Some Heuristics and Test Problems for Nonconvex Quadratic Programming over a Simplex

Ivo Nowak *

September 3, 1998

Abstract

In this paper we compare two methods for estimating a global minimizer of an indefinite quadratic form over a simplex. The first method is based on the enumeration of local minimizers of a so-called control polytope. The second method is based on an approximation of the convex envelope using semidefinite programming. In order to test the algorithms a method for generating random test problems is presented where the optimal solution is known and the number of binding constraints is prescribed. Moreover, it is investigated if some modifications of the objective function influence the performance of the algorithms. Numerical experiments are reported.

Keywords:global optimization, nonconvex quadratic programming, heuristics, Bézier methods, test problems

1 Introduction

Consider a nonconvex quadratic programming problem of the form

$$\begin{array}{rl} \text{global minimize} & x^T F x \\ \text{subject to} & x \in \Delta_n, \end{array} \tag{1}$$

where the admissible set is the n-dimensional standard simplex

$$\Delta_n := \{x \in I\!R^{n+1} : x_i \geq 0, \quad 1 \leq i \leq n+1, \quad e^T x = 1\},$$

$F \in I\!R^{(n+1,n+1)}$ is an indefinite symmetric matrix and $e \in I\!R^{n+1}$ is the vector of ones. Finding a global solution of (1) is known to be NP-hard (see Horst, Pardalos and Thoai (1995)). Problems of the type (1) occur for example in the search for a maximum (weighted) clique in an undirected graph. It is also strongly related to the general quadratic optimization problem (QP) which has numerous applications. Of course, Problem (1) can be considered as a general global optimization problem which has been investigated by many authors (see for example Horst and Pardalos (1995)). However, only few methods for (approximately) solving (1) are available which take advantage of the special structure of the problem. Bomze presents in Bomze (1998) applications and a heuristic for (1) using an evolutionary approach. Coleman and Hulbert use in Coleman and Hulbert (1989) an active set method for solving large scale sparse quadratic problems with box constraints. For linearly constrained quadratic programs in Han, Pardalos and Ye (1992) an interior point algorithm and in An and Tao (1996) a d.c. algorithm are presented.

In this paper we propose two algorithms for approximately solving (1). The first method, which is described in Section 2, is a sampling method. The sample points are local minimizers of a so-called

*Email:ivo@mathematik.hu-berlin.de, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

control polytope. These points are manipulated by a descent method to compute a candidate global minimizer. The second method is described in Section 3. This approach is based on the approximation of the convex envelope using semidefinite programming. This leads to a convex quadratic relaxation of the original problem. A minimizer of this relaxed problem is used to estimate a global minimizer of (1). In Section 4 a method for constructing random test problems with known optimal value and prescribed number of binding constraints at the global minimum point is proposed. Moreover, it is investigated in Section 5 whether some modifications of the objective function of (1) influence the performance of the algorithms. Finally we compare in Section 6 both algorithms using the random test case generator proposed in Section 4.

2 Enumerating local minimizers of the Bézier-net

In this section we describe a sampling method for approximately solving (1). The idea of the method is to generate a set of sampling points which are manipulated by means of a local search to yield a candidate global minimizer. We choose local minimizers of a so-called control polytope as sample points. A control polytope $P_f(x)$ of a quadratic form $f(x) := x^T F x$ over Δ_n is a piecewise linear interpolant of the Bézier-net defined by $\{(x_{ij}, F_{ij}) : 1 \le i \le j \le n+1\}$ where $x_{ij} := \frac{1}{2}(e_i + e_j)$ are lattice points and $e_i \in \mathbb{R}^{n+1}$ is the i-th unit vector. A control polytope has the following properties (see Dahmen (1986)):

Lemma 1 *Let $P_f(x)$ be a control polytope over Δ_n of the quadratic form $f(x) := x^T F x$ where $F \in \mathbb{R}^{(n+1,n+1)}$ is a symmetric matrix. Then*
(i) the function values of $P_f(x)$ are equal to $f(x)$ at the vertices of Δ_n, i.e. $f(e_i) = P_f(e_i) = F_{ii}$, $1 \le i \le n+1$;
(ii) the tangent plane of $P_f(x)$ coincides with the tangent plane of $f(x)$ at the vertices of Δ_n, i.e. $\partial_{e_j - e_i} f(e_i) = \partial_{e_j - e_i} P_f(e_i) = 2 \cdot (F_{ij} - F_{ii})$, $1 \le i, j \le n+1$, $i \ne j$;
(iii) if $P_f(x)$ is convex on Δ_J then $f(x)$ is convex on Δ_J where $J \subset \{1, .., n+1\}$ and Δ_J is the J-face of Δ_n, i.e. $\Delta_J := \{x \in \Delta_n : x_i = 0, \quad i \in \{1, .., n+1\} \setminus J\}$.

Lemma 1 shows that $P_f(x)$ is an approximation of $f(x)$ over Δ_n. A local minimizer of $P_f(x)$ over Δ_n can be considered therefore as an estimation of a local minimizer of $f(x)$. Since $P_f(x)$ is a piecewise linear function the set $\{x_{ij} : ij \in T\}$ of lattice points, where $T := \{ij : F_{ij} \le F_{ik}, \quad 1 \le i, j, k \le n+1\}$, are local minimizers of $P_f(x)$. We choose T as the set of sampling points hoping that among these sample points there are points which are in the region of attraction of a global minimizer. The region of attraction of a minimizer is defined as the set of points in Δ_n starting from which a local search will converge to the minimizer. The above observations lead to the following algorithm for computing an estimation for a global minimizer of (1) denoted by x_{est1}.

Algorithm 1

1. Set $f_{min} := F_{11}$, $x^{(0)} := e_1$
2. For $ij \in T$ do :
   Let $a_{ij} := \max_{1 \le k \le n+1, l \in \{i,j\}} F_{kl}$ and $b_{ij} := \min_{1 \le k \le n+1, l \in \{i,j\}} F_{kl}$.
   Compute an estimation of a local minimizer:
   $$x_{ij}^{(0)} := \sum_{1 \le k \le n+1, l \in \{i,j\}} \mu_{kl} x_{kl} \Big/ \sum_{1 \le k \le n+1, l \in \{i,j\}} \mu_{kl},$$
   where $\mu_{ij} := (a_{ij} - b_{ij})^6$.
   Refinement: Compute $\hat{x}_{ij}$ by performing some steps of a descent algorithm starting from $x_{ij}^{(0)}$.
   If $f(\hat{x}_{ij}) < f_{min}$ then set $f_{min} := f(\hat{x}_{ij})$ and $x^{(0)} := \hat{x}_{ij}$

3. Compute an estimation x_{est1} by locally minimizing $f(x)$ over Δ_n by a descent algorithm starting from $x^{(0)}$.

3 Approximating the convex envelope

The second method for approximately solving (1) is a relaxation method. The idea of this method is to replace the (nonconvex) objective function by a convex function. The solution of this simplified problem is used to estimate a global minimizer of the original problem. The relaxation of problem (1) is based on an approximation of the convex envelope. Let $f(x) = x^T F x$ be the objective function and let $f_{conv}(x)$ be the the convex envelope of $f(x)$ taken over Δ_n which is a function satisfying :

(i) $f_{conv}(x)$ is convex on Δ_n.

(ii) $f_{conv}(x) \leq f(x)$ for all $x \in \Delta_n$.

(iii) If $h(x)$ is any convex function defined on Δ_n such that $h(x) \leq f(x)$ for all $x \in \Delta_n$, then $h(x) \leq f_{conv}(x)$ for all $x \in \Delta_n$.

The convex envelope has the following properties (see Horst, Pardalos and Thoai (1995)):

Lemma 2 *Let f^* be the global minimum of (1). Then*

$$f^* = \min_{x \in \Delta_n} f_{conv}(x)$$

and

$$\{x \in \Delta_n : f(x) = f^*\} \subset \{x \in \Delta_n : f_{conv}(x) = f^*\}. \tag{2}$$

Equality holds in (2) if (1) has a unique global minimizer.

In general, computing the convex envelope is as difficult as solving (1) (see Horst, Pardalos and Thoai (1995)). However, it is possible to approximate the convex envelope in polynomial time hoping that the global minimum point of the related optimization problem is in the region of attraction of a global minimum point of problem (1). Consider the following optimization problem:

$$\begin{aligned} W := \ \text{argmin} \quad & \sum_{1 \leq i,j \leq n+1} |F_{ij} - G_{ij}| \\ \text{subject to} \quad & \text{diag}\, G = \text{diag}\, F, \quad G \leq F, \\ & x^T G x \text{ is convex on } \Delta_n \end{aligned} \tag{3}$$

and define the quadratic form $w(x) := x^T W x$. The constraint $\text{diag}\, G = \text{diag}\, F$ ensures that the function values of $w(x)$ and $f(x)$ are equal at the vertices of Δ_n. Obviously, $w(x)$ is convex on Δ_n. We have $w(x) \leq f(x) \quad$ for all $x \in \Delta_n$ (see Lemma 4 in Section 5). Therefore, $w(x)$ is an underestimator of the convex envelope $f_{conv}(x)$ and a minimizer of $w(x)$ over Δ_n is an estimation of a global minimizer of (1). Based on this observation we propose the following algorithm for computing an estimation of a global minimizer of (1) denoted by x_{est2}.

Algorithm 2

1. Compute W by solving (3).
2. Solve the convex quadratic program:

$$x^{(0)} := \arg\min_{x \in \Delta_n} w(x). \tag{4}$$

3. Starting from $x^{(0)}$ compute by a descent method a local minimizer x_{est2} of problem (1).

Problem (3) can be formulated as a semidefinite program (see Nowak (1997)). Since semidefinite programs and convex quadratic programs can be solved in polynomial time the estimation x_{est2} can be computed in polynomial time.

4 Constructing test problems with known solutions

In order to investigate the performance of both algorithms we propose a test case generator which produces problems of the form (1) with known optimal value and prescribed number of binding constraints at the solution point. The method is based on the following result (see also Hager, Pardalos, Roussos and Sahinoglou (1991) for a similar approach):

Lemma 3 *Let* $f_1(x) := \sum_{i=1}^{s} \lambda_i x_i^2$ *where* $\lambda_i < 0$ *(*$1 \le i \le s$*) and* $f_2(x) := \sum_{i=s+1}^{n} \lambda_i x_i^2$ *where* $\lambda_i > 0$ *(*$s+1 \le i \le n$*) be two quadratic forms and let* $v_i := \begin{pmatrix} u_i \\ w_i \end{pmatrix}$ *(*$1 \le i \le n+1$*) be the vertices of a simplex where* $u_i \in \mathbb{R}^s$ *and* $w_i \in \mathbb{R}^{n-s}$*. Consider the quadratic program:*

$$\begin{array}{rl} \textit{global minimize} & f_1(x) + f_2(x) \\ \textit{subject to} & x \in \text{ co } \{v_1, .., v_{n+1}\}. \end{array} \tag{5}$$

If $u_1 = u_2 = .. = u_k$*,* $0 \in \text{ co } \{w_1, .., w_k\}$*,* $1 \le k \le n$ *and* $f_1(u_1) \le f_1(u_i)$ *for* $i \in \{k+1, n+1\}$ *then* $x^* := \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$ *is a global minimizer and* $f^* := f_1(u_1)$ *is the global minimum of the quadratic program (5).*

Proof. Since $f_1(x)$ is concave $\binom{u_1}{0}$ is a global minimizer of $f_1(x)$ over co $\{v_1, .., v_{n+1}\}$. Since $f_2(x) \ge 0$ for $x \in \mathbb{R}^n$ we have $f(x) \ge f_1(x) \ge f_1(x^*)$ for $x \in$ co $\{v_1, .., v_{n+1}\}$. □

The following algorithm produces a quadratic optimization problem of the form (1) with a known optimal value f^* and a solution point x^* based on Lemma 3. The parameters n, s, k and δ ($1 \le s \le n-1$, $1 \le k \le n$, $\delta \in [0,1)$) denote the problem size, the number of negative eigenvalues of the Hessian of $f(x)$, the number of non-binding constraints at x^* and a measure how close local (not global) minima of (1) are to f^*.

random_qp (n, s, k, δ)

1. Choose random values $\lambda_i \in [-b_1, a_1]$, $1 \le i \le s$ and $\lambda_i \in [a_1, b_1]$, $s+1 \le i \le n$ ($0 < a_1 < b_1$).
2. Choose a random value $f^* \in [a_2, b_2]$ ($a_2 < b_2 < 0$).
3. Choose random vectors $\hat{v}_i \in [-a_3, a_3]$ where $\hat{v}_i := \begin{pmatrix} \hat{u}_i \\ \hat{w}_i \end{pmatrix}$, $\hat{u}_i \in \mathbb{R}^s$, $\hat{w}_i \in \mathbb{R}^{n-s}$ and $\hat{u}_1 = \hat{u}_2 = .. = \hat{u}_k$ ($a_3 \in \mathbb{R}^n$, $1 \le i \le n$).
4. Set $u_i := -|f^*/f_1(\hat{u}_i)|^{1/2} \cdot \hat{u}_i$ for $1 \le i \le k$ and $u_i := -|f^* \cdot (1-\delta)/f_1(\hat{u}_i)|^{1/2} \cdot \hat{u}_i$ for $k+1 \le i \le n+1$ where $f_1(x) := \sum_{i=1}^{s} \lambda_i x_i^2$.
5. Choose a random vector $\mu \in \Delta_{k-1}$ and set $x^* := \begin{pmatrix} \mu \\ 0 \end{pmatrix}$.
6. Set $w_i := \hat{w}_i - z$, ($1 \le i \le k$) and $w_i := \hat{w}_i$, ($k+1 \le i \le n+1$) where $z := \sum_{i=1}^{k} \mu_i \hat{w}_i$.
7. Set $v_i := \binom{u_i}{w_i}$ for $1 \le i \le n+1$, $V := (v_1, .., v_{n+1})$ and $F := V^T$ diag $(\lambda)V$ defining the objective function $f(x) = x^T F x$.

5 Modifying the objective function

In this section we discuss modifications of the objective function of (1) which do not change the optimal value. The modifications are based on the following results:

Lemma 4 *Let $F, W \in I\!R^{(n+1,n+1)}$ be symmetric matrices and $W \leq F$ (componentwise). Then*

$$x^T W x \leq x^T F x \quad \text{for} \quad x \in \Delta_n.$$

Proof. Since $x \geq 0$ on Δ_n and $F - W \geq 0$ we have

$$x^T F x - x^T W x = x^T (F - W) x \geq 0 \quad \text{for} \quad x \in \Delta_n$$

which proves the assertion. □

Lemma 5 *There exist a global minimizer $\hat{x}$ of (1) such that $f(x)$ is strictly convex on $\Delta_{\sigma(\hat{x})}$ where $\sigma(\hat{x})$ is the set of non-binding constraints defined by*

$$\sigma(\hat{x}) := \{i : 1 \leq i \leq n+1, \quad \hat{x}_i > 0\}$$

and Δ_I is the I-face of Δ_n, i.e.

$$\Delta_I := \{x \in \Delta_n : x_i = 0, \quad i \in \{1, .., n+1\} \setminus I\}.$$

Proof. Let X^* be the solution set of (1). It is well known that X^* is the union of polyhedra. Let $\hat{x}$ be a vertex of X^*. Then $\Delta_{\sigma(\hat{x})} \cap X^* = \{\hat{x}\}$ implying that $f(x)$ is strictly convex on $\Delta_{\sigma(\hat{x})}$.

□

We are now in the position to prove :

Proposition 1 *Let $f(x) := x^T F x$ be given and define the quadratic form $\tilde{f}(x) := x^T \tilde{F} x$ by*

$$\tilde{F}_{ij} := \min\{F_{ij}, \frac{1}{2}(F_{ii} + F_{jj})\}, \quad 1 \leq i, j \leq n+1. \tag{6}$$

Consider the following optimization problem :

$$\begin{aligned} \text{global minimize} \quad & \tilde{f}(x) \\ \text{subject to} \quad & x \in \Delta_n. \end{aligned} \tag{7}$$

Then the global minima of (7) and (1) are equal and every global minimizer of (7) is also a global minimizer of (1).

Proof. Let f^* and $\tilde{f}^*$ be the global minimum of (1) and (7) respectively. Let $\tilde{x}$ be a global minimizer of (7) such that according to Lemma 5 $\tilde{f}(x)$ is strictly convex on $\Delta_{\sigma(\tilde{x})}$. This implies $\partial^2_{e_i - e_j} \tilde{f}(x) > 0$ for $i, j \in \sigma(\tilde{x})$, $i \neq j$, which is equivalent to $\tilde{F}_{ij} < \frac{1}{2}(\tilde{F}_{ii} + \tilde{F}_{jj})$ for $i, j \in \sigma(\tilde{x})$, $i \neq j$. Since $\tilde{F}_{ii} = F_{ii}$ for $1 \leq i \leq n+1$ it follows $F_{ij} = \tilde{F}_{ij}$ for $i, j \in \sigma(\tilde{x})$ and therefore $\tilde{f}^* = \tilde{f}(\tilde{x}) = f(\tilde{x}) \geq f^*$. Since $\tilde{F}_{ij} \leq F_{ij}$ for $1 \leq i, j \leq n+1$ Lemma 4 yields $\tilde{f}^* \leq f^*$ which proves $\tilde{f}^* = f^*$.

Now let x^* be a global minimizer of (1). Then $\partial^2_{e_i - e_j} f(x) \geq 0$ for $i, j \in \sigma(x^*)$, $i \neq j$ implying $F_{ij} = \tilde{F}_{ij}$ for $i, j \in \sigma(x^*)$ since $F_{ii} = \tilde{F}_{ii}$. Therefore $\tilde{f}^* = f^* = f(x^*) = \tilde{f}(x^*)$. Hence x^* is also a global minimizer of problem (7).

□

A consequence from Proposition 1 is:

Corollary 1 *Let $f(x) := x^T F x$ be given and define the quadratic form $\hat{f}(x) := x^T \hat{F} x$ by*

$$\hat{F}_{ij} = \begin{cases} F_{ij} & \text{if} \quad F_{ij} < \frac{1}{2}(F_{ii} + F_{jj}) \text{ or } i = j \\ \frac{1}{2}(F_{ii} + F_{jj}) + \delta_{ij} & \text{else} \end{cases}$$

where $\delta_{ij} \geq 0$, $1 \leq i, j \leq n+1$. Then every global minimizer x^* of (1) which is strictly convex on $\sigma(x^*)$ is a global minimizer of $\hat{f}(x)$ over Δ_n. (According to Lemma 5 such global minimizers exist.)

Proof. From Lemma 4 and Proposition 1 it follows $f^* \leq \tilde{f}(x) \leq \hat{f}(x)$ for $x \in \Delta_n$ since $\tilde{F}_{ij} \leq \hat{F}_{ij}$, $1 \leq i, j \leq n+1$. Let x^* be a global minimizer of (1) which is strictly convex on $\Delta_{\sigma(x^*)}$. Then $F_{ij} < \frac{1}{2}(F_{ii} + F_{jj})$ for $i, j \in \sigma(x^*), i \neq j$, implying $F_{ij} = \hat{F}_{ij}$ for $i, j \in \sigma(x^*)$. Hence $f^* = f(x^*) = \hat{f}(x^*)$. □

Note that the modification (6) can lead to a reduction of the number of local minimizers of (1). Consider the quadratic form $f(x) := x^t F x$ where $F_{ii} := i, F_{ij} := \frac{1}{2}(i+j) + n + 1$, $1 \leq i, j \leq n+1$. Then $f(x)$ has $n+1$ local minimizers at the vertices of Δ_n and e_1 is a global minimizer. The modified objective function $\tilde{f}(x)$ corresponding to $f(x)$ is linear over Δ_n and has a unique (global) minimum over Δ_n at e_1. However, in the case of random test problems, which are generated with the procedure *random_qp* (see Section 4) the modification (6) reduces the number of minimizers of the Bézier-net (i.e. the number of elements of the set T, see Section 2) almost never. We observed that the modification (6) changes the computational time for solving random test examples up to dimension 30 by Algorithm 1 and by Algorithm 2 only slightly. In Bomze 1997 a modification of Corollary 1 was proved for the maximum clique problem in order to avoid spurious solutions.

6 Numerical results

In order to compare the performance of both algorithms we made several numerical experiments. We constructed random test problems using the procedure *random_qp* and modified the objective function according to Proposition 1. The results are shown in Table 1.

We made always 10 runs and averaged the quantities. The parameter n denotes the problem size, s denotes the number of negative eigenvalues of the Hessian of the objective function and *type* describes different values for k and δ. If *type* $= 1$ we chose $k = 4$ and $\delta = 0.5$, if *type* $= 2$ we chose $k = 4$ and $\delta = 0$ and if *type* $= 3$ we chose $k = n/2$ and $\delta = 0.5$. The percentage relative error of the estimations f_{est1} and f_{est2} is denoted by $e_1 := 100 \cdot \dfrac{|f_{est1} - f^*|}{|2F_{max}|}$ and $e_2 := 100 \cdot \dfrac{|f_{est1} - f^*|}{|2F_{max}|}$ respectively where $F_{max} := \max_{1 \leq i,j \leq n+1} |F_{ij}|$. We have $\max_{x \in \Delta_n} f(x) - f^* \leq 2F_{max}$ since $f(x) \in \text{ co } \{F_{ij} : 1 \leq i, j \leq n+1\}$ (see Dahmen (1986)). The percentage averaged number of cases where the relative errors e_1, e_2 and $\min(e_1, e_2)$ do not exceed 10^{-5} is denoted by opt_1, opt_2 and opt respectively. The CPU time in seconds for computing the estimations f_{est1} and f_{est2} is denoted by $time_1$ and $time_2$ respectively. All computations were performed on a HP J 280 workstation.

The results show that both algorithms find a global minimizer in many instances. In particular, the combination of Algorithm 1 and 2 leads almost always to the exact solution (see column *opt* in Table 1). The most important fact is that the estimations produced by Algorithm 1 are not much more inaccurate than the estimations produced by Algorithm 2, although Algorithm 1 is much faster. An acceleration of Algorithm 2 is probably possible by using advanced SDP-solvers which are currently under investigation. The choice of the parameters k and δ does not seem to influence the results very much.

Table 1: Comparison of Algorithm 1 and Algorithm 2

n	s/n	$type$	e_1	e_2	opt_1	opt_2	opt	$time_1$	$time_2$
10	0.2	1	0	0	100	100	100	0.082	0.531
10	0.2	2	0	0	100	100	100	0.085	0.503
10	0.2	3	0	0	100	100	100	0.071	0.54
10	0.5	1	0.0323	0	80	100	100	0.304	0.492
10	0.5	2	0.00156	0	90	100	100	0.473	0.531
10	0.5	3	0	0	100	100	100	0.166	0.524
10	0.8	1	0.0853	0	70	100	100	0.588	0.637
10	0.8	2	0.127	0.0525	40	70	80	0.71	0.832
10	0.8	3	0.367	0.000478	50	90	90	0.707	0.707
20	0.2	1	0	0	100	100	100	0.696	7.98
20	0.2	2	0	0	100	100	100	0.68	7.81
20	0.2	3	0	0	100	100	100	0.673	7.67
20	0.5	1	0	0	100	100	100	1.1	7.68
20	0.5	2	0.0137	0	90	100	100	1.12	7.86
20	0.5	3	0	0	100	100	100	1.09	8.22
20	0.8	1	0.0504	0	80	100	100	3.1	7.43
20	0.8	2	0.0675	0	50	100	100	3.96	7.53
20	0.8	3	0	0	100	100	100	1.33	8.9
30	0.2	1	0	0	100	100	100	2.95	53.8
30	0.2	2	0	0	100	100	100	2.89	52.2
30	0.2	3	0	0	100	100	100	2.57	67.8
30	0.5	1	0	0	100	100	100	3.32	52
30	0.5	2	0	0	100	100	100	3.26	51.5
30	0.5	3	0	0	100	100	100	4.08	60.3
30	0.8	1	0.0522	0	80	100	100	10.4	51.5
30	0.8	2	0.0482	0	70	100	100	4.87	49.3
30	0.8	3	0.0236	0	90	100	100	4.18	63.5

References

An, L.T.H. and Tao,.P.,D. (1996). "Solving a Class of Linearly Constrained Indefinite Quadratic Problems by D.C. Algorithms", J. Global Optimization 11, 253–285.

Bomze, I. (1997). "Evolution towards the Maximum Clique", J. Global Optimization 10, 143–164.

Bomze, I. (1998). "On standard quadratic optimization problems", to appear in J. Global Optimization 13.

Borchers, B. (1997). "CSDP, a C library for semidefinite programming", manuscript, http://www.nmt.edu/~borchers/csdp.html

Coleman, T.F. and Hulbert, L. A. (1989). "A Direct Active Set Algorithm for Large Sparse Quadratic Programs with Simple Bounds", Math. Program 45, 373–406.

Dahmen, W. (1986). "Bernstein-Bézier Representation of Polynomial Surfaces", proceedings of SIGGRAPH '86, 1–43.

Han, G., Pardalos, P.M., Ye., Y. (1992), " On the Solution of Indefinite Quadratic Problems using an Interior-Point Algorithm", Informatica 3, 474–496.

Horst, R.,Pardalos, P. (1995). 'Handbook of Global Optimization', Kluwer Academic Publishers.

Hager, W.W., Pardalos, P.M., Roussos, I.M. and Sahinoglou, H.D. (1991). "Active Constraints, Indefinite Quadratic Test Problems, and Complexity", J. of Optimization Theory and Applications

68, 499–511.

Helmberg, C., Rendl, F., Vanderbei, R.J. and Wolkowicz, H. (1996). "An interior-point method for semidefinite programming", SIAM J. on Opt. 6 (2), 342–361.

Horst, R., Pardalos, P. and Thoai, N., (1995). Introduction to Global Optimization, Kluwer Academic Publishers.

Nowak, I. (1997). "A new semidefinite programming bound for indefinite quadratic forms over a simplex", to appear in J. Global Optimization.