Polyhedral Annexation, Dualization and Dimension Reduction Technique in Global Optimization

HOANG TUY
*Institute of Mathematics, Hanoi, Vietnam**

(Received: 20 June 1990; accepted: 20 February 1991)

Abstract. We demonstrate how the size of certain global optimization problems can substantially be reduced by using dualization and polyhedral annexation techniques. The results are applied to develop efficient algorithms for solving concave minimization problems with a low degree of nonlinearity. This class includes in particular nonconvex optimization problems involving products or quotients of affine functions in the objective function.

Key words. Polyhedral annexation, dualization, dimension reduction technique, linearly constrained quasiconcave minimization.

1. Introduction

For minimizing a concave function over a polytope several methods have been developed in the past twenty years (see [5] and references therein). From the computational experiments reported in [3] it seems that one of the promising approaches is by polyhedral annexation ([15] and also [5]).

The term "polyhedral annexation" comes from the fact that this method builds up a sequence of expanding polytopes

$$P_1 \subset P_2 \subset \cdots \subset P_k \subset \cdots$$

by "annexing" more and more vertices of D to an initial polytope P_1 until an optimal vertex is identified. Although the basic idea of this method was put forward as early as in [14], an implementable version of it was developed only recently [15]. The new technique introduced in the polyhedral annexation method presented in the latter paper was to associate with the sequence of polytopes P_k the dual sequence of their polars

$$S_1 \supset S_2 \cdots \supset S_k \cdots$$

An advantage of this dualization procedure is that it converts the subproblem of computing the facets of P_k into that of computing the vertices of S_k, which is an

* This work was completed while the author was visiting the Department of Mathematics of Linköping University.

easier problem for which reasonably efficient solution procedures already exist ([4] and also [5]).

The aim of this paper is to demonstrate another significant advantage of this dualization procedure. Namely, it turns out that by dualization the original n-dimensional problem under certain conditions can be transformed into a problem of substantially smaller dimension. This is of course very important, considering the well known "curse of dimensionality" in global optimization.

This work is written under the stimulus of some recent results of Konno and Kuno [6] (followed by [7, 8]) and of Thach and Burkard [13]. In the paper [6] the authors give a surprisingly simple procedure for minimizing the product of two linear functions over a polytope. On the other hand, an efficient method is presented in [13] for treating a related problem, where the product of two linear functions appears in the constraints rather than in the objective function. Both approaches look very attractive, though the techniques used are quite different. While Konno and Kuno convert the original problem into a linear parametric program, Thach and Burkard reduce, by means of dualization, the original problem to a concave minimization problem of only two variables, which, due to its small size, can readily be solved by currently available methods.

In Section 2 we shall review the basic ideas of polyhedral annexation. Next, in Section 3 we shall show how the dualization technique underlying this procedure can be used to reduce the dimensionality of certain classes of global optimization problems. In Section 4 this dimension reduction technique is applied to concave minimization problems with a relatively mild degree of nonlinearity. Finally, Section 5 is devoted to linear multiplicative and fractional programming problems. These form a class of problems for the study of which dualization turns out to be particularly useful.

2. Basic Ideas of Polyhedral Annexation

In this section we review the basic ideas of polyhedral annexation as presented in [15] (see also [5]).

Consider the quasiconcave minimization problem

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

where D is a polytope (bounded polyhedron) in R^n and $f(x)$ is a quasiconcave function defined and continuous on a closed convex set Ω containing D.

As can easily be seen, the core of this problem is the following sub-problem of "transcending the incumbent":

Given a vertex $\bar{x}$ of D which represents the best feasible solution known so far (incumbent), find a point $x \in D$ such that $f(x) < f(\bar{x})$ or else establish that no such point exists, i.e. $\bar{x}$ is a global optimal solution.

Setting $G = \{x \in \Omega: f(x) \geqq f(\bar{x})\}$ we can also restate that subproblem in the following form:

(#P) Check whether $D \subset G$ and if not find a point $x \in D \setminus G$.

Since G is a closed convex set this is a special GCP (Geometric Complementarity Problem) studied in [17].

Without loss of generality we can assume that the origin 0 belongs to D and that $f(0) > f(\bar{x})$, i.e. $0 \in D \cap \operatorname{int} G$.

For solving (#P) the idea of polyhedral annexation [15] is to construct a sequence of expanding polyhedrons

$$P_1 \subset P_2 \subset \cdots P_k \subset \cdots \subset G, \tag{1}$$

such that eventually a polyhedron P_h will be obtained satisfying $D \subset P_h \subset G$, or else a point x^h will be found satisfying $x_h \in D \setminus G$.

The initial polyhedron P_1 is chosen so that $P_1 \subset G$, $0 \in \operatorname{int} P_1$ and the facets of P_1 are known or can readily be computed. Suppose that P_k has been constructed and is given by a system of linear inequalities of the form:

$$v^T x \leqq 1 \ (v \in V_k), \tag{2}$$

where $v \in V_k$ if and only if the hyperplane $v^T x = 1$ passes through a facet of P_k. We shall refer to (2) as the *defining system* for P_k.

For each $v \in V_k$ let $\mu(v) = \max\{v^T x : x \in D\}$. If $\mu(v) \leqq 1$ then $D \subset \{x : v^T x \leqq 1\}$, therefore, if $\mu(v) \leqq 1$ for all $v \in V_k$ then $D \subset P_k \subset G$. Otherwise, let

$$v^k \in \operatorname{argmax}\{\mu(v) : v \in V_k\}, \tag{3}$$

$$x^k \in \operatorname{argmax}\{(v^k)^T x : x \in D\}. \tag{4}$$

If $x^k \not\in G$ then we terminate: x^k solves (#P). Otherwise, we take the intersection $\hat{x}^k$ of the ray $\Gamma(x^k)$ from 0 through x^k with the boundary ∂G of G (if this intersection does not exist, we let $\hat{x}^k$ be the direction of $\Gamma(x^k)$). Now define

$$P_{k+1} = \operatorname{conv}(P_k \cup \{\hat{x}^k\}). \tag{5}$$

To obtain the defining system for P_{k+1} we consider the polar set S_k of P_k:

$$S_k = \{y : y^T x \leqq 1 \ \forall x \in P_k\}.$$

Since $0 \in \operatorname{int} P_k$ it is known that S_k is a polytope containing 0 ([11], Corollary 14.5.1) and from (5) it follows that

$$S_{k+1} = S_k \cap \{y : (\hat{x}^k)^T y \leqq 1^*\}, \tag{6}$$

where the star * indicates that 1 should be replaced by 0 if $\hat{x}^k$ is a direction. Furthermore, by the duality correspondence between facets of a polyhedron and vertices of its polar (see, e.g., [15, 5]) we know that V_k is precisely the set of nonzero vertices of S_k (for brevity we shall simply say "vertex set" to mean the set of nonzero vertices). In view of (6) the vertex set V_{k+1} of S_{k+1} can then be derived from V_k using for example the procedure of Horst–Thoai–de Vries [4, 5]. Then the defining system for P_{k+1} is

$$v^T x \leqq 1 \ (v \in V_{k+1}) .$$

The procedure can now be repeated with P_{k+1} replacing P_k.

In terms of the sequence $S_1 \supset S_2 \supset \cdots$ the above polyhedral annexation procedure can be summarized as follows.

PA ALGORITHM (for solving (#P). Start with a polyhedron P_1 such that $P_1 \subset G$, $0 \in \text{int } P_1$ and let S_1 be the polar of P_1 (the vertex set V_1 of S_1 should be readily available). Set $V'_1 = V_1$, $k = 1$.

Step k.1. For each $v \in V'_k$ solve the linear program

$$\text{LP}(v) \qquad \max\{v^T x : x \in D\} ,$$

obtaining its optimal value $\mu(v)$ and basic optimal solution $x(v)$.

Step k.2. Compute v^k according to (3). Let $x^k = x(v^k)$. If $\mu(v^k) \leqq 1$ then conclude that $D \subset G$. Otherwise continue.

Step k.3. If $x^k \not\subset G$ then terminate: x^k solves (#P). Otherwise find the intersection $\hat{x}^k$ of the ray $\Gamma(x^k)$ with ∂G(if $\Gamma(x^k) \subset G$ then let $\hat{x}^k$ be the direction of $\Gamma(x^k)$) and define S_{k+1} by (6). Compute the vertex set V_{k+1} of S_{k+1} (from knowledge of V_k). Let $V'_{k+1} = V_{k+1} \setminus V_k$. Set $k \leftarrow k + 1$ and return to Step k.1.

THEOREM 1. *The PA Algorithm terminates after finitely many steps.*

Proof. This follows from the fact that each x^k is a vertex of D and there is no repetition in the sequence $\{x^k\}$.

REMARK. The condition $0 \in \text{int } P_1$ is to ensure the boundedness of S_1. An alternative variant of the PA Algorithm allows for a weaker condition:

$$0 \in P_1 , \quad \text{int } P_1 \neq \emptyset ,$$

but then, since S_k may be unbounded, the defining system (2) for P_k is

$$v^T x \leqq 1 * (v \in V_k) ,$$

where V_k now denotes the set of all vertices and extreme directions of S_k and $1*$ stands for 1 if v is a vertex, 0 if v is an extreme direction (accordingly, in Step k.2, $D \subset G$ if $\mu(v^k) \leqq 1*$).

3. Dualization and Dimension Reduction Technique

We now discuss an important feature of the PA Algorithm that makes this procedure very useful for handling certain problems having a low degree of nonlinearity.

Denote by G^* the polar of the convex set G considered above:

$$G^* = \{y : y^T x \leqq 1 \quad \forall x \in G\} .$$

Since $0 \in \operatorname{int} P_1$, $P_1 \subset G$ and the sequence $\{P_k\}$ satisfies condition (1) we have

$$S_1 \supset \cdots \supset S_k \supset \cdots \supset G^* . \tag{7}$$

On the other hand, the function $v \to \mu(v) = \max\{v^T x : x \in D\}$ is convex as the pointwise supremum of a family of linear functions $v \to v^T x$. Therefore, v^k defined by (3) is also a maximizer of the convex function $v \to \mu(v)$ over the polytope S_k. Since the polytopes S_k form a decreasing sequence approximating G^* from the outside, we see that the PA Algorithm can be interpreted as an outer approximation procedure for checking whether or not the maximum of the convex function $\mu(v)$ over G^* is less than or equal to 1 (or equivalently, whether $G^* \subset D^*$). Thus, the PA Algorithm operates basically in the space spanned by G^* rather than in the original space.

This interpretation underlines an obvious potential advantage of the polyhedral annexation approach when the dimension of G^* is smaller than n. Indeed, in this case the original problem in R^n is transformed into a problem of lower dimension. In global optimization this usually results in a significant saving of computational effort.

To be specific assume that in problem (#P) the convex set G contains a cone

$$K = \{x : u^i x \leqq 0 (i \in I)\} , \tag{8}$$

where the system u^i, $i \in I$, contains exactly p linearly independent vectors. Then from the inclusion $K \subset G$ we deduce $G^* \subset K^*$, while from (8) K^* is the convex cone of dimension p generated by the vectors u^i, $i \in I$. Therefore, if $p < n$ the original problem (#P) in R^n has been converted into a problem in a space of reduced dimension.

As is well known from convex analysis, for a given closed proper concave function $f(x)$ with $\operatorname{dom} f = \Omega$ the recession cone and the lineality space of the (nonempty) level set $G = \{x \in \Omega : f(x) \geqq \gamma\}$ are the recession cone and the constancy space of $f(x)$, respectively ([11], Theorem 8.7). Furthermore, the rank of G, i.e. the number dimension G – lineality G (which measures the nonlinearity of G), is just equal to the dimension of G^* ([11], Corollary 14. 6.1). From the above we see that whenever the lineality space of G, i.e. the constancy space of $f(x)$, is not trivial (has at least dimension 1), so that G has not full rank, then the PA approach offers a method for lowering the dimension of the problem ($\neq$P).

For an effective implementation of this dimension reduction technique an important question which we now discuss is the construction of the initial polyhedron P_1. Clearly such a polyhedron must satisfy certain conditions, namely:

(1) $P_1 \subset G$ and $0 \in \operatorname{int} P_1$ (or at least $0 \in P_1 \subset G$ and P_1 has full dimension; see Remark at the end of Section 2);

(2) the polar of P_1 can be described by an explicit system of linear inequalities and its vertex set V_1 is simple and can readily be computed.

Denote by $u^1, \ldots, u^p$ the maximal set of linearly independent vectors among the $u^i, i \in I$, in (8). Then the linear space generated by $u^i, i \in I$, which contains K^*, can be identified with R^p via the isomorphism

$$t \in R^p \leftrightarrow \pi(t) = \sum_{i=1}^{p} t_i u^i . \tag{9}$$

(any $v = \Sigma_{j \in I} \lambda_j u^j$ of this space can be uniquely represented as a linear combination of the $u^1, \ldots, u^p$).

Let $u^0 = -(u^1 + \cdots + u^p)$ and denote by $\hat{u}^i$ the point where the boundary ∂G of $G = \{x \in \Omega : f(x) \geqq f(\bar{x})\}$ meets the ray from 0 through u^i. Define now

$$P_1 = M_1 + L , \tag{10}$$

where $L = \{x : u^i x = 0 \ (i = 1, \ldots, p)\}$ (the lineality space of K) and M_1 is the convex hull of the set $\{\hat{u}^0, \hat{u}^1, \ldots, \hat{u}^p\}$. Let

$$\alpha_{ij} = \langle u^i, \hat{u}^j \rangle (i = 1, \ldots, p ; \quad j = 0, 1, \ldots, p) .$$

PROPOSITION 1. *The polyhedron P_1 defined by* (10) *satisfies*: $P_1 \subset G$, $0 \in \operatorname{int} P_1$ *and its polar is $S_1 = \pi(T_1)$, where T_1 is a p-simplex in R^p determined by the system of inequalities*:

$$\sum_{i=1}^{p} \alpha_{ij} t_i \leqq 1 \ (j = 0, 1, \ldots, p) . \tag{11}$$

Proof. Obviously, $M_1 \subset G$ and $L \subset K$. Since L is a subspace of the recession cone of G it follows that $P_1 = M_1 + L \subset G$. On the other hand, we have $0 \in \operatorname{relint}$ $[u^0, u^1, \ldots, u^p]$, hence $0 \in \operatorname{relint} M_1$. Now any arbitrary point x of R^n can be decomposed as $x = x' + x''$ with $x' \in L$ and $x'' L^\perp$ (orthogonal complement of L). But clearly, $L^\perp$ is nothing but the linear space spanned by M_1 so that $x'' \in \lambda M_1$ for some $\lambda > 0$. Since obviously $x' \in \lambda L$, it follows that $x \in \lambda P_1$. Thus, for any $x \in R^n$ there exists $\lambda > 0$ such that $x \in \lambda P_1$. This shows that $0 \in \operatorname{int} P_1$.

Denote by M_1^* the polar of M_1. Since L is a subspace and $0 \in M_1$ it easily follows that the polar of P_1 is

$$S_1 = M_1^* \cap L^\perp .$$

hence $v \in S_1$ if and only if $v = \Sigma_{i=1}^{p} t_i u^i$ and $\langle v, \hat{u}^j \rangle \leqq 1 \ (j = 0, 1, \ldots, p)$. This is equivalent to saying that $v = \pi(t)$, where t satisfies (11).

It remains to show that the system (11) determines a p-simplex. Observe that any p of the vectors $\hat{u}^0, \hat{u}^1, \ldots, \hat{u}^p$ are linearly independent. Now consider the system of equations

$$\sum_{i=1}^{p} \alpha_{ij} t_i = 1 \ (j = 0, 1, \ldots, p-1) . \tag{12}$$

Let α^i be the p-vector with components $\alpha_{i0}, \alpha_{i1}, \ldots, \alpha_{ip-1}$. If these vectors were linearly dependent, i.e. if there were numbers λ_i, not all zero, such that

$\Sigma_{i=1}^{p} \lambda_i \alpha^i = 0$, then

$$\sum_{i=1}^{p} \lambda_i \alpha_{ij} = 0 \qquad \text{all } j = 0, 1, \ldots, p-1,$$

hence $\langle \Sigma_{i=1}^{p} \lambda_i u^i, \hat{u}^j \rangle = 0$ $(j = 0, 1, \ldots, p-1)$, so that the vector $\Sigma_{i=1}^{p} \lambda_i u^i$ would be orthogonal to all $\hat{u}^j (j = 0, 1, \ldots, p-1)$. Since this vector belongs to the space generated by $u^1, \ldots, u^p$, which is the same as the space generated by $\hat{u}^0, \hat{u}^1, \ldots, \hat{u}^{p-1}$, it would follow that $\Sigma_{i=1}^{p} \lambda_i u^i = 0$, which cannot hold because $u^1, \ldots, u^p$ are linearly independent. Therefore, the vectors $\alpha^i (i = 1, \ldots, p)$ must be linearly independent and the matrix $\alpha_{ij} (i = 1, \ldots, p;\ j = 0, 1, \ldots, p-1)$ must be nonsingular. This implies that the system (12) has a unique solution.

Analogously, for any $j_0 = 0, 1, \ldots, p$ the system

$$\sum_{i=1}^{p} \alpha_{ij} t_i = 1 (j \in \{0, 1, \ldots, p\} \setminus \{j_0\}) \tag{13}$$

has a unique solution. Clearly, this solution yields the j_0-th vertex of the polyhedron T_1. That is, T_1 is a p-simplex (a polytope with exactly $p + 1$ vertices). □

REMARKS. (i) If for certain $j = 0, 1, \ldots, p$ the ray $\Gamma_j = \{\theta u^j : \theta > 0\}$ does not meet ∂G (Γ_j entirely lies in G), then we define $\hat{u}^j$ to be the direction of Γ_j; but in that event the j-th inequality in (11) should have 0, instead of 1, on the right hand side. Alternatively we can set $\hat{u}^j = \theta_\infty u^j$ where θ_∞ is an arbitrarily large positive number.

(ii) The proof shows that the vertices of S_1 can be obtained by solving each of the $p + 1$ systems of equations (13).

In applying the PA Algorithm starting from P_1 (i.e. S_1) note that since each vertex v of S_k is of the form $v = \Sigma_{i=1}^{p} t_i u^i$, where $t = (t_i)$ is a vertex of $T_k = \pi^{-1}(S_k)$, the linear program LP(v) to be solved in Step k.1 is

$$\text{LP}(v) \qquad \max\left\{\sum_{i=1}^{p} t_i \langle u^i, x\rangle : x \in D\right\}.$$

(iii) When the vectors $u^i, i \in tI$, in (8) are linearly independent, it is more convenient to take $P_1 = K$ (then $0 \in P_1$ and int $P_1 \neq \emptyset$).

If a point $w \in -\text{int}\, K$ is available (e.g., w is the unique solution of the system $u^i w = 1, i \in I$) one can also take $P_1 = \hat{w} + K$, where $\hat{w} = \theta w \in G$ $(\theta > 0)$. The polar of P_1 is then $S_1 = \pi(T_1)$, with T_1 being the p-simplex determined by the system of linear inequalities:

$$\sum_{i \in I} t_i \leqq \frac{1}{\theta}, \quad t_i \geqq 0 \ (i \in I).$$

Indeed, since $u^i \hat{w} = \theta u^i w = \theta$ $(i \in I)$, we have that P_1 is the polyhedron

$$u^i x \leqq \theta \ (i \in I),$$

and the result follows.

4. Problems Where the Objective Function Has Not a Full Rank

To illustrate how the above proposed dimension reduction technique can be applied, in this and the next section we shall examine some classes of problems for which this technique seems to be efficient.

Consider first the class of concave minimization problems (P) of the form

$$\text{minimize } f(x) := f_1(y) + d^T z \quad \text{s.t.} \quad Ay + Bz \leqq c\,, \tag{14}$$

where $x = (y, z) \in R^p \times R^q$ $(p + q = n)$, $f_1 : R^p \to R$ is a concave function, $d \in R^q$, $c \in R^m$ and A, B are matrices of orders $m \times p$ and $m \times q$ respectively.

As previously, assume that $0 \in D := \{x = (y, z) : Ay + Bz \leqq c\}$, $f(0) > \gamma$ (so that 0 is interior to the level set $G = \{x : f(x) \geqq \gamma\}$). Moreover, assume that $d \neq 0$ (the case $d = 0$ is simpler). Since for any $y \in R^p$ such that $Ay \leqq 0$ the point $(y, 0) \in R^p \times R^q$ belongs to D (note that $0 \leqq c$ by the assumption $0 \in D$), if the cone $\{y : Ay \leqq 0\}$ has an extreme ray Γ over which $f_1(y)$ is unbounded below then $f(x)$ is unbounded below over the ray $\Gamma \times \{0\} \subset D$ and the problem has no finite optimal solution. Otherwise, if no such ray exists, then by a known property of concave functions ([11], Section 32) we must have $f_1(y) \geqq f_1(0)$ for every y in the cone $\{y : Ay \leqq 0\}$, i.e. $f(x) = f_1(y) + d^T z \geqq f_1(0) + 0 > \gamma$ for every $x = (y, z)$ such that $Ay \leqq 0$, $d^T z \geqq 0$. This shows that G contains the cone

$$K = \{x = (y, z) : Ay \leqq 0\,, \quad d^T z \geqq 0\}\,,$$

whose polar is

$$K^* = \{v = (r, s) \in R^p \times R^q : r = A^T \lambda\,, \quad s = -\alpha d,\ \lambda \in R^m_+\,, \quad \alpha \in R_+\}\,.$$

If $a^1, \ldots, a^h$ are a maximal set of linearly independent rows of A, then $(a^1, 0), \ldots, (a^h, 0)$ and $(0, d)$ are linearly independent vectors of R^n and we have that the lineality space of K is

$$L = \{x = (y, z) : a^i y = 0\ (i = 1, \ldots, h)\,, \quad dz = 0\}\,.$$

Therefore, G^* is contained in the space $L^\perp$ generated by $(a^1, 0), \ldots, (a^h, 0)$ and $(0, d)$. This space can be identified with R^{h+1} via the isomorphism

$$t \in R^{h+1} \leftrightarrow \pi(t) = \sum_{i=1}^{h} t_i(a^i, 0) - t_{h+1}(0, d)\,.$$

Thus the PA Algorithm applied to this problem will operate in a space of dimension $h + 1 \leqq p + 1$ (usually $h + 1$ is much smaller than $p + q$).

The initial polyhedron P_1 (and its polar S_1) can be constructed as indicated in the previous section. However, in the case where the rank of A is exactly p, it is simpler to proceed as follows.

Observe that G contains the cone

$$K = \{x = (y, z) : y = 0\,, \quad d^T z \geqq 0\}\,,$$

since $x \in K$ implies $f(x) \geqq f_1(0) > \gamma$. Therefore, G^* is contained in the orthogonal complement of the space $y = 0$, $d^T z = 0$, i.e. in the space which is the direct sum of R^p and the line $\{\theta d : \theta \in R\}$. By Remark (iii) in the previous section, as initial polyhedron for the PA Algorithm one can then take $P_1 = \hat{w} + K$, where $\hat{w} = \theta w$, $w = (-e, d) \in R^p \times R^q$ $(e = (1, \dots, 1))$ and $\theta > 0$ is chosen so that $f(\hat{w}) = \gamma$.

EXAMPLE 1. As an example consider in more detail the concave quadratic minimization problem

$$\text{minimize } f(x) := a^T x - \frac{1}{2} x^T U x \quad \text{s.t.} \quad x \in D\,, \tag{15}$$

where D is a polyhedron in R^n, $a \in R^n$ and U is a symmetric positive semi-definite matrix of rank $p < n$.

As previously, assumed $0 \in D$ and $\gamma < f(0)$ (e.g. $\gamma = f(\bar{x})$ where $\bar{x}$ is a vertex of D with $f(\bar{x}) < f(0)$).

Since rank $U = p$, using an affine transformation one could convert $f(x)$ into the form $f_1(y) + d^T z$ with $f_1 : R^p \to R$ a concave quadratic function of p variables $y_1, \dots, y_p$ and $z \in R^{n-p}$. Therefore this problem belongs to the class considered above.

In actual practice, however, there is no need to perform this transformation. Proceeding directly, we just observe that the level set $G = \{x : f(x) \geqq \gamma\}$ contains the cone

$$K = \{x : Ux = 0\,, \quad a^T x \geqq 0\}$$

because $x \in K$ implies that $f(x) \geqq 0 > \gamma$. The lineality space of K, i.e. the constancy space of $f(x)$ is $L = \{x : Ux = 0, a^T x = 0\}$. Let $u^1, \dots, u^h$ be the maximal set of linearly independent vectors among the n rows of U and the vector a ($h = p + 1$ if a is linearly independent from the rows of U, $h = p$ otherwise). Then the space $L^\perp$ that contains G^* can be identified with R^h

$$t \in R^h \leftrightarrow \pi(t) = \sum_{i=1}^{h} t_i u^i\,.$$

Let $u^0 = -(u^1 + \cdots + u^h)$ and let $\hat{u}^i = \theta u^i$ be the points such that $f(\theta u^i) = (i = 0, 1, \dots, h)$. Define $P_1 = M_1 + L$, where $M_1 = [\hat{u}^0, \hat{u}^1, \dots, \hat{u}^h]$. By Proposition 1, the polar of P_1 is $S_1 = \pi(T_1)$ with T_1 being the h-simplex

$$\sum_{i=1}^{h} \alpha_{ij} t_i \leqq 1 \; (j = 0, 1, \dots, h)\,, \tag{16}$$

where

$$\alpha_{ij} = \langle u^i, \hat{u}^j \rangle\,.$$

The detailed algorithm for solving (15) reads as follows.

ALGORITHM 2 (for solving (15)). By translating if necessary assume $0 \in D$. Let $\bar{x}$ be a vertex of D such that $\gamma = f(\bar{x}) < f(0)$ ($\bar{x}$ is the current best feasible solution;

if no such $\bar{x}$ is readily available, set $\bar{x}=0$, $\gamma=f(0)-\varepsilon$ where $\varepsilon>0$ is the tolerance, but then an optimal solution is understood within this tolerance).

0. Define T_1 by (16). Set $\mathcal{M}_1 = V_1' = V_1 =$ vertex set of T_1 (computed by solving each of the $h+1$ systems of equations obtained from (16) by omitting just one inequality and setting all the others to equalities). Set $k=1$.

k.1. For each $t \in V_k'$ solve the linear program

$$\mathrm{LP}(t) \qquad \text{maximize } \sum_{i=1}^{h} t_i \langle u^i, x\rangle \quad \text{s.t.} \quad x \in D\,,$$

obtaining its optimal value $\mu(t)$ and basic optimal solution $x(t)$.

k.2. Delete all $t \in V_k'$ for which $\mu(t) \leqq 1$. Let $\mathcal{R}_k$ be the collection of remaining elements of $\mathcal{M}_k$. If $\mathcal{R}_k = \emptyset$ then terminate: $\bar{x}$ is an optimal solution. Otherwise, update $\bar{x}$ and γ, using the $x(t)$, $t \in V_k'$. Go to k.3.

k.3. Let $t^k \in \operatorname{argmax}\{\mu(t): t \in \mathcal{R}_k\}$, $x^k = x(t^k)$. Compute the point $\hat{x}^k$ where the ray from 0 through x^k meets the surface $f(x)=\gamma$ (see Remark (i), Section 3). Define

$$T_{k+1} = T_k \cap \left\{t: \sum_{i=1}^{h} t_i \langle u^i, \hat{x}^k\rangle \leqq 1*\right\}$$

and compute the vertex set V_{k+1} of T_{k+1} (from knowledge of V_k).

Set $V_{k+1}' = V_{k+1} \backslash V_k$, $\mathcal{M}_{k+1} = \mathcal{R}_k \cup V_{k+1}'$, $k \leftarrow k+1$ and go to k.1.

REMARK. In Step k.2, if the new current best feasible solution $\bar{x}$ has improved then one can set

$$D \leftarrow D \cap \{x: l_k(x) \leqq 1\}\,,$$

where $l_k(x) \leqq 1$ is a γ-valid concavity cut for (f, D) (see [5]).

Also in Step k.3, when the vertex set V_{k+1} becomes too numerous, it is advisable to restart the whole procedure after translating the origin to a new vertex of D (but using the current best feasible solution $\bar{x}$, the value γ and the polyhedron D last obtained).

The above method seems to be quite efficient when the rank of u is small. In particular, if rank $U \leqq 5$ then even for fairly large n the procedure should not present difficulties since all the potentially difficult computations are done in at most 6 dimensions.

Concave minimization problems where the objective function has not full rank occur in many contexts. Let us mention two more examples:

EXAMPLE 2. Consider the well known bilinear programming problem:

$$\text{minimize } F(x, y) := h^T x + y^T U x + g^T y \quad \text{s.t.} \quad x \in X\,, \quad y \in Y\,,$$

where $h \in R^n$, $g \in R^m$, $U \in R^{m \times n}$, and X, Y are polyhedrons in R^n, R^m respectively. (see e.g. [5]). This problem is equivalent to the concave minimization problem

$$\text{minimize } f(x) \quad \text{s.t.} \quad x \in X ,$$

where

$$f(x) = h^T x + \inf\{y^T(g + Ux) : y \in Y\} .$$

But it can be seen that the constancy space of this concave function is

$$L = \{x : h^T x = 0, Ux = 0\} ,$$

therefore the dimension of the problem can be reduced if rank $U < n$.

EXAMPLE 3. In certain game situations, one has to minimize a function of the form

$$f(x) = \sup\{d^T y : Ax + By \leqq c\} ,$$

where $d \in R^q$, $A \in R^{m \times n}$, $B \in R^{m \times q}$, $c \in R^m$. To check that this function is concave denote by $R(x)$ the optimization problem whose optimal value gives $f(x)$. For any x', x'' and $0 < \alpha < 1$, let y', y'' solve $R(x')$, $R(x'')$ respectively. Then $\alpha y' + (1 - \alpha)y''$ is feasible to $R(\alpha x' + (1 - \alpha)x'')$, hence $f(\alpha x' + (1 - \alpha)x'') \geqq d^T(\alpha y' + (1 - \alpha)y'') = \alpha f(x') + (1 - \alpha)f(x'')$. The constancy space of this concave function contains $L = \{x : Ax = 0\}$, so if rank $A < n$ then rank $f < n$.

Before closing this section, it is worth noticing another positive feature of the present method in that all the linear subproblems LP(t) have the same constraint set (at least for each cycle of iterations, as long as the polyhedron D is unchanged). Since only the coefficient vector of the linear objective function changes, starting from a basic optimal solution of a LP(t) it is easy to find a basic optimal solution of the next.

5. Linear Multiplicative/Fractional Programming Problems

We now discuss another important class of nonconvex optimization problems amenable to the above dimension reduction technique, namely the so called generalized linear multiplicative programming problems which have the following formulation:

$$\text{(GLMP) minimize } f(x) := \prod_{i=1}^{p} (c_i^T x + d_i)^{\varepsilon_i} \text{ subject to } x \in D ,$$

where $p < n$, $\varepsilon_i \in \{1, -1\}$ and D is a polytope in R^n such that

$$c_i^T x + d_i > 0 \ (i = 1, \ldots, p) \text{ for all } x \in D . \tag{17}$$

Consider the set

$$\Omega = \{x \in R^n : c_i^T x + d_i > 0 \ (i = 1, \ldots, p)\}$$

which clearly is convex, open and contains the feasible region. Then the objective function $f(x)$ can be replaced by

$$F(x) = \log f(x) = F_1(x) - F_2(x) ,$$

with

$$F_1(x) = \sum\nolimits_1 \log(c_i^T x + d_i) , \quad F_2(x) = \sum\nolimits_2 \log(c_i^T x + d_i) ,$$

where Σ_1 indicates the sum over all i such that $\varepsilon_i = 1$ and Σ_2 the sum over all i such that $\varepsilon_i = -1$. Using the fact that $\log(t)$ is a concave and increasing function of t for $0 < t < +\infty$ it can easily be checked that both functions $F_1(x)$ and $F_2(x)$ are concave on Ω. Three cases are possible:

Case 1: $\varepsilon_i = -1$ for all i. Then F_1 is absent and the problem amounts to minimizing the convex function $-F_2(x)$ over D, which is an easy convex minimization problem.

Case 2: $\varepsilon_i = 1$ for all i. Then F_2 is absent and the problem is to minimize the concave function $F_1(x)$ over D.

By translating if necessary, assume that $0 \in D$. For any $\gamma < F_1(0)$ the level set $G = \{x \in \Omega: F_1(x) \geqq \gamma\}$ contains the cone

$$K = \{x: c_i^T x \geqq 0 \ (i = 1, \ldots, p)\} .$$

Indeed, if $x \in K$ then $c_i^T x + d_i \geqq d_i > 0$ for all i, hence $x \in \Omega$. Moreover,

$$F_1(x) = \sum \log(c_i^T x + d_i) \geqq \sum \log d_i = F_1(0) > \gamma ,$$

so that $x \in G$. Therefore, the polar G^* of G is contained in the convex cone K^* of dimension at most p generated by the vectors $-c_1, \ldots, -c_p$. That is, the problem is reduced to one in a space of at most dimension p.

Case 3: $\varepsilon_i = 1$ for certains i and -1 for the others, i.e. both $F_1(x)$ and $F_2(x)$ are present. For $p = 2$, i.e. $f(x) = (c^1 x + d_1)/(c^2 x + d_2)$ the problem has been extensively treated in the literature. For $p > 2$ we can rewrite the problem as the following concave minimization problem:

$$\text{minimize } F_1(x) - y \quad \text{s.t.} \quad x \in D , \quad F_2(x) - y \geqq 0 .$$

Assuming, as usual, that $0 \in D$, $F_2(0) \geqq 0$ (i.e. $x = 0$, $y = 0$ is a feasible solution) and $\gamma < F_1(0)$, we see that the level set

$$G = \{(x, y): x \in \Omega , \quad F_2(x) - y \geqq 0 , \quad F_1(x) - y \geqq \gamma\}$$

contains the cone

$$K = \{(x, y): c_i^T x \geqq 0 \ (i = 1, \ldots, p) , \quad y \leqq 0\}$$

because if $x \in K$ then $c_i^T x + d_i \geqq d_i > 0$ (all i), i.e. $x \in \Omega$, and furthermore, $F_2(x) - y \geqq \Sigma_2 \log(c_i^T x + d_i) \geqq \Sigma_2 \log d_i = F_2(0) \geqq 0$, $F_1(x) - y \geqq \Sigma_1 \log(c_i^T x + d_i) \geqq \Sigma_1 \log d_i = F_1(0) > \gamma$. Therefore, G^* is a convex subset of the convex cone (of dimension at most $p + 1$)

$$K^* = \{(v, s) \in R^n \times R_+ : v = -\sum_{i=1}^{p} \lambda_i c_i \ (\lambda_i \geqq 0 \, \forall i), s \in R_+\}$$

Thus, in this case the problem is reduced to a space of dimension $r + 1$, where r is the rank of the system $c_1, \ldots, c_p$.

REMARK. The more general situation when instead of the condition (17) we only assume $c_i^T x + d_i > 0 \ \forall x \in D$ for each i such that $\varepsilon_i = -1$ can be handled by splitting D into subpolytopes over each of which each function $c_i^T x + d_i$ keeps a constant sign.

EXAMPLE 4. Let us consider in more detail the GLMP problem when all $\varepsilon_i = 1$:

$$\text{minimize } f(x) := \prod_{i=1}^{p} (c_i^T x + d_i) \text{ subject to } x \in D \ , \tag{18}$$

where $c_i \in R^n$ and $c_i^T x + d_i > 0$ $(i = 1, \ldots, p)$ for all $x \in D$ (as usual, D is a polytope in R^n).

For $p = 2$ this problem which has applications in certain fields of economics (see [6, 7]) has been investigated by several authors [1, 6, 7, 8, 10, 12], and also [18], as it came to my knowledge just recently.

As seen above, the objective function can be replaced by $F(x) = \log f(x)$ which is concave on $\Omega = \{x : c_i^T x + d_i > 0 \ (i = 1, \ldots, p)\}$ (by hypothesis Ω contains D). Assuming $0 \in D$ and $\gamma < F(0)$ we have also seen that the polar G^* of the set $G = \{x \in \Omega : F(x) \geqq \gamma\}$ is contained in the space generated by $-c_1, \ldots, -c_p$.

Let $c_1, \ldots, c_h$ $(h \leqq p)$ be a maximal set of linearly independent vectors among $c_1, \ldots, c_p$.

(a) If $h = 1$, then G^* is a line segment (G^* is bounded because $0 \in \operatorname{int} G$). In this case, any $x \in R^n$ can be written as $x = \lambda c_1 + y$ (assuming $c_1 \neq 0$) with $c_1^T y = 0$ and $f(x) = \phi(\lambda)$ with $\phi(\lambda) = \Pi_{i=1}^{p} (\lambda c_i^T c_1 + d_i)$. Hence, the problem reduces to minimizing the function $\phi(\lambda)$ over the segment $[\lambda', \lambda'']$ where $\lambda' = \min\{\lambda : \lambda c_1 + y \in D,\ c_1^T y = 0\}$ and $\lambda'' = \max\{\lambda : \lambda c_1 + y \in D,\ c_1^T y = 0\}$. Since $\log \phi(\lambda)$ is concave in λ the minimum is attained either at λ' or at λ''.

(b) In the general situation when $h > 1$, the space that contains G^* can be identified with R^h via the isomorphism

$$t \in R^h \leftrightarrow \pi(t) = -\sum_{i=1}^{h} t_i c_i \ .$$

If $h < p$, the initial simplex T_1 for the PA Algorithm can be defined by a system of the form (16) described in Section 4 (using Proposition 1). When $h = p$ (i.e. the vectors $c_1, \ldots, c_p$ are linearly independent), T_1 can be constructed in a simpler way, as follows (see Remark (iii), Section 3).

Observe that since $c_1, \ldots, c_p$ are linearly independent we can always find a point w satisfying

$$c_i^T w = -1 \ (i = 1, \ldots, p) \ .$$

Next, noting that $F(0) > \gamma$ and $F(\lambda w) = \Sigma_{i=1}^{p} \log(-\lambda + d_i)$ is a concave decreasing function of λ, we can compute $\theta > 0$ such that $F(\theta w) = \gamma$. Under these conditions, $-w \in \operatorname{int} K$ and $\theta w \in G$, hence if we take $P_1 = \theta w + K$, i.e. $P_1 = \{x : c_i^T x \geqq -\theta$ $(i = 1, \ldots, p)\}$ then $0 \in \operatorname{int} P_1$, $P_1 \subset G$ and the polar of P_1 is $S_1 = \pi(T_1)$, with T_1 being defined by the system of inequalities:

$$\sum_{i=1}^{p} t_i \leqq \frac{1}{\theta}, \quad t_i \geqq 0 \ (i = 1, \ldots, p) . \tag{19}$$

The vertices of T_1 are obviously: $0 \in R^p$, $(1/\theta)e^i$ $(i = 1, \ldots, p)$ where e^i denotes the i-th unit vector in R^p.

We can state the following.

ALGORITHM 3 (for solving (18) when $c_1, \ldots, c_p$ are linearly independent). By translating if necessary assume $0 \in D$. Let $\bar{x}$ be a vertex of D such that $\gamma = F(\bar{x}) < F(0)$.

0. Define T_1 by (19). Set $\mathcal{M}_1 = V_1' = V_1 =$ vertex set of T_1. Set $k = 1$.

k.1. For each $t \in V_k'$ solve the linear program

$$\text{LP}(t) \qquad \text{maximize } -\sum_{i=1}^{p} t_i \langle c_i, x \rangle \quad \text{s.t.} \quad x \in D ,$$

obtaining its optimal value $\mu(t)$ and basic optimal solution $x(t)$.

k.2. Delete all $t \in V_k'$ for which $\mu(t) \leqq 1$. Let $\mathcal{R}_k$ be the collection of remaining elements of $\mathcal{M}_k$. If $\mathcal{R}_k = \emptyset$ then terminate: $\bar{x}$ is an optimal solution. Otherwise, update $\bar{x}$ and γ, using the $x(t)$, $t \in V_k'$. Go to k.3.

k.3. Let $t^k \in \operatorname{argmax}\{\mu(t) : t \in \mathcal{R}_k\}$, $x^k = x(t^k)$. Compute the point $\hat{x}^k = \theta_k x^k$ such that

$$\theta_k = \sup\left\{\lambda : \sum_{i=1}^{p} \log(\lambda c_i^T x^k + d_i) \leqq \gamma\right\} .$$

(see Remark (i) below).
Define

$$T_{k+1} = T_k \cap \left\{t : -\sum_{i=1}^{p} t_i \langle c_i, \hat{x}^k \rangle \leqq 1\right\}$$

and compute the vertex set V_{k+1} of T_{k+1} (from knowledge of V_k).

Set $V_{k+1}' = V_{k+1} \setminus V_k$, $\mathcal{M}_{k+1} = \mathcal{R}_k \cup V_{k+1}'$, $k \leftarrow k + 1$ and go to k.1.

REMARKS. (i) In Step k.3, since $\mu(t^k) > 1$, it follows that $x^k \not\in P_k$, hence $x^k \not\in P_1$, i.e. $\min\{c_i^T x^k : i = 1, \ldots, p\} < -\theta < 0$ and there must exist $\lambda > 0$ such that $\min\{c_i^T(\lambda x^k) + d_i : i = 1, \ldots, p\} = 0$. That is, the ray $\{\lambda x^k : \lambda > 0\}$ must intersect $\partial\Omega$, hence must intersect ∂G. Therefore, $\hat{x}^k$ always exists. When $p = 2$, θ_k can be computed by solving the quadratic equation

$$(\lambda c_1^T x^k + d_1)(\lambda c_2^T x^k + d_2) = e^\gamma .$$

(ii) As with Algorithm 2, the efficiency of the procedure can be enhanced by an appropriate use of concavity cuts (at the completion of Step k.3) and the restart strategy.

(iii) For the case $p = 2$ often studied in the literature, the computation of the sets V_k (which is the heart of the procedure and may set limitation to the method in high dimension) is quite easy (for example, in dimension 2 all vertices are nondegenerate). Therefore, in this special case the above method is remarkably simple. In a subsequent paper we will discuss in detail an implementation based on this property which will show that the procedure could perhaps compete with some existing efficient methods (for $p = 2$) as the recent parametric method of Konno and Kuno [6, 7].

(iv) Since the lineality space of G is $L = \{x: c_i^T x = 0 \ (i = 1, \ldots, p)\}$, if we represent each vector $x \in R^n$ as $x = \Sigma_{i=1}^p t_i c_i + y$, where y satisfies $c_i^T y = 0$ ($i = 1, \ldots, p$), then the problem (18) with linearly independent c_i can be seen to be equivalent to the following problem

$$\min\left\{\Phi(t): \sum_{i=1}^{p} t_i c_i + y \in D, \ c_i^T y = 0 \ (i = 1, \ldots, p)\right\}$$

where $\Phi(t) = \Pi_{i=1}^p [\Sigma_{j=1}^p t_j \langle c_i, c_j \rangle + d_i]$. The objective function of this problem is a quasiconcave function which only depends upon $t \in R^p$. Therefore, this problem could also be solved by a branch and bound algorithm operating basically in R^p (but using simplical partition of R^p), see [5, 16]. However, in this approach the linear subproblems for bounding will also involve $p + 1$ variables (each p-simplex in R^p has $p + 1$ vertices!) and, moreover, will have additional constraints generated by the conditions $c_i^T y = 0$ ($i = 1, \ldots, p$); on the other hand, it is not necessary to assume $c^i x + d_i > 0$ ($i = 1, \ldots, p$) $\forall x \in D$.

Conclusion

The difficulty of a global optimization problem depends to a large extent upon the degree of nonlinearity of the problem data (objective function, constraints). When this nonlinearity is relatively mild (for example, when a concave function to be minimized has a nontrivial lineality space), it is generally possible to transform the problem to a space of smaller dimension than the original one by using an appropriate dualization procedure. In this paper we have restricted ourselves to concave minimization but, as we will show in a subsequent paper, with some effort the method can actually be extended to a much wider class of global optimization problems.

References

1. Aneja, Y. P., V. Aggarwal, and K. P. K. Nair (1984), On a class of quadratic programming, *EJOR* **18**, 62–70.
2. Henderson, J. M. and R. E. Quandt (1871), *Microeconomics*, McGraw-Hill, New York.

3. Horst, R. and N. V. Thoai (1989), Modification, implementation and comparison of three algorithms for globally solving linearly constrained concave minimization problems, *Computing* **42**, 271–289.
4. 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.
5. Horst, R. and H. Tuy (1990), *Global Optimization: Deterministic Approaches*, Springer-Verlag, Berlin-New York.
6. Konno, H. and T. Kuno (1989), Linear multiplicative programming, IHSS Report 89–13, Institute of Human and Social Sciences, Tokyo Institute of Technology.
7. Konno, H. and T. Kuno (1989), Generalized linear multiplicative and fractional programming, IHSS Report 89–14, Institute of Human and Social Sciences, Tokyo Institute of Technology.
8. Konno, H., Y. Yajima, and T. Matsui (1990), Parametric simplex algorithms for solving a special class of nonconvex minimization problems, IHSS Report 90–16, Institute of Human and Social Sciences, Tokyo Institute of Technology.
9. Konno, H. and M. Inori (1988), Bond portfolio optimization by bilinear fractional programming, *J. Oper. Res. Soc. Japan* **32**, 14 3–158.
10. Pardalos, P. M. (1988), On the global minimization of the product of two linear functions over a polytope, preprint, Pennsylvania State University, 1988.
11. Rockafellar R. T. (1970), *Convex Analysis*, Princeton University Press.
12. Swarup K. (1966), Programming with indefinite quadratic function with linear constraints, *Cahier du Centre d'Etudes de Recherche Operationnelle* **9**, 133–136.
13. Thach P. T. and R. Burkard (1989), Reverse convex programs dealing with the product of two linear functions, preprint, Graz Institute of Mathematics, Graz University of Technology.
14. Tuy, H. (1964), Concave programming under linear constraints, *Soviet Math. Doklady* **5**, 1437–1440.
15. Tuy H. (1990), On polyhedral annexation method for concave minimization, in *Functional Analysis, Optimization, and Mathematical Economics*, ed. Lev J. Leifman, Oxford University Press 248–260.
16. Tuy, H. (1991), On nonconvex optimization problems with separated nonconvex variables, *Journal of Global Optimization* (to appear).
17. Tuy, H. and R. Horst (1990), The geometric complementarity problem and transcending stationarity in global optimization, preprint, Institute of Mathematics, Hanoi, 1990. (submitted).
18. Gabasov, R. and F. M. Kirillova (1980), *Linear Programming Methods*, Part 3 (Special problems), Minsk (in Russian).
19. Tuy, H. (1991), The complementary convex structure in global optimization. Proceedings, Conference on Recent Advances in Global Optimization, Princeton (to appear).