A Global Optimality Criterion for Nonconvex Quadratic Programming over a Simplex

Ivo Nowak *

Abstract

In this paper we propose a global optimality criterion for globally minimizing a quadratic form over the standard simplex, which in addition provides a sharp lower bound for the optimal value. The approach is based on the solution of a semidefinite program (SDP) and a convex quadratic program (QP). Since there exist fast (polynomial time) algorithms for solving SDP's and QP's the computational time for checking the global optimality criterion and for computing the lower bound is reasonable. Numerical experiments on random test examples up to 30 variables indicate that the optimality criterion verifies a global solution in almost all instances.

Keywords: global optimality criterion, nonconvex quadratic programming, semidefinite programming

AMS subject classification: 90C20, 90C30, 90C06, 65K05

1 Introduction

We consider the global quadratic optimization problem over the standard simplex :

$$\begin{array}{ll} \text{global minimize} & f(x) := x^T F x \\ \text{subject to} & x \in \Delta_n \end{array} \tag{1}$$

where the admissible set is the n-dimensional standard simplex

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

$F \in I\!R^{n+1 \times n+1}$ is an indefinite symmetric matrix and $e \in I\!R^{n+1}$ is the vector of ones. Problems of the type (1) occur for example in the search for a maximum (weighted) clique in an undirected graph. Problem (1) is also strongly related to the problem of minimizing a quadratic form over a polyhedron, which has numerous applications. Indeed, given a polyhedron $D \subset I\!R^n$ and a candidate global minimizer x^* it is possible by cutting planes to split of a simplex from the polyhedron which contains x^*. Finding a global solution - and even checking if a local solution is a global solution - of (1) is known to be NP-hard (see [7]). Several authors proposed branch-and-bound algorithms for solving (1) (see for example [13],[11],[12],[14]). Most of these algorithms use bounding techniques which produce lower bounds which are almost never exact. This can lead to a large number of iterations of a branch-and-bound procedure since the admissible set has to be subdivided very often in order to improve the bounds. It is therefore desirable to have lower bounds which are very accurate or even exact, so that only a moderate number of iterations of a branch-and-bound algorithm are needed to verify global optimality.

In this paper we present a lower bound for (1) which is very accurate and in many instances exact so that it can serve as a global optimality criterion. Since the approach is based on the solution of a semidefinite program (SDP) and a convex quadratic program (QP), which can be solved in polynomial

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

time, the numerical cost is reasonable. To our knowledge this is the first efficient global optimality criterion for (1) which can be checked in polynomial time. Only few global optimality criteria for (1) are available. In [2] a global optimality criterion for (1) is proposed based on the verification of a copositive condition which is again an NP-hard problem. In [8] a global optimality criterion is described which can be easily checked. However, no numerical results are reported which demonstrate the effectiveness of the criterion. In [1] a local optimality criterion for a d.c. formulation of problem (1) is used to compute estimates of global minima which are often exact. An overview of global optimality conditions for general global optimization problems is given in [6].
The paper is organized as follows. In Section 2 we introduce some notations and summerize some results on quadratic forms over the standard simplex. In Section 3 we describe a semidefinite programming bound for (1) which was obtained in [9]. Section 4 is devoted to the new global optimality criterion and lower bound for (1). Numerical experiments on random test examples, which are reported in Section 5, demonstrate the efficiency of the criterion and show that the new lower bound improves considerably a known SDP bound.

2 Notations and some properties of quadratic forms over the standard simplex

We use the following notations. Let $A \in I\!R^{(n+1,n+1)}$ be a given matrix and denote by $A^l \in I\!R^{(n+1,n+1)}$ and by $A^c \in I\!R^{(n+1,n+1)}$ the matrices defined by

$$A^l_{ij} := \frac{1}{2}(A_{ii} + A_{jj}), \quad A^c_{ij} := \frac{1}{2}(A_{ii} - 2A_{ij} + A_{jj}), \quad 1 \le i,j \le n+1. \tag{2}$$

We have obviously $A = A^l - A^c$. The quadratic form $x^T A^l x$ is linear over Δ_n and the entries of the matrix $2 \cdot A^c$ are the second order derivatives of the quadratic form $x^T Ax$ along the edges of Δ_n, i.e. $\partial^2_{e_i - e_j} x^t Ax = 2 \cdot A^c_{ij}$. The notation $B \succeq 0$ means that the matrix B is positive semidefinite. Let $I \subset \{1,..,n+1\}$ be a given index set. The I-face of a standard simplex Δ_n is defined by

$$\Delta_I := \{x \in \Delta_n : x_i = 0, \quad i \notin I\}.$$

The set of non-binding constraints at a point $x \in \Delta_n$ is defined by

$$\sigma(x) := \{i : 1 \le i \le n+1, \quad x_i > 0\}.$$

The complementary set of $\sigma(x)$ is denoted by $\bar{\sigma}(x) := \{1,...,n+1\} \setminus \sigma(x)$. In [10] the following results concerning quadratic forms over Δ_n were proved:

Lemma 1
(i) Let $F, G \in I\!R^{(n+1,n+1)}$ be symmetric matrices and $G \le F$ (componentwise). Then $x^T Gx \le x^T Fx$ for $x \in \Delta_n$.
(ii) Let $\Phi : I\!R^{(n+1,n+1)} \to I\!R^{(n,n)}$ be the linear map defined by $\Phi(G)_{ij} := G_{ij} + G_{n+1,n+1} - G_{n+1,i} - G_{n+1,j}$ $(1 \le i,j \le n)$ where $G \in I\!R^{(n+1,n+1)}$ is a symmetric matrix. A quadratic form $x^T Gx$ is convex on Δ_n if and only if $\Phi(G) \succeq 0$.
(iii) Let

$$E^c := \{ij : F^c_{ij} > 0, \quad 1 \le i < j \le n+1\} \tag{3}$$

be the edge set of edges of Δ_n where the objective function $f(x) = x^T Fx$ is strictly convex. If $\Phi(G) \succeq 0$ and $G_{ij} \le F_{ij}$ for $ij \in E^c$ then $G \le F$.

3 A semidefinite programming bound

In [10] a lower bound for the optimal value of problem (1) was proposed which can be computed by solving a semidefinite program (SDP) and a convex quadratic program (QP). The bound is based on Lemma 1. Consider the following SDP:

$$\begin{array}{rll} W_1 := & \text{argmin} & \text{tr } J(F-G) \\ & \text{subject to} & G_{ij} \le F_{ij}, \quad ij \in E^c \\ & & \text{diag } G = \text{ diag } F, \quad \Phi(G) \succeq 0, \end{array} \tag{4}$$

where Φ is defined as in Lemma 1 (ii) and $J \in I\!R^{(n+1,n+1)}$ is the matrix of ones. Then :

$$b_{sdp1} := \min_{x \in \Delta_n} x^T W_1 x \tag{5}$$

is a lower bound for (1). This follows from Lemma 1 (i) and (iii) since $W_1 \le F$. The constraint $\text{diag } G = \text{ diag } F$ ensures that the function values of $x^T W_1 x$ and $f(x)$ are equal at the vertices of Δ_n. From Lemma 1 (ii) it follows that (5) is a (convex) QP. Since SDP's and QP's can be solved in polynomial time b_{sdp1} can be solved in polynomial time. Preliminary computational experiments in [9] indicate that the bound b_{sdp1} improves considerably the semidefinite programming bound introduced by Shor and others (see [15]).

4 A global optimality criterion and an improved lower bound

In this section we propose a global optimality criterion and a new lower bound for (1). The following result provides a criterion for checking if a local minimizer of (1) is a global minimizer.

Lemma 2 *Let x^* be a local minimizer of (1) and let $g(x)$ be a quadratic form over Δ_n which satisfies:*

(a) $g(x) \le f(x)$ *for* $x \in \Delta_n$,

(b) $g(x) = f(x)$ *for* $x \in \Delta_{\sigma(x^*)}$,

(c) $\partial_{e_i - e_k} g(x^*) \ge 0$ *for* $i \in \bar{\sigma}(x^*)$, *where* $k \in \sigma(x^*)$,

(d) $g(x)$ *is convex on* Δ_n,

where $f(x)$ is the objective function of (1). Then x^* is a global minimizer of (1).

Proof. Because of properties (b) and (c) it follows that x^* is a local minimum point of $g(x)$ over Δ_n. Since $g(x)$ is convex x^* is also a global minimum point. Using (b) and (a) we infer

$$f(x^*) = g(x^*) \le g(x) \le f(x) \quad \text{for } x \in \Delta_n.$$

□

The global optimality criterion of Lemma 2 can also be formulated as an optimization problem which is similar to (4). This formulation provides in addition a lower bound for (1):

Proposition 1 *Let x^* be a local minimizer of (1), $\delta \in I\!R$ a positive penalty parameter and $k \in \sigma(x^*)$. Consider the mixed linear-semidefinite programming problem (LSDP):*

$$\begin{array}{rll} (W_2, s^*) := & \textit{argmin} & \sum_{i,j \in \sigma(x^*)} (F_{ij} - G_{ij}) + \delta \cdot s \\ & \textit{subject to} & -s \le \partial_{e_i - e_k}(x^T G x), \quad i \in \bar{\sigma}(x^*), \\ & & G_{ij} \le F_{ij}, \quad ij \in E^c \\ & & \textit{diag } G = \textit{ diag } F \\ & & \Phi(G) \succeq 0, \quad s \ge 0. \end{array} \tag{6}$$

Then

$$b_{sdp2} := \min_{x\in\Delta_n} x^T W_2 x$$

is a lower bound for (1). If the optimal value of (6) is zero then b_{sdp2} is the optimal value of (1) and x^* is a global minimizer of (1).

Proof. Lemma 1 (iii) implies $W_2 \leq F$ and therefore $b_{sdp2} \leq f^*$ by Lemma 1 (i) where f^* is the optimal value of (1). Let us now assume that the optimal value of (6) is zero. We show that $g(x)$ satisfies conditions (a)-(d) of Lemma 2 where $g(x) := x^T W_2 x$. Property (a) follows from $W_2 \leq F$. Since the optimal value of (6) is zero we have $W_{2,ij} - F_{ij} = 0$ for $i,j \in \sigma(x^*)$ implying (b) and since $s^* = 0$ we have $\partial_{e_i - e_{n+1}} g(x) \geq -s^* = 0$ implying (c). Condition (d) follows from $\Phi(W_2) \succeq 0$ via Lemma 1 (ii). □

We show that problem (6) is well defined. Let $q_{min} := \min\{-F^c_{ij} : 1 \leq i,j \leq n+1\}$ and let $\hat{W} \in I\!R^{(n+1,n+1)}$ be the matrix defined by

$$\hat{W}_{ij} := F^l_{ij} + \begin{cases} q_{min} & \text{if } i \neq j \\ 0 & \text{else} \end{cases}, \quad 1 \leq i,j \leq n+1.$$

Since $F^c_{ii} = 0$ it follows $q_{min} \leq 0$ and therefore $\hat{W} \leq F$. Note that $\Phi(\hat{W}) = -q_{min}(I + J)$ where I is the $n \times n$ identity matrix. Hence convexity of the quadratic form $x^T \hat{W} x$ over Δ_n follows immediately via Lemma 1 (ii) and $\hat{W}$ satisfies the constraints of problem (6).

The number of variables in the LSDP (6) can be reduced by eliminating the constraints $\operatorname{diag} G = \operatorname{diag} F$. Define the linear map $\Psi : I\!R^{(n,n)} \to \{U \in I\!R^{(n+1,n+1)} : \operatorname{diag} U = 0\}$ by

$$\begin{aligned} \Psi(X)_{i,n+1} = \Psi(X)_{n+1,i} &= -\frac{1}{2} X_{ii}, \quad 1 \leq i \leq n \\ \Psi(X)_{ij} &= X_{ij} - \frac{1}{2}(X_{ii} + X_{jj}), \quad 1 \leq i,j \leq n \\ \Psi(X)_{n+1,n+1} &= 0, \end{aligned}$$

where $X \in I\!R^{(n,n)}$. We have $\Phi(\Psi(X) + F^l) = X$, $\operatorname{diag}(\Psi(X) + F^l) = \operatorname{diag} F$ and $\partial_{e_i - e_k}(x^T(\Psi(X) + F^l)x) = 2e_i^T X\tilde{x} - X_{ii} + (e_i - e_k)^T F^l x$ where $\tilde{x} := (x_1, .., x_n)$. Substituting G by $\Psi(X) + F^l$ in problem (6) we obtain the following equivalent semidefinite program with the variable $X \in I\!R^{(n,n)}$:

$$\begin{aligned} W_2 := \Psi(X^*) + F^l, \quad (X^*, s^*) := \ &\text{argmin} && \sum_{i,j\in\sigma(x^*)} (-F^c_{ij} - \Psi(X)_{ij}) + \delta \cdot s \\ &\text{subject to} && -s \leq 2e_i^T X\tilde{x}^* - X_{ii} + (e_i - e_k)^T F^l x^*, \quad i \in \bar{\sigma}(x^*), \\ & && \Psi(X)_{ij} \leq -F^c_{ij}, \quad ij \in E^c, \\ & && X \succeq 0, \quad s \geq 0. \end{aligned} \tag{7}$$

5 Numerical results

We made several numerical experiments to compare the bounds b_{sdp1} and b_{sdp2}. The SDP's were solved using the implementation of Borchers [3] of the interior point algorithm of [5]. The LSDP (7) was transformed into an SDP by defining the matrix $\hat{X} := \begin{pmatrix} X & 0 \\ 0 & s \end{pmatrix}$ and claiming $\hat{X} \succeq 0$. The QP's were solved by a descent method. All computations were performed on a HP J 280 workstation. We used two kinds of test problems. The first test problems were generated using the procedure *random_qp* described in the Appendix with parameters $k = 4$ and $\delta = 0.5$. In [10] it was shown that this procedure generates problems of the form (1) where the optimal value f^* and a global minimizer x^* are known.

Table 1: Comparison of b_{sdp1} and b_{sdp2} using random test problems with known solutions

n	s/n	e_{sdp1}	e_{sdp2}	opt_{sdp1}	opt_{sdp2}	$time_{sdp1}$	$time_{sdp2}$
10	0.2	0.000482	0.000403	80	90	0.531	0.702
10	0.5	0.00241	0	50	100	0.548	0.552
10	0.8	0.0194	0	30	100	0.504	0.98
30	0.2	0.000425	0	90	100	53	75.4
30	0.5	0	0	100	100	51.8	78.1
30	0.8	0.000143	0	90	100	53.2	79.8

Table 2: Comparison of b_{sdp1} and b_{sdp2} using random test problems where a global minimizer is estimated by a heuristic

n	$dens$	e_{sdp1}	e_{sdp2}	opt_{sdp1}	opt_{sdp2}	$time_{sdp1}$	$time_{sdp2}$
10	0.25	0.85	0	40	100	0.153	0.189
10	0.5	2.27	0	0	100	0.231	0.298
10	0.75	2.68	0	10	100	0.257	0.35
30	0.25	3.86	0.73	0	90	6.47	8.66
30	0.5	2.64	0	0	100	10.6	20.7
30	0.75	2.42	0.256	0	80	18.8	39.5

Table 1 displays the numerical results. We made always 10 runs and averaged the quantities. The parameter n denotes the problem size and s denotes the number of negative eigenvalues of the Hessian of the objective function. The percentage relative error of the bound b_{sdp1} and b_{sdp2} is denoted by $e_{sdp1} := 100 \cdot \dfrac{f^* - b_{sdp1}}{|F_{max}|}$ and $e_{sdp2} := 100 \cdot \dfrac{f^* - b_{sdp2}}{|F_{max}|}$ 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 $f(x) \in$ co $\{F_{ij} : 1 \leq i,j \leq n+1\}$ (see [4]). The percentage averaged number of instances where b_{sdp1}, b_{sdp2} are exact is denoted by opt_{sdp1} and opt_{sdp2} respectively.
The second test problems were generated using the procedure *rand_qps* described in the Appendix. A global minimizer of these problems is not known in advance. We estimated a global minimizer using a heuristic which is presented is [10]. The results are displayed in Table 2. Apart from the parameter *dens* which denotes the density of edges of Δ_n where $f(x)$ is strongly convex, i.e. $dens := \frac{2\#E^c}{n(n+1)}$, the notation is as in Table 1.
The results of Table 1 and Table 2 show that b_{sdp2} is in general more accurate than b_{sdp1}. While b_{sdp1} is only exact for some instances, b_{sdp2} is almost always exact. This demonstrates that the new global optimality criterion is efficient. The results demonstrate in addition that the heuristic of [10] works well. It is not clear if the inaccuracies of b_{sdp2} in Table 2 come from inexact estimations of the heuristic. The computational time for computing b_{sdp2} is larger than for computing b_{sdp}. Solving the LSDP (7) directly (without transforming it into an SDP) would save probably computational time.

Appendix: Random test case generators

In order for the reader to be able to reproduce the numerical experiments we describe the methods for generating random test problems of the form (1). The first method is presented in [9]. It produces a quadratic optimization problem with known optimal value f^* and solution x^*. The parameters n, s, k and δ ($1 \leq s \leq n-1$, $1 \leq k \leq n$, $\delta \in [0,1)$) 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 $\lambda_i \in [-b_1, a_1]$, $1 \leq i \leq s$ and $\lambda_i \in [a_1, b_1]$, $s+1 \leq i \leq n$ $(0 < a_1 < b_1)$.
2. Choose a random values $f^* \in [a_2, b_2]$ $(a_2 < b_2 < 0)$.
3. Choose random vectors $\hat{v}_i \in [-a_3, a_3]$ where $\hat{v}_i := \begin{pmatrix} \hat{u}_i \\ \hat{w}_i \end{pmatrix}$, $\hat{u}_i \in I\!R^s$, $\hat{w}_i \in I\!R^{n-s}$ and $\hat{u}_1 = \hat{u}_2 = .. = \hat{u}_k$ $(a_3 \in I\!R^n$, $1 \leq i \leq n)$.
4. Set $u_i := -|f^*/f_1(\hat{u}_i)|^{1/2} \cdot \hat{u}_i$ for $1 \leq i \leq k$ and $u_i := -|f^* \cdot (1-\delta)/f_1(\hat{u}_i)|^{1/2} \cdot \hat{u}_i$ for $k+1 \leq i \leq n+1$ 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 $w_i := \hat{w}_i - z$, $(1 \leq i \leq k)$ and $w_i := \hat{w}_i$, $(k+1 \leq i \leq n+1)$ where $z := \sum_{i=1}^{k} \mu_i \hat{w}_i$.
7. Set $v_i := \binom{u_i}{w_i}$ for $1 \leq i \leq n+1$, $V := (v_1, .., v_{n+1})$ and $F := V^T$ diag $(\lambda) V$ defining the objective function $f(x) = x^T F x$.

The second type of test problems were generated by the following procedures which compute the entries of the matrices F, F^c and F^l:

```
void rand_qps(int n,double dens,double dvert,int &seed,rmatrix &Fc,rmatrix &Fl,
   rmatrix &F)
{
   int i,j,l;
   double r=(4.*double(seed)+1.)/16384./16384.;
   Fc=0.0; seed++;
   for(i=0;i<n;i++)
      for(j=i+1;j<=n;j++)
      if(random(r,0,1)<dens)
        Fc(i,j)=Fc(j,i)=random(r,0.,10.);
      else
        Fc(i,j)=Fc(j,i)=random(r,-10.,0.);
   for(i=0;i<=n;i++) Fl(i,i)=random(r,0,dvert);
   for(i=0;i<n;i++)
      for(j=i+1;j<=n;j++) Fl(i,j)=Fl(j,i)=0.5*(Fl(i,i)+Fl(j,j));
   for(i=0;i<=n;i++)
      for(j=i;j<=n;j++) F(i,j)=F(j,i)=Fl(i,j)-Fc(i,j);
}

double random(double &r,double a,double b)
{
   r=fmod(r*41475557.,1.);
   return(r*(b-a)+a);
}
```

The parameter `seed` is initialized by one.

References

[1] L.T.H An and P.D. Tao, "Solving a Class of Linearly Constrained Indefinite Quadratic Problems by D.C. Algorithms", *J. Global. Optimiz.* 11 (1996) 253-285.

[2] I.H. Bomze, "Copositive conditions for global optimality in indefinite quadratic programming problems", *Czechoslovak J. Oper. Res.* 1(1992) 7-19.

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

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

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

[6] J.B. Hiriart-Urruty, "Conditions for Global Optimality", In *Handbook of Global Optimization* (ed. R. Horst and P. Pardalos), (1995) 1-26.

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

[8] A. Neumaier, "An Optimality Criterion for Global Quadratic Optimization", *J. Global Optimization* 2 (1992) 201-208.

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

[10] I. Nowak, "Some Heuristics and Testproblems for Nonconvex Quadratic Programming over a Simplex", (1998) to appear.

[11] P. Pardalos, J.H. Glick, J.B. Rosen, "Global Minimization of Indefinite Quadratic Problems", *Computing* 39 (1987) 281-291.

[12] Phillips, A.T. and Rosen J.B., Guaranteed ϵ-approximate Solution for Indefinite Quadratic Global Minimization, *Naval Research Logistic* **37**, (1990) 499-514.

[13] T. Q. Phong, L.T.H. An, P.D. Tao, "On the global solution of linearly constrained indefinite quadratic minimization problems by decomposition branch and bound method", *RAIRO, Recherche Opérationnelle* 30 (1996) 31-49.

[14] H.D. Sherali and C.H. Tuncbilek, "A Reformulation-Convexification Approach for Solving Nonconvex Quadratic Programming Problems", *J. of Global Optimization* 7 (1995) 1-31.

[15] N.Z. Shor, *Nondifferentiable Optimization and Polynomial Problems* (Kluwer Academic Publishers, 1998).