On Nonconvex Optimization Problems with Separated Nonconvex Variables

HOANG TUY
Institute of Mathematics, Hanoi, Vietnam

(Received: 20 June 1990; accepted: 22 February 1991)

Abstract. A mathematical programming problem is said to have separated nonconvex variables when the variables can be divided into two groups: $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, such that the objective function and any constraint function is a sum of a convex function of (x, y) jointly and a nonconvex function of x alone. A method is proposed for solving a class of such problems which includes Lipschitz optimization, reverse convex programming problems and also more general nonconvex optimization problems.

Key words. Global optimization, separated nonconvex variables, reverse convex programming, Lipschitz optimization, decomposition, lower linearization, outer approximation, branch and bound, relief indicator method.

Introduction

It is common knowledge that the cost for solving a nonconvex global optimization problem generally increases exponentially as a function of the number of nonconvex variables. Therefore, from the efficiency point of view, it is essential, when dealing with nonconvex problems, to be able to isolate the nonconvexity part in order to treat it separately. Of course, this separation is only provisional and must be suitably adjusted through an "integration" process in order to gradually improve the accuracy and approach the optimal solution in a reliable way. This decomposition strategy is particularly useful when the given problem is so structured that its nonconvexity is concentrated on a small part of the variables, while the total number of variables may be fairly large. Specifically, in many problems of interest encountered in the applications, the variables can be separated into two groups: $x = (x_1, \ldots, x_p)$ and $y = (y_1, \ldots, y_q)$, with p usually small as compared to q, so that the objective function and any constraint function is the sum of a convex function of (x, y) jointly and a nonconvex function of x alone. A typical problem of this class is the following:

$$\min\{f(x): (x, y) \in \Omega,\ g_i(x) \leq h_i(y)(i = 1, \ldots, m)\}\,, \tag{P}$$

where $f: R^p \to R$, $g_i: R^p \to R$ are continuous functions, Ω is a compact convex set, while $h_i: R^q \to R$ are affine functions.

Note that problems with an objective function depending on both x and y can

Journal of Global Optimization **2**: 133–144, 1992.

easily be converted to this form. For example, the problem studied in [15]:

$$\min\{F(x, y) - g(x) : (x, y) \in \Omega\},$$

where $F: R^p \times R^q \to R$ is convex, $g: R^p \to R$ is continuous, and Ω is a polytope, can be rewritten as

$$\text{minimize } t - g(x) \quad \text{s.t.} \quad (x, y) \in \Omega, \quad F(x, y) \leqq t, \quad \alpha \leqq t \leqq \beta,$$

where $\alpha = \min\{F(x, y) : (x, y) \in \Omega\}$ and $\beta = \max\{F(x, y) : (x, y) \in \Omega\}$.

The aim of the present paper is to discuss methods for solving problem (P) that take advantage of the separability of the non-convex variables.

In Part I of the paper we shall treat the problem under an additional assumption on the functions $f(x)$ and $g_i(x)$ which is fulfilled in particular when these functions are concave or Lipschitz. Obviously, reverse convex programming problems (see [5], [7] and the references therein) are most noticeable examples of problems of this class.

The method to be proposed in Part I is one of branch and bound type characterized by two basic ideas. The first idea which was already contained in the works of Kalantari–Rosen [6] and Rosen–Pardalos [9] (see also [5], [7]) is to branch with respect to the nonconvex variables only (however, we use simplicial rather than rectangular subdivision as in the mentioned works – these are concerned only with quadratic objective functions). The second idea which is borrowed from the work of Horst–Thoai–Benson is to use for lower bounding an outer approximation scheme of the convex constraints of the original problem.

Part II of the paper treats the general case when $f(x)$ and $g_i(x)$ are continuous but may not be concave or Lipschitz. Then the problem is to minimize a continuous function over a compact subset of a convex set. It turns out that the relief indicator method ([12], [14]) for minimizing a continuous function over a compact set can be adapted to handle the additional convex constraints.

We assume that the reader is familiar with the concept of branch and bound and the technique of simplicial subdivision as used in many global optimization algorithms developed in recent years (otherwise a detailed presentation of these concepts can be found in [5]).

I. Lipschitz and Reverse Convex Programming

In this part we study problem (P) under an additional assumption which is fulfilled for Lipshitz and reverse convex programming problems.

I.1. A BRANCH AND BOUND METHOD

DEFINITION. A function $f: R^p \to R$ is said to be *locally lower linearizable* if for every p-simplex M contained in a fixed compact subset of R^p one can find an affine function $\phi_M(x)$ such that:

$$\phi_M(x) \leqq f(x) \forall x \in M, \sup_{x \in M}[f(x) - \phi_M(x)] \to 0 \text{ as } \mathrm{diam}(M) \to 0 .$$

(diam(M) denotes the diameter of M, i.e. the length of its longest edge). The function $\phi_M(x)$ is then called a *lower linearization* of $f(x)$ over M.

As established in [16], $f(x)$ is locally lower linearizable if it is concave or Lipschitz. In the former case, $\phi_M(x)$ is the affine function that agrees with $f(x)$ at the vertices of M. In the latter case, as $\phi_M(x)$ one can take the affine function that agrees with $f(x_M) - L\|x - x_M\|$ at the vertices of M, where L is the Lipschitz constant and x_M is an arbitrary point of M.

Now consider problem (P) where: (1) $\Omega = \{(x, y): G(x, y) \leqq 0\}$, with $G: R^p \times R^q \to R$ a convex function; (2) the functions $f(x)$, $g_i(x)$ are locally lower linearizable, with lower linearizations $\phi_M(x)$, $\phi_{iM}(x)$, respectively.

When $f(x)$, $g_i(x)$ are Lipschitz, (P) is a Lipschitz optimization problem; when $f(x)$, $g_i(x)$ are concave, (P) is a reverse convex programming problem. In recent time these problems have been studied by many authors (see the references in [5]). However, the new feature to be dealt with here is the presence of the additional variables y which enter the problem in a convex way.

To take advantage of this specific structure of the problem and of the lower linearizability property we propose for solving (P) a branch and bound method in which branching is performed with respect to the x-variables by subdividing the space into subsets of the form $M \times Y$ with M a p-simplex in R^p and $Y = R^p$. Furthermore, bounding is based on solving, for each simplex M in R^p, a relaxed form of (P) obtained by lower linearizing $f(x)$, $g_i(x)$ $(i = 1, \dots, m)$ and replacing the convex set Ω with an enclosing polyhedron. The latter is updated at each new iteration through an outer approximation process.

Specifically, we start with a p-simplex M_0 in R^p containing the projection of Ω on this space. At iteration k, we already have at hand a linear system:

$$L_j(x, y) \leqq 0 \quad (j \in J_k) \tag{1}$$

with $J_k \subset \{1, 2, \dots, k\}$, such that the polyhedron (1) contains Ω. Also we have a collection of subsimplices of M_0 that remain for investigation. For each simplex M of this collection a lower bound of $f(x)$ over all feasible points (x, y) with $x \in M$ is furnished by the optimal value $\beta(M)$ of the linear program:

$$\min\{\phi_M(x): L_j(x, y) \leqq 0 \ (j \in J_k), \quad \phi_{Mi}(x) \leqq h_i(y)(i = 1, \dots, m), x \in M\} . \tag{LP(M)}$$

If $\beta(M) \geqq \mathrm{CBV}$ (CBV := current best objective function value) then M is discarded from further consideration. As usual, the simplex M_k chosen for branching corresponds to the minimal value of $\beta(M)$. Let (x^k, y^k) be a basic optimal solution of LP(M_k). If $(x^k, y^k) \in \Omega$ then we set $J_{k+1} = J_k$ (so the system (1) is unchanged), otherwise, we set $J_{k+1} = J_k \cup \{k + 1\}$ and construct a cut $L_{k+1}(x, y) \leqq 0$ separating

(x^k, y^k) strictly from Ω, i.e. such that

$$L_{k+1}(x^k, y^k) > 0\,, \quad L_{k+1}(x, y) \leqq 0 \quad \forall (x, y) \in \Omega\,. \tag{2}$$

From the general theory of outer approximation (see [41]) we know how to construct $L_{k+1}(x, y)$ so as to eventually ensure that

$$\overline{\lim_{k\to\infty}}\{(x^k, y^k), \ldots, (x^k, y^k), \ldots\} \subset \Omega \tag{3}$$

($\overline{\lim}$ is the set of cluster points). For example, if we take

$$L_{k+1}(x, y) = \langle u^k, x - x^k\rangle + \langle v^k, y - y^k\rangle + G(x^k, y^k) \tag{4}$$

where (u^k, v^k) is a subgradient of $G(x, y)$ at (x^k, y^k), then (2) holds and (3) will be ensured (Lemma 1 below).

As for the subdivision of M_k, it may follow several alternative rules. To guarantee convergence of the method we use the following (N, ρ)-rule (see [15]):

(N, ρ are prechosen parameters: N is a natural number, $\rho \in (0, 1)$; usually $N \leqq 5$ and ρ is close to 1; $\nu(M)$ denotes the generation index of M computed by setting $\nu(M_0) = 0$, $\nu(M') = \nu(M) + 1$ whenever M' is a son of M).

(N, ρ)-rule. Let $M_k = [s^{k1}, \ldots, s^{k,p+1}]$. If $\max(\|x^k - s^{kr}\| : r = 1, \ldots, p + 1) \leqq \rho\,\mathrm{diam}(M_k)$ and $\nu(M_k)$ is not a multiple of N, then divide M_k with respect to x^k; otherwise, divide it with respect to the midpoint of a longest edge.

In a formal way the algorithm can be stated as follows:

ALGORITHM 1

Initialization.

Take a p-simplex M_0 in R^p containing the projection of Ω on R^p. If a feasible point $(\bar{x}^k, \bar{y}^k)$ is available, let CBV$= f(\bar{x}^k)$; otherwise CBV$= +\infty$. Set $J_1 = \emptyset$, $\mathcal{M}_1 = \mathcal{P}_1 = \{M_0\}$, $k = 1$.

Iteration $k = 1, 2, \ldots$

(1) For each $M \in \mathcal{P}_k$ solve the linear program LP(M) obtaining its optimal value $\beta(M)$. Update CBV and $(\bar{x}^k, \bar{y}^k)$ whenever the optimal solution of a LP(M) is feasible.

(2) Delete all $M \in \mathcal{M}_k$ such that $\beta(M) \leqq$ CBV (and all M that are known to contain no feasible solution). Let $\mathcal{R}_k$ be the collection of remaining simplices. If $\mathcal{R}_k = \emptyset$ then terminate: $(\bar{x}^k, \bar{y}^k)$ solves (P) (if CBV$< +\infty$), or (P) is feasible (if CBV$= +\infty$). Otherwise, go to (3).

(3) Select $M_k \in \mathrm{argmin}\{\beta(M) : M \in \mathcal{R}_k\}$ and subdivide M_k following the (N, ρ)-rule. Denote by $\mathcal{P}_{k+1}$ the partition of M_k.

(4) Let (x^k, y^k) be a basic optimal solution of LP(M_k). If $(x^k, y^k) \in \Omega$ then set $J_{k+1} = J_k$; otherwise, set $J_{k+1} = J_k \cup \{k + 1\}$ and construct $L_{k+1}(x, y)$ according to (4).

(5) Set $\mathcal{M}_{k+1} = (\mathcal{R}_k \setminus \{M_k\}) \cup \mathcal{P}_{k+1}$, $k \leftarrow k + 1$ and return to (1).

I.2. CONVERGENCE

We now prove the convergence of the above algorithm. Assume that the algorithm is infinite.

LEMMA 1. *Condition (3) holds.*

Proof. This follows from the general theory of outer approximation. We give here a direct proof which is very simple. Let $(x^*, y^*) = \lim(x^k, y^k)$ $(k \to \infty, k \in \Delta)$. We have to show that $G(x^*, y^*) \leqq 0$. Since this is obvious if $G(x^k, y^k) \leqq 0$ for infinitely many $k \in \Delta$, we can assume that $G(x^k, y^k) > 0$ (i.e. $(x^k, y^k) \not\in \Omega$) for all $k \in \Delta$ sufficiently large. Then $L_{k+1}(x, y)$ is defined for these k and since for $l > k+1$, (x^l, y^l) solves LP(M) we have $L_{k+1}(x^l, y^l) \leqq 0$, hence, by making $l \to \infty$, $L_{k+1}(x^*, y^*) \leqq 0$. But

$$L_{k+1}(x^*, y^*) = \langle u^k, x^* - x^k \rangle + \langle v^k, y^* - y^k \rangle + G(x^k, y^k) . \tag{5}$$

Since the sequence $\{(x^k, y^k)\}$ is bounded, it follows from a known property of the subgradients of a convex function ([8], Theorem 24.7) that the sequence $\{(u^k, v^k)\}$ is bounded too. Letting then $k \to \infty$ in (5) yields $G(x^*, y^*) = \lim L_{k+1}(x^*, y^*) \leqq 0$, as desired. □

LEMMA 2. *We have* $\mathrm{diam}(M_k) \to 0 (k \to \infty)$.

Proof. This follows from the exhaustiveness of the (N, ρ) subdivision process (see [16]). □

LEMMA 3. *Any cluster point* (x^*, y^*) *of the sequence* $\{(x^k, y^k)\}$ *is feasible to* (P).

Proof. Let $(x^*, y^*) = \lim(x^k, y^k)$ $(k \to \infty, k \in \Delta)$. Since $\phi_{M_k i}(x^k) \leqq h_i(y^k)$ and $\phi_{M_k i}(x)$ is a lower linearization of $g_i(x)$ over M_k, it follows from Lemma 2, by making $k \to \infty$ $(k \in \Delta)$, that $g_i(x^*) \leqq h_i(y^*)$ $(i = 1, \ldots, m)$. On the other hand, $(x^*, y^*) \in \Omega$ by Lemma 1. □

THEOREM 1. *If Algorithm 1 is infinite then any cluster point of the sequence* $\{(x^k, y^k)\}$ *solves* (P).

Proof. Let $(x^*, y^*) = \lim(x^k, y^k)$ $(k \to \infty, k \in \Delta)$. By Lemma 3, (x^*, y^*) is feasible. Furthermore, since $\mathrm{diam}(M_k) \to 0$ by Lemma 2, we have $f(x^k) - \beta(M_k) = f(x^k) - \phi_{M_k}(x^k) \to 0$ by virtue of the property of lower linearization. Hence $f(x^*) = \lim \beta(M_k)$. But for any feasible (x, y) we have $f(x) \geqq \beta(M) \geqq \beta(M_r) \geqq \beta(M_k)$ if x belongs to some M deleted at an iteration $r < k$, while $f(x) \geqq \beta(M) \geqq \beta(M_k)$ if (x, y) belongs to some $M \in \mathscr{R}_k$. In any case, $f(x) \geqq \beta(M_k)$ for all k, hence $f(x) \geqq f(x^*)$. □

REMARKS. (i) The linear program LP(S) can be written in a more convenient form. If $S = [s^1, \ldots, s^{p+1}]$ then LP(S) is equivalent to

$$\min \sum_{r=1}^{p+1} \lambda_r \phi_M(s^r) \quad \text{s.t.} \quad \sum_{r=1}^{p+1} \lambda_r L_j(s^r, y) \leqq 0 \quad (j \in J_k),$$

$$\sum_{r=1}^{p+1} \lambda_r \phi_{M_i}(s^r) \leqq h_i(y) \quad (i = 1, \ldots, m), \quad \sum_{r=1}^{p+1} \lambda_r = 1, \quad \lambda_r \geqq 0 \quad \forall r.$$

(ii) When $f(x)$ and $g_i(x)(i = 1, \ldots, m)$ are *concave* (as in reverse convex programming problems), then $\phi_M(x)$ and $\phi_{Mi}(x)$ agree with $f(x)$ and $g_1(x)$ resp. at the vertices of M. In this case, instead of the (N, ρ)-rule one can use a simpler and more efficient subdivision rule, namely:

(*) If $\nu(M_k)$ is a multiple of N then bisect M_k; otherwise, divide M_k with respect to x^k.

Denote by K the set of all $k = 1, 2, \ldots$ for which $\nu(M_k)$ is not multiple of N. Instead of Lemmas 2, 3 and Theorem 1 we have now:

LEMMA 2′. *Let* $M_k = [s^{k1}, \ldots, s^{k,p+1}]$. *Then* $min(\|x^k - s^{kr}\| : r = 1, \ldots, p + 1\} \to 0$ *as* $k \to \infty$ $(k \in K)$.

Proof. This follows from Theorem 2 in [15]. □

LEMMA 3′. *Any cluster point of the sequence* $\{(x^k, y^k), k \in K\}$ *is feasible to* (P).

Proof. Let $(x^*, y^*) = \lim(x^k, y^k)$ $(k \to \infty, k \in \Delta \subset K)$. In view of Lemma 2′, without loss of generality one can assume that $\|x^* - s^{k1}\| \to 0$ as $k \to \infty$ $(k \in \Delta)$. But $\phi_{M_k i}(x)$ coincides with $g_i(x)$ at the vertices of M_k. Therefore, $\phi_{M_k i}(x^k) - g_i(s^{k1}) \to 0$, i.e. $\phi_{M_k i}(x^k) \to g_i(x^*)$. Since $\phi_{Mi}(x^k) \leqq h_i(y^k)$, we then deduce $g_i(x^*) \leqq h_i(y^*)$ $(i = 1, \ldots, m)$, which, together with the fact $(x^*, y^*) \in \Omega$ (Lemma 1), implies that (x^*, y^*) is feasible. □

THEOREM 1′. *Assume that* $f(x)$ *and* $g_i(x)$ $(i = 1, \ldots, m)$ *are concave. If Algorithm* 1 *with the subdivision rule* (*) *instead of* (N, ρ) *is infinite then any cluster point* (x^*, y^*) *of the sequence* $\{(x^k, y^k), k \in K\}$ *solves* (P).

Proof. Reasoning as in the proof of Lemma 3′ we can show that $\beta(M_k) = \phi_{M_k}(x^k) \to f(x^*)$ as $k \to \infty$ $(k \in K)$. The proof can then be completed just as for Theorem 1. □

II. Continuous Optimization

We now consider problem (P) where $f(x)$ and $g_i(x)$ are assumed only to be continuous.

II.1. REDUCTION TO D.C. OPTIMIZATION

Introducing additional variables t_i we can rewrite (P) as

$$\min\{f(x) : (x, y) \in \Omega, t_i \leqq h_i(y), g_i(x) \leqq t_i, \alpha \leqq t_i \leqq \beta(i = 1, \ldots, m)\},$$

where $\alpha = \min\{g_i(x): i = 1, \ldots, m,\ (x, y) \in \Omega \text{ for some } y\}$, $\beta = \max\{h_i(x): i = 1, \ldots, m,\ (x, y) \in \Omega \text{ for some } y\}$. After a change of notation we thus obtain the problem:

$$\min\{f(x): x \in S,\ G(x, y) \leqq 0\}\,, \tag{Q}$$

where $f: R^n \to$ is a continuous function, S is a compact subset of R^n and $G: R^{n+q} \to R$ is a convex function.

In [12] and [13] a method called the Relief Indicator Method (RIM) has been developed for minimizing a continuous function over a compact set. Of course this method could be applied here to problem (Q) considered as a problem in R^{n+q}. However, this approach is not advisable since it would lead to solve a difficult problem of very large size. A better approach is to consider y as an intermediate variable and apply RIM to the problem in R^n:

$$\min\{f(x): x \in C\},\ C = \{x \in S: \min_y G(x, y) \leqq 0\}$$

Still, this approach may not be the best since it does not take advantage of the convexity of the constraint $G(x, y) \leqq 0$, making no distinction between the latter and the much more difficult constraint $x \in S$.

Below we propose to treat the constraint $G(x, y) \leqq 0$ as an additional convex constraint to the main problem

$$\min\{f(x): x \in S\}\,. \tag{R}$$

For each $\alpha \in (-\infty, +\infty]$ let

$$S_\alpha = \{x \in S: f(x) < \alpha\};\ \tilde{S}_\alpha = \{x \in S: f(x) \leqq \alpha\}\,.$$

and denote by $d(u, A)$ the distance from u to A.

Following [12] we say that a lower semi-continuous function $r(\alpha, u): (-\infty, +\infty] \times R^n \to R$ is a *separator* for the function f over the set S (briefly, for (f, S)) if it satisfies the following conditions:

(i) $0 \leqq r(\alpha, u) \leqq d(u, \tilde{S}_\alpha)$;

(ii) $r(\alpha_1, u) \geqq r(\alpha_2, u)$ whenever $\alpha_1 \leqq \alpha_2$;

(iii) $r(\alpha, u) > 0$ whenever $u \not\in \tilde{S}_\alpha$.

A trivial example of separators is the distance function $d(u, S_\alpha)$ but this separator is of little use since computing it is as difficult as solving the problem itself. More practical separators for the most usual cases have been given in [11].

Assuming that a separator $r(\alpha, u)$ for (f, S) is available whose value at any given point can be computed without difficulty, let us define the function

$$h(\alpha, x) = \sup\{r^2(\alpha, u) + 2\langle x, u\rangle - \|u\|^2: u \not\in S_\alpha\} \tag{6}$$

Clearly, for α fixed, $h(\alpha, x)$ is a closed convex function as the pointwise supremum of a family of affine functions. Consider now the following d.c. optimization problem which depends on the parameter α:

$$\min\{h(\alpha, x) - \|x\|^2 : G(x, y) \leqq 0\} . \tag{Q_α}$$

Assume the regularity condition:

$$\min(Q) = \inf\{f(x) : x \in \operatorname{int} S,\ G(x, y) \leqq 0\} . \tag{7}$$

THEOREM 2. (1) *If the optimal value of* (Q_α) *is positive then* $f(x) > \alpha$ *for all feasible solutions* (x, y) *to* (Q).

(ii) *If the optimal value of* (Q_α) *is negative then any* (x, y) *feasible to* (Q_α) *such that* $h(\alpha, x) - \|x\|^2 < 0$ *yields a feasible solution* x *to* (Q) *such that* $f(x) < \alpha$.

(iii) *If the optimal value of* (Q_α) *is 0 then any optimal solution* (x^*, y^*) *of* (Q_α) *yields an optimal solution* x^* *of* (Q).

Proof. (i) Suppose that the optimal value of (Q_α) is positive and consider any feasible solution (x, y) to (Q). Then $h(\alpha, x) - \|x\|^2 > 0$, hence from (6), there is a $u \not\in S_\alpha$ such that:

$$r^2(\alpha, u) + 2\langle x, u\rangle - \|u\|^2 > \|x\|^2 .$$

Consequently, $d^2(u, S_\alpha) \geqq r^2(\alpha, u) > \|x - u\|^2$. This implies $x \not\in \tilde{S}_\alpha$, for otherwise one would have $d(u, S_a) \leqq \|x - u\|$. Since $x \in S$ it follows that $f(x) > \alpha$.

(ii) Suppose that the optimal value of (Q_α) is negative and consider any (x, y) feasible to (Q_α) such that $h(\alpha, x) - \|x\|^2 < 0$. The latter inequality implies that if $x \not\in S_\alpha$ then $r^2(\alpha, x) + 2\langle x, x\rangle - \|x\|^2 < \|x\|^2$, i.e. $r^2(\alpha, x) < 0$. Since this is impossible, we must have $x \in S_\alpha$, i.e. $x \in S$, $f(x) < \alpha$.

(iii) Suppose that the optimal value of (Q_α) is 0 and consider any optimal solution (x^*, y^*) of (Q_α). Then for all $u \not\in S_a$ we have

$$r^2(\alpha, u) + 2\langle x^*, u\rangle - \|u\|^2 \leqq \|x^*\|^2, \text{ i.e. } r^2(\alpha, u) \leqq \|x^* - u\|^2 ,$$

hence, $x^* \in S_\alpha$, for otherwise, putting $u = x^*$ in the above inequality would yield $r^2(\alpha, x^*) \leqq 0$ and this would imply by virtue of property (iii) of a separator, $x^* \in S_\alpha$. Therefore, $x^* \in S$ and $f(x^*) \leqq \alpha$. On the other hand, if (Q) has a feasible solution (x, y) with $f(x) < \alpha$ then from the regularity condition (7) there exists a feasible solution (x', y') to (Q) such that $x' \in \operatorname{int} S_\alpha$, so that we can find a ball of centre x' and radius $\delta > 0$, contained in S_α. This implies that for all $u \not\in S_\alpha$: $d^2(u, S_\alpha) \leqq \|x' - u\|^2 - \delta^2$, hence $r^2(a, u) + 2\langle x', u\rangle - \|u\|^2 \leqq \|x'|^2 - \delta^2$, i.e. $h(\alpha, x') - \|x'\|^2 \leqq -\delta^2 < 0$, a contradiction. Therefore, $f(x^*) = \alpha$ and x^* is an optimal solution of (Q). □

II.2. GENERALIZED RELIEF INDICATOR METHOD

Theorem 2 reduces problem (Q) to solving the parametric d.c. optimization problem (Q_α), or more precisely, to finding α such that $\min(Q_\alpha) = 0$.

In the absence of the convex constraint $G(x, y) \leqq 0$ this problem was solved in [12] by an outer approximation method. Below we present a branch and bound procedure that extends the algorithm in [14].

We rewrite (Q_α) in the form of a concave minimization problem:

$$\min\{t - \|x\|^2 : h(\alpha, x) \leqq t,\ G(x, y) \leqq 0\}\,. \tag{8}$$

ALGORITHM 2

Initialization.
Take an n-simplex M_0 in R^n that is known to contain an optimal solution of (Q). If a feasible solution (x^0, y^0) is available then let $\alpha_0 = f(x^0)$; otherwise, let $\alpha_0 = +\infty$. Set $I_1 = J_1 = \emptyset$, $\mathcal{M}_1 = \mathcal{P}_1 = \{M_0\}$, $k = 1$.

Iteration $k = 1, 2, \ldots$

(1) For each $M \in \mathcal{P}_k$ solve the linear program

$$\min\{t - \phi_M(x) : L_i(x) \leqq t (i \in I_k)\,,\quad H_j(x, y) \leqq 0\ (j \in J_k), x \in M\}\,, \qquad \mathrm{LP}(M)$$

where $\phi_M(x)$ denotes the affine function that agrees with $\|x\|^2$ at the vertices of M. Let $\beta(M)$ be the optimal value and $(x(M), y(M), t(M))$ a basic optimal solution of LP(M).

(2) In $\mathcal{M}_k$ delete all M such that $\beta(M) > 0$. Let $\mathcal{R}_k$ be the collection of remaining simplices. If $\mathcal{R}_k = \emptyset$, terminate: (Q) is infeasible. Otherwise, go to (3).

(3) Select $M_k \in \mathrm{argmin}\{\beta(M) : M \in \mathcal{R}_k\}$. Let $x^k = x(M_k)$, $y^k = y(M_k)$, $t^k = t(M_k)$.

If $\beta(M_k) = 0$, terminate: if $\alpha_{k-1} < +\infty$ then (x^{k-1}, y^{k-1}) solves (Q), if $\alpha_{k-1} = +\infty$ then (Q) is infeasible. If $\beta(M_k) < 0$ then go to (4).

(4) Set

$$\alpha_k = \begin{cases} \alpha_{k-1} \text{ if } (x^k, y^k) \text{ is infeasible to (Q);} \\ \min\{\alpha_{k-1}, f(x^k)\} \text{ otherwise,} \end{cases}$$

and define (x^k, y^k) such that $f(x^k) = \alpha_k$.

If $\|x^k\|^2 \leqq t^k$ and $G(x^k, y^k) \leqq 0$, then set $I_{k+1} = I_k$, $J_{k+1} = J_k$.

Otherwise, if $\|x^k\|^2 - t^k > G(x^k, y^k)$ then set $J_{k+1} = J_k$,

$$I_{k+1} = I_k \cup \{k\},\ L_k(x) = r^2(\alpha_k, x^k) + 2\langle x, x^k\rangle - \|x^k\|^2\,.$$

If $\|x^k\|^2 - t^k \leqq G(x^k, y^k)$ then set $I_{k+1} = I_k$,

$$J_{k+1} = J_k \cup \{k\},\ H_k(x, y) = \langle u^k, x - x^k\rangle + \langle v^k, y - y^k\rangle + G(x^k, y^k)\,,$$

where (u^k, v^k) is a subgradient of $G(x, y)$ at (x^k, y^k).

(5) Subdivide M_k following the rule (*). let $\mathcal{P}_{k+1}$ be the partition of M_k.

(6) Set $\mathcal{M}_{k+1} = (\mathcal{R}_k \setminus \{M_k\}) \cup \mathcal{P}_{k+1}$, $k \leftarrow k + 1$ and return to (1).

The convergence of the above algorithm can be established by the same method as that used for the algorithm in [14] (which treats the problem when $G(x, y) = 0$, i.e. in the absence of the convex constraint).

As before let K denote the set of all natural numbers that are not multiple of N.

LEMMA 4. *$\phi_{M_k}(x^k) - \|x^k\|^2 \to 0$ as $k \to \infty$ $(k \in K)$.*

Proof. Let $M_k = [s^{k1}, \dots, s^{k,n+1}]$. Without loss of generality we can assume that $\|x^k - s^{k1}\| \to 0$ (see Lemma 2′). Then arguing as in the proof of Lemma 3′ we see that $\phi_{M_k}(x^k) - \|x^k\| = \phi_{M_k}(x^k) - \|s^{k1}\| + [\|s^{k1}\| - \|x^k\|] \to 0$ as $k \to \infty$ $(k \in K)$. □

LEMMA 5. (1) *$t^k = \max\{L_i(x^k) : i \in I_k\}$;*

(2) *$L_k(x^k) > t^k$, $L_k(x^l) \leqq t$ whenever L_k is defined and $l > k$; similarly, $H(x^k, x^k) > 0$ and $H_k(x^l, y^l) \leqq 0$ whenever H_k is defined and $l > k$.*

(3) *For any convergent subsequence $(x^k, y^k, t^k) \to (x^*, y^*, t^*)$ $(k \to \infty, k \in \Delta)$ we have $t^* - \|x^*\|^2 = 0$, $G(x^*, y^*) = 0$.*

Proof. The first point follows from the fact that (x^k, y^k, t^k) solves LP(M_k). To see the second point, observe that when L_k is defined, we must have $\|x^k\|^2 < t^k$, so that

$$L_k(x^k) = r^2(\alpha_k, x^k) + 2\langle x^k, x^k\rangle - \|x^k\|^2 \geqq \|x^k\|^2 > t^k .$$

Similarly, when $H_k(x, y)$ is defined, we must have $H_k(x^k, y^k) = G(x^k, y^k) > 0$. The third point is obvious if $\sigma_k := \max\{\|x^k\|^2 - t^k;\ G(x^k, y^k)\} \leqq 0$ for infinitely many $k \in \Delta$. Therefore, it suffices to consider the case when $\sigma_k < 0$ for all but finitely many $k \in \Delta$. Suppose that $\sigma_k = \|x^k\|^2 - t^k$ for infinitely many $k \in \Delta$. For $l > k$ we have $L_k(x^l) \leqq t$, hence fixing k and letting $l \to +\infty$ we obtain $L_k(x^*) \leqq t^*$. Then

$$\begin{aligned} 0 \leqq r^2(\alpha_k, x^k) &= L_k(x^*) - 2\langle x^k, x^*\rangle + \|x^k\|^2 \\ &\leqq t^* - 2\langle x^k, x^*\rangle + \|x^k\|^2 \to t^* - \|x^*\|^2 , \end{aligned}$$

which shows that $\|x^k\|^2 - t^k \to \|x^*\|^2 - t^* = 0$, and hence $\sigma^k \to 0$. Since $\sigma_k \leqq G(x^k, y^k) \leqq 0$, it follows that $G(x^k, y^k) \to G(x^*, y^*) = 0$.

If $\sigma_k = G(x^k, y^k)$ for infinitely many $k \in \Delta$, then by the same argument as that used for the proof of Lemma 1, we can see that $G(x^k, y^k) \to G(x^*, y^*) = 0$, hence $\sigma^k \to 0$, which in turn implies $\|x^k\|^2 - t^k \to \|x^*\|^2 - t^* = 0$. □

THEOREM 3. *If the algorithm is infinite then $\alpha_k \to \alpha^* = \min(Q)$ and any cluster point of the sequence $\{(x^k, y^k, t^k)\}$, $k \in K$, yields an optimal solution (x^*, y^*) of (Q).*

Proof. Let $(x^*, y^*, t^*) = \lim(x^k, y^k, t^k)$ $(k \to \infty, k \in \Delta)$. Since $\beta(M_k) = [t^k - \|x^k\|^2] - [\phi_{M_k}(x^k) - \|x^k\|^2]$ it follows from Lemmas 1 and 5 that $\beta(M_k) \to 0$. Now let $\alpha^* = \lim \alpha_k$. Noting that $r(\alpha_k, u) \leqq r(\alpha^*, u)$, we can write $L_k(x) \leqq r^2(\alpha^*, x^k) + 2\langle x, x^k\rangle - \|x^k\|^2 \leqq h(\alpha^*, x) \forall k, \forall x$, hence

$$\beta(M_k) \leqq \min\{t - \|x\|^2 : h(\alpha^*, x) \leqq t,\ G(x, y) \leqq 0\} = \inf(\mathrm{Q}_{\alpha^*}) .$$

Therefore, $\inf(Q_{\alpha*}) \geqq 0$. If $\inf(Q_{\alpha*}) > 0$ then by Theorem 2, $f(x) > \alpha^*$ for all feasible solutions to (Q), which implies, on the one hand, that $\alpha^* < +\infty$, on the other hand, that $f(x^k) > \alpha^*$, which is not the case since $f(x^k) = \alpha_k \downarrow \alpha^*$. Consequently, $\inf(Q_{\alpha*}) = 0$, and the conclusion follows from Theorem 2. □

Note that since the objective function of (Q_α) is quadratic separable, rectangular subdivision can be used instead of simplicial subdivision (see [15]).

Conclusion

In this paper we have considered a general method for handling a class of nonconvex mathematical programming problems where the nonconvexity is restricted to a relatively small number of variables and can be treated separately. This class includes all problems of the form

$$\text{minimize } F(x, y) + f(x) \quad \text{s.t.} \quad (x, y) \in \Omega,\ g(x) \leqq h(y),$$

where $F: R^p \times R^q \to R$ is a convex function, $f: R^p \to R$ is a continuous (in particular, concave or Lipschitz) function, Ω is compact convex set, $g: R^p \to R^n$ is continuous (in particular, concave or Lipschitz) mapping, $h: R^q \to R^m$ is an affine mapping. Indeed, the just stated problem is equivalent to minimizing $t + f(x)$ subject to $(x, y) \in \Omega$, $g(x) \leqq h(y)$, $F(x, y) \leqq t$. Thus, the method applies in fact to a wide class of nonconvex problems of large size (but with few nonconvex variables) encountered in the applications. Computational experiments to test the efficiency of this approach are under way and will be reported subsequently. Also we shall show in a subsequent paper how the method developed above can be applied to solve convex two-level optimization problems of the kind considered in [1].

References

1. J. F. Bard (1988), Convex Two-Level Optimization, *Mathematical Programming* **40**, 15–27.
2. H. B. Benson and R. Horst (1991), A Branch and Bound-Outer Approximation Algorithm for Concave Minimization over a Convex Set, *Journal of Computers and Mathematics with Applications* (forthcoming).
3. R. Horst (1988), Deterministic Global Optimization with Partition Sets Whose Feasibility Is Not Known. Application to Concave Minimization, Reverse Convex Constraints, D.C. Programming and Lipschitzian Optimization, *Journal of Optimization Theory and Applications* **58**, 11–37.
4. R. Horst, N. V. Thoai, and H. P. Benson (1990), Concave Minimization via Conical Partitions and Polyhedral Outer Approximation, *Mathematical Programming*, forthcoming.
5. R. Horst and H. Tuy (1990), *Global Optimization: Deterministic Approaches*, Springer-Verlag.
6. B. Kalantari and J. B. Rosen (1987), An Algorithm for Global Minimization of Linearly Constrained Concave Quadratic Functions, *Mathematics of Operations Research* **12**, 544–561.
7. P. M. Pardalos and J. B. Rosen (1987), Constrained Global Optimization: Algorithms and Applications, Lecture Notes in Computer Science 268, Springer.
8. R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton, N. J.

9. J. B. Rosen and P. M. Pardalos (1986), Global Minimization of Large Scale Constrained Concave Quadratic Problems by Separable Programming, *Mathematical Programming* **34**, 163–174.
10. P. T. Thach (1982), Concave Minimization Under Nonconvex Constraints with Special Structure, in *Essays in Nonlinear Analysis and Optimization Problems*, Institute of Mathematics, Hanoi, pp. 121–139.
11. P. T. Thach (1990), Convex Minimization under Lipschitz Constraints, *Journal of Optimization Theory and Applications*, forthcoming.
12. P. T. Thach and H. Tuy (1990), The Relief Indicator Method for Constrained Global Optimization, *Naval Research Logistics* **37**, 473–497.
13. H. Tuy (1985), Concave Minimization under Linear Constraints with Special Structure, *Optimization* **16**, 335–352.
14. H. Tuy (1989), The Relief Indicator Method as a New Approach to Constrained Global Optimization, IFIP Conference on System Modeling and Optimization, Leipzig.
15. H. Tuy (1990), Effect of Subdivision Strategy on Convergence and Efficiency of Some Global Optimization Algorithms, Preprint, Institute of Mathematics, Hanoi, Submitted.
16. H. Tuy (1990), On D.C. Programming and General Nonconvex Optimization Problems, Preprint, Institute of Mathematics, Submitted.