (C^*, D^*)-problem now reduces to checking whether $g(v) \leq 1$ for all $v \in C^*$, i.e., whether

$$\max\{g(v) : v \in C^*\} \leq 1 \, ? \tag{20}$$

This can be done as follows. First we note that since $0 \in \operatorname{int} C$ by (15) it follows from a well known property of polar sets (see, e.g. [26]) that C^* is compact. Let S_1 be a polytope in L such that

(1) $C^* \subset S_1 \subset K^*$; (21)

(2) the vertex set V_1 of S_1 is simple (i.e., is small and readily available).

(We shall comment later on the construction of such a polytope). Since $g(v)$ is convex, if $g(v) \leq 1$ for all $v \in V_1$ then $\max\{g(v): v \in S_1\} \leq 1$ and (20) will *a fortiori* hold. Therefore, we compute $v^1 \in \text{argmax}\{g(v): v \in V_1\}$. If

$$g(v^1) \leq 1 ,$$

then we are done: (20) holds, hence $C^* \subset D^*$ and so $\bar{x}$ is a global minimizer. Otherwise,

$$g(v^1) > 1 ,$$

then $v^1 \not\in D^*$. Since $v^1 \in K^*$, we have $v^1 = A^T u^1$ for some $u^1 \in R^q_+$ and referring to (19) let $x^1 \in G$ be such that

$$g(v^1) = \langle A^T u^1 , \quad \varphi(x^1) - \varphi(x^0)\rangle .$$

Then the point $t^1 = \varphi(x^1) - \varphi(x^0) = \varphi(x^1) - t^0$ satisfies

$$\langle t^1, v^1 \rangle = g(v^1) > 1 . \tag{22}$$

There are two possibilities:

(a) If $x^1 \not\in F$ then a point x^1 has been found such that $x^1 \in G \setminus F$, i.e., x^1 is better than $\bar{x}$((G, F)-problem is solved);

(b) If $x^1 \in F$, then from the definition of $\tilde{C}$, $\varphi(x^1) \in \tilde{C}$, hence $t^1 \in C$ and consequently $\langle t^1, v \rangle \leq 1$ for all $v \in C^*$.

Remembering (22) we see that in the latter case the cut

$$\langle t^1, v \rangle \leq 1$$

will eliminate v^1 without excluding any point of C^*. So the polytope

$$S_2 = S_1 \cap \{v: < t^1, v \leq 1\}$$

will still satisfy the same condition as (21) for S_1, i.e., $C^* \subset S_2 \subset K^*$. On the other hand, since it obtains from S_1 by just an additional linear constraint, its vertex set V_2 can be derived from their vertex set V_1 of S_1 by currently known methods (see, e.g., [27], or [23]). Therefore, we can repeat the procedure with S_2 in place of S_1. In this manner a nested sequence of polytope will be built

$$K^* \supset S_1 \supset S_2 \supset \cdots \supset S_h \supset \cdots \supset C^* \tag{23}$$

such that: if the sequence terminates at some S_h then either $\bar{x}$ is a global minimizer (when $S_h \subset D$, i.e., $g(v^h) \leq 1$) or a solution x^h better than $\bar{x}$ is obtained. In the next section we shall prove that the sequence can be infinite only when $\bar{x}$ is itself a global minimizer.

Note that each S_h can be viewed as the polar of some polytope P_h in R^p such that $K \subset P_h \subset C$ and $P_1 \subset P_2 \subset \cdots \subset P_h \subset \cdots$ Therefore, the method amounts to constructing a sequence of expanding polytopes in the hope that eventually some

P_h will cover all of D. This explains the name "Polyhedral annexation" given to this kind of procedure (see [16] for an alternative presentation).

Thus, given a feasible solution $\bar{x}^1$, we can solve the corresponding (G, F)-problem to check whether this is actually a global minimizer and if this is not, to find a better feasible solution x^2. In the latter case we can go to another (G, F)-problem corresponding to some $\bar{x}^2$ equal to x^2 or to a stationary point computed from x^2 by local inexpensive methods. And so on.

Of course, this scheme can be carried out only if the Generalized Rank k Property holds for every (G, F)-problem corresponding to successive current best solutions $\bar{x}^h$ $(h = 1, 2, \ldots)$. Fortunately, in many cases the problem (P) has the following property.

DEFINITION 2. We say that problem (P) has the *Generalized Rank k Property* if the following conditions hold:

(i) There exist a mapping $\varphi: R^n \to R^p$ $(p \leqslant n)$ and for each $\gamma \in f(G)$ a closed convex set $\tilde{C}_\gamma$ in R^p such that $\tilde{C}_\gamma \supset \tilde{C}'_\gamma$ for $\gamma < \gamma'$ and $f(x) \geqslant \gamma$ if and only if $\varphi(x) \in \tilde{C}_\gamma$;

(ii) For every γ the recession cone of $\tilde{C}_\gamma$ contains the cone

$$K = \{t \in R^p : At \leqslant 0\} ,$$

where A is a $q \times p$ matrix with rank $A = k$ $(k \leqslant p)$;

(iii) For every $u \in R^q_+$ the problem

$$\text{maximize}\langle A^T u, \varphi(x)\rangle \quad \text{subject to} \quad x \in G \tag{24}$$

can be solved by some known efficient algorithm.

Most often condition (i) holds in the following form:

(i′) There exists a mapping $\varphi: R^n \to R^p$ along with a continuous quasi-concave function ψ defined on some convex set $T \subset R^p$ containing $\varphi(G)$ and such that

$$f(x) = \psi(\varphi(x)) .$$

Indeed, it is easily verified that (i′) implies (i) with $\tilde{C}_\gamma = \{t \in T : \psi(t) \geqslant \gamma\}$. Also it is easily seen that each problem given in examples 1, 2, 4, 5 above has the Generalized Rank k Property, with (i′) holding.

When problem (P) itself has the Generalized Rank k Property as stated in Definition 2, then the sequence of (G, F)-problems corresponding to successive current best $\bar{x}^1$ $\bar{x}^2, \ldots$ can be incorporated into a unified process as follows.

ALGORITHM 1 (for solving (P))

Initialization.

Let $\bar{x}^1$ be the best feasible solution available. Take a feasible solution x^0 such that $f(x^0) > f(\bar{x}^1)$. Set $\gamma_1 = f(\bar{x}^1)$. Construct an initial polytope S_1 with a known vertex set V_1 and such that $C^*_{\gamma_1} \subset S_1 \subset K^*$. $(C_\gamma = \tilde{C}_\gamma - \varphi(x^0))$. Set $\tilde{V}_1 = V_1$, $h = 1$.

Iteration $h = 1, 2, \ldots$

Step 1. For each $v \in \tilde{V}_h$ solve the subproblem

$$SP(v) \qquad \max\{\langle v, \varphi(x)\rangle : x \in G\} \tag{25}$$

to obtain its optimal solution $x(v)$ and the value $g(v) = \langle v, \varphi(x(v)) - \varphi(x^0)\rangle$.

Step 2. Compute $v^h \in \operatorname{argmax}\{g(v) : v \in V_h\}$. If $g(v^h) \leq 1$ then terminate: $\bar{x}^h$ is a global minimizer. Otherwise,

Step 3. Let $x^h = x(v^h)$. If $f(x^h) < \gamma_h$ then set $\bar{x}^{h+1} = x^h$, $\gamma_{h+1} = f(x^h)$. Otherwise, set $\bar{x}^{h+1} = \bar{x}^h$, $\gamma_{h+1} = \gamma_h$. To S_h adjoin the new constraint

$$l_h(v) := \langle \varphi(x^h) - \varphi(x^0), v\rangle \leq 1 \tag{26}$$

to define S_{h+1}.

Step 4. Compute the vertex set V_{h+1} of S_{h+1} (from knowledge of V_h). Set $\tilde{V}_{h+1} = V_{h+1} \setminus V_h$. Go to iteration $h + 1$.

4. Convergence and Other Issues

(I) CONVERGENCE

THEOREM. *The above Algorithm either terminates after finitely many steps, yielding a global minimizer, or it generates an infinite sequence* $\{x^h\}$. *In the latter case, any cluster point of the sequence* $\{\bar{x}^h\}$ *is a global minimizer (in particular, if* $\bar{x}^h$ *remains unchanged for all* $h \geq h_0$ *then* $\bar{x}^{h^0}$ *is a global minimizer).*

Proof. We only sketch the proof, since it is much similar to the convergence proof in [13]. First it is easily checked that the coefficient vector of the cut (26), i.e., the vector $\varphi(x^h) - \varphi(x^0)$, is a subgradient of the convex function $g(v)$ at a point v^h, i.e., $\varphi(x^h) - \varphi(x^0) \in \partial g(v^h)$. Hence, applying a theorem on convergence of outer approximation methods (see [23]) to the set $D^* = \{v : g(v) \leq 1\}$, the sequence $\{v^h\}$ and the cuts (26), we conclude that every cluster point of the sequence $\{v^h\}$ must belong to D^*. This implies $g(v^h) \downarrow 1$. Now let $\hat{x} = \lim \bar{x}^{h_\nu}$ so that $f(\hat{x}) = \lim_{\nu\to\infty} \gamma_{h_\nu}$. Since for every h:

$$\max\{g(v) : v \in C_h^*\} \leq \max\{g(v) : v \in S_h\} = g(v^h) \downarrow 1$$

(here C_h stands for C_{γ_h}), it follows that $\max\{g(v) : v \in \cap_{\nu=1}^{\infty} C_{h_\nu}^*\} \leq 1$, hence $\cap_{\nu=1}^{\infty} C_{h_\nu}^* \subset D^*$ and consequently, $\cup_{\nu=1}^{\infty} C_{h_\nu} \supset D$. Thus, $\tilde{D} \subset \cup_{\nu=1}^{\infty} \tilde{C}_{h_\nu}$, i.e., for any $t \in \tilde{D}$ there exists h_ν such that $t \in \tilde{C}_{h_\nu}$. Hence, for any $x \in G$ there exists h_ν such that $f(x) \geq \gamma_{h_\nu} \geq \hat{\gamma}$, proving the global optimality of $\hat{x}$.

(II) ASSUMPTION $0 \in \operatorname{int} C_1$ (see (15); recall that C_1 stands for C_{γ_1})

We assumed $0 \in \operatorname{int} C_1$ in order that C_1^* be compact. This assumption amounts to requiring that a feasible point x^0 be available with $f(x^0) > f(\bar{x})$. However, since problems (8) are assumed to be solvable by some known efficient solution, a feasible solution (i.e., a point in G) can easily be found. Therefore, normally this assumption should be easy to satisfy.

In any case, the algorithm could be slightly modified to bypass this assumption. In fact, without the condition $0 \in \text{int } C_1$, C_1^* may be unbounded, so we must take S_1 to be a (possibly unbounded) polyhedron enclosing C_1^* (and contained in K^*). Then V_k should be understood as the set of all vertices and extreme directions of S_k. In Step 2 the condition $g(v^h) \leq 1$ should be replaced by: $g(v^h) \leq 1$ and v^h is a vertex, or $g(v^h) \leq 0$ and v^h is an extreme direction. Moreover, the right hand side of (26) should be 1 if v^h is a vertex, and 0 if v^h is an extreme direction.

(III) CONSTRUCTION OF THE INITIAL POLYHEDRON S_1

If $0 \in \text{int } C_1$ (see assumption (15)), then we can select $\alpha > 0$ so that $-\alpha e \in C_1$ (where $e = (1, \ldots, 1) \in R^p$). We can then prove that the polytope

$$S_1 = \left\{ v \in K^* : \sum_{j=1}^{p} v_j \leq \frac{1}{\alpha} \right\} \tag{27}$$

contains C_1^*. Indeed, since C_1 is closed and convex, $C_1 = (C_1^*)^*$, hence $-\alpha e \in C_1$ implies $-\langle \alpha e, v \rangle \leq 1$ for all $v \in C_1^*$.

If 0 may not be an interior point of C, then we simply take $S_1 = K^*$.

A special case of interest is when the matrix A has just k rows (which must then be linearly independent since rank $A = k$; recall that A is the matrix in (7)). In this case the system $Au = e$ is consistent, i.e., we can compute a point w such that $Aw = e$. Let $\alpha \geq 1$ be any number such that $\alpha w \in C$. Since obviously $-\alpha w \in$ int K it follows that $-\alpha w \in \text{int } C_1$, hence $-\langle \alpha w, v \rangle \leq 1$ for all $v \in C_1^*$. That is, the polytope

$$S_1 = \left\{ v \in K^* : \sum_{j=1}^{p} w_j v_j \leq \frac{1}{\alpha} \right\} \tag{28}$$

contains C_1^* and can be used to start the algorithm.

(IV) CONSTRUCTION OF THE CUT (26)

The cut (26) can be improved in most cases as follows.

Let $\theta_h \geq 1$ be any number such that $\theta_h(\varphi(x^h) - \varphi(x^0)) \in C_{h+1}$ (note that C_{h+1} may be larger than C_h). Then $\theta_h < \langle \varphi(x^h) - \varphi(x^0), v \rangle \leq 1$ for all $v \in C_{h+1}^*$. Therefore the cut $l_h(v) \leq 1/\theta_h$ eliminates v^h without excluding any point of C_{h+1}^* and so can be used instead of (26). In practice one can take

$$\theta_h = \sup\{\theta : \theta(\varphi(x^h) - \varphi(x^0)) \in C_{h+1}\} . \tag{29}$$

(V) COMPUTATION OF V_{h+1}

To compute the vertex set V_{h+1} from knowledge of V_h we can use, for example, the method in [27] (see also [23]). If k is not large then this method should not present difficulty. Especially, when $k=2$ (as it happens in certain problems of interest) the computation is very simple.

Anyway, to keep the size of the sets V_h within manageable limits, one can from time to time restart the whole procedure, using a different initial point x^0 and the best value γ_h so far obtained as the new γ_1. Also in certain cases it may be advisable to solve the (C^*, D^*)-problem by the Normal Conical Algorithm as developed in [28] (see also [23]).

5. Applications

The above approach seems to have many potential applications. In this section we only examine some particular classes of problems for which this approach has actually proven very successful.

I. First let us consider problems of the form

$$(Q) \qquad \text{minimize } f(x) \quad \text{subject to} \quad x \in G\,,$$

where G is a convex set in R^n and $f(x) = \psi(\varphi(x))$, satisfying the following conditions:

(a) $\varphi : \Omega \to R^p$ is a continuous mapping defined on some open set $\Omega \supset G$;

(b) $\psi : T \to R$ is a continuous quasiconcave function on some convex closed subset T of R^p containing $\varphi(\Omega)$;

(c) There exists a cone $K = \{t \in R^p : At \leq 0\}$ such that:

$$\forall t^0 \in \varphi(G) \quad t^0 + K \subset \{t \in T : \psi(t) \geq \psi(t^0)\}\,, \tag{30}$$

$$A \text{ is a } q \times p \text{ matrix with rank } A = k \tag{31}$$

and $A\varphi(x) = (r_1(x), \ldots, r_q(x))$ where

$$r_i(x) \quad (i = 1, \ldots, q) \text{ are concave functions on } G. \tag{32}$$

It is easily verified that problem (Q) has the Generalized Rank k Property. Indeed, here $\tilde{C}_\gamma = \{t \in T : \psi(t) \geq \gamma\}$ for any $\gamma \in f(G)$ and the subproblems (24) (with $u \in R^q_+$) are ordinary convex programs

$$\text{maximize} \sum_{i=1}^{p} u_i r_i(x) \quad \text{subject to} \quad x \in G \tag{33}$$

which can be solved efficiently by known standard algorithms. Therefore, the above approach can be applied to problem (Q).

PROPOSITION 2. *For each $\lambda \in R^q_+$ such that $\Sigma_{i=1}^q \lambda_i = 1$ let x_λ be an (arbitrary) optimal solution of the subproblem* (33) *with $u = \lambda$. Then a global optimal solution of (Q) is given by the vector $x_{\tilde{\lambda}}$, where*

$$\tilde{\lambda} \in \operatorname{argmin}\left\{ f(x_\lambda) : \lambda \in R^q_+\,, \ \sum_{i=1}^{q} \lambda_i = 1 \right\}. \tag{34}$$

Proof. This Proposition has been established in [11] for $k=2$ but the proof carries over to the case $k>2$. Specifically, let us apply the Algorithm of Section 3 to problem (Q), where the starting feasible solution $\bar{x}^1$ is taken to be a x_λ for some $\lambda=\lambda^1$, while for each $v \in V_h$ in Step 1, $x(v)$ is taken to be the x_λ that corresponds to $\lambda = u/(\Sigma_{i=1}^q u_i)$ where $u \in R_+^q$ is such that $v = A^T u$. Then at any stage the current best solution $\bar{x}^h$ is some x_λ. Hence, the conclusion.

CONSEQUENCE. To solve (Q) one can solve the *parametric convex program*

$$\text{maximize} \sum_{i=1}^{q} \lambda_i r_i(x) \quad \text{s.t.} \quad x \in G \quad \left(\lambda \in R_+^q\,, \quad \sum_{i=1}^{q} \lambda_i = 1\right) \tag{35}$$

and for each $\lambda \in R_+^q$, $\Sigma_{i=1}^q \lambda_i = 1$, take an optimal solution x_λ. Then a global optimal solution of (Q) is the x_λ that minimizes $f(x_\lambda)$.

Of course, solving the parametric program (35) is not easy when $q>2$ and G or the functions $r_i(x)$ are nonlinear. Algorithm 1 proposed in Section 3 is just a procedure that allows us to solve (35) only for an adaptively constructed sequence of values λ.

There is, however, a case when solving the parametric program (35) may be more efficient than using Algorithm 1. This is when $q=2$ and G is a polyhedron, while each $r_i(x)$ is a convex piecewise affine function. Indeed, in that case, if $r_i(x) = \max\{r_{ij}(x) \colon j \in J_i\}$ (r_{ij} affine functions), then (35) becomes

$$\begin{aligned} &\text{maximize} \quad \alpha s_1 + (1-\alpha)s_2 \\ &\text{s.t.} \qquad r_{ij}(x) \le s_i \quad (j \in J_i\,, \quad i=1,2)\,, \quad x \in G \end{aligned} \tag{36}$$

where $\alpha \in [0,1]$ and so (35) reduces to an ordinary *parametric linear program*. If $\alpha_1 = 0 < \cdots < \alpha_N = 1$ are the breaking values of the parameter α and x^j is a basic optimal solution of (36) for $\alpha \in [\alpha_{j-1}, \alpha_j]$ then a global optimal solution of (Q) is given by the vector x^j that minimizes $f(x^j)$ ($j=1,\ldots,N-1$). Note, however, that even in this case, if the computation of the value $f(x)$ is time consuming then Algorithm 1 may be preferred to the parametric approach since it may require computing less values of $f(x)$.

EXAMPLE 6. The problem considered in Example 1 (minimizing $f(x) := (c^1x + d_1) \times (c^2x + d_2)$ over a polytope $M \subset \{x \colon c^ix + d_i > 0 \ (i=1,2)\}$) is a problem (Q) with $G=M$, $\varphi(x)=x$ (identity mapping), $\psi(t) = (c^1t + d_1)(c^2t + d_2)$, $\hat{K} = \{t \in R^2 \colon c^1t \ge 0,\ c^2t \ge 0\}$, $r_i(x) = -c^ix$ $(i=1,2)$. Hence, as said in the Introduction, this problem reduces to solving the ordinary parametric linear program:

$$\text{minimize} \ \langle \alpha c^1 + (1-\alpha)c^2, x\rangle \quad \text{subject to} \quad x \in M \quad (\alpha \in [0,1])\,.$$

EXAMPLE 7. The problem considered in Example 2 (minimizing $f(x) := \Sigma_{j=1}^p q_j(h_j(x))$ over $x \in R^n$) also is a problem (Q) with $G=R^n$, $\varphi(x) = (h_1(x), \ldots, h_p(x))$, $\psi(t) = \Sigma_{j=1}^p q_j(t_j)$, $K = \{t \in R^p \colon t \ge 0\}$, $r_j(x) = -h_j(x)$ $(j=$

$1, \ldots, p$). Hence, if $p = 2$ and all the functions $h_j(x)$ are polyhedral then we can solve it by solving an ordinary parametric linear program

EXAMPLE 8. Consider the max-min problem studied in [29]:

$$\max_x \min_y (cx + dy) \quad \text{s.t.} \quad Ux + Vy \leq b\,, \quad x \geq 0\,, \quad y \geq 0\,. \tag{37}$$

Setting $P = \{x \geq 0 : \exists y \geq 0,\ Ux + Vy \leq b\}$, $h(x) = \min\{dy : Ux + Vy \leq b,\ y \geq 0\}$, we can rewrite the problem as

$$\max\{cx + h(x) : x \in P\}\,. \tag{38}$$

Since $h(x)$ is convex (and piecewise affine), this is a problem (Q) with $f(x) = -cx - h(x)$, $\varphi(x) = x$, $\psi(t) = -ct - h(t)$, $T = P$, $K = \{t : ct \leq 0,\ Ut \leq 0\}$, $A\varphi(x) = (Ux, cx)$. Indeed, denote by $M(x)$ the constraint set of the minimization problem that defines $h(x)$. Then for any $t^0 \in P$, if $y \in M(t^0)$ (i.e., $y \geq 0$ and $Ut^0 + Vy \leq b$) then for all t such that $U(t - t^0) \leq 0$ we will have $Ut + Vy \leq Ut^0 + Vy \leq b$, so that $y \in M(t)$; hence $M(t^0) \supset M(t)$, and consequently, $h(t) \leq h(t^0)$. Therefore, $t - t^0 \in K$ implies $ct \leq ct^0$, $h(t) \leq h(t^0)$, proving that $t^0 + K \subset \{t \in P : ct + h(t) \leq ct^0 + h(t^0)\}$. The verification of other conditions of (Q) is straightforward. It then follows from the above results that if the matrix

$$\tilde{U} = \begin{bmatrix} U \\ c \end{bmatrix}$$

has not a full rank (as it often happens) then the problem (37) can be solved by Algorithm 1 as a problem in smaller dimension. In particular, when U has a single row then by Proposition 2 the problem can be solved by solving an ordinary parametric linear program.

II. Problems of the form (Q) can be characterized as those in which the nonconvexity is concentrated in the objective function. Since, however, there is a duality relationship between objective and constraints (see [30]), it is natural that our results can also be applied to problems in which the *non-convexity is concentrated in the constraints*. These problems can be given the general formulation:

$$(R) \qquad \text{minimize } f(x) \quad \text{subject to} \quad x \in G \setminus \operatorname{int} F\,,$$

where F is a closed set in R^n, while $f(x)$ is a continuous convex function and G is a closed convex set (in R^n). Of course it is assumed here that the problem

$$\text{minimize } f(x) \quad \text{subject to} \quad x \in G \tag{39}$$

has no optimal solution feasible to (R) (otherwise the constraint $x \not\in \operatorname{int} F$ could simply be omitted). That is, a point x^0 is readily available such that

$$x^0 \in G \cap \operatorname{int} F\,, \quad f(x^0) < \inf\{f(x) : x \in G\} \tag{40}$$

(for example, x^0 is an optimal solution of (39)).

For any $x \notin F$, since $x^0 \in \text{int } F$, the intersection $[x^0, x] \cap \partial F$ of the line segment $[x^0, x]$ with the boundary ∂F of F is a nonempty compact set. Let $\pi(x) \in \text{argmin}\{f(x') : x' \in [x^0, x] \cap \partial F\}$. We will assume that the computation of $\pi(x)$ is straightforward (which is the case if F is convex or is such that the points where $[x^0, x]$ meets ∂F are finite in number and can easily be determined). The convexity of f, G and the condition (40) then imply the following

PROPOSITION 3. *Let $\bar{x}$ be any feasible solution of problem (R) and $G(\bar{x}) = G \cap \{x: f(x) \leq f(\bar{x})\}$. If $x \in G(\bar{x}) \setminus F$ then $\pi(x)$ is a better feasible solution than $\bar{x}$.*

Proof. Since f is convex and $f(x^0) < f(x)$ by (40), it follows that $f(\lambda x^0) + (1 - \lambda)(x) < f(x)$ $\forall \lambda \in (0, 1]$. Hence, $\pi(x) = \lambda x^0 + (1 - \lambda)x$ for some $\lambda \in (0, 1]$ and $f(\pi(x)) < f(x) \leq f(\bar{x})$. On the other hand, by convexity of G, $\pi(x) \in G$. Hence, $\pi(x) \in G \cap \partial F$, i.e., $\pi(x)$ is feasible to (R).

Problem (R) is said to be *regular* (stable) if any feasible point of it is the limit point of a sequence of interior feasible points (i.e., points that belong to the interior of the feasible set). It is easily verified that for any closed set $F_\varepsilon \subset \text{int } F$, the "perturbed" problem

$$(R_\varepsilon) \qquad \text{minimize } f(x) \quad \text{subject to} \quad x \in G \setminus \text{int } F_\varepsilon$$

is regular.

PROPOSITION 4. *Let $\bar{x}$ be any feasible solution of (R). If*

$$G(\bar{x}) \subset F \tag{41}$$

and if the problem is regular then $\bar{x}$ is a global optimal solution.

For a proof of this proposition, see, e.g., [30].

Thus, under the regularity assumption, transcending an incumbent $\bar{x}$ in the problem (R) amounts to solving the $(G(\bar{x}), F)$-problem. Therefore, the above approach can be applied to solve this problem as a problem in R^k whenever the Generalized Rank k Property holds for $(G(\bar{x}, F))$.

A feasible solution $\bar{x}$ is called a global ε-optimal solution of problem (R) if $f(x) > f(\bar{x}) - \varepsilon$ $\forall x \in G \setminus \text{int } F$. Clearly, without any regularity assumption, the condition

$$G_\varepsilon(\bar{x}) := G \cap \{x: f(x) \leq f(\bar{x}) - \varepsilon\} \subset \text{int } F \tag{42}$$

is always sufficient for $\bar{x}$ to be a global ε-optimal solution.

Let us examine in more detail problem (R) when

$$F = \{x \in \Omega: \psi(\varphi(x)) \geq 0\},$$

with Ω being some open set containing G and ψ, φ two mappings satisfying the conditions (a)(b)(c) stated at the beginning of this section. With the purpose of finding a global ε-optimal solution, we define

$$C = \tilde{C} - t^0\,, \quad D = \tilde{D} - t^0\,, \quad g(v) = \sup\{\langle v, t\rangle : t \in D\}\,,$$

where $\tilde{C} = \{t \in T : \psi(t) \geq 0\}$, $\tilde{D} = \varphi(G_\varepsilon(\bar{x}))$ and $t^0 = \varphi(x^0) \in D \cap \operatorname{int} C$. From Algorithm 1 we can then derive the following

ALGORITHM 2 (for solving (R) under assumption (43)).

Initialization.
If a feasible solution $\bar{x}^1$ is available, set $\gamma_1 = f(\bar{x}^1)$,

$$G_1 = G \cap \{x : f(x) \leq \gamma_1 - \varepsilon\}\,.$$

Otherwise, set $\gamma_1 = +\infty$, $D_1 = D$. Construct an initial polytope S_1 with a known vertex set V_1 and such that $C^* \subset S_1 \subset K^*$. Set $\tilde{V}_1 = V_1$, $r = 1$, $h = 1$ (r is the iteration counter, h is the cycle counter within the current iteration).

Iteration $r = 1, 2, \ldots$.
Step h.1. For each $v \in \tilde{V}_h$ solve

$$(SP(v)) \qquad \max\{\langle v, \varphi(x)\rangle : x \in G_r\}$$

(recall that $\langle v, \varphi(x)\rangle = \Sigma_{i=1}^q u_i r_i(x)$ for $v = A^T u$, see (33)). Let $x(v)$ be an optimal solution of $(SP(v))$ and $g(v) = \langle v, \varphi(x(v)) - \varphi(x^0)\rangle$.

Step h.2. Select $v^h \in \operatorname{argmax}\{g(v) : v \in \tilde{V}_h\}$. If $g(v^h) < 1$, then terminate (see Remark 3 below): if $\gamma_r < +\infty$ then $\bar{x}^r$ is a global ε-optimal solution: if $\gamma_r = +\infty$, then problem (R) is infeasible.

Step h.3a. Let $x^h = x(v^h)$. If $\psi(\varphi(x^h)) \geq 0$ (i.e., $x^h \in F$) then let

$$S_{h+1} = S_h \cap \left\{v : \langle \varphi(x^h) - \varphi(x^0), v\rangle \leq \frac{1}{\theta_h}\right\},$$

where

$$\theta_h = \sup\{\theta : \theta(\varphi(x^h) - \varphi(x^0)) \in C\}$$

(see (29)). Compute the vertex set V_{h+1} of S_{h+1}. Let $\tilde{V}_{h+1} = V_{h+1} \setminus V_h$.
Set $h \leftarrow h + 1$ and return to Step *h.1.*

Step h.3b. If $\psi(\varphi(x^h)) < 0$ (i.e., $x^h \in G_r \setminus F$) then compute $\pi(x^h)$ (see Proposition 3) and set $\bar{x}^{r+1} = \pi(x^h)$, $\gamma_{r+1} = f(\bar{x}^{r+1})$,

$$G_{r+1} = G_r \cap \{x : f(x) \leq \gamma_{r+1} - \varepsilon\}\,, \quad D_{r+1} = \varphi(G_{r+1}) - t^0\,.$$

Set $r \leftarrow r + 1$ and return to Step *h.1.*

REMARK 3. If $g(v^h) < 1$ then $g(v) < 1$ $\forall v \in S_h$ and from this it follows that $D_r \subset \operatorname{int} C$, hence $G_r \subset \operatorname{int} F$ and $\bar{x}^r$ is a global ε-optimal solution (if $\bar{x}^r$ exists, i.e., $\gamma_r < +\infty$). Indeed, if $D_r \not\subset \operatorname{int} C$ so that there is a $\bar{t} \in D_r \setminus \operatorname{int} C$, then since $\bar{t} \in D_r$

and $C^* \subset S_h$, we would have $\langle v, \bar{t}\rangle \leq g(v) < 1\ \forall v \in C^*$; on the other hand, since $\bar{t} \not\in \operatorname{int} C$, by the separation theorem, there would exist a $\bar{v}$ such that $\langle \bar{v}, t\rangle \leq 1\ \forall t \in C$, while $\langle \bar{v}, \bar{t}\rangle = 1$, i.e., a $\bar{v} \in C^*$ but $\langle \bar{v}, \bar{t}\rangle = 1$, a contradiction.

REMARK 4. Algorithm 2 involves a number of iterations during each of which the incumbent $\bar{x}^r$ remains unchanged. Since from one iteration to the next the value γ_r decreases at least by $\varepsilon > 0$, the number of iterations is finite. On the other hand, each iteration solves a (G_r, F)-problem by Algorithm 1, therefore it either terminates after finitely many steps, yielding a global ε-optimal solution, or generates an infinite sequence of x^h. In the latter case the current incumbent $\bar{x}^r$ is actually a global ε-optimal solution (provided the problem is regular): indeed, by Theorem in Section 4, we then have $G_r \subset F$.

As illustrations, let us consider some examples.

EXAMPLE 9. Let $F = \{x \in \Omega : \Pi_{j=1}^p \varphi_j(x) \geq 1\}$, where Ω is an open set in R^n such that $G \subset \Omega \subset \{x : \varphi_j(x) > 0\}$. That is, the problem is

$$\text{minimize } f(x) \quad \text{subject to} \quad x \in G\,, \quad \prod_{j=1}^{p} \varphi_j(x) \leq 1\,. \tag{43}$$

Since the function $\psi : R_+^p \to R$ defined by $\psi(t) = \Pi_{j=1}^p t_j - 1$ is quasiconcave and $F = \{x \in \Omega : \psi(\varphi(x)) \geq 0\}$ with $\varphi(x) = (\varphi_1(x), \ldots, \varphi_p(x))$, conditions (a) and (b) are satisfied. It is also immediate that for any $t^0 \in R_+^p$ and any $t \in R_+^p$ we have $\psi(t^0 + t) \geq \psi(t^0)$. Hence the cone $K = R_+^p = \{t \in R^p : t_j \geq 0\ (j = 1, \ldots, p)\}$ satisfies condition (30) with $A = -I$, I being the identity matrix, $k = p$. Here $A\varphi(x) = (-\varphi_1(x), \ldots, -\varphi_p(x))$, so that the problems (33) are

$$\text{minimize } \sum_{j=1}^{p} u_j \varphi_j(x) \quad \text{subject to} \quad x \in G\,. \tag{44}$$

Therefore, if the functions $\varphi_j(x)$ are convex then condition (c) is satisfied as well, and Algorithm 2 can be applied to solve this problem. Since p is usually very small compared to n, this method which basically reduces to solving a sequence of convex programs of the form (44) should be quite efficient. Especially, when $p = 2$ (problem studied by Kuno and Konno in [31]), the underlying global procedure, i.e., the construction of the polytopes S_h and the computation of their vertex sets V_h, is very simple. If, in addition, $f(x)$ is linear, G is a polyhedron and the $\varphi_j(x)$ are linear, too, as in the problem studied by Thach and Burkard in [21], then the method furnished by Algorithm 2 is quite near to that of these authors. In a subsequent paper we will present an adaptation of Algorithm 2 to the case $p = 2$ which allows the problem to be solved through ordinary parametric linear programming.

EXAMPLE 10. The two level optimization problem considered in Example 3 in the Introduction can be reformulated as

$$\text{minimize } c^1x + d^1y \quad \text{subject to} \quad A_1x + B_1y \leq g_1 ,$$

$$A_2x + B_2y \leq g_2 , \quad x \in R^p_+ , \quad y \in R^q_+ , \quad d^2y \leq \omega(x) ,$$

where $\omega(x)$ is the optimal value of the linear program (depending on x)

$$(L(x)) \qquad \min\{d^2y : A_2x + B_2y \leq g_2 , \quad y \geq 0\} .$$

Since $\omega(x)$ is obviously a convex function, this appears to be a problem (R) with $f(z) = c^1x + d^1y$, $G = \{z = (x, y) : A_ix + B_iy \leq g_i \ (i = 1, 2),\ x \in R^p_+,\ y \in R^p_+\}$, $F = \{z = (x, y) : d^2y - \omega(x) \geq 0\}$. The conditions (a)(b)(c) are fulfilled with $\varphi(z) = z$ (identity mapping), $\psi(z) = d^2y - \omega(x)$ (concave function), $K = \{z = (x, y) : A_2x \leq 0,\ d^2y \geq 0\}$ and k being the rank of the matrix

$$\begin{pmatrix} A_2 & 0 \\ 0 & d^2 \end{pmatrix} \tag{45}$$

(k is usually smaller than $p + q$). Indeed, for any $z^0 = (x^0, y^0) \in R^{p+q}_+$ and any $w = (u, v) \in L$ we have $d^2(y^0 + v) \geq d^2y^0$; furthermore, since $A_2u \leq 0$, the constraint set of the program $L(x^0 + u)$ contains that of $L(x^0)$, hence $\omega(x^0 + u) \leq \omega(x^0)$ and so $\psi(z^0 + w) \geq \psi(z^0)$, proving that

$$z^0 + K \subset \{z : \psi(z) \geq \psi(z^0)\} .$$

Therefore, the above method can be applied to this problem. It turns out that because of the specific structure of this problem (linearity of f and G), Algorithm 2 as applied to it can be made finite. For the details, we refer the interested reader to the paper of Tuy, Migdalas, and Värbrand [15]. When $k = 2$ (the matrix A_2 has a single row, say $a^2 \in R^q$) one can even show that the problem can be solved through ordinary parametric linear programming.

Appendix

THEOREM (see [1]). *Any nonempty closed set S in R^n is the projection on R^n of a set in R^{n+1} which is a difference of two convex sets.*

Proof. (This proof is different from that in [1]). We first show that $S = \{x \in R^n : g(x) - |x|^2 \leq 0\}$, where $g(x)$ is a convex function. For any $y \not\in S$ denote by $d(y, S)$ the distance from y to S and let $B(y, r)$ be the ball around y with radius r. Then it is easily seen that

$$S = \bigcap_{y \not\in S} (R^n \setminus B(y, d(y, s))) .$$

That is, $x \in S$ if and only if $|x - y|^2 \geq d^2(y, S)$ $\forall y \not\in S$, i.e., if and only if $|x|^2 + |y|^2 - 2xy \geq d^2(y, S)$ $\forall y \not\in S$, i.e., if and only if

$$|x|^2 \geq \sup\{2xy - |y|^2 + d^2(y, S) : y \not\in S\} .$$

Denote by $g(x)$ the function on the right hand side. Clearly $g(x)$ is convex since for each fixed y the function $x \to 2xy - |y|^2 + d^2(y, S)$ is affine. Hence, $S = \{x: g(x) - |x|^2 \leq 0\}$. Now $S = \{x: \exists(x, t) \in R^{n+1};\ g(x) \leq t, t \leq |x|^2\}$, therefore S is the projection of $D = C \backslash B$ with $C = \{(x, t): g(x) \leq t\}$ and $B = \{(x, t): |x|^2 < t\}$. The convexity of C, B is obvious.

References

1. Thach, P. T. (1987), D. C. Sets, D. C. Functions and Systems of Equations, Preprint, Institute of Mathematics, Hanoi. To appear in *Mathematical Programming*.
2. Tuy, H. and R. Horst (1991), The Geometric Complementarity Problem and Transcending Stationarity in Global Optimization in "Applied Geometry and Discrete Mathematics", *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* **4**, 341–354.
3. Henderson, J. M. and R. E. Quant (1971), *Microeconomic Theory*, McGraw-Hill.
4. Maling, K., S. H. Mueller, and W. R. Heller (1982), On Finding Most Optimal Rectangular Package Plans, Proceedings of the 19th Design Automation Conference, 663–670.
5. Konno H. and M. Inori (1988), Bon Portfolio Optimization by Bilinear Fractional Programming, *J. of Oper. Res. of Japan* **32**, 143–158.
6. Pardalos, P. M. (1988), Polynomial Time Algorithms for Some Classes of Constrained Non-Convex Quadratic Problems, Preprint, Computer Science Department, the Pennsylvania State University.
7. Forgo, F. (1975), *The Solution of a Special Quadratic Problem*, Szigma, 53–59 (in Hungarian).
8. Gabasov, R. and F. M. Kirillova (1980), *Linear Programming Methods, Part 3 (Special Problems)*, Minsk (in Russian).
9. Pardalos, P. M. (1988), On the Global Minimization of the Product of Two Linear Functions over a Polytope, Preprint, Computer Science Department, The Pennsylvania State University.
10. Konno H. and T. Kuno (1989), Linear Multiplicative Programming, IHSS Report 89–13, Institute of Human and Social Sciences, Tokyo Institute of Technology.
11. Tuy, H. and B. T. Tam (1990), An Efficient Solution Method for Rank Two Quasiconcave Minimization Problems, to appear in *Optimization*.
12. Idrissi, H, P. Loridan, and C. Michelot (1988), Approximation of Solutions for Location Problems, *Journal of Optimization Theory and Applications* **56**, 127–143.
13. Tuy, H. and Faiz A. Al-Khayyal (1991), A Class of Global Optimization Problems Solvable by Sequential Unconstrained Convex Minimization, to appear in C. Floudas and P. Pardalos (eds), *Recent Advances in Global Optimization*, Princeton University Press.
14. Ben-Ayed, O. and C. E. Blair (1990), Computational Difficulties of Bilevel Linear Programming, *Operations Research* **38**, 556–560.
15. Tuy, H., A. Migdalas, and P. Värbrand (1990), A Global Optimization Approach for the Linear Two Level Program, Preprint, Department of Mathematics, Linköping University (submitted).
16. Tuy H. (1990), On Polyhedral Annexation Method for Concave Minimization, in Lev. J. Leifman (ed.), *Functional Analysis, Optimization, and Mathematical Economics* (volume dedicated to the memory of Kantorovich), Oxford University Press, New York, 248–260.
17. Konno, H. (1976), A Cutting Plane Algorithm for Solving Bilinear Programs, *Mathematical Programming* **11**, 14–27.
18. Tuy, H. (1987), Convex Programs with an Additional Reverse Convex Constraint, *Journal of Optimization Theory and Applications* **52**, 463–485.
19. Tuy, H. (1991), Polyhedral Annexation, Dualization, and Dimension Reduction Technique in Global Optimization, *Journal of Global Optimization* **1**, 229–244.
20. Thach, P. T. (1991), Quasiconjugates of Functions and Duality Correspondence between Quasiconcave Minimization under a Reverse Convex Constraint and Quasi Convex Maximization under a Convex Constraint, *Journal of Mathematical Analysis and Applications*.
21. Thach, P. T., and R. E. Burkard (1990), Reverse Convex Programs Dealing with the Product of Two Linear Functions, Preprint, Institute of Mathematics, Graz Technical University.

22. Thach, P. T. and H. Tuy (1990), Dual Solution Methods for Concave Programs and Reverse Convex Programs, Preprint, IHSS, Tokyo, Institute of Technology.
23. Horst, R. and H. Tuy (1990), *Global Optimization (Deterministic Approaches)*, Springer-Verlag.
24. Kuno, T. and H. Konno (1990), Parametric Successive Underestimation Method for Convex Multiplicative Programming Problems, Preprint IHSS, Tokyo Institute of Technology.
25. Konno H. and T. Kuno (1991), Generalized Linear Multiplicative and Fractional Programming, to appear in *Annals of Operations Research.*
26. Rockafella, R. T. (1970), *Convex Analysis,* Princeton University Press.
27. Horst, R., N. V. Thoai, and J. de Vries (1988), On Finding New Vertices and Redundant Constraints in Cutting Plane Algorithms for Global Optimization, *Operations Research Letters* **7**, 85–90.
28. Tuy, H. (1991), Normal Conical Algorithm for Concave Minimization over Polytopes, *Mathematical Programming* **51**, 229–245.
29. Falk, J. E. (1973), A Linear Max-Min Problem, *Mathematical Programming* **5**, 169–188.
30. Tuy H. and N. V. Thuong (1988), On the Global Minimization of a Convex Function under General Nonconvex Constraints, *Applied Mathematics and Optimization* **18**, 119–142.
31. Kuno T. and H. Konno (1990), A Parametric Successive Underestimation Method for Convex Programming Problems with an Additional Convex Multiplicative Constraint, Preprint, IHSS, Tokyo Institute of Technology.