New Results on Quadratic Minimization *

Yinyu Ye[†] Shuzhong Zhang[‡]

April 2001

Abstract

In this paper we present several new results on minimizing an indefinite quadratic function under quadratic/linear constraints. The emphasis is placed on the case where the constraints are two quadratic inequalities. This formulation is known as *the extended trust region subproblem* and the computational complexity of this problem is still unknown. We consider several interesting cases related to this problem and show that for those cases the corresponding SDP relaxation admits no gap with the true optimal value, and consequently we obtain polynomial time procedures for solving those special cases of quadratic optimization. For the extended trust region subproblem itself, we introduce a parameterized problem and prove the existence of a trajectory which will lead to an optimal solution. Combining with a result obtained in the first part of the paper, we propose a polynomial-time solution procedure for the extended trust region subproblem arising from solving nonlinear programs with a single equality constraint.

Keywords: Quadratic minimization, SDP relaxation, parameterization.

*Technical Report SEEM2001-03, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong.

[†]Department of Management Sciences, University of Iowa, Iowa City, USA and Visiting Professor of the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

[‡]Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong.

1 Introduction

This paper is concerned with solving quadratic optimization problems by means of semidefinite programming (SDP). In particular, we focus on indefinite quadratic optimization with two or more quadratic constraints.

In the literature, quadratic optimization has received much attention. It is a fundamental problem in optimization theory and practice. Economic equilibrium, combinatorial optimization, numerical partial differential equations, and general nonlinear programming are all sources of quadratic optimization.

Recently, there were several results on quadratic optimization. For examples, Bellare and Rogaway [1] established several negative result on approximating this problem; Goemans and Williamson [4], using an SDP relaxation, proved an approximation result for the Maxcut problem which is a special quadratic optimization problem; Nesterov [6] and Ye [12] extended their SDP relaxation to approximate quadratic optimization with simple bound and diagonally homogeneous quadratic constraints; Nesterov [7] and Nemirovskii et al. [5] established a quality bound when the constraints are convex and homogeneous; Fu et al. [3] constructed a quality bound for approximating quadratic optimization for general convex quadratic constraints.

More recently, Sturm and Zhang [11] proposed a quite different approach to quadratic optimization. They introduced a concept called matrix co-positivity over a domain; that is a set of matrices, which, in the quadratic form, is nonnegative over the given domain. For several specific choices of the domain, Sturm and Zhang [11] proved that such matrix set can be characterized using Linear Matrix Inequalities (LMI). Examples of such domains are: (1) the level set of an arbitrary quadratic function; (2) the contour of an arbitrary quadratic function at zero level; (3) the intersection of the level set of a convex quadratic function and a half-space. The key techniques used in [11] include a dual cone representation approach, and a specific matrix rank-one decomposition scheme. As a consequence of the results in [11], optimizing an indefinite quadratic function under a single (non-convex) quadratic constraint (equality or inequality), or under a convex quadratic inequality constraint and a linear inequality constraint, can be solved in polynomial time, by, first solving a specific form of SDP relaxation, followed by a matrix decomposition procedure.

In the current paper, we consider quadratic optimization directly. It turns out that there are more classes of quadratic optimization problems for which the SDP relaxation is exact, in the sense that its optimal value is equal to the true optimal value, and an optimal solution for the original problem can be obtained from the optimal solution of the SDP relaxation. More specifically, we extend in Section 2 the matrix decomposition idea to solve the following classes of non-convex quadratic minimization problems with two quadratic constraints:

(1) one of the two constraints in the SDP relaxation is not binding;

(2) the two constraint functions and the objective are all homogeneous quadratic functions;

(3) one ellipsoidal and a linear complementarity constraint.

It was unknown whether or not that these problems were polynomially solvable.

The problem of minimizing an indefinite quadratic function with two (general) convex quadratic constraints arises from applying the trust region method to solve equality constrained nonlinear programs. The method of such type was first proposed by Celis, Dennis, and Tapia in [2]. To distinguish it from the usual trust region subproblem, which is minimizing an indefinite quadratic function over a unit ball, we call the above problem *the extended trust region subproblem.* Although some of the cases discussed in the previous paragraph can be considered as special cases of the extended trust region subproblem, the computational complexity of the latter problem is still unknown. In Section 3 of the current paper, we introduce a parameterized problem, and show that by following a trajectory generated by the parameterized problem, one will arrive at the optimal solution of the original problem. Some examples are worked out in the same section to show how the method works.

Then, we consider the extended trust region subproblem for nonlinear programming with one equality constraint. By combining results in Section 2 and 3, we present in Section 4 a polynomial-time procedure for solving the subproblem. Some discussions and conclusions can be found in the same section.

Notations: We let $\|\cdot\|$ denote the Euclidean norm. e_i is the unit vector where the ith component is 1 and others are all 0. '$X \succeq 0$' stands for the fact that the symmetric matrix X is positive semidefinite. '$X \bullet Y := \text{tr }(X^T Y)$' is the usual matrix inner product. For a quadratic function $q(x) = x^T Q x - 2b^T x + c$ we denote the matrix representation of the function q be $M(q(\cdot)) = \begin{bmatrix} Q, & -b \\ -b^T, & c \end{bmatrix}$. '$SOC$' stands for the second order cone.

2 Exact SDP relaxations

This section is concerned with quadratic optimization whose SDP relaxation admit no gap with the true optimal value, and whose optimal solution can be found in polynomial time using the SDP optimal solution.

Formally we consider the following general quadratic optimization problem:

$$\begin{array}{lll} (Q) & \text{minimize} & x^T Q_0 x - 2b_0^T x \\ & \text{subject to} & x^T Q_i x - 2b_i^T x + c_i \leq 0,\ i = 1, ..., m. \end{array}$$

Let $q_i(x) = x^T Q_i x - 2b_i^T x + c_i$, $i = 1, ..., m$.

We assume throughout the paper that the Slater regularity condition is satisfied, i.e., there exists x_0 such that $q_i(x^0) < 0$ for all $i = 1, ..., m$.

For convenience, we adopt the following notation. For a quadratic function $q(x) = x^T Qx - 2b^T x + c$ we denote its matrix representation by

$$M(q(\cdot)) = \begin{bmatrix} c, & -b^T \\ -b, & Q \end{bmatrix}.$$

The homogenized version of (Q) is

$$\begin{array}{lll} (HQ) & \text{minimize} & x^T Q_0 x - 2b_0^T xt \\ & \text{subject to} & x^T Q_i x - 2b_i^T xt + c_i t^2 \leq 0, \ i = 1, ..., m \\ & & t^2 = 1. \end{array}$$

Clearly, if $\begin{bmatrix} t \\ x \end{bmatrix}$ solves (HQ), then x/t solves (Q).

The so-called semidefinite programming relaxation of (HQ) is

$$\begin{array}{lll} (SP) & \text{minimize} & M(q_0(\cdot)) \bullet X \\ & \text{subject to} & M(q_i(\cdot)) \bullet X \leq 0, \ i = 1, ..., m \\ & & X_{00} = 1, \ X \succeq 0, \end{array}$$

where $X = \begin{bmatrix} X_{00}, & x_0^T \\ x_0, & \bar{X} \end{bmatrix}$, $q_i(x) = x^T Q_i x - 2b_i^T x + c_i$ for $i = 0, 1, ..., m$.

The SDP problem (SP) has a dual, which is given by

$$\begin{array}{lll} (SD) & \text{maximize} & y_0 \\ & \text{subject to} & Z = M(q_0(\cdot)) - y_0 \begin{bmatrix} 1, & 0 \\ 0, & 0 \end{bmatrix} + \sum_{i=1}^m y_i M(q_i(\cdot)) \\ & & Z \succeq 0, \ y_i \geq 0, \ i = 1, ..., m. \end{array}$$

Since (Q) satisfies the Slater condition, it follows that (SP) satisfies the Slater condition too.

Additionally, we assume that (SD) satisfies the Slater condition as well. This is true at least for following two interesting cases, as shown by the next proposition.

Proposition 1 *The problem (SD) satisfies the Slater regularity condition, either when at least one of the m constraints is ellipsoidal, or, when the objective function is strictly convex.*

Proof. In the first case, let us assume without loss of generality, that the first constraint is ellipsoidal. In mathematical terms, this means that $Q_1 \succ 0$ and $c_1 - b_1^T A_1^{-1} b_1 < 0$.

By fixing $y_2 = \cdots = y_m = 1$ and letting $y_1 > 0$ be sufficiently large, we will have

$$Q_0 + \sum_{i=1}^{m} y_i Q_i \succ 0.$$

Then, we let $y_0 < 0$ be sufficiently large in absolute value to obtain

$$M(q_0(\cdot)) - y_0 \begin{bmatrix} 1, & 0 \\ 0, & 0 \end{bmatrix} + \sum_{i=1}^{m} y_i M(q_i(\cdot)) \succ 0.$$

In the second case, the objective function is strictly convex, i.e. $Q_0 \succ 0$. In that case, we let $y_i = \epsilon > 0$ be sufficiently small, $i = 1, ..., m$, and $y_0 < 0$ be sufficiently large in absolute value. The Slater condition follows from the fact that $M(q_0(\cdot)) - y_0 \begin{bmatrix} 1, & 0 \\ 0, & 0 \end{bmatrix} \succ 0$.

Q.E.D.

Following a well known result in optimization (see, e.g. [8]), if both (SP) and (SD) satisfy the Slater condition, then they have complementary optimal solutions.

In the subsequent discussion we are mainly concerned with the case where $m = 2$, and we assume that the assumptions in Proposition 1 are satisfied.

Before proceeding we first quote a matrix decomposition result proved in [11]:

Lemma 1 *Let G be an arbitrary symmetric matrix. Let X be a positive semidefinite matrix with rank r. Suppose that $G \bullet X \leq 0$. Then there exists a rank-one decomposition for X such that*

$$X = \sum_{i=1}^{r} x_i x_i^T$$

and $x_i^T G x_i \leq 0$ for all $i = 1, ..., r$. If, in particular, $G \bullet X = 0$, then $x_i^T G x_i = 0$ for all $i = 1, ..., r$.

2.1 Non-binding SDP relaxation

In this subsection we consider (Q) with $m = 2$, and at least one of the two constraints $M(q_i(\cdot)) \bullet X \leq 0$, $i = 1, 2$, is not binding at the optimality. Without loss of generality, suppose that $M(q_2(\cdot)) \bullet X < 0$. This implies, by complementarity, that $y_2 = 0$ at optimality. Let X^* be an optimal solution of

(SP). By applying Lemma 1 we get a rank-one decomposition of X^* such that

$$X^* = \sum_{j=1}^{r} x_j^* (x_j^*)^T, \text{ and } (x_j^*)^T M(q_1(\cdot)) x_j^* = 0, \text{ for all } j = 1, ..., r, \tag{1}$$

where

$$0 \neq x_j^* = \begin{bmatrix} t_j^* \\ \bar{x}_j^* \end{bmatrix}, \, j = 1, ..., r.$$

Since $M(q_2(\cdot)) \bullet X^* = \sum_{j=1}^{r} (x_j^*)^T M(q_2(\cdot)) x_j^* < 0$, there must exist k with $1 \leq k \leq r$ such that

$$(x_k^*)^T M(q_2(\cdot)) x_k^* \leq 0. \tag{2}$$

Since (SD) satisfies the Slater condition, we have $t_k^* \neq 0$, because otherwise the primal optimal set will be unbounded, which is impossible due to the dual Slater condition. (For a detailed account for the duality relations for conic optimization, one is referred to the Ph.D. thesis of Sturm; see [10]).

It follows from (1) and (2) that

$$M(q_1(\cdot)) \bullet \left(\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix} \cdot \left[1, (\bar{x}_k^*/t_k^*)^T\right] \right) = 0 \text{ and } M(q_2(\cdot)) \bullet \left(\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix} \cdot \left[1, (\bar{x}_k^*/t_k^*)^T\right] \right) \leq 0. \tag{3}$$

Let (y^*, Z^*) be an optimal solution for (SD). By complementarity we have $X^* Z^* = 0$. It follows therefore

$$\sum_{j=1}^{r} (x_j^*)^T Z^* x_j^* = 0,$$

and consequently,

$$(x_j^*)^T Z^* x_j^* = 0$$

for all $j = 1, ..., r$. In particular,

$$(x_k^*)^T Z^* x_k^* = 0$$

and so

$$Z^* \bullet \left(\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix} \cdot \left[1, (\bar{x}_k^*/t_k^*)^T\right] \right) = 0. \tag{4}$$

Combining (3) and (4), noting that $M(q_1(\cdot)) \bullet \left(\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix} \cdot \left[1, (\bar{x}_k^*/t_k^*)^T\right] \right) = 0$ and $y_2^* = 0$, we conclude that $\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix} \cdot \left[1, (\bar{x}_k^*/t_k^*)^T\right]$ is an optimal solution for (SP) as well. Note that (SP) is a relaxation of (Q). Therefore, $\begin{bmatrix} 1 \\ \bar{x}_k^*/t_k^* \end{bmatrix}$ must be an optimal solution for (Q). All the procedures described above, including solving the SDP relaxation (SP) and the rank-one decomposition procedure in Sturm and Zhang [11], are polynomial. This leads to the following result:

Theorem 1 *Suppose that (SP) and (SD) both satisfy the Slater condition and $m = 2$. Furthermore, suppose the primal problem (SP) has at least one non-binding constraint at optimality. Then, (Q) can be solved in polynomial time.*

One consequence of Theorem 1 is the following:

Corollary 1 *Suppose that (SP) and (SD) both satisfy the Slater condition and $m = 2$. Furthermore, suppose that, either $q_1(x) > 0$ for all x satisfying $q_2(x) = 0$, or $q_1(x) < 0$ for all x satisfying $q_2(x) = 0$. Then, (Q) can be solved in polynomial time.*

Proof.

Let us assume that $q_1(x) > 0$ for all $q_2(x) = 0$. The other case can be treated similarly. Obviously, we may also assume that there are x_1 and x_2 such that $q_2(x_1) < 0$ and $q_2(x_2) > 0$, for otherwise the constraint $q_2(x) \le 0$ will become redundant, in which case the polynomial-time solvability is already known.

Now we use Corollary 6 in [11] to conclude that there is t such that

$$M(q_1(\cdot)) + tM(q_2(\cdot)) \succeq 0. \tag{5}$$

By the Slater condition of (Q) we know that $M(q_1(\cdot)) \not\succeq 0$. Thus $t \neq 0$.

Now we wish to show that the SDP relaxation (SP) cannot be binding at any feasible solution. Suppose by contradiction that there is a feasible $X \succeq 0$ for (SP) such that $M(q_1(\cdot)) \bullet X = 0$ and $M(q_2(\cdot)) \bullet X = 0$. Then,

$$(M(q_1(\cdot)) + tM(q_2(\cdot))) \bullet X = 0. \tag{6}$$

By Lemma 1, we can get a rank-one decomposition of X, $X = \sum_{i=1}^{r} x_i x_i^T$, such that

$$x_i^T M(q_1(\cdot)) x_i = 0,$$

for all $i = 1, ..., r$. By (5) and (6) we also have

$$x_i^T (M(q_1(\cdot)) + tM(q_2(\cdot))) x_i = 0$$

for $i = 1, ..., r$. Because $t \neq 0$, it follows that

$$x_i^T M(q_2(\cdot)) x_i = 0,$$

for all $i = 1, ..., r$.

Since $X_{00} = 1$ there must exist $x_j = \begin{bmatrix} t_j \\ \bar{x}_j \end{bmatrix}$ such that its first component t_j is nonzero. Then, we have

$$q_1(x_j/t_j) = x_j^T M(q_1(\cdot)) x_j / t_j^2 = x_j^T M(q_2(\cdot)) x_j / t_j^2 = q_2(x_j/t_j) = 0.$$

This contradicts the assumption that $q_1(x) > 0$ for all $q_2(x) = 0$.

Q.E.D.

As an application, we consider the following quadratic program

$$\begin{array}{ll} \text{minimize} & q_0(x) \\ \text{subject to} & l \le q_1(x) \le u \end{array}$$

where $l < u$. This problem was analyzed thoroughly by Stern and Wolkowicz in [9].

This problem clearly satisfies the conditions in Corollary 1, because

$$q_1(x) - u < 0 \text{ for all } x \text{ satisfying } q_1(x) - l = 0.$$

Therefore, according to Corollary 1 it is solvable in polynomial time.

Interestingly, this lends to solving the following problem

$$\begin{array}{ll} \text{minimize} & |q_0(x)| \\ \text{subject to} & q_1(x) \le 0. \end{array}$$

The key is to rewrite the problem as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & q_1(x) \le 0 \\ & -t \le q_0(x) \le t, \end{array}$$

and observe that for any fixed $t \ge 0$, the feasibility check of the above problem reduces to

$$\begin{array}{ll} \text{minimize} & q_1(x) \\ \text{subject to} & -t \le q_0(x) \le t. \end{array}$$

If the optimal value of this problem is positive, then the original problem is infeasible for that given t; otherwise, it is feasible.

For $t = 0$, this problem can be solved by the SDP relaxation method, as it reduces to one quadratic equality constraint. Otherwise, $t > 0$, and we may resort to Theorem 1 for its solution.

Since the feasibility check can be done in polynomial time for any given objective value t, we may solve the optimization problem using bi-section on the objective value.

There are other nontrivial domains that are claimed by Corollary 1, such as the whole space with two non-intersecting ellipsoids taken away. Minimizing an indefinite quadratic function over such a domain, claims Corollary 1, is easy.

2.2 Homogeneous quadratic functions

Another polynomially solvable special case of (Q) is $m = 2$ and all the functions involved, $q_0(x)$, $q_1(x)$ and $q_2(x)$, are homogeneous, i.e., there are no linear terms. Hence, the problem can be simply written as

$$\begin{array}{ll} \text{minimize} & x^T Q_0 x \\ \text{subject to} & x^T Q_1 x \leq 1 \\ & x^T Q_2 x \leq 1. \end{array}$$

Since the problem is already homogeneous, we do not need to further homogenize it. Therefore, the SDP relaxation is

$$\begin{array}{ll} \text{minimize} & Q_0 \bullet X \\ \text{subject to} & Q_1 \bullet X \leq 1 \\ & Q_2 \bullet X \leq 1 \\ & X \succeq 0. \end{array}$$

Its dual problem is

$$\begin{array}{ll} \text{maximize} & y_1 + y_2 \\ \text{subject to} & Z = Q_0 - y_1 Q_1 - y_2 Q_2 \\ & Z \succeq 0,\ y_1 \leq 0,\ y_2 \leq 0. \end{array}$$

Suppose that the primal-dual problems have a pair of complementary optimal solutions. Again, a sufficient condition to ensure this is that one of the Q_1, Q_2 matrix is positive definite.

If one of the two constraints in the primal SDP relaxation is not binding at the optimality, then the results in Subsection 2.1 apply, and the problem is solved.

Consider the case where they are all binding at the optimality. Let the primal optimal solution be X^*, and the dual optimal solution be (y_1^*, y_2^*, Z^*).

We now apply Lemma 1 to generate

$$X^* = \sum_{i=1}^{r} x_i^* (x_i^*)^T$$

such that $(x_i^*)^T (Q_1 - Q_2) x_i^* = 0$ for all $i = 1, ..., r$.

Since $\sum_{i=1}^r (x_i^*)^T Q_1 x_i^* = 1$, we may select x_j^*, $1 \leq j \leq r$, such that $(x_j^*)^T Q_1 x_j^* =: \tau > 0$. By our construction, $(x_j^*)^T Q_2 x_j^* = \tau$.

Let

$$x^* = x_j^* / \sqrt{\tau}.$$

We see that $x^*(x^*)^T$ is a primal feasible solution for the SDP relaxation. Moreover, it is optimal, because

$$0 \leq (x^*)^T Z^* x^* \leq X^* \bullet Z^* = 0$$

and $[1-(x^*)^TQ_ix^*]y_i^* = 0 \times y_i^* = 0$ for $i = 1, 2$, and hence the primal-dual complementarity conditions are satisfied.

This shows that the SDP relaxation admits no gap with the true optimal value, and an optimal solution for the original quadratic optimization problem can be constructed in polynomial time.

2.3 Complementary linear constraints

In this subsection, we consider the following special case of (Q):

$$\begin{array}{lll} (CL) & \text{minimize} & q_0(x) \\ & \text{subject to} & \|x\|^2 \le 1 \\ & & \bar{a}^T x \le a_0 \\ & & \bar{b}^T x \le b_0 \\ & & (a_0 - \bar{a}^T x)(b_0 - \bar{b}^T x) = 0. \end{array}$$

The last constraint is a complementarity condition.

In [11], the above problem with $b_0 = 0$ and $\bar{b} = 0$ is solved via a special type of SDP relaxation. We now extend the method to solve (CL). Let

$$J = \begin{bmatrix} 1, & 0 \\ 0, & -I \end{bmatrix},\ a = \begin{bmatrix} a_0 \\ -\bar{a} \end{bmatrix},\ b = \begin{bmatrix} b_0 \\ -\bar{b} \end{bmatrix}.$$

Recall that we denote the standard second order cone in $\Re^{n+1}$ to be

$$SOC = \left\{ \begin{bmatrix} t \\ x \end{bmatrix} \,\middle|\, t \ge \|x\| \right\}.$$

Consider the following SDP relaxation for the homogenized version of (CL):

$$\begin{array}{lll} (CLSP) & \text{minimize} & M(q_0(\cdot)) \bullet X \\ & \text{subject to} & J \bullet X \ge 0 \\ & & Xa \in SOC \\ & & Xb \in SOC \\ & & a^T X b = 0 \\ & & X_{00} = 1 \\ & & X = \begin{bmatrix} X_{00}, & x_0^T \\ x_0, & \bar{X} \end{bmatrix} \succeq 0. \end{array}$$

Clearly, $(CLSP)$ is a relaxation of (CL), since, if X is rank one, then its eigenvector is simply a solution of (CL). This problem has a dual, given as follows

$$
\begin{array}{lll}
(CLSD) & \text{maximize} & y_1 \\
 & \text{subject to} & Z = M(q_0(\cdot)) - y_0 J - y_1 e_1 e_1^T \\
 & & \qquad -(a y_a^T + y_a a^T) - (b y_b^T + y_b b^T) - y_2(ab^T + ba^T) \\
 & & y_a \in SOC \\
 & & y_b \in SOC \\
 & & y_0 \geq 0 \\
 & & Z \succeq 0.
\end{array}
$$

Let us now assume that the regularity condition is satisfied so that $(CLSP)$ and $(CLSD)$ have complementary optimal solutions, denoted by X^* and $(y_0^*, y_1^*, y_2^*, y_a^*, y_b^*, Z^*)$ respectively. That is,

$$X^* Z^* = 0,\ y_0^*(J \bullet X^*) = 0,\ (y_a^*)^T X^* a = 0,\ (y_b^*)^T X^* b = 0. \tag{7}$$

The main result in this subsection is the following assertion:

Theorem 2 *Suppose that the SDP relaxation $(CLSP)$ and it dual problem $(CLSD)$ have complementary optimal solutions. Then, the optimal value of $(CLSP)$, which equals that of $(CLSD)$ by strong duality, is identical to the optimal value of (CL). In other words, the relaxation admits no gap. Moreover, an optimal solution for (CL) can be obtained in polynomial time, provided that the solution for its SDP relaxation $(CLSP)$ is available.*

Proof. The proof below uses the matrix rank-one decomposition procedure proposed by Sturm and Zhang in [11]. The idea is to construct a rank-one feasible solution for $(CLSP)$, based on X^*, such that the complementarity conditions are still satisfied, thus ensuring the optimality. We proceed by considering several possible cases regarding the status of X^*.

Case 1. Either $X^* a = 0$ or $X^* b = 0$.

Let us assume $X^* a = 0$.

Applying Lemma 1, we obtain a decomposition for X^*:

$$X^* = \sum_{i=1}^{r} x_i^* (x_i^*)^T$$

where r is the rank of X^*, such that $J \bullet [x_i^*(x_i^*)^T] \geq 0$ for all $i = 1, ..., r$. Moreover, $J \bullet [x_i^*(x_i^*)^T] = 0$ for all $i = 1, ..., r$, if $J \bullet X^* = 0$. We may choose the sign of the first component in x_i^* to ensure that $x_i^* \in SOC$, $i = 1, ..., r$.

By linear independence of x_i^*'s, we get $a^T x_i^* = 0$, $i = 1, ..., r$. Let $x_i^* = \begin{bmatrix} t_i^* \\ \bar{x}_i^* \end{bmatrix}$, $i = 1, ..., r$. Since $x_i^* \in SOC$ and $x_i^* \neq 0$, we have $t_i^* > 0$, $i = 1, ..., r$. Furthermore, $X^* b = \sum_{i=1}^r (b^T x_i^*) x_i^* \in SOC$. This implies that there must exist j, $1 \leq j \leq r$, such that $b^T x_j^* \geq 0$. It follows that $\begin{bmatrix} 1 \\ \bar{x}_j^*/t_j^* \end{bmatrix} [1, (\bar{x}_j^*/t_j^*)^T]$ is optimal for $(CLSP)$.

Case 2. $J \bullet X^* > 0$.

In this case, $y_0^* = 0$, and $X^* a \neq 0$. Let $x_a^* := X^* a = \begin{bmatrix} t_a^* \\ \bar{x}_a^* \end{bmatrix} \neq 0$. Since $x_a^* \in SOC$, by feasibility, we know that $t_a^* > 0$. Moreover, $J \bullet [x_a^* (x_a^*)^T] = (t_a^*)^2 - \|\bar{x}_a^*\|^2 \geq 0$, $x_a^* (x_a^*)^T b = 0$, and $x_a^* (x_a^*)^T a = (a^T X^* a) X^* a \in SOC$. Therefore, $x_a^* (x_a^*)^T / (t_a^*)^2$ is optimal for $(CLSP)$ as well, as it is feasible and satisfies the complementarity conditions stipulated in (7), after replacing X^* by $x_a^* (x_a^*)^T / (t_a^*)^2$.

Case 3. $J \bullet X^* = 0$.

Again, denote $x_a^* = X^* a$ and $x_b^* = X^* b$. In this particular case, $x_a^* \neq 0$ and $x_b^* \neq 0$.

Observe that

$$\tilde{X} := X^* - \frac{X^* a a^T X^*}{a^T X^* a} - \frac{X^* b b^T X^*}{b^T X^* b} \succeq 0.$$

Moreover, $\tilde{X} a = 0$ and $\tilde{X} b = 0$.

Consider two more possibilities.

Case 3.1. Either $J \bullet [x_a^* (x_a^*)^T] = 0$ or $J \bullet [x_b^* (x_b^*)^T] = 0$.

Let us assume $J \bullet [x_a^* (x_a^*)^T] = 0$. In this particular case, $x_a^* \neq 0$. Hence, just as in the previous case, $x_a^* (x_a^*)^T / (t_a^*)^2$ is optimal for $(CLSP)$.

Case 3.2. $J \bullet [x_a^* (x_a^*)^T] > 0$ and $J \bullet [x_b^* (x_b^*)^T] > 0$.

In this case,

$$J \bullet \tilde{X} = J \bullet X^* - J \bullet [x_a^* (x_a^*)^T] / a^T X^* a - J \bullet [x_b^* (x_b^*)^T] / b^T X^* b < 0. \quad (8)$$

Now let us decompose $\tilde{X}$ as

$$\tilde{X} = \sum_{i=1}^{s} \tilde{x}_i \tilde{x}_i^T$$

where $s = \text{rank } (\tilde{X}) > 0$. Since $\tilde{X} a = 0$ and $\tilde{X} b = 0$ we have

$$\tilde{x}_i^T a = 0 \text{ and } \tilde{x}_i^T b = 0$$

for all $i = 1, ..., s$. Choose j such that

$$J \bullet [\tilde{x}_j(\tilde{x}_j)^T] < 0.$$

Such j must exist due to (8).

Consider the following quadratic equation

$$J \bullet [(x_a^* + \alpha\tilde{x}_j)(x_a^* + \alpha\tilde{x}_j)^T] = 0.$$

This equation has two distinct real roots with opposite signs. Choose α with an appropriate sign so that the first component in $x_a^* + \alpha\tilde{x}_j$ is positive. Denote

$$x_a^* + \alpha\tilde{x}_j =: \begin{bmatrix} t^* \\ \bar{x}^* \end{bmatrix}.$$

In this case, $\begin{bmatrix} 1 \\ \bar{x}^*/t^* \end{bmatrix} [1, \bar{x}^*/t^*]$ is optimal for $(CLSP)$.

Q.E.D.

We remark here that the solution methodology readily extends to the following more general setting

$$\begin{array}{ll} \text{minimize} & q_0(x) \\ \text{subject to} & \|x\|^2 \le 1 \\ & \bar{a}_i^T x \le a_{i0},\ i = 1, ..., m \\ & (a_{i0} - \bar{a}_i^T x)(a_{j0} - \bar{a}_j^T x) = 0,\ \text{for all } i \neq j. \end{array}$$

3 Two convex quadratic constraints

Now we move on to consider the problem of minimizing a non-convex quadratic function with two convex quadratic constraints. We assume that one of the constraints is simply ellipsoidal. More specifically, without losing generality, we assume it to be a unit spherical constraint.

Let

$$\begin{array}{l} q_0(x) = \frac{1}{2}x^T Q_0 x - b_0^T x \\ q_1(x) = \frac{1}{2}x^T Q_1 x - b_1^T x + \frac{c_1}{2} \end{array}$$

where Q_0 is indefinite and $Q_1 \succeq 0$. Hence, $q_1(x)$ is convex.

The problem that we consider in this section is

$$\begin{array}{lll} (P) & \text{minimize} & q_0(x) \\ & \text{subject to} & \|x\|^2 \le 1 \\ & & q_1(x) \le 0. \end{array}$$

As we discussed in Section 1, this problem arises from the application of the trust region method for equality constrained nonlinear programming. More discussions on this subject can be found in Section 4.

Throughout our discussion we assume that the above problem satisfies the Slater condition, i.e., there is x such that $q_1(x) < 0$ and $\|x\|^2 < 1$. Let us denote the feasible region of (P) to be Ω. Obviously, Ω is a compact convex set, with a non-empty interior.

Let the optimal value of (P) be v^*.

Consider the following parameterized problem

$$\begin{array}{lll} (H_\lambda) & \text{minimize} & q_0(x) + \lambda q_1(x) \\ & \text{subject to} & \|x\|^2 \leq 1 \\ & & q_1(x) \leq 0 \end{array}$$

with $\lambda \geq 0$. Let the optimal value of (H_λ) be $h(\lambda)$.

Lemma 2 *The value function $h(\lambda)$ is non-increasing and is concave. Moreover, $h(\lambda) \leq h(0) = v^*$ for all $\lambda \geq 0$.*

Proof. The concavity of $h(\lambda)$ follows from the fact that for any fixed x, $q_0(x) + \lambda q_1(x)$ is linear, and hence concave, in λ. Moreover, it is non-increasing since $q_1(x) \leq 0$ for all $x \in \Omega$. The second assertion is obvious.

Q.E.D.

We may introduce a perturbation if necessary, $[\epsilon_1, \epsilon_2, \cdots, \epsilon_n] > 0$, on the diagonal elements of Q_0, so that the matrix $Q_0 + \lambda Q_1$ will always have at most two identical eigenvalues for any $\lambda \geq 0$. In the rest of the paper, we assume such is the case.

Consider another relaxed problem

$$\begin{array}{lll} (F_\lambda) & \text{minimize} & q_0(x) + \lambda q_1(x) \\ & \text{subject to} & \|x\|^2 \leq 1 \end{array}$$

with $\lambda \geq 0$. Let the optimal value of (F_λ) be $f(\lambda)$.

Using a similar argument as in Lemma 2, the following relation is readily seen,

Lemma 3 *The function $f(\lambda)$ is concave, and furthermore, it holds that*

$$f(\lambda) \leq h(\lambda) \leq v^*$$

for all $\lambda \geq 0$.

For any fixed λ, (F_λ) can be easily solved, e.g., by solving its SDP relaxation followed by a matrix decomposition procedure; see [11]. Among others, this implies that $f(\lambda)$ can be evaluated in polynomial time. In particular, for fixed λ, the optimality condition for (F_λ) is

$$\begin{cases} (Q_0 + \lambda Q_1 + \mu I)x = b_0 + \lambda b_1, \\ \mu(\|x\|^2 - 1) = 0,\ \mu \geq 0,\ \|x\|^2 - 1 \leq 0, \\ Q_0 + \lambda Q_1 + \mu I \succeq 0, \end{cases}$$

where the first two conditions are simply KKT, and the last one follows from the SDP duality.

Let X_λ be the set of optimal solutions for (F_λ). In our case, $|X_\lambda| \leq 2$. Then,

$$\partial f(\lambda) = \text{conv } \{q_1(x) \mid x \in X_\lambda\}.$$

Let

$$\hat{\lambda} = \text{argmax } \{f(\lambda) \mid \lambda \geq 0\}.$$

We remark here that due to the Slater condition, we have $f(\lambda) \to -\infty$ as $\lambda \to \infty$. Hence $\hat{\lambda}$ exists and is finite. Moreover, since $f(\lambda)$ is concave, using bisection one can find $\hat{\lambda}$ in polynomial time with any given precision.

By concavity of f, we conclude that X_λ only contains infeasible solutions ($q_1(x) > 0$) of (F_λ) for $\lambda < \hat{\lambda}$, and only contains feasible solutions of (F_λ) for $\lambda > \hat{\lambda}$.

Now that $\hat{\lambda}$ is a maximum point for $f(\lambda)$ we have

$$0 \in \partial f(\hat{\lambda}).$$

If $X_{\hat{\lambda}}$ is a singleton, then its element is also optimal for (P), and we are done. If $X_{\hat{\lambda}}$ contains two elements, say $\{x^+, x^-\}$, then we must have

$$q_1(x^-) \leq 0 \leq q_1(x^+).$$

If any of the above two inequalities is actually an equality, then, again, the corresponding solution is optimal to (P), and we are done. Next we are concerned with the remaining case, i.e.,

$$q_1(x^-) < 0 < q_1(x^+).$$

In that case, $x^- \in \text{int } \Omega$ and $x^+ \notin \Omega$.

According to Lemma 3 we know that $h(\hat{\lambda}) \geq f(\hat{\lambda})$. Due to the fact that $(F_{\hat{\lambda}})$ is a relaxation of $(H_{\hat{\lambda}})$ we conclude that x^- is optimal to $(H_{\hat{\lambda}})$, and consequently $0 > q_1(x^-) \in \partial h(\hat{\lambda})$. It can be shown that $\|x^-\| = 1$.

Now consider a set of solutions x, denoted by S_λ, such that the following conditions are satisfied

$$\begin{cases} (Q_0 + \lambda Q_1 + \mu I)x = b_0 + \lambda b_1, \\ \|x\|^2 - 1 = 0,\ \mu \geq 0, \\ Q_0 + \lambda Q_1 + \mu I \text{ has at most one negative eigenvalue.} \end{cases}$$

One can easily verify that, if $Q_0 + \lambda Q_1$ has distinct eigenvalues then $|S_\lambda| \leq 5$. Furthermore, let

$$x_\lambda = \arg \min_{\substack{x \in S_\lambda \\ q_1(x) \leq 0}} q_0(x) + \lambda q_1(x).$$

Our first result is the following.

Theorem 3 *For any λ such that $h(\lambda) < h(0) = v^*$, x_λ is an optimal solution of (H_λ).*

Proof. Let y_λ be an optimal solution of (H_λ). Since $h(\lambda) < h(0)$, i.e., λ is not a maximum point for h, it follows that $q_1(y_\lambda) < 0$. By the local optimality of y_λ we have

$$q_0(y_\lambda + d) + \lambda q_1(y_\lambda + d) \geq q_0(y_\lambda) + \lambda q_1(y_\lambda) \tag{9}$$

for all $\|y_\lambda + d\| \leq 1$, and $\|d\|$ sufficiently small. Moreover, y_λ must be a KKT point, i.e.,

$$(Q_0 + \lambda Q_1 + \mu I)y_\lambda = b_0 + \lambda b_1 \tag{10}$$

for some $\mu \geq 0$. Equations (9) and (10) imply that

$$\frac{1}{2} d^T (Q_0 + \lambda Q_1) d \geq \mu d^T y_\lambda$$

for all $\|y_\lambda + d\| \leq 1$, and $\|d\|$ sufficiently small. We may rewrite this relation as

$$\begin{aligned} \frac{1}{2} d^T (Q_0 + \lambda Q_1 + \mu I) d &\geq \mu d^T y_\lambda + \frac{\mu}{2} d^T d \\ &= \frac{\mu}{2} (y_\lambda + d)^T (y_\lambda + d) - \frac{\mu}{2} y_\lambda^T y_\lambda \end{aligned}$$

for all $\|y_\lambda + d\| \leq 1$, and $\|d\|$ sufficiently small. In particular,

$$\frac{1}{2} d^T (Q_0 + \lambda Q_1 + \mu I) d \geq 0 \tag{11}$$

for $\|y_\lambda + d\| = 1$, and $\|d\|$ sufficiently small.

Consider a fixed $\bar{d}$ satisfying $\bar{d}^T y_\lambda = 0$. Let $\epsilon > 0$ be a small number. Let $y_\lambda + d_\epsilon$ be the projection of $y_\lambda + \epsilon \bar{d}$ onto the unit sphere $\{y \mid \|y\| \leq 1\}$. Since $\|y_\lambda + \epsilon \bar{d}\| > 1$, we conclude that $\|y_\lambda + d_\epsilon\| = 1$. Let $\Delta d_\epsilon = \epsilon \bar{d} - d_\epsilon$. It follows

$$\|\Delta d_\epsilon\| = o(\epsilon). \tag{12}$$

Using (11) we have

$$\begin{aligned} 0 &\le d_\epsilon^T (Q_0 + \lambda Q_1 + \mu I) d_\epsilon \\ &= (\epsilon \bar{d})^T (Q_0 + \lambda Q_1 + \mu I)(\epsilon \bar{d}) - 2(\Delta d_\epsilon)^T (Q_0 + \lambda Q_1 + \mu I)(\epsilon \bar{d}) \\ &\quad + (\Delta d_\epsilon)^T (Q_0 + \lambda Q_1 + \mu I)(\Delta d_\epsilon). \end{aligned}$$

Dividing ϵ^2 on the both sides of the above inequality, and let $\epsilon \to 0$ we get

$$\bar{d}^T (Q_0 + \lambda Q_1 + \mu I) \bar{d} \ge 0. \tag{13}$$

Note that the inequality (13) holds for any $\bar{d}$ with $\bar{d}^T y_\lambda = 0$. Hence

$$Q_0 + \lambda Q_1 + \mu I$$

can have at most one negative eigenvalue. Together with (10), we conclude $y_\lambda = x_\lambda$.

Q.E.D.

Theorem 3 suggests a scheme to solve (P) by means of following x_λ while reducing λ until $q_1(x_\lambda) = 0$, or until $\lambda = 0$. More precisely, the procedure works as follows. We follow x_λ starting from $\lambda = \hat{\lambda}$, and stop either $q_1(x_\lambda) = 0$, or $\lambda = 0$. This can be accomplished by using Newton's method to solve the parameterized equation

$$(E_\lambda) \begin{cases} (Q_0 + \lambda Q_1 + \mu I)x = b_0 + \lambda b_1 \\ \|x\|^2 = 1. \end{cases}$$

The equation (E_λ) may have multiple solutions and we need to follow the path(s) generated by x_λ.

Theorem 4 *Let $\bar{\lambda}$ be either $q_1(x_{\bar{\lambda}}) = 0$, or $\bar{\lambda} = 0$ and $q_1(x_{\bar{\lambda}}) < 0$. Then, $x_{\bar{\lambda}}$ is an optimal solution for (P).*

Proof. For any $\lambda > \bar{\lambda}$, we have $h(\lambda) < h(0)$. Hence, applying Theorem 3 and taking limit yield that $x_{\bar{\lambda}}$ is an optimal solution for $(H_{\bar{\lambda}})$.

In the case that $q_1(x_{\bar{\lambda}}) = 0$, we have $h(\bar{\lambda}) = v^*$, and so $x_{\bar{\lambda}}$ is optimal for (P). Otherwise, if $\bar{\lambda} = 0$, then (H_0) is simply (P), and the assertion holds as well.

Q.E.D.

As a consequence of Theorem 4, we arrive at a simpler proof for the following result, which was established in Yuan [13].

Corollary 2 *The Hessian matrix of the Lagrangian function for (P), when evaluated with optimal Lagrangian multipliers at the optimal solution, can have at most one negative eigenvalue.*

To illustrate how these results can be understood, let us consider two examples, the first one being

$$
\begin{array}{lll}
(EX_1) & \text{minimize} & -x_1^2 + x_1 + 4x_2^2 \\
 & \text{subject to} & x_1^2 + x_2^2 \le 4 \\
 & & x_1^2 - 4x_1 + \frac{1}{4}x_2^2 \le 0.
\end{array}
$$

Its SDP relaxation is

$$
\text{minimize} \quad \begin{bmatrix} 0 & 0.5 & 0 \\ 0.5 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \bullet X
$$

$$
\text{subject to} \quad \begin{bmatrix} 0 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0.25 \end{bmatrix} \bullet X \le 0, \quad \begin{bmatrix} -4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \bullet X \le 0
$$

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \bullet X = 1, \; X \succeq 0.
$$

The optimal solution is $X^* = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ with the optimal value $v^* = -3$. The functions are

$$
f(\lambda) = \begin{cases} 12\lambda - 6, & \text{if } 0 \le \lambda \le 0.25 \\ -4\lambda - 2, & \text{if } \lambda \ge 0.25, \end{cases}
$$

and $h(\lambda) = -4\lambda - 2$ for all $\lambda \ge 0$. We see that $\hat{\lambda} = 0.25$, and $x^- = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. We then follow the trajectory while reducing λ. In this case, $x_\lambda \equiv x^-$ and this leads us to $x^* = x^-$ at $\lambda = 0$. The true optimal value of the original problem is -2. At the optimality, $\mu = 0.25$ and the Hessian matrix of the Lagrangian function is

$$
\begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix} + \mu \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
$$

which has a negative eigenvalue.

The second example we study is

$$
\begin{array}{lll}
(EX_2) & \text{minimize} & -x_1^2 + x_1 + x_2^2 \\
 & \text{subject to} & x_1^2 + x_2^2 \le 4 \\
 & & (x_1 + x_2)^2 + x_2^2 - 2x_1 \le 0.
\end{array}
$$

The corresponding value function $f(\lambda)$ attains its maximum at $\hat{\lambda} = 0.5$, and $f(0.5) = -2.3851$. Moreover, $x^- = \begin{bmatrix} 1.9639 \\ -0.3782 \end{bmatrix}$ and $x^+ = \begin{bmatrix} -1.9639 \\ 0.3782 \end{bmatrix}$. At x^-, $\lambda = \hat{\lambda} = 0.5$ and $\mu = 1.1926$. In our case, Newton's equation amounts to

$$\begin{bmatrix} Q_0 + \lambda Q_1 + \mu I, & x \\ x^T, & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \mu \end{bmatrix} = \begin{bmatrix} (Q_0 + \lambda Q_1 + \mu I)x - b_0 - \lambda b_1 \\ \frac{1}{2}\|x\|^2 - 2 \end{bmatrix},$$

with $Q_0 = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$ and $Q_1 = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$. Applying Newton's method, the trajectory can be computed as follows

$$x_{0.5} = \begin{bmatrix} 1.9639 \\ -0.3782 \end{bmatrix}, x_{0.4} = \begin{bmatrix} 1.9731 \\ -0.3267 \end{bmatrix}, x_{0.3} = \begin{bmatrix} 1.9823 \\ -0.2656 \end{bmatrix},$$

$$x_{0.2} = \begin{bmatrix} 1.9907 \\ -0.1925 \end{bmatrix}, x_{0.1} = \begin{bmatrix} 1.9973 \\ -0.1048 \end{bmatrix}, x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

The optimal solution is found by tracing this trajectory until $\lambda = 0$, i.e., $x^* = x_0$ with $v^* = -2$. The q_1 values at these points are:

$$q_1(x_{0.5}) = -1.2703,\ q_1(x_{0.4}) = -1.1289,\ q_1(x_{0.3}) = -0.9469,$$

$$q_1(x_{0.2}) = -0.7107,\ q_1(x_{0.1}) = -0.4023,\ q_1(x_0) = 0.$$

4 An extended trust region subproblem and conclusions

In [2], a trust region method was proposed for solving the nonlinear program

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) = 0 \end{array}$$

where $c(x) : \Re^n \mapsto \Re^m$, i.e., there are m equality constraints.

The subproblem to be solved at each iterative point x^k amounts to

$$\begin{array}{ll} \text{minimize} & d^T \nabla f(x^k) + \frac{1}{2} d^T B_k d \\ \text{subject to} & \|c(x^k) + \nabla c(x^k)^T d\| \le \xi_k \\ & \|d\| \le \Delta_k, \end{array}$$

where $\nabla c(x^k)$ stands for the Jacobian matrix of c evaluated at x^k.

In case $m = 1$, the above problem can be formally written as

$$\begin{array}{lll} (TR) & \text{minimize} & q_0(x) \\ & \text{subject to} & \|x\|^2 \leq 1 \\ & & -1 \leq \bar{a}^T x - a_0 \leq 1. \end{array}$$

The last constraint is equivalent to $q_1(x) = (\bar{a}^T x - a_0)^2 - 1 \leq 0$.

The optimal solution for (TR) may be binding at the constraint $q_1(x) \leq 0$, or it may not be. However, these two possibilities can be separately treated using the techniques developed in Section 3 and Subsection 2.3.

The binding case can be immediately dealt with by solving the following quadratic optimization with complementary linear constraints:

$$\begin{array}{ll} \text{minimize} & q_0(x) \\ \text{subject to} & \|x\|^2 \leq 1 \\ & -1 \leq \bar{a}^T x - a_0 \leq 1 \\ & (\bar{a}^T x - a_0 - 1)(\bar{a}^T x - a_0 + 1) = 0. \end{array}$$

As we discussed in Subsection 2.3, this can be solved by an SOC-based SDP relaxation in polynomial time.

The remaining task now is to consider the possibility that the constraint $q_1(x) \leq 0$ may not be binding at optimality. If that happens then the value function $h(\lambda)$ as defined in Section 3 attains its maximum value only at $\lambda = 0$. Using Theorem 4, we need only to consider solutions generated by the following equation

$$(E_0) \begin{cases} (Q_0 + \mu I)x = b_0 \\ \|x\|^2 = 1 \end{cases}$$

for given μ so that $Q_0 + \mu I$ has at most one negative eigenvalue, where we assume $q_0(x) = \frac{1}{2} x^T Q_0 x - b_0^T x$.

The key to note here is that any solution of (E_0) yields the same objective value under q_0, as shown below.

Lemma 4 *Suppose that x and x' both satisfy (E_0). Then, $q_0(x) = q_0(x')$.*

Proof. Multiplying x^T on both sides of the first equation and re-arrange yield

$$q_0(x) = -\frac{\mu}{2} - \frac{1}{2} b_0^T x.$$

On the other hand, we have

$$b_0^T x = b_0^T x' = x^T (Q_0 + \mu I) x'.$$

Hence $q_0(x) = q_0(x')$ as desired.

Q.E.D.

The procedure for finding the solution works as follows. We first compute the μ values such that (E_0) has a solution and $Q_0 + \mu I$ has at most one negative eigenvalue. This will result in at most 3 different μ values. Then, for each of these μ's, solve the following quadratic optimization problem

$$\begin{array}{ll} \text{minimize} & q_1(x) \\ \text{subject to} & (Q_0 + \mu I)x = b_0 \\ & \|x\|^2 = 1. \end{array}$$

This problem, after variable reduction if necessary, can be solved easily using the SDP relaxation plus decomposition approach; see [11]. If the optimal value of q_1 is positive for every computed μ, then we simply take the solution generated under the binding assumption. Otherwise, we take the solution with the lowest q_0 value among the selected μ's. Summarizing, we have shown the following result:

Theorem 5 *The trust region subproblem arising from a single equality constraint nonlinear programming can be solved in polynomial time.*

We remark here that a single equality constraint is not restrictive, as multiple equality constraints $c_i(x) = 0$, $i = 1, ..., m$, can be equivalently posed as a single constraint: $\sum_{i=1}^m c_i^2(x) = 0$.

The computational complexity for minimizing an indefinite quadratic function subject to two convex quadratic constraints remains unsettled. However, as shown in Section 3, there exist effective solution procedures to solve the problem. As we see in Section 2.1, there are interesting cases of quadratic optimization with indefinite objective function that can be solved in polynomial time using the SDP relaxation approach. There are several other related unsolved problems. For instance, how to minimize a non-homogeneous indefinite quadratic function with two homogeneous quadratic constraints? Is there an exact SDP relaxation? (cf. Section 2.2). Another problem one may attempt to solve is: Can one formulate an exact SDP relaxation for (TR)? (cf. Section 4).

References

[1] M. Bellare and P. Rogaway, The complexity of approximating a nonlinear program, *Mathematical Programming* 69 (1995) 429–442.

[2] M.R. Celis, J.E. Dennis, and R.A. Tapia, A trust region algorithm for nonlinear equality constrained optimization, in R.T. Boggs, R.H. Byrd, and R.B. Schnabel eds., *Numerical Optimization*, SIAM, Philadelphia, pp. 71–82, 1985.

[3] M. Fu, Z.-Q. Luo and Y. Ye, Approximation algorithms for quadratic programming, *Journal of Combinatorial Optimization* 2(1) (1998) 29–50.

[4] M.X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of ACM* 42 (1995) 1115–1145.

[5] A. Nemirovski, C. Roos, and T. Terlaky. *On maximization of quadratic form over intersection of ellipsoids with common center*, Technical report, Delft University of Technology, Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands, 1998.

[6] Yu.E. Nesterov, Semidefinite relaxation and nonconvex quadratic optimization, *Optimization Methods and Software* 9 (1998) 141–160. Special Issue Celebrating the 60th Birthday of Professor Naum Shor.

[7] Yu. E. Nesterov, *Global quadratic optimization via conic relaxation*, CORE Working Paper, Catholic University of Louvaine, Belgium, 1998.

[8] Yu. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Philadelphia, 1994.

[9] R. Stern and H. Wolkowicz, Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations, *SIAM Journal on Optimization* 5 (1995) 286–313.

[10] J.F. Sturm, Theory and algorithms of semidefinite programming, in H. Frenk, C. Roos, T. Terlaky and S. Zhang eds., *High Performance Optimization*, Kluwer Academic Publishers, Dordrecht, pp. 3–194, 2000.

[11] J.F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, SEEM2001-01, Chinese University of Hong Kong, March 2001.

[12] Y. Ye, Approximating quadratic programming with bound and quadratic constraints, *Mathematical Programming* 84 (1999) 219–226.

[13] Y. Yuan, On a subproblem of trust region algorithms for constrained optimization, *Mathematical Programming* 47 (1990) 53–63.