ACTA MATHEMATICA VIETNAMICA
Vol. 8 No. 2 (1983)

ON OUTER APPROXIMATION METHODS FOR SOLVING CONCAVE MINIMIZATION PROBLEMS

HOANG TUY
Institute of Mathematics
Hanoi

1. INTRODUCTION

Despite their obvious practical importance, for a long time nonlinear global optimization problems have been receiving little attention from mathematicians. For one thing, these problems have an entirely different character from local optimization problems one is accustomed to consider in mathematical programming. In fact, mathematicians like regularity. In the optimization problems with which they have been concerned until recent years, a local criterion is usually available by which any optimal solution can be recognized as such. Moreover, when a given feasible solution fails to satisfy this local criterion, then a rule can be indicated for proceeding to a better solution. Under certain conditions which are satisfied in many cases of interest, one can then devise an iterative procedure which is guaranteed to eventually lead to an exact optimal solution, or at least to an approximate optimal solution with any desired accuracy.

By contrast, in global optimization problems there is, in general, no local criterion for deciding whether a given feasible solution is actually a global optimum. There can exist, in fact quite often, a large number of local optima, and even if one has been lucky enough to find one of these local optima, one will normally have no clue as how to proceed further. Because of these unpleasant features, nonlinear global optimization problems have been regarded by most researchers as hopelessly difficult.

There is, however, in the class of global optimization problems a subclass which is more tractable than the others. Namely, this is the subclass constituted by the *concave minimization* problems which can be formulated in the following general form:

(P) Minimize $f(x)$, subject to $x \in D$, where $f : R^n \to R$ is a real-valued *concave function* defined throughout R^n, and D is a *closed convex subset* of R^n.

(For the practical origin, the applications, and the theoretical interest of this problem, the reader is referred to [12] and [13]).

According to a basic property of concave functions, either the global minimum of f over D is achieved at some extreme point of D, or f is unbounded below over some extreme ray of D (see e.g. [8]). Therefore, the search for an optimal solution to (P) can be restricted to the set of extreme points and extreme rays of D. To a certain extent this is a nice property which makes the problem easier, even though the set of extreme points and extreme rays of D is, in general, infinite and may still be very hard to handle.

A first attempt to solve Problem (P) was made in [12], where a cone splitting procedure of the type earlier developed in [11] was incorporated into a branch and bound scheme. Under certain general conditions, it was proved that this branch and bound scheme converges to an optimal solution.

In the present paper we shall discuss alternative methods for solving (P), which can be called *outer approximation* methods. Basically, these methods consist in approximating the given constraint set D by a polyhedral convex set containing it, whose vertices (including the vertices «at infinity» which determine the unbounded edges) are known or can be computed practically. Then the minimum of the objective function f over this polyhedral convex set (which is attained at one of the vertices) is offered as an approximate optimal solution to the original problem (P). If the accuracy of the solution attained is not yet satisfactory, the approximation is refined further, and the whole process of successive approximations can be arranged so that it will converge to an optimal solution of (P).

It should be noted that the basic ideas of outer approximation methods in nonlinear programming are not quite new. In fact, these ideas underlie the cutting plane method of Kelley for solving convex programs [7]. Outer approximation procedures have been combined with underestimation technique in the algorithm of Falk and Hoffman [3] for concave minimization over a polytope. More recently, they have been developed in the algorithm of Hoffman [5] and that of Thieu-Tam-Ban [9] for concave minimization over a compact convex set. Our aim in the sequel is to provide a unifying scheme which would include these previous algorithms as special cases and, more importantly, could be applied to a wider class of problems, where the constraint set may be unbounded or even nonconvex of a certain type.

The paper is divided into 8 sections. After the Introduction, in Section 2 the general scheme of outer approximation is described and the convergence theorem is proved for the case where the constraint set D is compact. In Section 3, some specific procedures for generating the sequence of polyhedral convex sets approximating D are discussed. In Section 4, the scheme is extended to the case of an unbounded constraint set. Section 5 is devoted to procedures for solving the relaxed problems involved in the steps of the scheme. Section 6 deals with the specialization of the method to the case where D is a polyhedral convex set. In Section 7 we extend the method to problems with a nonconvex constraint set of the form $D = C \setminus G$, where C is a closed convex set, and G an open convex set. Finally, in Section 8 we show how the outer approximation approach can be combined with other methods to produce more efficient algorithms.

2. THE GENERAL SCHEME OF OUTER APPROXIMATION METHODS

(case of a bounded constraint set).

Recall that the problem of concern is the following

(P) Minimize $f(x)$, subject to $x \in D$,

where $f : R^n \to R$ is a concave function, defined throughout R^n (hence continuous), and D is a closed convexsubset of R^n. Sometimes we assume that D isgiven by a system of inequalities of the form:

$$g_i(x) \leqslant 0 \quad (i = 1, \ldots, m) \tag{1}$$

with $g_i : R^n \to R$ $(i = 1, \ldots, m)$ being convex functions, defined throughout R^n.

In this section, we shall also assume that D is bounded, so that it can be enclosed in a polytope $S_1 \supset D$. The problem

(Q_1) Minimize $f(x)$, s.t. $x \in S_1$

is then a *relaxed* form of (P), whose optimal value gives a lower bound for the optimal value of (P). Since S_1 is a polytope, (Q_1) can be solved, for example by a search through the vertices of S_1 and taking the vertex x^1 that corresponds to the smallest value of f (the latter problem is not hard, provided, S_1 is chosen properly). If it happens that $x^1 \in D$ (or, more generally, $f(x^1) = f(y^1)$ for some $y^1 \in D$), then x^1 is obviously an optimal solution to Problem (P). Otherwise, $x^1 \notin D$, and using the convexity and closedness of D, we can always find a hyperplane H, strictly separating x^1 from D. That hyperplane gives rise to an inequality

$$h_1(x) \leqslant 0 \tag{2}$$

which is satisfied by all $x \in D$ but is violated by x^1. So, by adding (2) to the system of inequalities defining S_1 we « cut » away the point x^1 and determine a new polytope S_2 which better approximates D than S_1. The process can now be repeated starting with S_2 instead of S_1, and so on, until an approximate solution x^k is obtained which satisfies the constraints (1) to a sufficient degree.

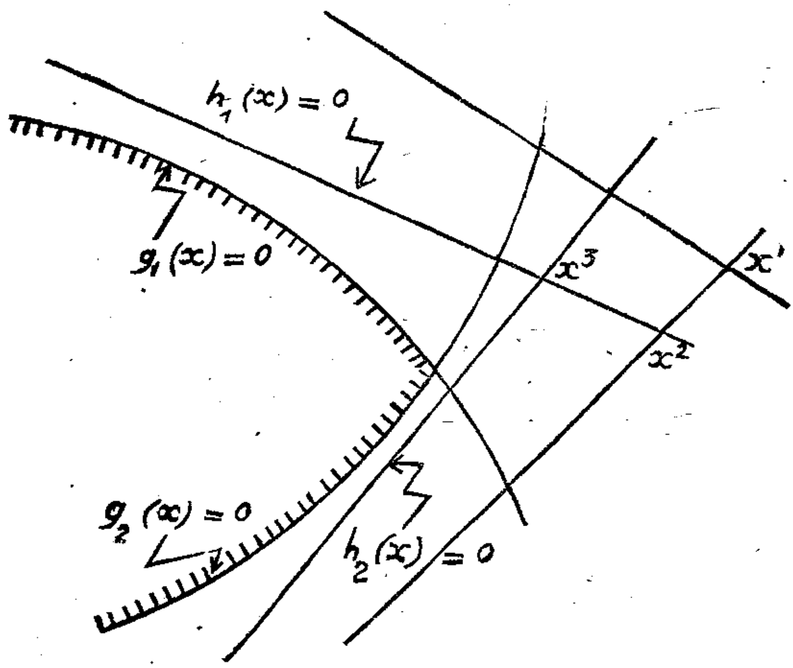

Fig. 1

Any method for solving Problem (P) that proceeds along this line will be called an *outer approximation* method. We can formulate the general scheme of outer approximation methods in the following precise form:

ALGORITHM 1 (D bounded)

Start from a polytope $S_1 \supset D$. Set $k = 1$.

a) Solve the relaxed problem

(Q_k) Minimize $f(x)$, s.t. $x \in S_k$

(for example by taking the minimum of f over the vertex set of S_k).

Let x^k be an optimal solution of (Q_k).

If $x^k \in D$ (or if $f(x^k) = f(y)$ for some $y \in D$), terminate: x^k (or y, resp.) solves (P). Otherwise, go to b).

b) Construct a hyperplane H_k strictly separating x^k from D, i.e. such that if $h_k(x) = 0$ is its equation, then

$$h_k(x^k) > 0, \quad h_k(x) \leqslant 0 \;\; \forall\, x \in D \tag{3}$$

Form the polytope

$$S_{k+1} = S_k \cap \{x: h_k(x) \leqslant 0\}$$

Set $k \leftarrow k+1$ and return to a). □

To determine this scheme completely one must specify how to construct the hyperplane H_k. But before discussing this question, it is expedient to indicate first the conditions ensuring the convergence of the process, i. e. the conditions under which one is assured that the sequence $\{x_k\}$ has at least one cluster point $\bar{x}$ such that

$$\bar{x} \in D, \ f(\bar{x}) = \min \{ f(x) : x \in D \}. \tag{4}$$

A hyperplane $H_k = \{x : h_k(x) = 0\}$ where $h_k(x) = \langle p^k, x \rangle + \eta_k$ is said to tend to a limit $\bar{H} = \{x : \bar{h}(x) = 0\}$ as $k \to \infty$, in symbols $H_k \to \bar{H}$ $(k \to \infty)$, if $\bar{h}(x) = \langle \bar{p}, x \rangle + \bar{\eta}$, with

$$\frac{p^k}{\| p^k \|} \to \bar{p}, \quad \frac{\eta_k}{\| p_k \|} \to \bar{\eta}.$$

THEOREM 2. 1. *Suppose that the following condition (*) holds.*

(*) Whenever a subsequence $\{x^{k_\nu}\}$ tends to $\bar{x} \notin D$, and the hyperplane H_{k_ν} tends to $\bar{H}$, then $\bar{H}$ strictly separates $\bar{x}$ from D.

Then the above scheme either terminates after finitely many steps with an optimal solution of (P) or generates an infinite decreasing sequence $\{S_k\}$. In the latter case, the sequence $\{x^k\}$ has at least one cluster point and any cluster point of this sequence is an optimal solution to (P).

Proof. Suppose that the process is infinite, and let $\bar{x}$ be any cluster point of $\{x^k\}$, e. g. $\bar{x} = \lim_{\nu\to\infty} x^{k_\nu}$. (That a cluster point exists follows from the boundedness of $\{x^k\} \subset S_1$). Let $H_k = \{x : h_k(x) = 0\}$ with $h_k(x) = \langle p^k, x \rangle + \eta_k$. We may of course assume $\| p^k \| = 1$ $(\forall k)$. By hypothesis

$$0 < h_k(x^k) = \langle p^k, x^k \rangle + \eta_k,$$

hence

$$\eta_k > -\langle p^k, x^k \rangle. \tag{5}$$

Further,

$$h_k(x) = \langle p^k, x \rangle + \eta_k \leqslant 0 \ (\forall x \in D),$$

hence for any point $x^o \in D$:

$$\eta_k \leqslant - \langle p^k , x^o \rangle . \tag{6}$$

From (5) and (6) it follows that the sequence $\{ \eta_k \}$ is bounded. Therefore, by taking a subsequence if necessary, we may assume that

$$p^{k_\nu} \rightarrow \overline{p} , \quad \eta_{k_\nu} \rightarrow \overline{\eta} ,$$

i. e. H_{k_ν} tends to $\overline{H} = \{ \overline{h}(x) = 0 \}$ with $\overline{h}(x) = \langle \overline{p} , x \rangle + \overline{\eta}$.

We contend that $\overline{x} \in D$. Indeed, if it were not so, we would have in view of condition (*),

$$\overline{h} (\overline{x}) > 0. \tag{7}$$

But since for all $l > k_\nu$, $x^l \in S^l \subset S^{k_\nu}$, it follows that $h_{k_\nu}(x^l) \leqslant 0$. Hence fixing ν and letting $l = k_\mu$, $\mu \rightarrow \infty$, we get $h_{k_\nu} (\overline{x}) \leqslant 0$. This in turn implies, by letting $\nu \rightarrow \infty$,

$$\overline{h} (\overline{x}) \leqslant 0 ,$$

contradicting (7). Therefore, $\overline{x} \in D$, and we have

$$f (\overline{x}) \geqslant \min \{ f(x) ; x \in D \}.$$

But, since $S_k \supset D$, $f(x^k) \leqslant \min \{ f(x): x \in D \}$ and hence,

$$f(x^k) \rightarrow f(\overline{x}) = \min \{ f(x): x \in D \} ,$$ as was to be proved. □

Remark 2.1. Let D be given by a system (1) and let

$$g(x) = \max_{i=1,\dots,m} g_i (x).$$

Since every $g_i (\cdot)$ is continuous (as a convex function which is finite throughout R^n), the function $g(x)$ is continuous. Clearly

$$D = \{x : g(x) \leqslant 0 \}.$$

When $x^{k_\nu} \rightarrow \overline{x}$ we have, by continuity, $g(x^{k_\nu}) \rightarrow g(\overline{x})$. But as proved above, $\overline{x} \in D$, i.e. $g(\overline{x}) \leqslant 0$. Hence, given any $\varepsilon > 0$, we shall have for large enough ν:

$$g(x^{k_\nu}) \leqslant \varepsilon .$$

A point x^{k_ν} satisfying this condition can be considered as an approximate optimal solution, since it is nearly feasible, and $f(x^{k_\nu}) \leqslant \min \{ f(x) : x \in D \}$.

If we want to have an approximate optimal solution which is feasible, then we can find the point y^{k_ν} of D that is the nearest to x^{k_ν}. Since the set D is convex, this amounts to solving the convex program

$$\text{Minimize } \| y - x^{k_\nu} \| , \text{ subject to } y \in D.$$

Remark 2.2. Close scrutiny of the proof of Theorem 2.1 shows that we have in fact established a more general fact, namely the following « *cutting plane convergence principle* » :

Let D be an arbitrary (not necessarily convex) *subset of R^n. Let $\{x^k\} \subset R^n$ be a bounded sequence and for every k let $h_k(x)$ be an affine function such that*

$$h_k(x) \leqslant 0 \;\; \forall x \in D, \quad h_k(x^k) > 0 \; ; \tag{8}$$

$$h_j(x^k) \leqslant 0 \;\; (j = 1, \ldots, k-1). \tag{9}$$

Assume furthermore, that

(*) *Whenever* $x^{k_\nu} \to \overline{x} \notin D$ *and* $(\forall x)$ $h_{k_\nu}(x) \to \overline{h}(x)$, *one must have* $\overline{h}(\overline{x}) > 0$.

Then every cluster point of $\{x^k\}$ belongs to D.

In the above proposition one can even replace (8), (9) by a weaker condition:

$h_k(x^0) \leqslant 0$ for some $x^0 \in D$, $\quad h_k(x^k) > 0$;

$h_j(x^k) \leqslant 0$ for all $j < k$ such that $h_j(x^i) > 0$ for at least N indices $i < j$, where N is a given natural number.

The proof of the proposition under this weaker condition is almost the same as before, with only some minor changes. On the basis of this generalized cutting plane convergence principle, one could, in forming the current relaxed problem (Q_{k+1}), drop all the earlier cuts $h_j(x) \leqslant 0$ $(j < k)$ which satisfy $h_j(x^i) > 0$ for less than N indices $i < j$ (i. e. all the cuts that had eliminated less than N earlier generated points x^i, $i < j$). This constraint dropping strategy is advisable when the number of constraints of the relaxed problems quickly increases as the algorithm proceeds.

3. GENERATING THE ENCLOSING POLYTOPES

We turn to the question of how to construct, for any given point $x^k \notin D$, a hyperplane H_k strictly separating x^k from D, such that condition (*) holds.

METHOD I. (Hoffman [5]). This method requires the prior knowledge of an interior point x^0 of D:

$$x^0 \in \text{int } D,$$

(so $g(x^0) < 0$). Since $x^k \notin D$, the line segment $[x^o, x^k]$ cuts the boundary ∂D of D at a unique point $z^k \neq x^k$, which is determined by solving the equation in λ:

$$g(\lambda x^0 + (1 - \lambda)\, x^k) = 0.$$

Now select any $p^k \in \partial g(z^k)$ for this it suffices to take any $p^k \in \partial g_{i_0}(z^k)$, where $i_o = i_o(k)$ is an index such that $g_{i_0}(z^k) = \max g_i(z^k)$). Let

$$h_k(x) = \langle p^k, x - z^k \rangle.$$

$$= \langle p^k, x \rangle - \langle p^k, z^k \rangle.$$

(so $\eta_k = - \langle p^k, z^k \rangle$). Observe that by the definition of a subgradient.

$$g(x) - g(z^k) \geqslant \langle p^k, x - z^k \rangle \, (\forall x). \tag{10}$$

In particular,

$$g(x^o) - g(z^k) \geqslant \langle p^k, x^o - z^k \rangle,$$

hence, noting that $g(z^k) = 0, g(x^o) < 0$:

$$0 > g(x^o) \geqslant \langle p^k, x^o - z^k \rangle.$$

This implies $p^k \neq 0$. Further, since $x^k - z^k = -\alpha(x^o - z^k)$ for some $\alpha > 0$, it follows that

$$h_k(x^k) = \langle p^k, x^k - z^k \rangle = - \alpha \langle p^k, x^o - z^k \rangle > 0.$$

On the other hand, from (10) we obtain for all $x \in D$, i. e. for all x satisfying $g(x) \leqslant 0$:

$$h_k(x) = \langle p^k, x - z^k \rangle \leqslant g(x) - g(z^k) \leqslant 0.$$

Thus, the hyperplane $H_k = \{x : h_k(x) = 0\}$ strictly separates x^k from D.

Let us verify now the condition (*). Suppose that

$$x^{k_\nu} \to \overline{x} \notin D;\ H_{k_\nu} \to \overline{H} \tag{11}$$

with $h(x) = \langle \overline{p}, x \rangle + \overline{\eta}$. Since $z^{k_\nu} = \lambda_{k_\nu} x^o + (1 - \lambda_{k_\nu})\, x^{k_\nu}$,

$0 < \lambda_{k_\nu} < 1$, we may assume $\lambda_{k_\nu} \to \lambda \in [0,1]$, i. é. $z^{k_\nu} \to z = \lambda\, x^o + (1 - \lambda)\overline{x}$,

and since $z^k \in \partial D$ and the set ∂D is closed, we must have $\overline{z} \in \partial D$. So $\overline{z}$ is the point where the line segment $[x^o, \overline{x}]$ meets ∂D. Passing to the limit as $\nu \to \infty$, the inequality $h_{k_\nu}(x) \leqslant 0$ ($\forall x \in D$) yields

$$\overline{h}(x) \leqslant 0 \ (\forall x \in D) \tag{12}$$

Similarly, the relations $h_k(z^k) = 0$, $h_k(x^k) > 0$, give in the limit $\bar{h}(\bar{z}) = 0$, $\bar{h}(\bar{x}) \geqslant 0$. But if $\bar{h}(\bar{x}) = 0$, this together with $\bar{h}(\bar{z}) = 0$ would imply $\bar{h}(\bar{x}^o) = 0$, and hence, in view of (12) and the fact $x^o \in int\ D$, we would have $\bar{h}(x) = 0$ $(\forall x \in D)$, i.e. $D \subset \bar{H}$, which is impossible since $int\ D \neq \varnothing$. Therefore $\bar{h}(\bar{x}) > 0$ and this, along with (12), shows that $\bar{H}$ strictly separates $\bar{x}$ from D.

Remark 3.1. The previous proof does not need in any way the continuous differentiability of g_i $(i = 1,..., m)$ as required by Hoffman in [5]. It is enough to assume that g_i are defined on D and subdifferentiable on the boundary points of D.

METHOD II (Thiêu-Tâm-Ban [9]). This method was first used by Kelley [7] in the context of convex minimization. It does not require the prior knowledge of an interior feasible point, but, instead, assumes the functions g_i $(i = 1,..., m)$ to be continuous throughout R^n, or at least on some open neighbourhood of D.

Since $x^k \notin D$, we have $g(x^k) > 0$. Let $p^k \in \partial g(x^k)$, and

$$h_k(x) = \langle p^k, x - x^k \rangle + g(x^k)$$

(so $\eta_k = -\langle p^k, x^k \rangle + g(x^k)$). Clearly $h_k(x^k) = g(x^k) > 0$, and since

$$g(x) - g(x^k) \geqslant \langle p^k, x - x^k \rangle,$$

we have for all $x \in D$ (i. e. satisfying $g(x) \leqslant 0$).

$$h_k(x) = \langle p^k, x - x^k \rangle + g(x^k) \leqslant g(x) \leqslant 0.$$

Thus $H_k = \{x : h_k(x) = 0\}$ strictly separates x^k from D.

To verify the condition (*), assume that (11), (12) hold. In view of the continuity of g we have from (11): $g(x^{k_\nu}) \to g(\bar{x}) > 0$. Further, since $\{x^k\} \subset S_1$ and $\partial g(S_1)$ is bounded ([8], Th. 24.7), by taking a subsequence if necessary, we may suppose $p^{k_\nu} \to \bar{p} \in \partial g(\bar{x})$. Therefore, $\bar{h}(x) = \langle \bar{p}, x - \bar{x} \rangle + g(\bar{x})$ and, as previously, it is easily seen that $\bar{h}(\bar{x}) > 0$, while $\bar{h}(x) \leqslant 0$ for all $x \in D$; i.e. $\bar{H}$ strictly separates $\bar{x}$ from D. □

Remark 3.2. 1) The continuity assumption about the functions g_i $(i = 1,..., m)$ can be weakened as follows:

The functions $g_i : R^n \to (-\infty, +\infty]$ are convex, continuous relative to a polytope $S_1 \supset D$ and such that for every $x \in S_1$ one can select for the function $g(x) = \max_i g_i(x)$ a vector $p(x) \in \partial g(x)$ so that the set $\{p(x), x \in S_1\}$ is bounded.

It can easily be verified that the previous convergence proof remains valid under these assumptions, provided one takes $p^k = p(x^k)$ in each iteration k of the Algorithm.

2) In certain problems it may happen that some of the functions g_i are given in such a form that for the solution of these problems Method II is much more convenient to apply than Method I. For example, let us consider the problem

$$\text{Minimize } f_1(x) + f_2(y), \text{ s.t.}$$
$$Ax + By + c \leqslant 0,$$
$$x \geqslant 0,\ y \geqslant 0,$$

where $f_1 : R^{n_1} \to R$ is concave, $f_2 : R^{n_2} \to R$ is convex and A, B are matrices of appropriate orders. Setting

$$X = \{x \geqslant 0 : \exists\, y \geqslant 0\ Ax + By + c \leqslant 0\},$$
$$g(x) = \inf \{f_2(y) : Ax + By + c \leqslant 0,\ y \geqslant 0\},$$

we convert the problem into the following one:

$$\text{Minimize } f_1(x) + t, \qquad \text{s.t. } g(x) - t \leqslant 0,\ x \in X,$$

where the function g is convex. It can be shown that at every x where $g(x) > -\infty$ the subdifferential $\partial g(x)$ consists precisely of vectors $A^T q$, where q is a Kuhn-Tucker vector of the convex program defining $g(x)$. Therefore, the application of Method II to the above problem is relatively easy, whereas Method I would involve solving in each iteration an equation of the form $g(\lambda x^0 + (1-\lambda)x^k) = \lambda t^0 + (1-\lambda)t^k$ — which may be very hard in view of the form in which the function g is defined.

Remark 3.3. In some applications we may encounter the Problem (*P*), where only certain functions g_i $(i \in I \subset \{1,.., m\})$ are convex, while the other g_i $(i \in \{1,..., m\} \setminus I)$ are concave. That is, only the constraints with indices $i \in I$ in (1) are convex, the other are *reverse convex*. Then to apply the scheme described in Section 2, we can modify the construction of H_k as follows.

Let $x^k \notin D$, so that there is $i_0 = i_0(k) \in \{1,..., m\}$ such that

$$g_{i_0}(x^k) > 0.$$

If $i_0 \in I$ (the function g_{i_0} is convex), we proceed as in Method I or Method II above.

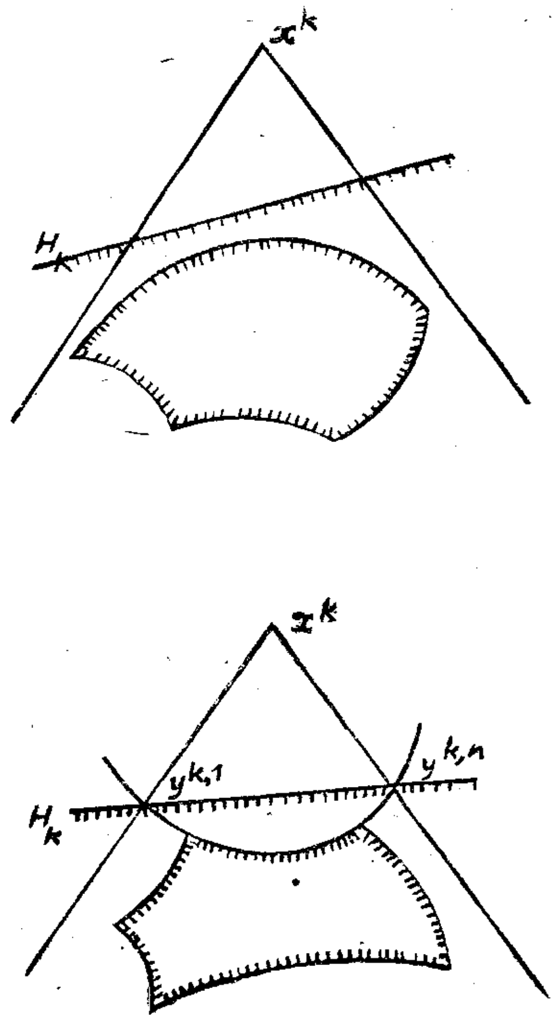

Fig. 2

If $i_o \notin$ I (the function g_{i_o} is concave) we construct a cutting plane H_k of the type originally used by Tuy in [11] (and later called « convexity cuts » by Glover). Namely, assume x^k is a nondegenerate vertex of S_k (since x^k is an optimal solution of (Q_k), it can be assumed to be a vertex of S_k; a slight perturbation, if necessary, will make this vertex nondegenerate). On each of the n edges of S_k emanating from x^k we take a point, say $y^{k,j}$ ($j = 1, \ldots, n$) such that $g_{i_o}(y^{k,j}) = 0$. Then H_k is taken to be the hyperplane through $y^{k,1}, \ldots,$

$y^{k,n}$. Clearly H_k strictly separates x^k from D. If x^k is a degenerate vertex of S_k, we can also use as H_k a cut of the type proposed by Carvajal — Moreno [2].

However, it is an open question whether this method satisfies the convergence condition (*) or not.

4. CASE OF AN UNBOUNDED CONSTRAINT SET

To extend the previous method to the case where the constraint set D is unbounded, we must construct a decreasing sequence of generalized polytopes (unbounded polyhedral convex sets) S_k enclosing D:

$$S_1 \supset S_2 \supset \dots \supset S_k \supset \dots \supset D$$

such that the relaxed problems

(Q_k) Minimize $f(x)$, s . t. $x \in S_k$

approximate the original Problem (P) more and more closer, in the sense that

$$\inf \{f(x) : x \in S_k\} \to \inf \{f(x) : x \in D\}.$$

Furthermore, the scheme described in Section 2 must be modified, because now certain relaxed problems (Q_k) may have no finite optimal solution. In fact, in solving (Q_k) we may discover an extreme ray of S_k over which f is unbounded below (for the sake of convenience, we shall call the direction of such a ray an « optimal direction »).

ALGORITHM 2 (D possibly unbounded)

Select an interior point x^o of D, and a generalized polytope $S_1 \supset D$. Set $k = 1$.

a) Solve the relaxed problem (Q_k). If an optimal direction u^k is obtained go to b). If an optimal solution x_k is obtained go to c).

b) Let Γ_k be the halfline emanating from x^o in the direction u^k (because of the concavity of f, f is also unbounded below over Γ_k; see e.g. [8], Th. 8.6). If $\Gamma_k \subset D$, terminate : the function f is unbounded below over D. Otherwise, let z^k be the point where Γ_k meets ∂D, the boundary of D. Construct a supporting hyperplane H_k to D at z^k and go to d).

c) If $x^k \in D$ or $f(x^k) = f(x^o)$, terminate : x^k (or x^o, resp.) is an optimal solution to (P). Otherwise, construct a hyperplane H_k strictly separating x^k from D and go to d).

d) Let $h_k(x) = 0$ be the equation of H_k (such that $h_k(x^o) < 0$). Form the generalized polytope

$$S_{k+1} = S_k \cap \{x : h_k(x) \leqslant 0\}. \tag{13}$$

Set $k \leftarrow k+1$ and return to a). □

Remark 4.1. Obviously, if c) occurs in some iteration it will occur in all subsequent iterations.

THEOREM 4.1. *Assume that for some $\alpha < f(x^0)$ the set $\{x : \alpha = f(x)\}$ is bounded, and that the condition (*) in Theorem 2. 1 holds. Then the above scheme either terminates after finitely many steps with a finite optimal solution or a halfline in D over which f is unbounded below, or it is infinite. In the latter case, the scheme generates either a bounded sequence $\{x^k\}$, whose every cluster point x is an optimal solution to (P), or a sequence $\{u^k\}$, whose every cluster point u is the direction of a halfline in D over which f is unbounded below.*

The proof of this Theorem results from three following lemmas.

LEMMA 4.1. *Assume that for some $\alpha < f(x^0)$ the set $\{x : \alpha = f(x)\}$ is bounded. If $\{u^k\}$ is a sequence such that f is unbounded below over each halfline $\Gamma_k = \{x^o + \lambda u^k : \lambda \geqslant 0\}$ and if $u^k \to u$ $(k \to \infty)$, then f is also unbounded below on the halfline $\Gamma = \{x^o + \lambda u : \lambda \geqslant 0\}$.*

Proof. Take on each Γ_k a point y^k such that $f(y^k) = \alpha$. By hypothesis the sequence $\{y^k\}$ is bounded. Therefore, by taking a subsequence if necessary, we may assume that the sequence $\{y^k\}$ converges to some y. Because of the continuity of f we then have $f(y) = \alpha < f(x^o)$. Since $y \in \Gamma$, it follows from the concavity of f, that f is unbounded below on Γ (see e.g. [8], Corollary 32.3.4).

LEMMA 4.2. *Under the hypothesis of Theorem 4. 1, if the scheme generates an infinite sequence* $\{x^k\}$, *then this sequence is bounded and any cluster point of it is an optimal solution to (P).*

Proof. Suppose that the sequence $\{x^k\}$ is unbounded. Then there is a subsequence $\{x^{k_\nu}\}$ such that $\|x^{k_\nu}\| > \nu$. Let $\widehat{x}$ be a cluster point of the sequence $\frac{x^{k_\nu}}{\|x^{k_\nu}\|}$. Since $f(x^{k_\nu}) < f(x^o)$, the concavity of f implies that f is unbounded below on the halfline from x^o through x^{k_ν}. Hence, by the previous Lemma, f is unbounded below on the halfline from x^o through $\widehat{x}$, and we can find on this halfline a point y such that $f(y) < f(x^1)$. Let U be a ball around y, such that $f(x) < f(x^1)$ for all $x \in U$. Then for all large enough ν, the halfline from x^o through x^{k_ν} meets U at some point y^ν such that $f(y^\nu) < f(x^1) \leqslant f(x^{k_\nu})$. Because of the concavity of f, this implies that x^{k_ν} lies in the line segment $[x^o, y^\nu]$. Consequently, all x^{k_ν} with large enough ν lie in the convex hull of x^o and U. This conflicts with $\|x^{k_\nu}\| > \nu$. Therefore, the sequence $\{x^k\}$ is bounded. But then the same argument as that used in the proof of Theorem 2.1 shows that any cluster point of $\{x^k\}$ is an optimal solution to (P).

LEMMA 4.3. *Under the hypotheses of Theorem 4.1, if the scheme generates an infinite sequence* $\{u^k\}$, *then every cluster point u of this sequence is an optimal direction for (P) (i.e. a direction of recession of D over which f is unbounded below).*

Proof. Denote by Γ_k (Γ) the halfline emanating from x^o in the direction u^k (u, resp.). Suppose that Γ is not entirely contained in D, and let z be the point where it meets ∂D. . Let $u = \lim_{\nu\to\infty} u^{k_\nu}$. It is easily seen that $z^{k_\nu} \to z$. Indeed denoting by φ the gauge of the convex set $D - x^o$, we have, by the continuity of φ:

$$z^{k_\nu} - x^o = \frac{u^{k_\nu}}{\varphi(u^{k_\nu})} \to \frac{u}{\varphi(u)} = z - x^o,$$

since $\varphi(z^{k_\nu} - x^o) = \varphi(z - x^o) = 1$. Now let $h_k(x) = \langle p^k, x - z^k\rangle$ be the affine function defining the supporting hyperplane H_k to D at z^k. Here $p^k \neq 0$ and since p^k is normal to D we may assume $p^k \in \partial g(z^k)$ (see e.g. [8], Corollary 23. 7.1). Therefore $g(x) - g(z^k) \geqslant \langle p^k, x - z^k\rangle$ for all x, and hence, by taking $x = x^o$ an noting that $g(z^k) = 0$:

$$\langle p^k, x^o - z^k\rangle \leqslant g(x^o) < 0. \tag{14}$$

But the sequence $\{z^{k_\nu}\}$ being bounded (because convergent), so must be the sequence $\{p^{k_\nu}\}$ (see e.g. [8], Theorem 24.7). Thus by taking a subsequence if necessary, we may assume $p^{k_\nu} \to p \in \partial g(z)$.

Let $W = [w^1, \ldots, w^n]$ be an $(n-1)$-simplex with barycentre at z and lying in the hyperplane $H = \{x : \langle p, x - z\rangle = 0\}$. Since $\langle p, x^o - z\rangle \leqslant g(x^o) - g(z) < 0$, we can choose $w^1, \ldots, w^n$ so close to z that

$$\langle p, x^o - w^i\rangle < 0 \qquad (i = 1, \ldots, n).$$

Noting that $p^{k_\nu} \to p$, we shall have for large enough ν

$$\langle p^{k_\nu}, x^o - w^i\rangle < 0 \qquad (i = 1, \ldots, n).$$

From this an the inequality (.4) it follows that for each $i = 1, \ldots, n$ there exists $\lambda = \lambda(i) > 0$ such that

$$\langle p^{k_\nu}, x^o + \lambda(w^i - x^o) - z^{k_\nu}\rangle = \langle p^{k_\nu}, x^o - z^{k_\nu}\rangle + \lambda\langle p^{k_\nu}, w^i - x^o\rangle = 0,$$

i.e. $h_{k_\nu}(x^o + \lambda(w^i - x^o)) = 0$. That is, the hyperplane H_{k_ν} cuts all the n half-lines from x^o through $w^1, \ldots, w^n$., resp. But, since $u^{k_\mu} \to u$, for all large enough μ, Γ_{k_μ} lies in the cone K spanned by the n halflines from x^o through $w^1, \ldots, w^n$. Hence Γ_{k_μ} cuts H_{k_ν} and is not contained in the halfspace $\{x : h_{k_\nu}(x) \leqslant 0\}$. On the other hand, for $\mu > \nu$ we have $\Gamma_{k_\mu} \subset S_{k_\mu} \subset S_{k_{\nu+1}} \subset \{x : h_{k_\nu}(x) \leqslant 0\}$. This contradiction shows that the hypothesis that Γ is not contained in D is untenable.

Therefore $\Gamma \subset D$. Since f is unbounded below on each Γ_k, it follows from Lemma 4.1 that f is also unbounded below on Γ. □

Remark 4.2. If the constraints are linear (i.e D is a polyhedral convex set), it is not necessary for starting the algorithm to know an interior point of D. This question will be discussed later, when we specialize the algorithm to the case of linear constraints (see Section 6).

5. SOLVING THE RELAXED PROBLEMS

The relaxed problem in step $k+1$, (Q_{k+1}), differs from that in step k, (Q_k), only by an additional constraint. Therefore, any reasonably efficient method should exploit this fact by using the information obtained in solving (Q_k) for the solution of (Q_{k+1}).

Since a problem (Q_k) is the minimization of the concave function f over a (generalized) polytope S_k, one way to solve it is to take the minimum of f over the finite set V_k consisting of all vertices and all unbounded edges of S_k. For $k = 1$, the polytope S_1 is chosen to be simple enough, so that V_1 can be considered as known or easily computable (usually $D \subset R^n_+$, so one can take $S_1 = R^n_+$ if D is unbounded, or $S_1 = \{ x \in R^n_+ : \Sigma x_j \leqslant c \}$ where c is sufficiently large, otherwise). Thus, the question is now reduced to finding a more or less efficient procedure to derive V_{k+1} from V_k.

Let us first consider the case where D is *bounded*.

The following property is basic for both methods to be presented.

LEMMA 5.1. *Let S be a polytope with vertex set V. Let $h(x) \leqslant 0$ be a new linear constraint, and let V' be the vertex set of the polytope*

$$S' = S \cap \{ x : h(x) \leqslant 0 \}$$

Then any vertex $v \in V' \setminus V$ must be the intersection of the hyperplane $h(x) = 0$ with an edge $[u, w]$ *of S, such that* $h(u) < 0$, $h(w) > 0$.

Proof. Since $v \in V'$, there are among the constraints defining S n linearly independent constraints binding for v. Further, since $v \notin V$, one of these n binding constraints must be $h(v) = 0$. Let $l_i(v) = 0$, $i \in I$, $(|I| = n-1)$ be the remaining $n-1$ binding constraints. Then

$$F = \{ x \in S : l_i(x) = 0, i \in I \}$$

is a face of S, containing v. We cannot have $F = \{v\}$, for then v would be a vertex of S, i.e. $v \in V$. Therefore, $\dim F = 1$, and F is an edge $[u, w]$ of S, such that $v \in [u, w]$, but $v \neq u$, $v \neq w$. The latter implies $h(u) \neq 0$, $h(w) \neq 0$. Since $h(v) = 0$, we must have $h(u) . h(w) < 0$, as was to be proved. □

On the basis of this Lemma, two alternative methods can be used for deriving V_{k+1} from V_k.

METHOD I (Falk-Hoffman [3], [5]).

From Lemma 5.1, it follows that if we denote

$$\tilde{S}_{k+1} = S_k \cap \{h_k(x) \geqslant 0\},$$

then for any point $v \notin V_k$ we have $v \in V_{k+1}$ if and only if $h_k(v) = 0$ and v is a vertex of $\tilde{S}_{k+1}$ neighbouring to some vertex $w \in V_k$ such that $h_k(w) > 0$.

(Indeed, by the previous Lemma v lies on some edge $[u, w]$ of S_k, such that $w \in V$, $h_k(w) > 0$).

Based on this remark, Falk and Hoffman proposed the following procedure for deriving V_{k+1} from V_k.

Take any vertex $w \in V_k$ with $h_k(w) > 0$. Considering w as a vertex of

$$\tilde{S}_{k+1} = S_k \cap \{h_k(x) \geqslant 0\},$$

generate all the vertices v of S_{k+1} that are neighbouring to w and satisfy $h_k(v) = 0$ (this is done by performing pivots on the simplex tableau corresponding to the vertex w of $\tilde{S}_{k+1}$). Perform this procedure for all vertices $w \in V_k$ with $h_k(w) > 0$. Then the set of all different points v that can be obtained in that way are the new elements of V_{k+1} (the old elements being all those elements of V_k that satisfy $h_k(x) \leqslant 0$).

This method requires forming the simplex tableaux corresponding to all the vertices $w \in V_k$ with $h_k(w) > 0$. This may be computationnally expensive, since the simplex tableaux will increase in size as the algorithm progresses.

METHOD II (Thiêu-Tâm-Ban) [9]).

An alternative method utilizes the observation that, by Lemma 5.1, $v \in V_{k+1} \setminus V_k$ if and only if v is the intersection of the hyperplane $h_k(x) = 0$ with an edge $[u, w]$ of S_k such that $h_k(u) < 0$, $h_k(w) > 0$. Therefore, to find all the elements of $V_{k+1} \setminus V_k$ one can consider all pairs $u, w \in V_k$ with $h_k(u) < 0$, $h_k(w) > 0$, such that $[u, w]$ is an edge of S_k (i.e. such that $n-1$ linearly independent constraints are simultaneously binding at u and v). Then for each of these pairs, we take the point v on the line segment $[u, w]$ such that $h_k(v) = 0$.

Since the verification that $[u, w]$ is an edge of S_k (or, equivalently, that v is a vertex of S_{k+1}) may be time-consuming, one can simply consider all pairs $u, w \in V_k$ with $h_k(u) < 0$, $h_k(w) > 0$, such that $n - 1$ constraints are simultaneously binding at u and v, no matter whether these constraints are linear independent or not. Then we may have in general a larger set than V_{k+1}, which will require a bit more computations to perform when finding the minimum of f over S_{k+1}, but the cost to be paid for this may sometimes be less than checking the linear independence of certain systems of constraints.

Turning to the general case where D may be *unbounded*, we observe that any unbounded edge of a polyhedral convex set may be viewed as an edge joining an ordinary vertex to a vertex at infinity. With this observation in mind, the above method can be extended to the unbounded case in the following way.

Without loss of generality we assume $S_1 \subset R^n_+$. To each point $x \in S_1$ let us associate the point $\pi(x) \in R^{n+1}_+$ where the ray from $0 \in R^{n+1}$ through $(x, 1) \in R^{n+1}_+$ meets the n-simplex.

$$\Sigma_o = \{(x, t) \in R^{n+1}_+ : x_1 + \dots + x_n + t = 1\}.$$

To each direction of recession u of S_1 let us associate the point $\pi(u) \in R^{n+1}_+$ where the ray from $0 \in R^{n+1}$ in the direction $(u, 0)$ meets Σ_o. In this way there is a $1 - 1$ correspondence between the points and directions of recession of S_k with the points of the polytope $\pi(S_k) \subset \Sigma_o$.

A vertex (x, t) of $\pi(S_k)$ corresponds to an (ordinary) vertex $\frac{x}{t}$ of S_k if $t > 0$, or to a vertex at infinity (an extreme direction x) of S_k if $t = 0$. Furthermore, if S_k is defined by the inequalities

$$x \in S_1,\ h_i(x) = \langle p^i, x \rangle + \eta_i \leqslant 0 \qquad (i = 1, \dots, k - 1)$$

then $\pi(S_k)$ is defined by the inequalities

$$(x, t) \in \pi(S_1),\ \tilde{h}_i(x,t) = \langle p^i, x \rangle + \eta_i t \leqslant 0 \qquad (i = 1 \dots, k - 1).$$

Denote by W_k the vertex set of $\pi(S_k)$. Since the polytope $\pi(S_{k+1})$ obtains from $\pi(S_k)$ just by adding a new constraint, W_{k+1} can be derived from W_k by the method presented above. Therefore, V_{k+1}, the set of vertices and extreme directions of S_{k+1}, can be computed when V_k and the new constraint are known.

6. CASE OF LINEAR CONSTRAINTS

An interesting feature of the general approach developed above is that when specialized to the case where the constraints are linear, i. e. the set D is a polyhedral convex set, it yields a finite algorithm. In this section we shall discuss this specialized algorithm, which is due to T. V. Thiêu (for the case where D is bounded this algorithm was first developed in [9]).

Consider Problem (P) in which the set D is a polyhedral convex set given by the system:

$$g_i(x) = \langle a^i, x \rangle - \alpha_i \leqslant 0 \qquad (i = 1, \ldots, m) \tag{15}$$

Let us apply Algorithm 2 to this problem, using, for example, Method II for the construction of H_k in each iteration k. Because of the linearity of the constraints, we can avoid the use of an interior point x^0 of D by modifying step b) as follows.

Suppose that u^k is an optimal direction of the relaxed problem (Q_k). Since the recession cone of D is merely the cone

$$\langle a^i, u \rangle \leqslant 0 \ (i = 1, \ldots, m), \tag{16}$$

in order to test whether u^k belongs to this cone it will suffice to look at the inequality:

$$\max \{ \langle a^i, u^k \rangle : i = 1, \ldots, m \} \leqslant 0. \tag{17}$$

If this inequality holds, u^k belongs to the recession cone of D, and hence the algorithm terminates: u^k is a direction of recession of D over which f is unbounded below. Otherwise,

$$\langle a^{i_k}, u^k \rangle = \max_i \langle a^i, u^k \rangle > 0, \tag{18}$$

and so by taking

$$h_k(x) = \langle a^{i_k}, x \rangle - \alpha_{i_k},$$

the inequality $h_k(x) \leqslant 0$ will exclude u^k from the recession cone of the generalized polytope S_{k+1} to be formed in step d).

Further, in step c) if $x_k \notin D$ then

$$g_{i_k}(x^k) = \max_i g_i(x^k) > 0 \tag{19}$$

and the affine function defining the hyperplane H_k to be constructed is

$$h_k(x) = \langle p^k, x - x^k \rangle - g_{i_k}(x^k),$$

with $p^k \in \partial g_{i_k}(x^k)$. But since the functions g_i are affine, $p^k = a^k$, and hence,

$$h_k(x) \equiv g_{i_k}(x). \tag{20}$$

Thus, every function h_k generated during the algorithm coincides with one of the functions g_i that define D (see (15)). But, as will be seen later, all these functions h_k $(k = 1, 2, ...)$ are distinct. Therefore, the algorithm will involve at most m iterations, and no interior point x^0 of D is needed (in the case where g_i are convex x^0 was needed to ensure the convergence of the algorithm when step b) occurs infinitely often).

We are led to the following specialized algorithm for the case of linear constraints.

ALGORITHM 3 (linear constraints)

Select a generalized polytope S_1 enclosing D.

Find the set V_1^1 of vertices and the set V_1^0 of extreme directions of S_1. Set $k = 1$.

a) For every direction $u \in V_k^0$ check whether $f(x)$ is unbounded below over the halfline emanating from x^1 in the direction u (where x^1 is an arbitrary point of S_k). If such an extreme direction, u^k, exists, go to b). Otherwise, go to c).

b) If $\langle a^i, u^k \rangle \leqslant 0$ $(i = 1, ..., m)$, terminate: the problem has no finite optimal solution and u^k is a direction of recession of D over which $f(x)$ is unbounded below.

Otherwise, compute $i_k = \arg\min_i \langle a^i, u^k \rangle$.

Go to d).

c) Find

$$x^k = \arg\min \{f(x) : x \in V_k^1\}.$$

If $\langle a^i, x^k \rangle - \alpha_i \leqslant 0$ for all $i \notin \{i_1, ..., i_{k-1}\}$, terminate: x^k is an optimal solution of the problem. Otherwise, find

$$i_k = \arg\max \{\langle a^i, x^k \rangle - \alpha_i\}$$

Go to d).

d) Form the new generalized polytope

$$S_{k+1} = S_k \cap \{x : \langle a^{i_k}, x \rangle - \alpha_{i_k} \leqslant 0\}.$$

Find the set V^1_{k+1} of vertices and the set V^0_{k+1} of extreme directions of S_{k+1}.

Set $k \leftarrow k + 1$ and return to a).

THEOREM 6. 1. *The above algorithm terminates after at most m iterations.*

Proof. We have

$$S_k = \{x : g_{i_\nu}(x) \leqslant 0\} \qquad (\nu = 1, ..., k - 1)$$

But the functions $g_{i_1}, ..., g_{i_\nu}, ...$ are all different, because if (Q_ν) has an optimal solution x^ν then

$$\langle a^{i_\nu}, x^\nu \rangle - \alpha_\nu > 0, \langle a^{i_\mu}, x^\nu \rangle - \alpha_\mu < 0 \qquad (\nu > \mu),$$

while if (Q) has an optimal direction u, then

$$\langle a^{i_\nu}, u^\nu \rangle > 0, \qquad \langle a^{i_\mu}, u^\nu \rangle \leqslant 0 \qquad (\nu > \mu).$$

(the latter inequality simply expresses the fact that u^ν is an extreme direction of S_ν, hence a recession direction of S_μ for $\nu > \mu$). Therefore, we must terminate after at most m iterations. □

Remark 6. 1. The above Algorithm 3 is quite different from the recent algorithm of V.T. Ban [1]. It also differs from an earlier algorithm of Falk and Hoffman [3], who treated only the case where the constraint set is a polytope. The difference is in at least two points: 1) the method for deriving V_{k+1} from V_k; 2) the subproblem to be solved at the iteration k.

In Algorithm 3, this subproblem is min $\{f(x): x \in S_k\}$, while in Falk-Hoffman's algorithm it is min $\{F_k(x) : x \in D\}$, with $F_k(x)$ being the convex envelope of f over S_k. Since min $\{F_k(x) . x \in S_k\} =$ min $\{f(x): x \in S_k\}$, we have min $\{f(x): x \in S_k\} \leqslant$ min $\{F_k(x) : x \in D\} \leqslant$ min $\{f(x) : x \in D\}$. Thus, the subproblem in Falk-Hoffman's algorithm is in general a finer approximation to (P) than the subproblem in Algorithm 3, but the latter is much easier to solve.

7. CASE OF ONE ADDITIONAL REVERSE CONVEX CONSTRAINT

It turns out that the above method can be extended to certain problems with nonconvex constraint sets. In this section, we proceed to show how the method can be modified to solve Problem (P) when the constraint set D is of the form

$$D = C \setminus G \tag{21}$$

where C is a closed convex set given by

$$C = \{x : g_i(x) \leqslant 0\} \qquad (i = 1,..., m) \tag{22}$$

and G is an *open* convex set:

$$G = \{x : g_{m+1}(\bar{x}) < 0\}, \tag{23}$$

$g_i : R^n \to R$ $(i = 1,..., m + 1)$ being given convex functions.

Note that problems of this kind have been considered earlier, e.g. in [4], [13].

For the sake of simplicity, we shall assume that G is *bounded*. From this assumption it follows that any halfline $\Gamma \subset C$ has an unbounded part in D.

Since the convergence of convexity cuts of the type described in Section 3 is not assured, the idea is to use just the surface $g_{m+1}(\bar{x}) = 0$ to separate any point $x \in G$ from the feasible set D.

The algorithm then looks like the following.

ALGORITHM 4 (one additional reverse convex constraint)

Take a point $x^0 \in \text{int } D$ and a generalized simplex $S_1 \supset C$. Set $k = 1$.

a) Solve the relaxed problem

(Q_k) Minimize $f(x)$, s.t. $x \in S_k \setminus G$.

If an optimal direction u^k is obtained go to b). If an optimal solution x^k is obtained go to **c**).

b) Let Γ_k be the halfline emanating from x^0 in the direction u^k. If $\Gamma_k \subset C$, terminate: the function f is unbounded below over D (indeed, Γ_k has an unbounded part in D). Otherwise, let z^k be the point where Γ_k meets ∂C, the boundary of C. Construct a supporting hyperplane H_k to C at z^k and go to d).

c) If $x^k \in C$ or $f(x^k) = f(x^0)$, terminate: x^k (or x^0, resp.) is an optimal solution to (P). Otherwise construct a hyperplane H_k strictly separating x^k from C and go to d).

d) Let $h_k(x) = 0$ be the equation of H_k, such that $h_k(x^0) < 0$. Form the generalized polytope

$$S_{k+1} = S_k \cap \{x : h_k(x) \leqslant 0\}$$

Set $k \leftarrow k + 1$ and return to a). □

Obviously this Algorithm reduces to Algorithm 2 when $G = \varnothing$ (so that $C = D$).

By exactly the same argument as that used in Section 4, one can easily prove for this Algorithm a convergence theorem identical to Theorem 4.1.

The implementation of this Algorithm requires obviously the availability of a finite method for solving the subproblems (Q_k). Therefore, we now examine how these subproblems can be solved.

First note the following property which is an immediate corollary of Proposition 2 in [13].

LEMMA. 7.1. *The set* $\overline{co}\,(S_k \setminus G)$ *is a polyhedral convex set, whose extreme directions are extreme directions of* S_k *and whose vertices are endpoints of sets of the form* $co\,(E \setminus G)$, *where* E *is any edge of* S_k.

Thus, the extreme directions of $\overline{co}\,(S_k \setminus G)$ are the same as those of S_k, while the vertices of $\overline{co}\,(S_k \setminus G)$ lie on the edges of S_k. Furthermore, G being bounded, any unbounded edge of S_k has an unbounded part in $S_k \setminus G$. From these results and Corollary 1 in [13], we see that every subproblem (Q_k) is equivalent to the following

$$(Q'_k) \qquad \text{Minimize } f(x), \text{ s.t. } x \in S'_k$$

with $S'_k = \overline{co}\,(S_k \setminus G)$. Hence to solve (Q_k) it suffices to investigate the set W'_k of generalized vertices of S'_k i. e. the set W'_k of all vectors $(x, 1)$ or $(u, 0)$ where x is a vertex of S'_k, u an extreme direction of S'_k. If for some $(u^k, 0) \in W'_k$, the function f is unbounded below in the direction u^k then u^k is an optimal direction of S_k, hence an optimal direction of $S_k \setminus G$. Otherwise, let x^k be the vertex of S'_k with smallest value of f : then x^k is an optimal solution of (Q_k).

In this way the question is now reduced to deriving W'_{k+1} from W'_k. From Lemma 7.1 we can derive

LEMMA 7.2. *A point* $x \in S_k$ *is a vertex of* S'_k *if and only if it satisfies either of the following conditions ;*

(*i*) x *is a vertex of* S_k *and* $x \notin G$;

(*ii*) x *is the intersection point of* ∂G *with an edge of* S_k *having one endpoint* (or an unbounded part) *in* G.

Proof. By the previous Lemma, a point $x \in S_k$ is a vertex of S_k if and only if x is an endpoint of a set of the form $co\,(E \setminus G)$, where E is an edge of S_k. But, because of convexity, the only case where co $(E \setminus G)$ does not coincide with E is that in which E has one endpoint (or an unbounded part) in G, and in that case the point where E meets ∂G is just an endpoint of co $(E \setminus G)$. The Lemma then follows immediately from this observation. □

Denote by V_k, V'_k the vertex set of S_k, S'_k resp.

On the basis of this result, one can propose the following procedure for deriving V'_{k+1} from V'_k and V_k, assuming that S_k is bounded (the case of an unbounded S_k can be treated using the method indicated at the end of Section 5).

a) Derive the set V_{k+1} from V_k using Method II of Section 5.

b) For every pair u, w such that $u \in V_{k+1} \setminus V_k$, $w \in V_{k+1}$ and $g_{m+1}(u)\ g_{m+1}(w) < 0$, compute the point v of the line segment $[u, w]$ where $g_{m+1}(v) = 0$.

c) Add to V'_k all $v \in V_{k+1} / V'_k$ such that $g_{m+1}(v) \geqslant 0$, and all v that are obtained from b). □

Remark 7.1. If the functions $g_i(x)$ $(i = 1, \ldots m)$ are affine (i. e. C is a polyhedral convex set) Algorithm 4 is finite, because each function $h_k(x)$ coincides just with one of the functions $g_1(x), \ldots, g_m(x)$. In this case one can also use the following variant of Algorithm 4:

ALGORITHM 5.

Stage 1: Apply Algorithm 3 to the problem without the additional reverse convex constraint, i. e. the problem

Minimize $f(x)$, s. t. $g_i(x) \leqslant 0$ $(i = 1, \ldots m)$ (24)

(here g_i are affine)

If an optimal direction is obtained for this problem, terminate: this is also an optimal direction for (P).

Otherwise, go to Stage 2.

Stage 2: Let S_k be the generalized polytope obtained at the completion of Stage 1.

a) Solve the relaxed problem

$$(Q_k) \qquad \text{Minimize } f(x), \text{ s.t. } x \in S_k \setminus G$$

obtaining an optimal solution x^k. If $g_i(x^k) \leqslant 0$ $(i = 1, \ldots, m)$ terminate: x^k is an optimal solution of (P). Otherwise go to b).

b) Compute $i_k = \arg\max \{g_i(x^k) : i = 1, \dots, m\}$. Let
$$S_{k-1} = S_k \cap \{x : g_{i_k}(x) \leqslant 0\}$$
Set $k \leftarrow k+1$ and return to a).

8. COMBINING THE OUTER APPROXIMATION APPROACH WITH OTHER METHODS

An apparent disadvantage of the outer approximation approach is that the number of vertices of the enclosing polyhedral convex set S_k quickly increases with k. On the other hand, other available methods for solving (P) also have their own difficulties. It seems that a combined approach would enable us to partially circumvent the difficulties inherent to each method when used separately.

Aside from the above outer approximation approach, there are two other approaches to solving (P): the «cone splitting and cutting» method, and the «cone splitting and bounding» method.

In this section we shall examine how these methods can be improved by combining them with outer approximation algorithms.

I. CONE SPLITTING AND CUTTING APPROACH

This approach relies upon the following notion of cut, already mentioned at the end of Section 3.

Let M be a cone in R^n with vertex at x^o and with exactly n edges; let γ be a real number such that $f(x^o) > \gamma$. Then, by definition, the cut $H(\gamma, M)$ is the hyperplanne passing through the n points where the edges of M meet the surface $f(x) = \gamma$; $H^-(\gamma, M)$ is the open halfspace determined by $H(\gamma, M)$ that contain x^o (by convexity, $f(x) \geqslant \gamma$ for all $x \in M \cap H^-(\gamma, M)$).

Now the «cone splitting and cutting» method for solving (P) can be described as follows.

ALGORITHM 6.

Select an interior point x^o of D and n + 1 rays emanating from x^o, such that any n of these rays generate a solid cone in R^n. (by translating if necessary one can assume $x^o = 0$). Let $\mathcal{R}_1$ be the collection of $n+1$ cones constructed this way. For each of the $n+1$ rays constructed find the point where it meets ∂D, the boundary of D. Let x^1 be the best among all these points, let $\alpha_1 = f(x)$. Set $k = 0$.

a) If $\mathcal{R}_k = \emptyset$, stop: x^k is an optimal solution to (P). Otherwise, go to b).

b) Take any $M \in \mathcal{R}_k$ and construct the cut $H(\alpha_k, M)$, then find the point of $D \cap M$ that stands the farthest beyond the cut $H(\alpha_k, M)$. (this amounts to maximizing a linear function over the convex set $D \cap M$). If no such point exists, set $\mathcal{R}_k \leftarrow \mathcal{R}_k \setminus \{M\}$ and return to a). Otherwise, let y be the point found. Go to c).

c) Let $v^1, \ldots, v^n$ be the directions of the n edges of M. Then $y = \sum_{i \in I} \lambda_i v_i$ with $I \subset \{1, \ldots, n\}$, $\lambda_i > 0$ $(i \in I)$. For each $i \in I$ consider the cone M_i whose set of edges obtains from that of M by replacing the edge of direction v^i with the ray of direction y. Let $\mathcal{R}_{k+1} = \mathcal{R}_k \setminus \{M\} \cup \{M_{(i)}, i \in I\}$, $\alpha_{k+1} = \min\{f(y), \alpha_k\}$, $x^{k+1} = \text{argmin}\{f(x^k), f(y)\}$. Set $k \leftarrow k+1$ and return to a). □.

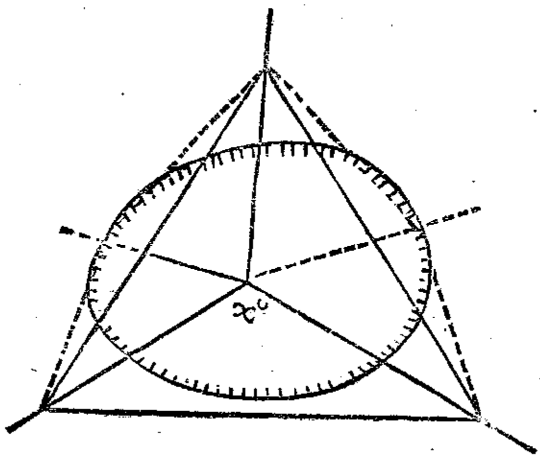

Fig. 3

If D is a polyhedral convex set, then in the above algorithm one can start from a nondegenerate vertex x^o of D (preferably, a local minimum of f over D). Then $\mathcal{R}_1$ consists of a single cone, which is the cone vertexed at x^o, with exactly n edges passing through the n vertices of D neighbouring to x^o. In that case, Algorithm 6 has been proved to be convergent [6]. It is conjectured that in the general case we are considering, the convergence is still ensured. Anyway, the cardinality of $\mathcal{R}_k$, i. e. the number of cones to be explored, may quickly increase. To circumvent this difficulty, we can modify the algorithm as follows.

Stage 1. Apply Algorithm 6, with a) modified in the following way:

a) If $\mathcal{R}_k = \emptyset$, stop: x^k is an optimal solution to (P). If $0 < |\mathcal{R}_k| < N$ (where N is a prescribed integer), go to b). If $|\mathcal{R}_k| \geq N$, go to Stage 2.

Stage 2.

a) Take any $M \in \mathcal{R}_k$. Apply Algorithm 2 to the subproblem

Minimize $f(x)$, s.t. $x \in D \cap M \setminus H^-(\alpha, M)$ (25)

If the optimal value of this subproblem is $\geqslant \gamma_k$, go to b). Otherwise, let $\bar{x}$ be an optimal solution of the subproblem. Let $x^{k+1} = \bar{x}$, $\gamma_{k+1} = f(x^{k+1})$. Set $k \leftarrow k+1$ and go to b).

b) If $\mathcal{R}_k \setminus \{M\} = \emptyset$, stop: x^k is an optimal solution to (P). Otherwise, set $\mathcal{R}_k \leftarrow \mathcal{R}_k \setminus \{M\}$ and return to a).

(Note: In the process of solving the subproblem (25) by Algorithm 2, as soon as an approximate solution is obtained which corresponds to a value of f at least equal to γ_k then stop solving the subproblem and go to b)).

II. « CONE SPLITTING AND BOUNDING » APPROACH

In this approach (see [12]) two basic operations are used:

1) Cone splitting. Given a cone M vertexed at x^0 and having exactly n edges passing through $v^1, \ldots, v^n$ we select a longest edge $[v^i, v^j]$ of the simplex $[v^1, \ldots, v^n]$(*). Let u be the midpoint of this edge, and denote by $M_{(i)}$ ($M_{(j)}$, resp.) the cone whose set of edges obtains from that of M by replacing the edge through v^i (through v^j, resp.) with the ray from x^0 through u. Then M is split into $M_{(i)}$, $M_{(j)}$.

2) Bounding. Given a cone M vertexed at $x^0 \in D$, and having exactly n edges, such that any of these edges intersects D in a nondegenerate line segment, we associate with M a number $\mu(M) \leqslant \min\{f(x) : x \in D \cap M\}$ which is defined as follows.

(i) If for every edge of M we have $f(x) \geqslant f(x^0)$ for all x on this edge, then set

$$\mu(M) = f(x^0) \tag{26}$$

(ii) Otherwise, there is at least one edge of M such that f is unbounded below on this edge. If this edge lies entirely in D, the algorithm stops. Assuming this is not the casé, let $[x^0, z]$ be the intersection of D with this edge. Then there is a constraint g_i such that

$$g_i(x^0) < 0, \quad g_i(z) = 0.$$

(*) We assume that all v^i lie in a hyperplane which is kept fixed throughout the Algorithm.

Select any $l \in \partial g_i(z)$ and consider any polyhedral convex set S such that

$$D \cap M \subseteq S \subseteq \{ x \in M : \langle l, x-z \rangle \leqslant 0 \}.$$

Then set

$$\mu(M) = \min \{ f(x) : x \in S \}.$$

(Note: if M has been obtained by splitting some cone M' (for which $\mu(M')$ has been computed), and if $\mu(M') > \min \{f(x) : x \in S\}$ then we let $\mu(M) = \mu(M')$).

The «cone splitting and bounding» algorithm can now be described.

ALGORITHM 7.

Start with a collection $\mathcal{M}_1$ of cones vertexed at x^0 such that:

1) Each cone $M \in \mathcal{M}_1$ has exactly n edges, any of which intersects D in a nondegenerate line segment.

2) Any two of the cones of $\mathcal{M}_1$ intersect in a joint face.

3) The union of all these cones covers D.

For each edge of the cones, find its intersection point with ∂D. Let x^1 be the best among all these points. Let $\gamma_1 = f(x^1)$.

Compute $\mu(M)$ for each $M \in \mathcal{M}_1$, Set $k = 1$.

a) Delete all $M \in \mathcal{M}_k$ with $\mu(M) \geqslant \gamma_k$. Let $\mathcal{R}_k$ be the collection of all remaining cones

If $\mathcal{R}_k = \varnothing$, stop: x^k is an optimal solution. Otherwise, go to b).

b) Let $M_k = \arg \min \{ \mu(M) : M \in \mathcal{R}_k \}$. Split M_k into two subcones and compute $\mu(M)$ for each of these subcones. These splitting and bounding operations generate some new rays contained in M_k (for example, the new edge of the new subcones is such a ray). For each of these newly generated rays find its intersection point $\widehat{x}$ with, ∂D. If for a certain ray this point does not exist (i. e. the ray lies entirely in D) and $f(x) < f(x^o)$ for some x on this ray, stop: the objective function is unbounded below over this ray. Otherwise, let x^{k+1} be the best among x^k and all the points $\widehat{x}$ that correspond to all the newly generated rays. Let $\gamma_{k+1} = f(x^{k+1})$, $\mathcal{M}_{k+1} = (\mathcal{R}_k \setminus \{ M_k \}) \cup \{ M_{k,1}, M_{k,2} \}$. Set $k \leftarrow k + 1$ and return to a).

THEOREM 8.1. *Assume that for some $\alpha < f(x^o)$ the set $\{x : f(x) = \alpha\}$ is bounded. Then the above Algorithm either terminates after finitely many steps,* *with an optimal solution or a halfline in D over which f is unbounded below, or it generates an infinite sequence* $\{x^k\}$ *such that*

$$f(x^k) \searrow \min\{f(x) : x \in D\}.$$

Note that this Theorem improves upon a similar result in [12], where we assumed a stronger hypothesis, namely that the set $\{x : f(x) \geqslant f(x^o)\}$ was bounded (see Lemma 4 of [12]).

For the proof of this Theorem, we need the following Lemmas which have been established in [12] and [10].

LEMMA 8.1. *The splitting operation is exhaustive, in the sense that for any infinite decreasing sequence of cones* $\{M_{k_q}\}$ *(i. e. any sequence such that* $M_{k_{q+1}}$ *is a descendant of* M_{k_q} : $M_{k_{q+1}} \subset M_{k_q}$*), the intersection of all* M_{k_q} *is a ray.*

LEMMA 8.2. *The bounding operation is consistent, in the sense that for any infinite decreasing sequence of cones* $\{M_{k_q}\}$ *whose intersection is a ray not entirely contained in D we have* $\gamma_{k_q} - \mu(M_{k_q}) \to 0$ *as* $q \to \infty$.

LEMMA 8.3. *Suppose* $\gamma^* = \lim_{k\to\infty} \gamma_k > -\infty$. *If the algorithm generates an infinite decreasing sequence of cones* M_{k_q} *whose intersection is a ray* $\Gamma \subset D$, *then the minimum of f over Γ is achieved at the origin* x^o *of Γ.*

LEMMA 8.4. *Assume that for some* $\alpha < f(x^o)$ *the set* $\{x : f(x) = \alpha\}$ *is bounded. Then for every cone* $M \in \bigcup_{p=1}^{\infty} \bigcap_{k=p}^{\infty} \mathcal{R}_k$ *we have*

$$\inf\{f(x) : x \in D \cap M\} \geqslant \gamma^* = \lim_{k\to\infty} \gamma_k.$$

Proof. Since $\inf\{f(x) : x \in D \cap M\} \geqslant \mu(M)$, it suffices to show that if $\gamma^* > -\infty$ then $\mu(M) \geqslant \gamma^*$. Observe that the process being infinite, at least one member of $\mathcal{R}_1$ has infinitely many descendants. Let it be M_{k_1} for some k_1 (i. e. let it be split at some step k_1). Then at least one of the two subcones of M_{k_1} has infinitely many descendants; let it be M_{k_2} for some $k_2 > k_1$. And so on. Thus, there exists an infinite decreasing sequence $\{M_{k_q}\}$. By Lemma 8. 1, the intersection of this sequence of cones is a ray Γ. Suppose for a moment that $\Gamma \subset D$. We claim that there is one q_o such that

$$f(x) \geqslant f(x^0) \tag{27}$$

for any x on any edge of $M_{k_{q_0}}$. Indeed, if this were not so, then for every q the function f would be unbounded below on some edge of M_{k_q}. Then for every q, we could find on some edge of M_{k_q} a point y^q such that $f(y^q) = \alpha$. Since by hypothesis the set $\{y^q\}$ is bounded, it has a cluster point y^*. Obviously, $y^* \in \Gamma$, and since, by the continuity of f, $f(y^*) = \alpha < f(x^0)$, it follows that f would be unbounded below on Γ, contradicting Lemma 8.3. Therefore, if $\Gamma \subset D$ then (27) must hold for all x on every edge of some $M_{k_{q0}}$. This implies $\mu(M_{k_{q_0}}) = f(x^0) \geq \gamma_{k_{q_0}}$ (see (26)), and hence $M_{k_{q_0}} \notin \mathcal{R}_{k_{q_0}}$. Since this contradicts the definition of $M_{k_{q}}$, we conclude that Γ is not contained in D. But then, by Lemma 8.2. $\gamma_{k_q} - \mu(M_{k_q}) \to 0$, i.e. $\mu(_{k_q}) \to \gamma^*$. On the other hand, by the definition of M_{k_q}, we have $\mu(M_{k_q}) \leq \mu(M)$, since $M \in \mathcal{R}_{k_q}$. Therefore, $\mu(M) \geq \gamma^*$, as was to be proved. □

Proof of Theorem 8.1. Lemma 8.4 shows that, under the conditions of Theorem 8.1, the candidate selection (i.e. the selection of the cone to be split in each step) is ultimately complete in the sense of [12]. Since the splitting operation is exhaustive by Lemma 8.1, and the bounding operation consistent by Lemma 8.2, the conclusion follows from the convergence theorem established in [12]. □

As shown in [12], a weakness of Algorithm 7 is that the maximal number of cones to be stored (the maximal cardinality of $\mathcal{R}_k$) may quickly increase with n, the dimension of the space, and that to speed the convergence it is important that the bounds $\mu(M)$ be sufficiently tight. In particular one should try to avoid having $\mu(M) = -\infty$ whenever possible. All this can be done by applying the outer approximation method to the subproblem $\min\{f(x): x \in D \cap M\}$ and by taking as $\mu(M)$ a reasonably good approximate solution to this subproblem. Since $D \cap M$ is a relatively small piece of D, it is expected that solving the corresponding subproblem will not be hard; furthermore, since it is not necessary to attain a high accuracy, only a reasonable number of iterations will suffice.

To prevent an excessive growth of $\mathcal{R}_k$, one can also modify step a) in Algorithm 7 as follows.

a) Delete all $M \in \mathcal{M}_k$ with $\mu(M) \geqslant \gamma_k$. Let $\mathcal{R}_k$ be the collection of all remaining cones.

If $\mathcal{R}_k = \varnothing$, stop: x^k is an optimal solution.

If $0 < |\mathcal{R}_k| < N$, go to b).

If $|\mathcal{R}_k| \geqslant N$, go to c).

c) Let $M_k = \arg\max \{\mu(M) : M \in \mathcal{R}_k\}$. Apply Algorithm 2 to the problem

$$\text{Minimize } f(x), \text{ s.t. } x \in D \cap M_k. \tag{28}$$

If at some iteration, an optimal solution to the current relaxed problem is obtained with the corresponding function value $\geqslant \gamma_k$, then delete M_k. Set $\mathcal{R}_k \setminus \{ M_k^* \} \leftarrow \mathcal{R}_k$ and return to a).

Otherwise, an optimal solution x^{k+1} to (28) is obtained such that $f(x^{k+1}) < \gamma_k$. Then let $\mathcal{M}_{k+1} = \mathcal{R}_k \setminus \{ M_k^* \}$, $\gamma_{k+1} = f(x^{k+1})$. Set $k \leftarrow k+1$ and return to a).

(Note that here and else where in this section, the « feasibility » of apoint is understood within a given error tolerance).

Acknowledgment. The final version of this work has been completed during the visit of the author to the University of Bremen. The author should like to express his thanks to. Prof. D. Hinrichsen and the colleagues in the Forschungsschwerpunkt « Dynamische Systeme » for useful discussions during this stay.

Received April 10, 1983

REFERENCES

1. V.T. Ban. *A finite algorithm for minimizing a concave function under linear constraints and its applications*. Preprint. Institute of Mathematics, Hanoi 1982.
2. R. Carvajal-Moreno. *Minimization of concave functions subject to linear constraints*: Operations Research Center Report ORC 72—3, University of California, Berkeley, 1972.
3. J. Falk, K.R. Hoffman. *A successive underestimation method for concave minimization problems* Math. Oper Res. 1 (1976), 251—259.
4. J. R. Hillestad, S. E. Jacobsen. *Reverse convex programming*. Appl. Math. Optim. 6 (1980), 63—78.
5. K. R. Hoffman. *A method for globally minimizing concave functions over convex sets* Mathematical Programming. 20 (1981), 22—32.
6. S. E. Jacobsen. *Convergence of a Tuy-type algorithm for concave minimization subject to linear inequality constraints*. Appl. Math. Optim. 7. (1981), 1—9.

7. J. E. Jr. Kelley. *The cutting-plane method for solving convex programs.* J. Soc. Indust. Appl. Math. *8 (1960), 703–712.*

8. R. T. Rockafellar. *Convex analysis.* Princeton University Press, *1970.*

9. T. V. Thieu, B. T. Tam, V. T. Ban. *An outer approximation method for globally minimizing a concave function over a compact convex set.* Presented at the IFIP Working Conference on Recent Advances on System Modelling and Optimization. Hanoi, January *1983.*

10. Ng. V. Thoai, H. Tuy. *Convergent algorithms for minimizing a concave function.* Math. Oper. Res. *5 (1980), 556–566.*

11. H. Tuy. *Concave-programming under linear constraints.* Doklady Akad, Nauk *159 (1964), 32–35.* Translated *Soviet Math. 5 (1964), 1437–1440.*

12. H. Tuy. *Global maximization of a convex function over a closed, convex, not necessarily bounded set.* Cahiers de Mathématiques de la Décision. n⁰ *8223,* Université Paris Dauphine, *1982.*

13. H. Tuy. *Global minimization of a concave function subject to mixed linear and reverse convex constraints.* Preprint Institute of Mathematics. Hanoi *1982.*