The Trust Region Subproblem with Non-Intersecting Linear Constraints

Samuel Burer* Boshi Yang†

February 21, 2013

Abstract

This paper studies an extended trust region subproblem (eTRS) in which the trust region intersects the unit ball with m linear inequality constraints. When $m = 0$, $m = 1$, or $m = 2$ and the linear constraints are parallel, it is known that the eTRS optimal value equals the optimal value of a particular convex relaxation, which is solvable in polynomial time. However, it is also known that, when $m \geq 2$ and at least two of the linear constraints intersect within the ball, i.e., some feasible point of the eTRS satisfies both linear constraints at equality, then the same convex relaxation may admit a gap with eTRS. This paper shows that the convex relaxation has no gap for arbitrary m as long as the linear constraints are non-intersecting.

Keywords: trust-region subproblem, second-order cone programming, semidefinite programming, nonconvex quadratic programming.

Mathematics Subject Classification: 90C20, 90C22, 90C25, 90C26, 90C30.

1 Introduction

The classical *trust region subproblem* minimizes a nonconvex quadratic objective over the unit ball:

$$v(T_0) := \min_x \{x^T Q x + c^T x : \|x\| \leq 1\}. \qquad (T_0)$$

(T_0) is an important subproblem in trust region methods for nonlinear optimization and has drawn intense research interest [3, 4, 5, 6, 8, 11]. In particular, even though (T_0) is nonconvex, it can be solved efficiently both in theory and practice. Several extensions of (T_0) have been proposed that enforce additional constraints on the trust region, e.g., parallel

*Department of Management Sciences, University of Iowa, Iowa City, IA, 52242-1994, USA. Email: `samuel-burer@uiowa.edu`.

†Department of Mathematics, University of Iowa, Iowa City, IA, 52242-1419, USA. Email: `boshi-yang@uiowa.edu`.

linear constraints [12] or a second full-dimensional ellipsoidal constraint [2]. An important theoretical and practical issue is whether such extensions can still be solved efficiently. In this paper, we investigate the theoretical tractability of the following extension of (T_0), which enforces m additional linear inequality constraints:

$$v(T_m) \ := \ \min_x \left\{ x^TQx + c^Tx \ : \ \begin{array}{c} \|x\| \leq 1 \\ a_i^Tx \leq b_i \ (i = 1, \ldots, m) \end{array} \right\}. \qquad (T_m)$$

A natural starting point is of course (T_0), which has the following polynomial-time solvable semidefinite programming (SDP) relaxation:

$$v(T_0) \geq v(R_0) := \min_{x,X} \left\{ Q \bullet X + c^Tx : \text{trace}(X) \leq 1, \ X \succeq xx^T \right\}, \qquad (R_0)$$

Here, X is a symmetric matrix, $Q \bullet X$ is the matrix inner product, and the constraint $\text{trace}(X) \leq 1$ makes the feasible region of (R_0) compact. One can show—by appealing to Pataki's theory concerning the rank of extreme points of semidefinite feasibility systems [7], for example—that every extreme point of (R_0) satisfies $X = xx^T$, which guarantees that $v(R_0)$ actually equals $v(T_0)$ since some optimal solution of (R_0) must occur at an extreme point.

Sturm and Zhang [10] and Burer and Anstreicher [1] study the following polynomial-time solvable relaxation of (T_1), in this case (R_0) with an added second-order cone (SOC) constraint:

$$v(T_1) \geq v(R_1) := \min_{x,X} \left\{ Q \bullet X + c^Tx : \begin{array}{c} \text{trace}(X) \leq 1, \ X \succeq xx^T \\ \|b_1x - Xa_1\| \leq b_1 - a_1^Tx \end{array} \right\}. \qquad (R_1)$$

The SOC constraint is constructed [10] by relaxing the valid quadratic SOC constraint $\|(b_1 - a_1^Tx)x\| = (b_1 - a_1^Tx)\|x\| \leq b_1 - a_1^Tx$ and is called an *SOC-RLT constraint* [1] since its construction is closely related to the reformulation-linearization technique of [9]. Sturm and Zhang prove $v(T_1) = v(R_1)$, extending the case for $m = 0$, while Burer and Anstreicher show further that every extreme point of (R_1) satisfies $X = xx^T$.

Ye and Zhang [12] and Burer and Anstreicher [1] also studied (T_2). Ye and Zhang provided a polynomial-time "trajectory following" procedure for solving (T_2), whereas Burer and Anstreicher studied the following polynomial-time solvable relaxation, in this case (R_1) with the SOC-RLT constraint for $a_2^Tx \leq b_2$ and a new linear *RLT* constraint reflecting the valid quadratic inequality $(b_1 - a_1^T x)(b_2 - a_2^T x) \geq 0$:

$$v(T_2) \geq v(R_2) := \min_{x,X} \left\{ Q \bullet X + c^T x : \begin{array}{c} \text{trace}(X) \leq 1,\ X \succeq xx^T \\ \|b_i x - X a_i\| \leq b_i - a_i^T x \ (i = 1, 2) \\ b_1 b_2 - b_2\, a_1^T x - b_1\, a_2^T x + a_1^T X a_2 \geq 0 \end{array} \right\}. \qquad (R_2)$$

The authors showed that, when $a_1 \parallel a_2$, the extreme points of (R_2) satisfy $X = xx^T$, and so $v(T_2) = v(R_2)$ in this case. On the other hand, they gave a counter-example for which $v(T_2) > v(R_2)$ when $a_1 \nparallel a_2$. This example had the property that $a_1^T x \leq b_1$ and $a_2^T x \leq b_2$ intersected inside the unit ball—more precisely, some x in the unit ball simultaneously made both linear constraints tight—leaving open the possibility that $v(T_2) = v(R_2)$ could still hold when the two inequalities are non-intersecting.

In this paper, we study the following polynomial-time solvable relaxation of (T_m), which includes all possible SOC-RLT and RLT constraints:

$$\begin{array}{rrll} v(T_m) \geq v(R_m) & := & \min\limits_{x,X} & Q \bullet X + c^T x \\ & & \text{s.t.} & \text{trace}(X) \leq 1,\ X \succeq xx^T \\ & & & \|b_i x - X a_i\| \leq b_i - a_i^T x \qquad\qquad i \leq m \\ & & & b_i b_j - b_j a_i^T x - b_i a_j^T x + a_i^T X a_j \geq 0 \qquad i < j \leq m \end{array} \qquad (R_m)$$

In light of the results in [1], we focus only on the non-intersecting case, that is, when there exists no x feasible for (T_m) satisfying $a_i^T x = b_i$ and $a_j^T x = b_j$ for some $i < j$. Note that the RLT constraints corresponding to $i = j$ are implied by $X \succeq xx^T$, and by the symmetry of X, the RLT constraints for $i > j$ are automatically satisfied.

For the non-intersecting case, we will prove that every extreme point of (R_m) satisfies $X = xx^T$, which immediately implies $v(T_m) = v(R_m)$. This is our main goal. Combined with the results of [1], we thus achieve a very clear demarcation of when the relaxation (R_m) achieves $v(T_m) = v(R_m)$ generally: precisely when the linear constraints $a_i^T x \leq b_i$ are non-intersecting in the unit ball. We also discuss a slight extension in the last section of the paper.

2 Preliminaries

We discuss some preliminary items in preparation for Section 3. In terms of notation, $\mathcal{F}(T_m)$ and $\mathcal{F}(R_m)$ denote the feasible sets of (T_m) and (R_m), respectively. In addition, given (x, X), we define

$$Y(x, X) := \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

and remark that $X = xx^T$ if and only if $Y(x, X)$ is rank-1. Accordingly, we also say that (x, X) is rank-1 if $X = xx^T$.

We assume throughout that (T_m) is feasible and formally state our assumption that no two hyperplanes defined by $a_i^T x = b_i$ and $a_j^T x = b_j$ intersect within the unit ball.

Assumption 1. *For all $i < j$, there exists no $x \in \mathcal{F}(T_m)$ such that $a_i^T x = b_i$ and $a_j^T x = b_j$.*

We will also need the following result, which has been discussed in the Introduction:

Lemma 1. *Every extreme $(x, X) \in \mathcal{F}(R_0)$ satisfies $X = xx^T$.*

Another important result for Section 3 demonstrates how any $(x, X) \in \mathcal{F}(R_m)$ with $a_i^T x < b_i$ gives rise to a special vector $z_i \in \mathcal{F}(T_m)$.

Lemma 2. *Suppose $(x, X) \in \mathcal{F}(R_m)$ with $a_i^T x < b_i$ for some i. Define*

$$z_i := (b_i - a_i^T x)^{-1}(b_i x - X a_i) \tag{1}$$

Then $z_i \in \mathcal{F}(T_m)$.

Proof. First, the i-th SOC-RLT constraint of (R_m) guarantees $\|z_i\| \leq 1$. Furthermore, $X \succeq xx^T$ implies

$$\begin{aligned}(b_i - a_i^T x)(b_i - a_i^T z_i) &= (b_i - a_i^T x)b_i - a_i^T(b_i x - X a_i) = b_i^2 - 2b_i a_i^T x + a_i^T X a_i \\ &\geq b_i^2 - 2b_i a_i^T x + a_i^T x x^T a_i = (b_i - a_i^T x)^2 > 0.\end{aligned}$$

Finally, for all $j \neq i$, the RLT constraints of (R_m) and symmetry f X imply

$$\begin{aligned}(b_i - a_i^T x)(b_j - a_j^T z_i) &= (b_i - a_i^T x)b_j - a_j^T(b_i x - X a_i) \\ &= b_i b_j - b_j a_i^T x - b_i a_j^T x + a_i^T X a_j \\ &\geq 0.\end{aligned}$$

This completes the proof. □

Finally, Section 3 requires a simple fact about the extreme points of the intersection of a compact convex set with a half-space.

Lemma 3. *Let C be a compact convex set, and let H be a half-space. Every extreme point of $C \cap H$ may be expressed as the convex combination of at most two extreme points in C.*

Proof. Suppose that H is defined by the linear inequality $\alpha^T x \leq \beta$, and let $\bar{x} \in C \cap H$ be extreme. Since $\bar{x} \in C$, we may write $\bar{x} = \sum_{k \in K} \bar{\lambda}_k \bar{x}_k$, where K is some index set, each $\bar{x}_k$ is extreme in C, each $\bar{\lambda}_k > 0$, and $\sum_{k \in K} \bar{\lambda}_k = 1$.

For a general vector λ with the same length as $\bar{\lambda}$, define the linear function $x(\lambda) := \sum_{k \in K} \lambda_k \bar{x}_k$ and the linear system $L := \{\lambda \geq 0 : \alpha^T x(\lambda) \leq \beta, \sum_{k \in K} \lambda_k = 1\}$. Clearly $\lambda \in L$ if and only if $x(\lambda) \in C \cap H$. In addition, by standard polyhedral theory, every extreme point $\lambda \in L$ has at most two positive entries.

So $\bar{\lambda} > 0$ is feasible for L since $\bar{x} = x(\bar{\lambda}) \in C \cap H$. Hence, we can write $\bar{\lambda} = \sum_j \rho_j \lambda^j$, where $\sum_j \rho_j = 1$, each $\rho_j > 0$, and each λ^j is extreme in L. By expanding, this means $\bar{x} = \sum_j \rho_j x(\lambda^j)$ with each $x(\lambda^j) \in C \cap H$. Since $\bar{x}$ is extreme in $C \cap H$, it holds that every $x(\lambda^j) = \bar{x}$. Since λ^j has at most two positive entries, this completes the proof. $\square$

3 The Result

We would like to prove $v(R_m) = v(T_m)$ under Assumption 1, and we will accomplish this by showing that every extreme point (x, X) of the compact $\mathcal{F}(R_m)$ satisfies $X = xx^T$. Our proof is by induction on m, where Lemma 1 with $m = 0$ serves as the base case, i.e., every extreme $(x, X) \in \mathcal{F}(R_0)$ satisfies $X = xx^T$. The induction hypothesis is thus as follows:

Assumption 2. *Given $1 \leq \ell \leq m$, every extreme $(x, X) \in (R_{\ell-1})$ satisfies $X = xx^T$.*

Furthermore, our proof is broken down into two cases:

Case 1. *Some constraint $a_i^T x \leq b_i$ is redundant for $\mathcal{F}(T_m)$.*

Case 2. *No constraint $a_i^T x \leq b_i$ is redundant for $\mathcal{F}(T_m)$.*

Case 1 can be handled immediately.

Theorem 1. *For Case 1, every extreme $(x, X) \in \mathcal{F}(R_m)$ satisfies $X = xx^T$.*

Proof. Without loss of generality, assume $a_m^T x \leq b_m$ is redundant. Because $\mathcal{F}(R_{m-1})$ is a relaxation of $\mathcal{F}(R_m)$, $(x, X) \in \mathcal{F}(R_{m-1})$. Hence, by Assumption 2, $Y(x, X)$ may be expressed as the convex combination of rank-1 matrices of the form $\binom{1}{\bar{x}}\binom{1}{\bar{x}}^T$, where each $\bar{x} \in \mathcal{F}(T_{m-1})$. Such $\bar{x}$ also satisfy $a_m^T \bar{x} \leq b_m$ by redundancy, and so (x, X) is the convex combination of points $(\bar{x}, \bar{x}\bar{x}^T) \in \mathcal{F}(R_m)$. Since (x, X) is extreme, this implies $X = xx^T$. $\square$

Now assume Case 2. Propositions 1 and 2 are the key results leading to Theorem 2 below, but first we state a critical lemma that applies in Case 2.

Lemma 4. *For Case 2, suppose x satisfies $\|x\| \le 1$ and $a_i^T x = b_i$ for some i. Then $x \in \mathcal{F}(T_m)$.*

Proof. Because $a_i^T x \le b_i$ is non-redundant and $\mathcal{F}(T_m)$ is convex, there exists $\bar{x} \in \mathcal{F}(T_m)$ such that $a_i^T \bar{x} = b_i$. Now suppose $x \notin \mathcal{F}(T_m)$, i.e., $a_j^T x > b_j$ for all $j \in \mathcal{J}$, where $\mathcal{J} \neq \emptyset$ is some collection of indices not including i. Then some convex combination of $\bar{x}$ and x, say $\hat{x}$, is feasible for (T_m) and satisfies $a_i^T \hat{x} = b_i$ and $a_j^T \hat{x} = b_j$ for some $j \in \mathcal{J}$. However, this contradicts Assumption 1, so x is in fact feasible for (T_m). $\square$

Proposition 1. *For Case 2, let $(x, X) \in \mathcal{F}(R_m)$ be extreme such that some SOC-RLT constraint is active. Then $X = xx^T$.*

Proof. Assume without loss of generality that $\|b_1 x - X a_1\| = b_1 - a_1^T x$, and consider (R_1) based on the single constraint $a_1^T x \le b_1$. Since (x, X) is also in $\mathcal{F}(R_1)$, Assumption 2 implies

$$Y := Y(x, X) = \sum_k \lambda_k \binom{1}{x_k} \binom{1}{x_k}^T,$$

where $\sum_k \lambda_k = 1$ and, for each k, $x_k \in \mathcal{F}(T_1)$ and $\lambda_k > 0$. Note that

$$\begin{pmatrix} b_1 - a_1^T x \\ b_1 x - X a_1 \end{pmatrix} = Y \begin{pmatrix} b_1 \\ -a_1 \end{pmatrix} = \sum_k \lambda_k \begin{pmatrix} 1 \\ x_k \end{pmatrix} \begin{pmatrix} 1 \\ x_k \end{pmatrix}^T \begin{pmatrix} b_1 \\ -a_1 \end{pmatrix} = \sum_k \lambda_k (b_1 - a_1^T x) \begin{pmatrix} 1 \\ x_k \end{pmatrix}. \tag{2}$$

Since the left-hand side of (2) is on the boundary of the SOC, each summand $\lambda_k (b_1 - a_1^T x_k) \binom{1}{x_k}$ on the right is either 0 or parallel to $\binom{b_1 - a_1^T x}{b_1 x - X a_1}$ (which itself could be 0). As $\lambda_k > 0$ and $\binom{1}{x_k} \neq 0$, we can thus separate the indices k into two groups (using new indices j and ℓ to distinguish the groups):

$$Y = \sum_{j \,:\, a_1^T x_j = b_1} \lambda_j \binom{1}{x_j} \binom{1}{x_j}^T + \sum_{\ell \,:\, a_1^T x_\ell < b_1} \lambda_\ell \binom{1}{x_\ell} \binom{1}{x_\ell}^T.$$

Each x_ℓ, if there exist any, must be equal to $z_1 := (b_1 - a_1^T x)^{-1} (b_1 x - X a_1)$ since $\binom{1}{x_\ell}$ is parallel to $\binom{b_1 - a_1^T x}{b_1 x - X a_1}$. Note that the existence of at least one ℓ implies $a_1^T x < b_1$ so that z_1 is well-defined. By (1) and Lemma 2, $z_1 \in \mathcal{F}(T_m)$. In addition, each $x_j \in \mathcal{F}(T_m)$ by Lemma 4. Overall, we see that Y is the convex combination of rank-1 solutions to (T_m). Therefore, Y is rank-1 because (x, X) is extreme. $\square$

Proposition 2. *For Case 2, let $(x, X) \in \mathcal{F}(R_m)$ be extreme such that no SOC-RLT constraint is active. Then $X = xx^T$.*

Proof. Suppose that all RLT constraints corresponding to the pairs $(1, m), \ldots, (m-1, m)$ are inactive at (x, X). Then (x, X) is also extreme for (R_{m-1}), and so $X = xx^T$ by Assumption 2. On the other hand, suppose the (i, m)-th RLT constraint is tight; say $i = 1$ without loss of generality. We will derive a contradiction that (x, X) is extreme to complete the proof.

We first claim that the other RLT constraints corresponding to $(2, m), \ldots, (m-1, m)$ are inactive. Let z_m be given by (1); it is feasible for (T_m) by Lemma 2. Moreover, the proof of Lemma 2 shows that $a_k^T z_m = b_k$ if and only if the (k, m)-th RLT constraint is tight. Since at most one $a_k^T z_m \leq b_k$ can be tight by Assumption 1, at most one of the (k, m)-th RLT constraints can be active, as claimed.

Let G denote the intersection of $\mathcal{F}(R_{m-1})$ with the single RLT constraint $b_1 b_m - b_m a_1^T x - b_1 a_m^T x + a_1^T X a_m \geq 0$. So (x, X) is extreme for G. Then Lemma 3 implies (x, X) can be expressed as a convex combination of at most two extreme points of (R_{m-1}). By Assumption 2, this means $\operatorname{rank}(Y) \leq 2$, where $Y := Y(x, X)$.

Defining $s := \binom{b_1}{-a_1}$ and $t := \binom{b_m}{-a_m}$, we see

$$\begin{aligned} s^T Y s &= \begin{pmatrix} b_1 \\ -a_1 \end{pmatrix}^T \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \begin{pmatrix} b_1 \\ -a_1 \end{pmatrix} = b_1^2 - 2b_1 a_1^T x + a_1^T X a_1 \\ &\geq b_1^2 - 2b_1 a_1^T x + a_1^T x x^T a_1 = (b_1 - a_1^T x)^2 \\ &> 0 \end{aligned}$$

and similarly $t^T Y t > 0$. Notice also that the tight RLT constraint can be expressed as $s^T Y t = 0$. Next consider the equation

$$W := \begin{pmatrix} s^T \\ t^T \\ I \end{pmatrix} Y \begin{pmatrix} s & t & I \end{pmatrix} \begin{pmatrix} s^T Y s & s^T Y t & s^T Y \\ t^T Y s & t^T Y t & t^T Y \\ Y s & Y t & Y \end{pmatrix} = \begin{pmatrix} s^T Y s & 0 & s^T Y \\ 0 & t^T Y t & t^T Y \\ Y s & Y t & Y \end{pmatrix}.$$

We have $W \succeq 0$ and $\operatorname{rank}(W) \leq \operatorname{rank}(Y) \leq 2$. Then the Schur complement theorem implies $M := Y - (s^T Y s)^{-1}(Y s s^T Y) - (t^T Y t)^{-1}(Y t t^T Y) \succeq 0$ and $\operatorname{rank}(M) = \operatorname{rank}(W) - 2 \leq 2 - 2 = 0$, i.e., $M = 0$ or

$$Y = (s^T Y s)^{-1}(Y s)(Y s)^T + (t^T Y t)^{-1}(Y t)(Y t)^T.$$

We now prove several properties of

$$Y s = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \begin{pmatrix} b_1 \\ -a_1 \end{pmatrix} = \begin{pmatrix} b_1 - a_1^T x \\ b_1 x - X a_1 \end{pmatrix}.$$

First, $Y s$ lies in the interior of the SOC because $\|b_1 x - X a_1\| < b_1 - a_1^T x$. In particular, z_1 in

(1) is well-defined with $\|z_1\| < 1$. Furthermore, $t^T Y s = 0$ implies $t^T \binom{1}{z_1} = 0$, or equivalently $a_m^T z_1 = b_m$. Hence, Lemma 4 implies $z_1 \in \mathcal{F}(T_m)$. In a similar manner, we can prove from Yt that z_m defined by (1) is feasible for (T_m) with $a_1^T z_m = b_1$. Also, we see $z_1 \neq z_m$ by Assumption 1.

Summarizing, $Y = \alpha \binom{1}{z_1} \binom{1}{z_1}^T + \beta \binom{1}{z_m} \binom{1}{z_m}^T$ for appropriate positive scalars α, β and $z_1, z_m \in \mathcal{F}(T_m)$ with $z_1 \neq z_m$. Since the top-left entry in Y equals 1, this is a proper convex combination of points in $\mathcal{F}(R_m)$. However, this contradicts that (x, X) is extreme. □

We are now ready to state the desired theorem regarding Case 2 and the main result of the paper (Corollary 1).

Theorem 2. *For Case 2, every extreme $(x, X) \in \mathcal{F}(R_m)$ satisfies $X = xx^T$.*

Proof. Proposition 1 covers the case when (x, X) has an active SOC-RLT constraint, while Proposition 2 handles when (x, X) has no such active constraint. □

Corollary 1. *Under Assumption 1, $v(R_m) = v(T_m)$.*

Proof. Theorem 1 covers Case 1, and Theorem 2 covers Case 2. □

In [1], Burer and Anstreicher gave a counter-example for which $v(R_2) < v(T_2)$ when Assumption 1 is violated:

$$Q = \begin{pmatrix} 2 & 3 & 12 \\ 3 & -19 & 6 \\ 12 & 6 & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 14 \\ 14 \\ 9 \end{pmatrix}, \quad -x_1 \leq \tfrac{1}{2}, \quad x_1 + \tfrac{6}{5} x_2 \leq 0.$$

In this instance, $v(T_2) \approx -12.9419$ with $x^* \approx (-0.8536, 0.2947, 0.4294)^T$, while $v(R_2) \approx -13.8410$ with optimal

$$\bar{x} \approx \begin{pmatrix} -0.3552 \\ 0.3881 \\ -0.2119 \end{pmatrix}, \ \bar{X} \approx \begin{pmatrix} 0.2595 & -0.2248 & -0.0913 \\ -0.2248 & 0.4495 & -0.0694 \\ -0.0913 & -0.0694 & 0.2911 \end{pmatrix},$$

and the numerical rank of $\bar{X}$ is 3. We wish to examine this counter-example from the viewpoint of our proof. One can verify that $Y(\bar{x}, \bar{X})$ makes both SOC-RLT constraints active in (R_2), and so considering that $(\bar{x}, \bar{X})$ is likely to be extreme in (R_m) from the numerical point of view, Proposition 1 is violated in this case. Of course, Proposition 1 is based on Lemma 4, which heavily uses Assumption 1.

4 An Extension

Consider the following assumption, which is slightly relaxed compared to Assumption 1.

Assumption 3. *For all $i < j$, there exists no $x \in \mathcal{F}(T_m)$ such that $\|x\| < 1$, $a_i^T x = b_i$, and $a_j^T x = b_j$.*

In comparison to Assumption 1, Assumption 3 allows the linear constraints to intersect on the boundary of the unit ball. Every such intersection point on the boundary must clearly be an extreme point of the polyhedron $P := \{x : a_i^T x \leq b_i \, (i = 1, \ldots, m)\}$. For example, when the dimension of x is 2, Assumption 3 allows (T_m) to model polytopes P inscribed in the unit disk.

We have the following extension of Corollary 1.

Proposition 3. *Under Assumption 3, $v(R_m) = v(T_m)$.*

Proof. For $\epsilon > 0$ small, tighten the constraint $\|x\| \leq 1$ of (T_m) to $\|x\| \leq 1 - \epsilon$ in order to form a new extended trust region subproblem (T_m^ϵ) satisfying Assumption 1. Relative to (R_m), a suitably modified convex relaxation (R_m^ϵ) can be derived such that $v(R_m^\epsilon) = v(T_m^\epsilon)$ by Corollary 1. The result follows because $v(R_m) = \lim_{\epsilon \to 0} v(R_m^\epsilon) = \lim_{\epsilon \to 0} v(T_m^\epsilon) = v(T_m)$ as all involved feasible sets are compact. $\square$

We end with an application of Proposition 3. Consider $\min\{x^T Q x + c^T x : x \in P\}$, where P is the regular pentagon inscribed in the unit disk with $(1, 0)$ as an extreme point and

$$Q = \begin{pmatrix} -8 & 8 \\ 8 & -14 \end{pmatrix}, \qquad c = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

For example, P could be represented by the system

$$\begin{pmatrix} 1.3764 & 1 \\ -0.3249 & 1 \\ -1.0000 & 0 \\ -0.3249 & -1 \\ 1.3764 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 1.3764 \\ 0.8507 \\ 0.8090 \\ 0.8507 \\ 1.3764 \end{pmatrix}.$$

Since the redundant constraint $\|x\| \leq 1$ is not given explicitly, a reasonable approach would be to solve the following relaxation that only contains the RLT constraints:

$$\begin{aligned}\min_{x,X}\quad & Q \bullet X + c^T x \\ \text{s.t.}\quad & X \succeq xx^T \\ & b_i b_j - b_j a_i^T x - b_i a_j^T x + a_i^T X a_j \ge 0 \qquad i < j.\end{aligned}$$

This yields optimal

$$\bar{x} \approx \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \bar{X} \approx \begin{pmatrix} 0.8090 & 0 \\ 0 & 0.8090 \end{pmatrix}, \qquad Q \bullet \bar{X} + c^T \bar{x} \approx -17.7984.$$

Note that the numerical rank of $\bar{X}$ is 2. According to Proposition 3, however, we can obtain the exact optimal value by solving (R_m). Doing so yields

$$x^* \approx \begin{pmatrix} 0.3090 \\ -0.9511 \end{pmatrix}, \qquad X^* \approx \begin{pmatrix} 0.0955 & -0.2939 \\ -0.2939 & 0.9045 \end{pmatrix}, \qquad Q \bullet X^* + c^T x^* \approx -17.4873.$$

Indeed, the numerical rank of X^* is 1, showing that x^* is a global minimizer of $x^T Q x + c^T x$ over P.

Acknowledgments

The authors would like to thank Kurt Anstreicher for numerous discussions and for suggesting first that the case $m = 2$ could be extended to arbitrary m.

References

[1] S. Burer and K. Anstreicher. Second-order cone constraints for extended trust-region subproblems. Manuscript, University of Iowa, March 2011. Revised July 2012 and November 2012. To appear in *SIAM Journal on Optimization*.

[2] M. R. Celis, J. E. Dennis, and R. A. Tapia. A trust region strategy for nonlinear equality constrained optimization. In *Numerical Optimization, 1984 (Boulder, Colo., 1984)*, pages 71–82. SIAM, Philadelphia, PA, 1985.

[3] A. R. Conn, N. I. M. Gould, and P. L. Toint. *Trust-Region Methods*. MPS/SIAM Series on Optimization. SIAM, Philadelphia, PA, 2000.

[4] M. Fu, Z.-Q. Luo, and Y. Ye. Approximation algorithms for quadratic programming. *J. Combinatorial Optimization*, 2:29–50, 1998.

[5] N. I. M. Gould, S. Lucidi, M. Roma, and P. L. Toint. Solving the trust-region subproblem using the Lanczos method. *SIAM J. Optim.*, 9(2):504–525 (electronic), 1999.

[6] J. J. Moré and D. C. Sorensen. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, 4(3):553–572, 1983.

[7] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23:339–358, 1998.

[8] F. Rendl and H. Wolkowicz. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Math. Programming*, 77(2, Ser. B):273–299, 1997.

[9] H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems.* Kluwer, 1997.

[10] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Math. Oper. Res.*, 28(2):246–267, 2003.

[11] Y. Ye. A new complexity result on minimization of a quadratic function with a sphere constraint. In C. Floudas and P. Pardalos, editors, *Recent Advances in Global Optimization.* Princeton University Press, Princeton, NJ, 1992.

[12] Y. Ye and S. Zhang. New results on quadratic minimization. *SIAM J. Optim.*, 14(1):245–267 (electronic), 2003.