(REVISED): May, 1986

AN ALGORITHM FOR GLOBAL MINIMIZATION OF LINEARLY CONSTRAINED CONCAVE QUADRATIC FUNCTIONS

B. Kalantari* and J.B. Rosen**

DCS-TR-147
(Revised)

*Department of Computer Science
Hill Center for the Mathematical Sciences
Busch Campus, Rutgers University
New Brunswick, NJ 08903

**Computer Science Department
University of Minnesota
Minneapolis, MN 55455

An ALGORITHM FOR GLOBAL MINIMIZATION OF LINEARLY CONSTRAINED CONCAVE QUADRATIC FUNCTIONS[1]

B. KALANTARI
Department of Computer Science, Rutgers University, New Brunswick, NJ 08903.

J. B. ROSEN
Computer Science Department, University of Minnesota, Minneapolis, MN 55455.

We present an algorithm for the global minimization of a quadratic function $\psi(x,y) = -\frac{1}{2}x^TQx + h^Tx + d^Ty$, over a polytope $\Omega = \{(x,y) \in R^{n+k} : Ax + By \leq b, \quad x \geq 0, \quad y \geq 0\}$, where $x, h \in R^n$, $y, d \in R^k$, $b \in R^m$, and where A and B are m x n and m x k matrices, respectively, and Q is an n x n symmetric positive definite matrix.

We first consider the case where k=0 and construct a "tight" parallelepiped R, containing Ω by using an arbitrary set of Q-conjugate directions and by solving 2n linear programs. We show that the convex envelope of ψ with respect to R is linear and obtain an explicit formula for it. We then describe a branch and bound algorithm in which the linearity of the convex envelope allows efficient lower bounding for the subproblems. For each subproblem we obtain the linear convex envelope of ψ over a smaller parallelepiped, R', with facets parallel to those of R. Then, we obtain a lower bound by minimizing this convex envelope over $R' \cap \Omega$. If ψ^* is the optimal objective value, for each $\epsilon \geq 0$, the algorithm generates a sequence of feasible points z_j with nonincreasing function values ψ_j and a nondecreasing sequence Γ_j, of lower bounds to ψ^*. Moreover if $\epsilon > 0$, there exist j_0 such that $(\psi_j - \Gamma_j) \leq \epsilon$ for all $j \geq j_0$, and if $\epsilon = 0$, $(\psi_j - \Gamma_j)$ converges to 0. The results are then generalized to the case where k is nonzero and possibly much larger than n. Preliminary computational results are also presented.

Keywords : Concave minimization, Global optimization, Quadratic functions, Convex envelope.

1. Introduction

In this paper we consider the global minimization problem :

$$(P_{XY}): \quad \min_{(x,y)\in\Omega} \psi(x,y) = \phi(x) + d^Ty,$$

where $\phi(x) = -\frac{1}{2}x^TQx + h^Tx$, $x, h \in R^n$, $y, d \in R^k$, Q is an n x n symmetric positive definite

---

[1] This research was supported in part by the National Science Foundation under Grant No.MCS-8101214 and in part by the Department of Computer Science at Rutgers University.

matrix, and $\Omega = \{(x,y) \in \mathcal{R}^{n+k}: Ax + By \leq b, x \geq 0, y \geq 0\}$, and where A and B are m x n and m x k matrices, respectively, and $b \in \mathcal{R}^m$. We assume that Ω is nonempty and bounded. We note that $\psi(x,y)$ is concave.

The *convex envelope* of $\psi(x,y)$ over a polytope is the *supremum* of all convex functions which underestimate $\psi(x,y)$ over that polytope. Although the minimization of the convex envelope of an arbitrary concave quadratic function over an arbitrary parallelepiped is NP-hard (see Remark 2, section 2), in this paper we show that over specific parallelepiped, it may be done in polynomial time. We first consider the case where k=0 and construct a "tight" parallelepiped R, containing Ω by using an arbitrary set of Q-conjugate directions and by solving 2n linear programs. We show that the convex envelope of ψ with respect to R is linear and obtain an explicit formula for it. We then describe a branch and bound algorithm in which the linearity of the convex envelope allows efficient lower bounding for the subproblems. For each subproblem we obtain the linear convex envelope of ψ over a smaller parallelepiped, R', with facets parallel to those of R. Then, we obtain a lower bound by minimizing this convex envelope over $R' \cap \Omega$. If ψ^* is the optimal objective value, for each $\epsilon \geq 0$, the algorithm generates a sequence of feasible points z_j with nonincreasing function values ψ_j and a nondecreasing sequence Γ_j, each of each is a lower bounds to ψ^*. Moreover if $\epsilon > 0$, there exist j_0 such that $(\psi_j - \Gamma_j) \leq \epsilon$ for all $j \geq j_0$, and if $\epsilon=0$, $(\psi_j - \Gamma_j)$ converges to 0. The results are then generalized to the case where k is nonzero and possibly much larger than n. Preliminary computational results are also presented.

In section 2 we present several mathematical results for the special case k = 0. In section 3 we extend these results to the general case. In section 4 we develop a *branch and bound* algorithm for the special case as well as the general case. In section 5 we consider the computational aspect of the algorithm and in section 6 we present some computational results.

1.1. Review of Previous Work

The general "concave programming" minimization problem is defined as follows :

(P) $\min_{x \in \Omega} f(x)$,

where f(x) is any concave function and Ω is a polytope. Before presenting our results we wish to review previous research on concave programming which was initiated by Tuy [49]. Since then much research has been devoted to the development of algorithms for obtaining an optimal solution to problem (P). Many applied optimization problems may be formulated as a concave programming problem. Raghavachari [35] formulates the zero-one linear programming as a concave programming with a quadratic objective function (also see Charnes and Cooper [9] and Kalantari and Rosen [21]). The fixed charge problem (see

Murty [34], Bod [4], and Steinburg [42]) is another special case of problem (P), where the objective function is piecewise linear and concave. The quadratic assignment problem (see Koopmans and Beckman [28], Lawler [29], Graves and Winston [16]), and more generally any quadratic zero-one programming can be formulated as a special case of the concave programming problem. In fact it is possible to convert any zero-one programming problem into a concave minimization problem (see Kalantari and Rosen [24], Kalantari [25]). Furthermore, bilinear programming and concave minimization are known to be equivalent (see Carvajal-Moreno [8] and Thieu [46]). The class of concave programs also includes the linear complementarity problem (see Mangasarian [31] and Thoai and Tuy [48]). Bilinear programming approach for solving the linear complementarity problem has been considered by Al-Khayyal [2] who specializes an earlier branch and bound approach for biconvex programming Al-Khayyal and Falk [1].

It is well known that the global minimum of problem (P) is attained at a vertex of Ω. This property is used in most algorithms. The major difficulty in the development of efficient algorithms is due to the fact the problem may have many local solutions and thus local solution seeking methods such as those developed by Rosen [38] and Goldfarb [15] may not lead to an optimal solution. From the pathological point of view in [22] it is shown that even when the objective function is differentiable, given an arbitrary polytope, in the worst case there will be as many strong local solutions as there are vertices. From the complexity point of view, problem (P) is NP-hard. This is seen by the fact that the zero-one linear programming is a special case of problem (P) and that the former problem is NP-hard.

The algorithms for concave programming may be characterized as general-purpose and special-purpose algorithms. By general-purpose algorithms we mean those algorithms which make no assumption on the objective function or the constraint set, other than the ones given in problem (P) (see Tuy [49], Hu [20], Taha [44], Cabot [6], Zwart [51], Majthay and Whinston [30], Carillo [7], Falk and Hoffman [11], and Sung [43]). In the special-purpose algorithms, assumptions are made on either the objective function (see Konno [27], Ritter [36], Balas [3], Steinburg [42], Murty [34], Bod [4], Rosen [40]), Zilverberg [50], or both the objective function and the constraint set (see Lawler [29], Graves and Whinston [16], and Koopmans and Beckman [28]). For a survey of various algorithms, see Hoffman [19].

Two of the most important techniques in concave programming algorithms are branch and bound and cutting plane method (see Thoai and Tuy [47] for general approaches). Tuy [49] shows how to construct a cut which excludes a known local solution from the feasible set. The cut is known as the Tuy cut and was generalized by Glover [14]. Application of Tuy cut has proven to be useful in convergence of many algorithms. Another useful idea is the use of an underestimating function; i.e., a function $g(x)$ such that $g(x) \leq f(x)$ for all $x \in \Omega$. Given such an underestimating function, g^*, the global minimum

value of g(x) over Ω, serves as a lower bound to f^*, the optimal value of problem (P). The practicality of an underestimating function lies in the amount of work needed to compute g^*. Murty [34], Cabot and Francis [5], and Taha [44] use a linear underestimating function. A linear underestimating function gives rise to a cut in a natural way : If f_b is the current incumbent value for problem (P) (the best known objective function value thus far), the cut defined by $g(x) > f_b$ may reduce the feasible set. The use of convex envelope provides for a more powerful underestimater. Given a polytope Ω', the convex envelope of f(x) is the supremum of all linear underestimating functions of f(x) over Ω'. Now assuming that $\Omega \subset \Omega'$, the minimum of the convex envelope over Ω provides a lower bound to f^*. The convex envelope has been employed by Falk and Hoffman [12] for concave minimization and by Falk and Soland [10] who apply it to separable nonconvex programming. Also see Al-Khayyal and Falk [1] and Al-Khayyal [2]. If the objective function is concave, then the convex envelope over a polytope is piecewise linear and its minimization generally requires explicit knowledge of the enveloping polytope Ω'. For the explicit representation of the convex envelope of an arbitrary function see Rockafellar [37, p 157]. Other applications are discussed in McCormick [33].

Rosen [40] considers problem (P) with the assumption that the objective function is quadratic, given by

$$\phi(x) = -\frac{1}{2} x^T Q x + h^T x, \qquad (1)$$

where $x, h \in R^n$, and Q an nxn positive definite matrix. He considers the use of eigenvectors of Q in the determination of an initial enveloping polytope : Let $u_1, \ldots, u_n$ be the set of eigenvectors of Q and solve the 2n linear programs $\min_{x \in \Omega} \pm u_i^T x$, i=1, . . .,n. The 2n solutions of these linear programs lead to a rectangular region R, containing Ω. The minimum of f(x) over R can be computed in a trivial way, leading to an initial lower bound, $\underline{\phi}$. Next he considers partitioning the feasible region and applying the Falk and Hoffman method [11] to each partition. Based on this approach, Rosen [39] describes a parametric method for the more general objective function given in (P_{XY}).

In this paper we first consider the case where the objective function is given by (1). We show that once the containing rectangle, R, is determined by the 2n linear programs, the convex envelope of ϕ over R is linear and that it may be constructed through an explicit formula. Hence the minimum of the convex envelope over Ω can be obtained by solving a linear program. More generaly, we show that if the above 2n linear programs are replaced by $\min_{x \in \Omega} \pm u_i^T Q x$ where the u_i's are an arbitrary set of Q-conjugate directions, a containing parallelepiped may be obtained over which the linearity of convex envelope remains true. We then extend these results to the more general objective function of problem (P_{XY}), and

describe a branch and bound algorithm which also applies to the special case. The linearity of the convex envelope, together with the fact that an explicit formula for it exists, leads to a practical lower bounding scheme in the algorithm.

2. Mathematical Results

In this section we consider the special case, problem (P_X),

$$(P_X): \quad \min_{x \in \Omega_X} \phi(x) = -\frac{1}{2} x^T Q x + h^T x,$$

where as before $x, h \in \mathbb{R}^n$, Q a symmetric positive definite matrix, and $\Omega_X = \{x: Ax \leq b, x \geq 0\}$ with $b \in \mathbb{R}^m$, and A an $m \times n$ matrix. We assume Ω_X is bounded and non-empty.

Let $u_1, \ldots, u_n$ be a set of Q-conjugate directions, i.e., they are independent and $u_i^T Q u_j = 0$, for $i \neq j$. Since Q is positive definite, $u_i^T Q u_i > 0$, for all $i = 1, \ldots, n$. Consider the following $2n$ linear programs :

$$(P_i): \quad \max_{x \in \Omega_X} u_i^T Q x, \qquad (\bar{P}_i): \quad \min_{x \in \Omega_X} u_i^T Q x, \quad i = 1, \ldots, n.$$

Let x_i and $\bar{x}_i$ be optimal solutions to problems (P_i) and $(\bar{P}_i)$, respectively. For $i = 1, \ldots, n$ define

$$\lambda_i = u_i^T Q u_i ,$$

$$\beta_i = (u_i^T Q x_i)/\lambda_i , \qquad \bar{\beta}_i = (u_i^T Q \bar{x}_i)/\lambda_i ,$$

$$\alpha_i = \phi(\beta_i u_i) , \qquad \bar{\alpha}_i = \phi(\bar{\beta}_i u_i) ,$$

and the half-spaces

$$H_i(\beta_i) = \{x: \ u_i^T Q x \leq \lambda_i \beta_i\}, \qquad \bar{H}_i(\bar{\beta}_i) = \{x: u_i^T Q x \geq \lambda_i \bar{\beta}_i\}.$$

Let $R = \bigcap_{i=1}^{n} [H_i(\beta_i) \cap \bar{H}_i(\bar{\beta}_i)]$. Note that R is parallelepiped in $\mathbb{R}^n$.

Lemma 2.1 $\Omega_X \subset R$.

Proof. Let $x \in \Omega_X$. Since u_i's form a basis for $\mathbb{R}^n$, there exists $\delta_i \in \mathbb{R}$, $i = 1, \ldots, n$, such that

$$x = \sum_{i=1}^{n} \delta_i u_i.$$

For each i = 1, . . ., n we have

$$\lambda_i \bar{\beta}_i \le u_i^T Qx = \lambda_i \delta_i \le \lambda_i \beta_i. \blacksquare$$

Let $\alpha_i^* = \min\{\alpha_i, \bar{\alpha}_i\}$, i = 1, . . ., n. Define $\underline{\phi} = \sum_{i=1}^{n} \alpha_i^*$.

Theorem 2.1 $\min_{x \in R} \phi(x) = \underline{\phi}$.

Proof. Since ϕ is concave, the minimum value of ϕ over R is attained at a vertex w^*. Let the binding constraints of w^* be given by

$$u_i^T Qx = \lambda_i \tilde{\beta}_i, \qquad i = 1, \ldots, n \tag{2.1}$$

where $\tilde{\beta}_i = \beta_i$ or $\bar{\beta}_i$. Since u_i's form a basis, there exist $\delta_i \in R$, i = 1, . . ., n, such that

$$w^* = \sum_{i=1}^{n} \delta_i u_i.$$

By setting $x = w^*$ in (2.1), we get

$$u_i^T Q\left(\sum_{i=1}^{n} \delta_i u_i\right) = \delta_i u_i^T Q u_i = \lambda_i \tilde{\beta}_i.$$

Hence, $\delta_i = \tilde{\beta}_i$. Using this and that u_i's are Q−conjugate, we have

$$\phi(w^*) = \phi\left(\sum_{i=1}^{n} \delta_i u_i\right) = \sum_{i=1}^{n} \phi(\delta_i u_i) = \sum_{i=1}^{n} \phi(\tilde{\beta}_i u_i) = \sum_{i=1}^{n} \tilde{\alpha}_i,$$

where $\tilde{\alpha}_i = \alpha_i$ or $\bar{\alpha}_i$. Since w^* minimizes ϕ over R, $\tilde{\alpha}_i = \alpha_i^*$, i = 1, . . ., n. $\blacksquare$

For the special case where u_i's are the normalized eigenvectors of Q, the above theorem was proved in [40].

Definition 2.1 The convex envelope of a function f over a polytope P is a function $\Gamma(x)$ defined over P such that :

(i) $\Gamma(x)$ is convex over P.

(ii) $\Gamma(x) \le f(x)$, $\forall x \in P$.

(iii) If g(x) is any function satisfying (i) and (ii), then $g(x) \le \Gamma(x)$, $\forall x \in P$.

Theorem 2.2 (Falk and Hoffman [11]) Let $\{v_1, \ldots, v_k\}$ be the set of vertices of P. Suppose f is concave over P, then the convex envelope of f(x) over P can be expressed as :

$$\Gamma(x) = \min_{\alpha} \sum_{i=1}^{k} \alpha_i f(v_i)$$

subject to

$$\sum_{i=1}^{k} \alpha_i v_i = x , \quad \sum_{i=1}^{k} \alpha_i = 1 , \quad \alpha_i \geq 0 , \quad \text{for all } i = 1, \ldots, k,$$

where $\alpha = (\alpha_1, \ldots, \alpha_k)^T$. ■

For the explicit representation of the convex envelope of an arbitrary function see Rockafellar [37, p 157]. The function $\Gamma(x)$ is piecewise linear and agrees with f for each vertex v_i, i = 1, . . ., k. Falk and Hoffman [11] consider the global minimization problem :

$$\min_{x \in \Omega} f(x)$$

where f is concave and Ω a polytope. If $f^* = \min_{x \in \Omega} f(x)$, $\Omega \subset P$, and $\Gamma(x)$ the convex envelope of f over P, then $\Gamma^* = \min_{x \in \Omega} \Gamma(x) \leq f^*$. Thus Γ^* provides a lower bound to f^*. Although $\Gamma(x)$ is piecewise linear, Γ^* can be obtained by solving the single linear program :

$$\min_{\alpha} \sum_{i=1}^{k} \alpha_i f(v_i)$$

subject to

$$\sum_{i=1}^{k} \alpha_i Av_i \leq b , \quad \sum_{i=1}^{k} \alpha_i \geq 0 , \quad \text{for all } i = 1, \ldots, k.$$

Falk and Hoffman [11] then continue to reduce the containing polytope P, and minimize the new convex envelope over Ω to obtain tighter lower bounds. The difficulty in calculation of Γ^* is that at each stage all the vertices of the containing polytope have to be used explicitly. Returning to problem (P_x), we have :

Theorem 2.3 The convex envelope of $\phi(x)$ over the parallelepiped R is linear and given by

$$\Gamma(x) = c^T x + h^T x + c_0,$$

where

$$c = QUv, \; U = [u_1, \ldots, u_n],$$

$$v^T = (-\tfrac{1}{2}(\beta_1 + \bar{\beta}_1), \ldots, -\tfrac{1}{2}(\beta_n + \bar{\beta}_n)), \text{ and}$$

$$c_0 = \sum_{i=1}^{n} \frac{1}{2}\beta_i\bar{\beta}_i\lambda_i.$$

Proof. $\Gamma(x)$ is linear and hence convex. First we show that $\phi(x)$ and $\Gamma(x)$ agree on each vertex of R. Let w be a vertex of R, then

$$w = \sum_{i=1}^{n} \tilde{\beta}_i u_i,$$

where $\tilde{\beta}_i = \beta_i$ or $\bar{\beta}_i$. Note that

$$U^T Q u_i = \lambda_i e_i,$$

where e_i is the n-vector with 1 as the i-th element and zero for all other elements. For $i = 1, \ldots, n$ we have

$$c^T u_i = v^T U^T Q u_i = \lambda_i v^T e_i = -\tfrac{1}{2}(\beta_i + \bar{\beta}_i)\lambda_i.$$

Thus we have

$$\Gamma(w) = \sum_{i=1}^{n} [-\tfrac{1}{2}\tilde{\beta}_i(\beta_i + \bar{\beta}_i) + \frac{1}{2}\beta_i\bar{\beta}_i]\lambda_i + h^T w$$

$$= \sum_{i=1}^{n} (-\tfrac{1}{2}\tilde{\beta}_i^2\lambda_i) + h^T w = \phi(w).$$

Let $x \in R$ and assume $\{w_i\}_{i=1}^{n_1}$ is the set of vertices of R. We have

$$x = \sum_{i=1}^{n_1} \alpha_i w_i, \quad \sum_{i=1}^{n_1} \alpha_i = 1, \quad \alpha_i \geq 0, \quad \text{for } i = 1, \ldots, n_1.$$

Using this we get

$$\Gamma(x) = \sum_{i=1}^{n_1} \alpha_i \Gamma(w_i) = \sum_{i=1}^{n_1} \alpha_i \phi(w_i) \leq \phi\left(\sum_{i=1}^{n_1} \alpha_i w_i\right) = \phi(x).$$

Thus $\Gamma(x) \leq \phi(x)$ for all $x \in R$. Now let g(x) be any convex function satisfying conditions (i) and (ii) of Definition 2.1. We have

$$g(x) \leq \sum_{i=1}^{n_1} \alpha_i h(w_i) \leq \sum_{i=1}^{n_1} \alpha_i \phi(w_i) = \sum_{i=1}^{n_1} \alpha_i \Gamma(w_i) = \Gamma(x). \blacksquare$$

Corollary 2.1 $\underline{\phi} = \min_{x \in R} \phi(x) = \min_{x \in R} \Gamma(x) \leq \Gamma^* = \min_{x \in \Omega_X} \Gamma(x) \leq \phi^* = \min_{x \in \Omega_X} \phi(x). \blacksquare$

The lower bound $\underline{\phi}$ can be obtained at the completion of the 2n LP's. By solving an additional linear programming we can obtain Γ^* which most likely will result in a better lower bound.

In the following theorem we give an upper bound on the absolute error of $\phi(x)$ and $\Gamma(x)$.

Theorem 2.4 $\max_{x \in \Omega_X} (\phi(x) - \Gamma(x)) \leq \frac{1}{8} \sum_{i=1}^{n} (\beta_i - \bar{\beta}_i)^2 \lambda_i$.

Proof. Let $x \in \Omega_X$, then

$$x = \sum_{i=1}^{n} \delta_i u_i,$$

with $\bar{\beta}_i \leq \delta_i \leq \beta_i$, for all $i = 1, \ldots, n$. We have

$$\phi(x) - \Gamma(x) = -\frac{1}{2} x^T Q x - c^T x - c_0$$

$$= -\frac{1}{2} \sum_{i=1}^{n} \delta_i^2 \lambda_i + \frac{1}{2} \sum_{i=1}^{n} \delta_i (\beta_i + \bar{\beta}_i) \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \beta_i \bar{\beta}_i \lambda_i$$

$$= \frac{1}{2} \sum_{i=1}^{n} -\lambda_i (\delta_i - \beta_i)(\delta_i - \bar{\beta}_i).$$

Maximizing the above for $\delta_i \in [\bar{\beta}_i, \beta_i]$, $i = 1, \ldots, n$, we get $\tilde{\delta}_i = \frac{1}{2} (\beta_i + \bar{\beta}_i)$, for all $i = 1, \ldots, n$. Let $\tilde{x} = \sum_{i=1}^{n} \tilde{\delta}_i u_i$ then

$$\phi(\tilde{x}) - \Gamma(\tilde{x}) = \frac{1}{8} \sum_{i=1}^{n} (\beta_i - \bar{\beta}_i)^2 \lambda_i. \blacksquare$$

We note that $\tilde{x}$ is the center of R and unless it is a vertex of Ω_X, the error will be smaller if the maximum in the above theorem is taken over only the vertices of Ω_X. In the branch and bound method to be described later, the error bound will be used in defining new subproblems. Error analysis for separable functions has been considered by Thakur [45]. The following theorem on monotonicity of the convex envelope is from Falk and Hoffman [11].

Theorem 2.5 Let f be concave. Suppose P_1 and P_2 are polytopes with $P_2 \subseteq P_1$. Let Γ_1 and Γ_2 denote the convex envelopes of f over P_1 and P_2, respectively; then $\Gamma_1(x) \leq \Gamma_2(x)$, for all $x \in P_2$. $\blacksquare$

Returning to the quadratic case, let $i \in \{1, \ldots, n\}$ and let R′ be the parallelepiped obtained from R by replacing the facet $H_i(\beta_i) = \{u_i^T Q x \leq \lambda_i \beta_i\}$ by $H_i(\hat{\beta}_i)$, with $\bar{\beta}_i < \hat{\beta}_i < \beta_i$

(we have assumed $\bar{\beta}_i < \beta_i$). Let $\Gamma'(x)$ be the convex envelope of $\phi(x)$ over R'.

Theorem 2.6 $\Gamma'(x) = \Gamma(x) + \frac{1}{2}(\beta_i - \hat{\beta}_i)u_i^T Qx + \frac{1}{2}(\hat{\beta}_i - \beta_i)\bar{\beta}_i\lambda_i$. Moreover, $\Gamma'(x) \geq \Gamma(x)$ for all $x \in R'$.

Proof. The proof is based on the observation that

$$\Gamma'(x) = v'^T U^T Qx + h^T + c_0' ,$$

where

$$v'^T = v^T + \frac{1}{2}(\beta_i - \hat{\beta}_i)e_i , \text{ and } c_0' = c_0 + \frac{1}{2}(\hat{\beta}_i - \beta_i)\bar{\beta}_i\lambda_i . \blacksquare$$

From the above theorem we see that if $x = \sum_{i=1}^{n} \delta_i u_i \in R'$, with $\bar{\beta}_i < \delta_i \leq \hat{\beta}_i$, then

$$\begin{aligned}\Gamma'(x) &= \Gamma(x) + \frac{1}{2}(\beta_i - \hat{\beta}_i)\lambda_i\delta_i + \frac{1}{2}(\hat{\beta}_i - \beta_i)\bar{\beta}_i\lambda_i \\ &= \Gamma(x) + \frac{1}{2}\lambda_i(\beta_i - \hat{\beta}_i)(\delta_i - \bar{\beta}_i) > \Gamma(x).\end{aligned}$$

Thus $\Gamma'(x)$ is strictly greater than $\Gamma(x)$ except for points on the facet determined by $\bar{H}_i(\bar{\beta}_i)$. The theorem also indicates that new convex envelope can be obtained trivially. As will be seen in the algorithm, this property suggests partitioning the feasible region by adjustment of the facets R and in a parallel fashion.

Remark 1. The convex envelope of $\phi(x)$ depends on the enveloping polytope. Thus, if R is obtained by a different set of conjugate directions, the quantities $\underline{\phi}$ and Γ^* will be different.

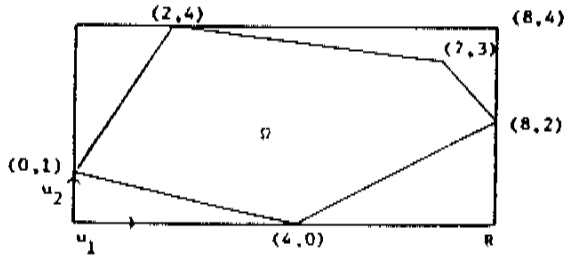
Remark 2. The minimization of the convex envelope of $\phi(x)$ over an arbitrary parallelepiped is NP-hard. To see this we note that to find the minimum of a concave function over a polytope is the same as finding the minimum of its convex envelope over that polytope. Thus, to find the minimum of the convex envelope of a concave quadratic function over the unit hypercube is equivalent to a quadratic zero-one minimization. Conversely, any quadratic zero-one minimization, by a simple transformation due to Rubin and Hammer [41], is equivalent to concave quadratic minimization over the unit hypercube. Finally, the quadratic zero-one minimization problem is equivalent to the minimum capacity cut (see Hammer [17]) which is NP-hard (see Garey et al. [13]). However, in view of Khachian's result [26], the minimization of the convex envelope of $\phi(x)$ can be done in polynomial time over R.

2.1. An Example

Let $\phi(x) = -\frac{1}{2}(2x_1^2 + 8x_2^2)$. Let $u_1 = \begin{bmatrix}1\\0\end{bmatrix}$ and $u_2 = \begin{bmatrix}0\\1\end{bmatrix}$, i.e., the normalized eigenvectors of Q. We have $\lambda_1 = 2$, $\lambda_2 = 8$. Let Ω be as in Figure 2-1, given by the constraints :

$$\begin{aligned} x_1 + x_2 &\le 10 \\ x_1 + 5x_2 &\le 22 \\ -3x_1 + 2x_2 &\le 2 \\ -x_1 - 4x_2 &\le -4 \\ x_1 - 2x_2 &\le 4 \end{aligned}$$

Figure 2-1



We have

$\beta_1 = 8$, $\bar{\beta}_1 = 0$, $\beta_2 = 4$, $\bar{\beta}_2 = 0$,

$\alpha_1 = \phi(8,0) = -64$, $\bar{\alpha}_1 = \phi(0,0) = 0$, $\alpha_1^* = -64$

$\alpha_2 = \phi(0,4) = -64$, $\bar{\alpha}_2 = \phi(0,0) = 0$, $\alpha_2^* = -64$

$\underline{\phi} = -128 = \phi(8,4)$.

The convex envelope is given by $\Gamma(x) = -8x_1 - 16x_2$. The vertices of Ω are $x^1 = (7, 3)$, $x^2 = (8, 2)$, $x^3 = (2, 4)$, $x^4 = (0, 1)$ and $x^5 = (4, 0)$. The following table is self-explanatory.

i	$\phi(x^i)$	$\Gamma(x^i)$	$\phi(x^i)-\Gamma(x^i)$
1	-85	-104	19
2	-80	-96	16
3	-68	-80	12
4	-4	-16	12
5	-16	-32	16

Table 2-1

Note that $\Gamma^* = \min_{x \in \Omega_X} \Gamma(x) = -104$ and $\phi^* = -85$. Also note that the bound in Theorem 2.4 is 32 attained at (4,2) while the error when maximization is restricted to the vertices is 19.

3. The General Case

In this section we consider the general problem :

$$(P_{XY}): \quad \min_{(x,y)\in\Omega} \psi(x,y) = \phi(x) + d^T y,$$

and extend the result of the previous section. As before let $u_1, \ldots, u_n$ be a set of Q-conjugate directions and consider the 2n linear programs :

$$(P_i): \quad \max_{(x,y)\in\Omega} u_i^T Qx, \qquad (\bar{P}_i): \quad \min_{(x,y)\in\Omega} u_i^T Qx, \quad \text{for all } i=1, \ldots, n.$$

Let $z_i = (x_i, y_i)$ and $\bar{z}_i = (\bar{x}_i, \bar{y}_i)$, $i = 1, \ldots, n$, be optimal solutions to problems (P_i) and $(\bar{P}_i)$, respectively. Let λ_i, β_i, $\bar{\beta}_i$, α_i, $\bar{\alpha}_i$, $H_i(\beta_i)$, $\bar{H}_i(\bar{\beta}_i)$, $\underline{\phi}$, R, and $\Gamma(x)$ be as in the previous section. Note that R is a subset of the X-space, i.e., $R \subset \mathbb{R}^n$. Let $\Omega_X = \{x: (x,y) \in \Omega, \text{ for some } y \in \mathbb{R}^k\}$ and $\Omega_Y = \{y: (x,y) \in \Omega, \text{ for some } x \in \mathbb{R}^n\}$, i.e., the projection of Ω into the X and Y-space, respectively. We note that $\Omega_X \subset R$. Consider the Cartesian product $P = R \times \Omega_Y$ and note that P contains Ω.

Theorem 3.1 The convex envelope of $\psi(x,y)$ over P is given by $\Gamma_{XY}(x,y) = \Gamma(x) + d^T y$

Proof. As in the special case, since Γ_{XY} is linear, we show that if z_0 is a vertex of P, then $\Gamma_{XY}(z_0) = \psi(z_0)$. Let $z_0 = (x_0, y_0)$ be a vertex of P, then x_0 is a vertex of R. Thus $\Gamma_{XY}(z_0) = \Gamma(x_0) + d^T y_0 = \phi(x_0) + d^T y_0 = \psi(z_0)$. Using this fact and in a manner similar to the special case it can be shown that Γ_{XY} satisfies conditions (ii) and (iii) of Definition 2.1. ■

The projection Ω_y in the above theorem is simply used in demonstrating that once the 2n LP's are solved we can immediately construct a linear underestimating function for ψ which is its convex envelope over a relatively "tight" polytope, namely P.

If $\Gamma^* = \min_{x\in R} \Gamma(x)$, $\ell^* = \min_{y\in\Omega_Y} d^T y = \min_{(x,y)\in\Omega} d^T y$ and $\underline{\psi} = \underline{\phi} + \ell^*$, we have the following string of inequalities;

$$\underline{\psi} \le \Gamma^*_{XY} = \min_{(x,y)\in\Omega} \Gamma(x,y) \le \psi^* = \min_{(x,y)\in\Omega} \psi.$$

Since $\psi(x,y) - \Gamma_{XY}(x,y) = \phi(x) - \Gamma(x)$, as in the previous section we have

Theorem 3.2 $\max_{(x,y)\in\Omega} (\psi(x,y) - \Gamma_{XY}(x,y)) = \max_{x\in R} (\phi(x) - \Gamma(x)) \le \frac{1}{8} \sum_{i=1}^{n} (\beta_i - \bar{\beta}_i)^2 \lambda_i$ ■

Thus, the absolute error is independent of the linear term.

4. The Algorithm

In this section, based on the result of the previous sections, we develop a branch and bound algorithm for problem (P_{XY}) which also applies to the special case, problem (P_X). The algorithm initially solves the 2n LP's described earlier. The linear convex envelope Γ_{XY} is then constructed and minimized over Ω to obtain an initial lower bound Γ^*_{XY}. While solving the 2n LP's, as well as the one which minimizes Γ_{XY}, the current incumbent value (best objective function value encountered) is updated. The algorithm then defines new subproblems by subdividing the feasible region into smaller regions. The subdivision is done by adding constraints which are parallel to facet of R. Thus each subproblem is of the form :

$$\min_{(x,y)\in R' \cap \Omega} \psi(x,y)$$

where

$$R' = \bigcap_{i=1}^{n} (H_i(\beta_i') \cap \bar{H}_i(\bar{\beta}_i')) ,$$

For each such subproblem the corresponding convex envelope based on R' is constructed according to Theorem 2.6. The lower bound for such a subproblem is obtained by minimizing its convex envelope over $R' \cap \Omega$. Each time an LP is solved the current incumbent value, ψ_b, is updated. Any subproblem with lower bound greater than or equal to ψ_b is fathomed. The algorithm continues the process until all subproblems are fathomed. The sequence of incumbent values and their corresponding points (which are not necessarily vertices in Ω will converge to the optimal value and an optimal solution, respectively. Finite convergence is guaranteed if the fathoming rule is modified as follows : Fathomed any subproblem whose lower bound exceeds $\psi_b - \epsilon$, where $\epsilon > 0$ is a specified tolerance. In terms of a branch and bound tree, each subproblem will be represented by a node. An open node is an unfathomed subproblem. A more formal description of the algorithm is as follows :

<u>Step 0</u>. Choose $\epsilon > 0$. Solve the 2n LP's described earlier. Let the first open node correspond to the original problem with lower bound $\Gamma^*_{XY} = \min_{(x,y)\in\Omega} \Gamma_{XY}(x,y)$. Let ψ_b and z_b be the current incumbent value and solution, respectively. Go to step 1.

<u>Step 1</u>. Fathom each open node with lower bound greater than or equal to ψ_b. If all nodes are fathomed stop. Otherwise select a node with smallest lower bound and go to step 2.

<u>Step 2</u>. Let i_0 be such that $\max_{1\le i\le n} \lambda_i(\beta_i^s - \bar{\beta}_i^s)^2 = \lambda_{i_0}(\beta_{i_0}^s - \bar{\beta}_{i_0}^s)^2$, where we have used the superscript s on β_i and $\bar{\beta}_i$ to indicate that they correspond to those of the selected

subproblem (node). Let the corresponding parallelepiped be R^s. Define two subproblems s_1 and s_2 :

(s_1) : $\min_{(x,y)\in R^s_1 \cap \Omega} \psi(x,y)$, (s_2) : $\min_{(x,y)\in R^s_2 \cap \Omega} \psi(x,y)$

where $R^s_1 = R^s \cap H_{i_0}(\frac{1}{2}(\beta^s_{i_0} + \bar{\beta}^s_{i_0}))$, and $R^s_2 = R^s \cap \bar{H}_{i_0}(\frac{1}{2}(\beta^s_{i_0} + \bar{\beta}^s_{i_0}))$. Construct the convex envelope of ψ with respect to $R^s_1 \cap \Omega$ and $R^s_2 \cap \Omega$, and while updating ψ_b and z_b, minimize over the respective region to obtain lower bounds for these subproblems. Append these subproblems as left and right children of the current node and go to step 1.

4.1. Convergence

In this section we prove the convergence of the algorithm. The search tree is a binary tree with each subproblem represented by a node. The root of the tree corresponds to the initial problem solved in step 0. The depth of a node on this tree will be defined in the usual sense. Let ψ^* be the optimal objective value. For each $\epsilon \geq 0$, let s_j ,j=1,2,..., be the selected subproblems in step 1 of the algorithm. For each such subproblem s_j , let z_j be the current incumbent solution known to the subproblem, $\psi_j=\psi(z_j)$, the current incumbent value, and Γ_j the minimum of the corresponding convex envelope over the feasible region of s_j , and hence the best upper bound to ψ^* known to s_j.

Theorem 4.1

(i) $\Gamma_1 \leq \Gamma_2 \leq \ldots \leq \psi^* \leq \ldots \leq \psi_2 \leq \psi_1$

(ii) Let s_j be a subproblem of depth N in the search tree and $s_{j'}$ a descendant of s_j of depth N+n, then $(\psi_{j'}-\Gamma_{j'}) \leq \frac{1}{4}(\psi_j-\Gamma_j)$.

(iii) If $\epsilon > 0$, there exist j_0 such that $(\psi_j-\Gamma_j) \leq \epsilon$ for all $j \geq j_0$. If $\epsilon=0$, then $(\psi_j-\Gamma_j)$ converges to 0.

Proof. The proof of (i) follows trivially by the monotonicity property of the convex envelope (Theorem 2.5), the selection strategy in step 1, and the fact that the updated incumbent values in step 2 are nonincreasing.

To prove (ii), let s be a subproblem selected in step 1. If s is so that $\frac{1}{8}(\beta^s_i - \bar{\beta}^s_i)^2\lambda_i \leq \epsilon/n$, for all i = 1, . . .,n, then the error bound of Theorem 3.2 is less that ϵ, and the subproblem will be fathomed. If not, let $i_0 \in \{1,...,n\}$ be the selected index in step 2 and s_1 and s_2 the new subproblems. The i_0−th term of the error bound for s is

$\frac{1}{8}(\beta^s_{i_0} - \bar{\beta}^s_{i_0})^2\lambda_{i_0}$. The i_0-th term of the error bound for s_1 and s_2 is $\frac{1}{32}(\beta^s_{i_0} - \bar{\beta}^s_{i_0})^2\lambda_{i_0}$. Thus the i_0-th term of the error bound improves by a factor of four.

Now assume that s is of depth N and let s′ be a descendant of s of depth N+n. Consider the binary search tree. There is a unique path from s to s′, say s_j, j=1,...,n+1 with s_1=s and s_{n+1}=s′, and where s_{j+1} is a child of s_j. Let ψ_j and Γ_j be the corresponding upper and lower bounds for s_j. We show that $(\psi_{n+1} - \Gamma_{n+1}) \leq \frac{1}{4}(\psi_1 - \Gamma_1)$. For i=1,...,n, let $a_i = \frac{1}{8}(\beta^s_i - \bar{\beta}^s_i)^2\lambda_i \geq 0$ be the i-th term of the error bound for s=s_1. Thus the error bound corresponding to s_1 is $\sum_{i=1}^{n} a_i$. Let σ_1 denote this bound. To get the corresponding error bound for s_2, in step 2 of the algorithm, the maximum term of σ_1 will be replaced with $\frac{1}{4}$ of its value to give σ_2. Repeating this process we get the error bounds σ_j corresponding to the subproblems s_j, j=1,...,n+1. We now prove by induction on n that $\sigma_{n+1} \leq \frac{1}{4}\sigma_1$. Clearly, this is true for n=1. Assume true for n−1. If the indices of the maximum terms selected are all distinct, then clearly $\sigma_{n+1} \leq \frac{1}{4}\sigma_1$. Assume there exists a term in all σ_i's which remains unchanged. Without loss of generality, let a_n be such term and let the σ_{n+1} be obtained from σ_n by replacing its first term by a $\frac{1}{4}$ of its value. Thus, the largest term of σ_n is $(\frac{1}{4})^{k-1}a_1$ for some k ≥ 1. In particular $(\frac{1}{4})^{k-1}a_1 \geq a_n$.

By induction hypothesis, $(\sigma_n - a_n) \leq \frac{1}{4}(\sigma_1 - a_n)$. We have

$$\sigma_{n+1} = (\sigma_n - 3(\tfrac{1}{4})^k a_1) \leq \frac{1}{4}\sigma_1 + \frac{3}{4}a_n - 3(\tfrac{1}{4})^k a_1 \leq \frac{1}{4}\sigma_1.$$

The proof of (iii) follows immediately from (ii), since for each integer N, there are finitely many nodes of depth N. For $\epsilon > 0$, the algorithm will terminate in a finite number of steps and in the worst case R will be partitioned into finitely many parallelepipeds, where over each such parallelepiped, the error bound is less than ϵ. ■

We now apply the algorithm to the simple example considered in section 1.2 :

Step 0. In step 0 we obtain the following information :

$$\Gamma^* = -104, \ \psi_b = -85.$$

Step 1. The open node corresponds to the initial node.

Step 2. $\lambda_1(\beta_1^s - \bar{\beta}_1^s)^2 = 128$, $\lambda_2(\beta_2^s - \bar{\beta}_2^s)^2 = 128$. In this case there is a tie. We break the tie arbitrarily. Pick $i_0 = 2$. The subproblems s_1 and s_2 are generated by adding to Ω the constraints $u_2^T x \leq 2$ and $u_2^T x \geq 2$, respectively. We note that $u_2^T x = x_2$. The corresponding feasible regions are Ω_1 and Ω_2 in Figure 4-1. The corresponding convex envelopes Γ_1 and Γ_2 are given by

$$\Gamma_1 = -8x_1 - 8x_2$$

$$\Gamma_2 = -8x_1 - 24x_2 + 32.$$

The corresponding lower bounds are

$$\Gamma_1^* = \min_{x \in \Omega_1} \Gamma_1(x) = -80, \quad \Gamma_2^* = \min_{x \in \Omega_2} \Gamma_2(x) = -96,$$

s_1 is fathomed in the second pass to step 1, and the algorithm continues to define new subproblems. In Figure 4-2, we give the tree generated by the algorithm. The numbers in the circles represent the lower bounds, and "F" denotes a fathomed node.

Figure 4-1

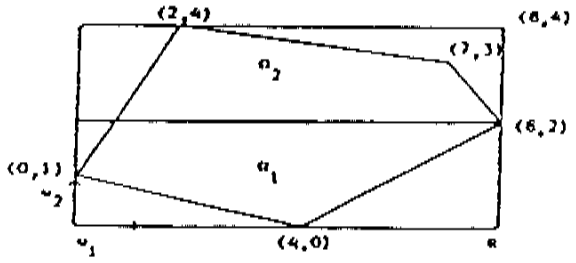
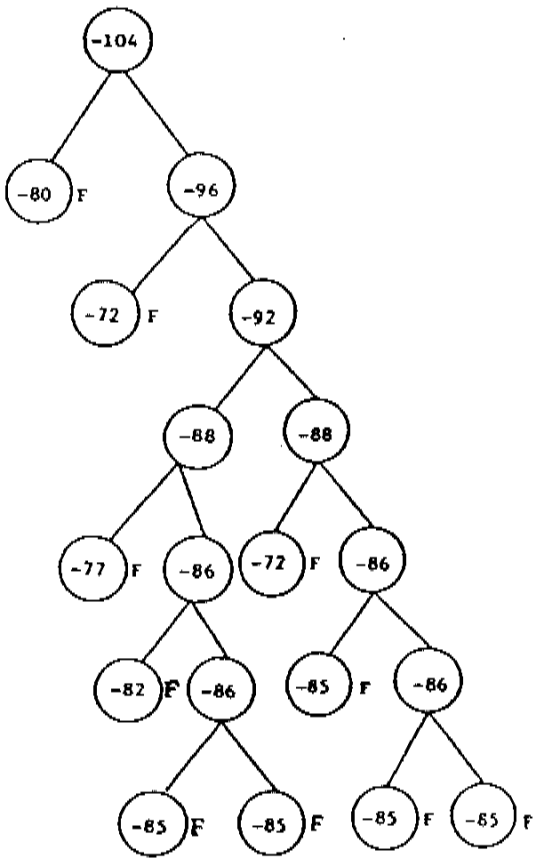


Figure 4-2



5. Computational Considerations

In this section we consider several computational issues regarding the algorithm of this paper.

5.1. Conjugate Directions

We recall that at the outset, the algorithm requires a set of conjugate directions. This can be done in $O(n^3)$ steps (see Hestenes [18]) via the conjugate Gram-Schmidt process. It is important to note that the convex envelope depends on the enveloping parallelepiped which in turn depends on a given set of conjugate directions. The relative merit of one set of conjugate directions to another depends on $\phi(x)$ and the constraint set. An alternate approach for the generation of a conjugate set is the use of conjugate gradient algorithm (see Hestenes [18, p 120]) in finding the unconstrained maximum of $\phi(x)$. As a by product of this maximization one obtain a set of conjugate directions. The difficulty in using the conjugate gradient method is that the method may not produce a complete set of conjugate directions which is required of our algorithm.

5.2. Solving The 2n LP's

To construct the initial convex envelope, the algorithm solves the 2n linear programs

$$\max_{(x,y)\in\Omega} \pm u_i^T Qx \ , \ i = 1, \ldots, n.$$

The computational efficiency of the above multiple cost LP depends on the order in which they are solved. One natural order is to maximize $u_i^T Qx$ for $i = 1, \ldots, n$, and in that order. Then, minimize using the same order. This has the effect of exploiting the polytope in a systematic way.

5.3. Affine Transformation

The feasible region of each subproblem consist of the initial feasible set, Ω, and at most 2n constraints of the type

$$u_i^T Qx \le \lambda_i \tilde{\beta}_i \ , \text{ or } u_i^T Qx \ge \lambda_i \tilde{\beta}_i.$$

Such constraints only involve the X-space. If $\phi(x)$ is separable, then it may be advantageous to use the standard basis as the conjugate direction set. This is because each subproblem reduces to a bounded variable LP. In general it is possible to replace problem (P_{XY}) to an equivalent problem with a separable objective function. Suppose in problem (P_{XY}) we replace x by Uw, where U is the matrix of the vectors $u_1, \ldots, u_n$ defined earlier and $w \in \mathbb{R}^n$. Since $U^T QU = D = \text{diag}(\lambda_1, \ldots, \lambda_n)$, we have

$$x^TQx = w^TU^TQUw = w^TDw = \sum_{i=1}^{n} \lambda_i w_i^2$$

By such affine transformation, it can be shown that problem (P_{XY}) is equivalent to the following separable program :

$$(P'_{XY}): \quad \min_{(w,y)\in\Omega'} \gamma(w,y) = -\frac{1}{2} w^TDw + h^TUw + d^Ty$$

where $\Omega' = \{(w,y): AUw + By \leq b,\ Uw \geq 0,\ y \geq 0\}$. Let us consider the result of this transformation on each subproblem. The constraints on each subproblem are now those of Ω' and constraints of the type

$$u_i^T QUw \leq \lambda_i\tilde{\beta}_i \text{ , or } u_i^T QUw \geq \lambda_i\tilde{\beta}_i.$$

But now these constraints will take the simple form

$$w_i \leq \tilde{\beta}_i \text{ , or } \quad w_i \geq \tilde{\beta}_i.$$

Thus the effect of such an affine transformation is that each subproblem can be reduced to a bounded variable linear problem. A major disadvantage, however, is that the non-negativity constraints $x \geq 0$ are now replaced by the more difficult constraints $Uw \geq 0$. Indeed the major attractiveness of our algorithm is in its ability to obtain a linear convex envelope by solving linear programs over the initial polytope.

5.4. The Use of Tuy Cuts

One of the difficult aspects of concave programming algorithms is the verification of optimality. It is possible for an algorithm to encounter an optimal solution at an early stage, yet to prove its optimality the algorithm may require an enormous amount of computation. Thus, it is advantageous to eliminate a local optimal point which is a candidate for the global optimal solution. Suppose $z_0 = (x_0, y_0)$ is such a candidate solution. Assume it is non-degenerate. Thus, z_0 has n+k adjacent vertices which may be obtained by pseudo-pivoting. Having obtained the adjacent vertices, since ψ is concave, we can exclude from further consideration the simplex obtained by the vertices z_i, i= 1, . . ., n+k. This can be achieved by finding r and ρ so that $r^Tz=\rho$ satisfies z_i, for all i=1, . . ., n+k. The determination of r and ρ requires the solution of a system of n+k linear equations. Such a cut with appropriate inequality sign, when added to the feasible region will eliminate z_0. Obviously, the above cut can be obtained for any concave function. In the quadratic case however we can obtain a deeper cut (Tuy cut) by extending the adjacent vertices and in an easy fashion :

Each edge emanating from z_0 is of the form

$$z_0 + \alpha(z_i - z_0)\ ,\ 0 \le \alpha \le 1.$$

Let $z_0 = (x_0, y_0)$ and $z_i = (x_i, y_i)$, i=1, . . ., n+k be the adjacent vertices of z_0. Assume $\psi(z_0) \le \psi(z_i)$, for all i=1, . . ., n+k. Consider the equation $\psi(z_0 + \alpha(z_i - z_0)) = \psi(z_0)$. Equivalently,

$$-\frac{1}{2}\,[x_0 + \alpha(x_i - x_0)]^T Q[x_0 + \alpha(x_i - x_0)] + h^T[x_0 + \alpha(x_i - x_0)] + d^T[y_0 + \alpha(y_i - y_0)] = \psi(z_0).$$

Expanding and simplifying we get the linear equation in α :

$$a_i\alpha + b_i = 0,$$

where

$$a_i = \frac{1}{2}(x_i - x_0)^T Q(x_i - x_0)\ ,\quad b_i = x_0^T Q(x_i - x_0) - h^T(x_i - x_0) - d^T(y_i - y_0).$$

Since $\psi(z_0) \le \psi(z_i)$, we have $h^T(x_i - x_0) + d^T(y_i - y_0) \ge -\frac{1}{2}\,x_0^T\,Qx_0 + \frac{1}{2}\,x_0^T\,Qx_i$. Since $(x_i - x_0)^T Q(x_i - x_0) \ge 0$, $-\frac{1}{2}\,x_0^T\,Qx_0 + \frac{1}{2}\,x_0^T\,Qx_i \ge x_0^T Q(x_i - x_0)$. Thus, $b_i \le 0$. Now if $(x_i - x_0) \ne 0$, then $a_i \ne 0$ and the linear equation in α has a positive solution α_i. For i = 1, . . .,n+k define $z_i' = z_0 + \alpha_i(z_i - z_0)$. We now find r and ρ so that the linear equation $r^T z = \rho$ satisfies z_i', for all i=1, . . ., n+k and add the cut with appropriate inequality so as to exclude z_0.

6. Computational Results

In order to obtain an efficient implementation of the above algorithm in addition to the above considerations many other issues have to be addressed, e.g., the use of appropriate data structure for the branch and bound, the effectiveness of various branching strategies. The data structure can clearly affect the efficiency of the algorithm as a whole. As an example the efficiency in recovery of the constraint set of a given subproblem as well the calculation of its lower bound depends heavily on the use of selected data structure. We are currently studying these issues. In this section however, we will report on some computational result and based on a preliminary FORTRAN program which makes use of XMP, a library of subroutines for linear programming, developed by R.E Marsten [32]. Computing was performed on a Sperry 1100/60 series computer.

Although it is possible to generate test problems with known global solutions and over arbitrary polytopes (see [23]), in our preliminary testing we generated problems of reasonable difficulty but with unknown global solutions and with the following format :

$$\min_{(x,y)\in\Omega}\ \psi(x,y) = -\frac{1}{2}\,(x-\bar{x})^T Q(x-\bar{x}) + d^T y,$$

where $\bar{x} \in R^n$, is a fixed vector and with $n \le 100$, and $n+k \le 300$. In these problems we

generated different $\bar{x}$'s, d's, and Q matrices. We first obtained an orthogonal matrix U through the Gram–Schmidt orthogonalization process and then formed $Q=U^TDU$ with $D=\text{diag}(\delta_1, \ldots, \delta_n)$ with $\delta_i > 0$. We used the orthogonal vectors of U as the set of conjugate direction. For more detail on the generation of the test problem including the underlying polytope see [25]. The algorithm was run and tested on a variety of problems.

6.1. Summary of The Computational Results

In this section we present our preliminary computational results.

In the first set of tables below we report various statistics once the 2n LP's are solved. These statistics for a given size m and n are averaged over various problems. For each constraint set the initial 2n LP's were solved. The total number of iterations required is summarized in Table 6–1. Table 6–2 gives the average number of iterations per LP.

m	n+k	n=10	n=30	n=50	n=100
20	30	--	1304	--	--
20	50	598	1778	2905	--
25	100	1462	3692	6022	--
35	200	1566	4276	6940	19204
45	300	1832	4658	7570	23338

Table 6-1: Total Number of Simplex Iterations for Initial 2n Linear Programs

m	n+k	n=10	n=30	n=50	n=100
20	30	--	22	--	--
20	50	30	30	30	--
25	100*	73	62	60	--
35	200	78	71	69	96
45	300	82	78	76	117

Table 6-2: Average Number of Iterations for Initial 2n Linear Programs

*The sharp increase in the average number of iterations is due to the block structure of the constraint matrix.

If we let T(n,k,m) represent the total number of simplex iterations required of the 2n linear programs, from table 6–2, we see that for n fixed, it depends linearly on m and independent of k which conforms to the expected computational behavior of linear programs.

Table 6–3 gives the total computing time for the 2n LP's. This includes data inputs, function

evaluations, as well as other preliminary calculations.

m	n+k	n=10	n=30	n=50	n=100
20	30	--	24	--	--
20	50	10	34	58	--
25	100	38	98	166	--
35	200	50	147	240	697
45	300	71	185	310	962

Table 6-3: Total Time for 2n Linear Programs (in seconds)

We now discuss the bounds obtained initially, and their improvements after additional iterations. Let ψ_b be the incumbent value after the solution of the 2n LP's and the one which minimizes the initial linear convex envelope. Let Γ^* be the initial minimum of the convex envelope over Ω. In order to estimate the rate of convergence of the algorithm we first generated 40 nodes of the branch and bound tree for various size problems. Denoting the new current incumbent value by ψ_b^1, and the new lower bound by Γ^1, we defined $\rho = |(\psi_b^1 - \Gamma^1)/(\psi_b - \Gamma^*)|^{1/40}$, which gives the average decrease in the error bound per node based on 40 nodes. Tables 6–4 and 6–5 give the best and worst case for ρ^{1000} respectively. These quantities estimate the error bound which would be obtained after generation of 1000 nodes.

m	n+k	n=10	n=30	n=50	n=100
20	50	$.17\times10^{-21}$	$.38\times10^{-11}$	$.15\times10^{-5}$	--
25	100	$.68\times10^{-13}$	$.22\times10^{-13}$	$.11\times10^{-6}$	--
35	200	$.48\times10^{-13}$	$.67\times10^{-10}$	$.84\times10^{-8}$	—
45	300	$.15\times10^{-8}$	$.43\times10^{-3}$	$.18\times10^{-3}$	$.5\times10^{-4}$

Table 6-4: Best Case of ρ^{1000}

In a number of problems we generated up to 1000 nodes, while updating the actual rate of convergence after every 40 iterations which had the tendency to become slower as more nodes where generated. In almost all the problems, the current incumbent value obtained after the initial (2n+1) LP's (minimization of the 2n LP's as well as the linear convex envelope) either did not change or if it did, the change was "small." This property has often been observed in branch and bound methods for integer linear programming. In a different experiment after a good incumbent solution had been obtained, we employed

m	n+k	n=10	n=30	n=50	n=100
20	50	$.33\times10^{-9}$	$.31\times10^{-3}$	$.74\times10^{-2}$	--
25	100	$.75\times10^{-2}$	$.60\times10^{-4}$	$.80\times10^{-2}$	--
35	200	$.18\times10^{-8}$	$.33\times10^{-5}$	$.22\times10^{-3}$	--
45	300	$.19\times10^{-8}$	$.43\times10^{-3}$	$.22\times10^{-3}$	$.24\times10^{-2}$

Table 6-5: Worst Case of ρ^{1000}

the cuts described in the previous section and found it to be quite useful. In case of degeneracy we would perturb the right-hand side vector and then generate the cut with respect to the perturbed solution. For more detail on computational results see [25].

7. Concluding Remarks and Future Work

In this paper we have presented an algorithm for the global minimization of a concave quadratic functions over a polytope. The major contribution of this paper is in effective construction of a linear convex envelope over a parallelepiped which contains the given polytope but based on an arbitrary set of Q-conjugate directions. The linearity of the convex envelope allows efficient calculation of an initial lower bound to the optimal objective value. A branch and bound scheme is then described in which the feasible region of each subproblem is the intersection of the original feasible set and smaller parallelepiped with facets parallel to those of the initial one. Lower bounding for each subproblem consist of minimization of a corresponding linear convex envelope over its feasible set. While our preliminary computational result with this algorithm is quite encouraging, there is need for further investigation of several computational issues which would result in more efficient implementation. In the near future, we hope to carry out more extensive computation and report on the results. We conclude by remarking that the algorithm of this paper can be generalized to solve the more difficult global minimization problem where the Q matrix may be indefinite [25]. We will treat this more general case in a separate paper.

Acknowledgments. We are thankful to the referees for their remarks and in particular to one of them who provided us with an extremely instructive list of remarks and corrections.

8. References

1. Al-Khayyal, F.A. and Falk, J.E. (1983). Jointly Constrained Biconvex Programming. Mathematics of Oper. Res. 8 273-286.

2. Al-Khayyal, F.A. (1983). Linear, Quadratic, and Bilinear Programming Approaches to Linear Complementarity Problem. Industrial and system Engineering series No. J-83-5. To appear in European Journal of Operations Research.

3. Balas, E. and Burdet, C.A. (1973). Maximizing a Convex Quadratic Function Subject to Linear Constraints. MSRR 299, GSIA, Carnegie-Mellon University.

4. Bod, P. (1970). Solution of a Fixed Charge Linear Programming Problem. Proceedings of Princeton Symposium on Mathematical Programming, Princeton University Press, Princeton, New Jersey.

5. Cabot, A.V. and Francis, R.L. (1970). Solving Nonconvex Quadratic Minimization Problems by Ranking the Extreme Points. Oper. Res. 18 82-86.

6. Cabot, A.V. (1974). Variations on a Cutting Plane Method for Solving Concave Minimization Problems with Linear Constraints. Naval Res. Logist. Quart. 21 265-274.

7. Carillo, J.J. (1977). A Relaxation for the Minimization of a Quasi-Concave Function on a Convex Polyhedron. Mathematical Programming 13 69-80.

8. Carvajal-Moreno, R. (1972). Minimization of Concave Functions Subject to Linear Constraints. Report ORC-3, Operations Research Center, University of California, Berkeley, CA.

9. Charnes, A., and Cooper, W.W. (1961). Management Models and Industrial Applications of Linear Programming. John Wiley and Sons, New York.

10. Falk, J.E. and Soland, R.M. (1969). An Algorithm for Solving Separable Nonconvex Programming Problems. Management Sci. 15 550-569.

11. Falk, J.E. and Hoffman, K.L. (1976). A Successive Underestimation Method for Concave Minimization Problems. Mathematics of Oper. Res. 1 251-259.

12. Falk, J.E. and Hoffman, K.L. (1982). Concave Minimization via Collapsing Polytopes. Rept. T-438, Institute for Management Science and Engineering, George Washington University, Washington, D.C.

13. Garey, M.R., Johnson, D.S. and Stockmeyer, L. (1976). Some Simplified NP-Complete graph Problems. Theor. Comp. Sci. 1 237-267.

14. Glover, F. (1970). Convexity Cuts and Cut Search. Management Science Report No. 70-2, University of Colorado, Boulder, Colorado.

15. Goldfarb, D. (1969). Extensions of Davidson's Variable Metric Method to Maximization Under Linear Inequality and Equality Constraints. SIAM J. 17 739-764.

16. Graves, G.W. and Whinston, A.B. (1970). An Algorithm for the Quadratic Assignment Problem, Management Sci. 17 453-471.

17. Hammer, P.L. (1965). Some Network Flow Problems Solved with Pseudo-Boolean Programming. Oper. Res. 13 388-399.

18. Hestenes, M.R. (1980). Conjugate Gradient Methods in Optimization, Springer-Verlag, New York Heidelberg Berlin.

19. Hoffman, K.L. (1975). A Successive Underestimating Method for Concave Minimization Problems. Ph.D. thesis, George Washington University.

20. Hu, T.C. (1969). Minimizing a Concave Function in a Convex Polytope. MRC Technical Summary Report No. 1011, Mathematics Research Center, University of Wisconsin.

21. Kalantari, B. and Rosen, J.B. (1982). Penalty For Zero-One Integer Equivalent Problems. Mathematical Programming 24 229-232.

22. Kalantari, B. (1985). Construction of Difficult Linearly Constrained Concave Minimization Problems. Oper. Res. 33 222-227.

23. Kalantari, B. and Rosen, J.B. (1986). Construction of Large-Scale Global Minimum Concave Quadratic Test Problems. Journal of Optimization Theory and Applications 48 303-313.

24. Kalantari,B. and Rosen, J.B. (1984). Equivalent Penalty Formulation for Zero-One Concave Minimization. To appear in Discrete Applied Mathematics.

25. Kalantari, B. (1984). Large Scale Global Minimization of Linearly Constrained Concave Quadratic Functions and Related Problems. Ph.D. thesis, Computer Science Department, University of Minnesota.

26. Khachian, L.G. (1979). A Polynomial Algorithm for Linear Programming. Doklady Akad. Nauk USSR, 244 (1979), 1093-1096. Translated in Soviet Math. Doklady 20 191-194.

27. Konno, H. (1976). Maximization of a Convex Quadratic Function Under Linear Constraints. Mathematical Programming 11 117-127.

28. Koopmans, T.C. and Beckman, M.J. (1957). Assignment Problems and the Location of Economic Activity. Econometrics 25 52-76.

29. Lawler, E.L. (1963). The Quadratic Assignment Problem. Management Sci. 9 586-599.

30. Majthay, A. and Whinston, A. (1974). Quasiconcave Minimization Subject to Linear Constraints. Discrete Mathematics 1 35-39.

31. Mangasarian, O.L. (1978). Characterization of Linear Complementarity Problems as Linear Programs. Mathematical Programming Study 7 74-87.

32. Marsten, R.E. (1978). XMP: A Structured Library of Subroutines for Experimental Mathematical Programming. Tech. Rept. 351, Dept. of MIS, University of Arizona, Tucson, Arizona.

33. McCormick, G.P. (1983). Nonlinear Programming: Theory, Algorithms, and Applications, John Wiley and Sons, New York.

34. Murty, K.G. (1969). Solving the Fixed Charge Problem by Ranking the Extreme Points. Oper. Res. 16 268-279.

35. Raghavachari, M. (1969). On Connections Between Zero-One Integer Programming and Concave Programming Under Linear Constraints. Oper. Res. 17 680-684.

36. Ritter, K. (1966). A Method for Solving Maximum Problems with a Nonconcave Quadratic Objective Functions. A fur Wahrscheinlichkeits Theorie und Verwandte Gebiete 4 340-351.

37. Rockafellar, R.T. (1970). Convex Analysis. Princeton University Press, Princeton.

38. Rosen, J.B. (1960). The Gradient Projection Method for Nonlinear Programming, Part I-linear constraints, SIAM J. Appl. Math., 8, 181-217.

39. Rosen, J.B..(1981). Parametric Global Minimization for Large-Scale Problems. Tech. Rept. 81-40, Computer Science Dept., University of Minnesota.

40. Rosen, J.B. (1983). Minimization of a Linearly Constrained Concave Function by Partition of Feasible Domain. Mathematics of Oper. Res. 8 215-230.

41. Rubin, A.A. and Hammer, P.L. (1969). Quadratic Programming with 0-1 Variables. Oper. Res, Statistical, and Economics Mimeograph Series No. 53, Technion, Israel.

42. Steinburg, D.I. (1970). The Fixed Charge Problem. Naval Res. Logist. Quart. 17 217-235.

43. Sung, Y.Y. (1983). Computational Algorithms for Global Minimization. Ph.D. thesis, Computer Science Dept., University of Minnesota.

44. Taha, H.A. (1973). Concave Minimization over a Convex Polyhedron. Naval Res. Logist. Quart. 20 533-548.

45. Thakur, L.S. (1978). Error Analysis for Convex Separable Programs. SIAM Journal of Applied Mathematics 704-714.

46. Thieu, T.V. (1978). Relationship Between Bilinear Programming and Concave Minimization Under Linear Constraints. Acta-Mathematica Vietnamica. 5 106-113.

47. Thoai N.V. and Tuy, H. (1980). Convergent Algorithms for Minimizing a Concave Function. Mathematics of Oper. Res. 5 556-566.

48. Thoai N.V. and Tuy, H. (1983). Solving the Linear Complementarity Problem Through Concave Programming. Zh. Vychisl. Mat. i. Mat. Fiz. 23 602-608,

49. Tuy, H. (1964). Concave Programming Under Linear Constraints. Soviet Math. Dokl. 5 1437-1440.

50. Zilverberg, N.D. (1983). Global Minimization for Large Scale Linear Constrained Systems. Ph.D. thesis, Computer Science Department, University of Minnesota.

51. Zwart, P.B. (1974). Global Maximization of a Convex Function with Linear Inequality Constraints. Oper. Res. 22 602-609.