Reformulations, Relaxations and Algorithms for Nonconvex Quadratically Constrained Quadratic Programming

LI, Duan

Department of Management Sciences
City University of Hong Kong

Joint Work with Rujun Jiang, Fudan University

May 24, 2017

This talk is based on our previous works:

- Jiang, R. and Li, D. Convex relaxations with second order cone constraints for nonconvex quadratically constrained quadratic programming. *arXiv preprint arXiv:1608.02096*.
- Jiang, R. and Li, D. Simultaneous diagonalization of matrices and its application in quadratic constrained quadratic programming. *SIAM J. Optim., 26.3 (2016): 1649-1668.*
- Jiang, R., Li, D., and Wu, B. SOCP reformulation for the generalized trust region subproblem via a canonical form of two symmetric matrices. *Mathematical Programming*, pages 1–33,2017.
- Jiang, R. and Li, D. Novel reformulations and efficient algorithm for the generalized trust region subproblem. *arXiv preprint arXiv:1707.08706*.

Problem formulation

Quadratically constrained quadratic programming (QCQP) problems arise in various areas, for example, combinatorial optimization, portfolio selection problems, economic equilibria, 0–1 integer programming and various applications in engineering.

Consider the following QCQP problem:

$$
\begin{array}{lll}
\text{(P)} \qquad & \min & x^T Q_0 x + c_0^T x \\
& \text{s.t.} & x^T Q_i x + c_i^T x + d_i \le 0,\ i = 1, \ldots, l, \\
& & a_j^T x \le b_j,\ j = 1, \ldots, m,
\end{array}
$$

where $Q_i \in \mathbb{S}^n$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, and $a_j \in \mathbb{R}^n,\ b_j \in \mathbb{R}$.

- When $Q_i \succeq 0$ for each i, problem (P) is convex.
  - semidefinite programming (SDP) relaxation is exact;
  - off-the-shelf software for second order cone programming (SOCP) is more efficient.
- In general, QCQP is NP–Hard.
- This talk mainly considers
  - a convex relaxation for general QCQP, and
  - reformulations and algorithms for singly constrained QCQP.

Literature review: Basic SDP relaxation

The basic SDP relaxation of problem (P) is

$$
\begin{aligned}
\text{(SDP)} \quad \min \quad & Q_0 \cdot X + c_0^T x \\
\text{s.t.} \quad & Q_i \cdot X + c_i^T x + d_i \le 0, i = 1, \dots, l, && (1) \\
& a_j^T x \le b_j, j = 1, \dots, m, && (2) \\
& \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0. && (3)
\end{aligned}
$$

- However, the SDP relaxation is too loose in many cases.
- Valid inequalities are considered in the literature to strengthen the basic SDP relaxation to make the bound tighter.

Literature review: RLT constraints

- [Sherali and Adams, 2013] developed the reformulation-linearization technique (RLT)

Linearizing the product of any two linear constraints, i.e.,

$$(a_i^T x - b_i)(a_j^T x - b_j) \geq 0$$

yields the following RLT constraints,

$$a_i a_j^T \cdot X + b_i b_j - b_j a_i^T x - b_i a_j^T x \geq 0. \quad (4)$$

Literature review: SOC-RLT constraints

[Sturm and Zhang, 2003, Burer and Saxena, 2012, Burer and Yang, 2015] Multiplying the linear term $b_j - a_j^T x$ to both sides of the second order cone representation of a convex quadratic constraint yields a valid inequality,

$$(b_j - a_j^T x)\left(\frac{1}{2}(1 - d_i - c_i^T x) - \left\| \begin{array}{c} B_i x \\ \frac{1}{2}(1 + d_i + c_i^T x) \end{array} \right\|\right) \geq 0.$$

Then, its linearization gives rise an SOC-RLT constraint:

$$\begin{aligned} & \left\| B_i(b_j x - X a_j), \frac{1}{2}(-c_i^T X a_j + (b_j c_i^T - d_i a_j^T - a_j^T)x + (1 + d_i)b_j) \right\| \\ \leq \quad & \frac{1}{2}(c_i^T X a_j + (d_i a_j^T - a_j^T - b_j c_i^T)x + (1 - d_i)b_j), \qquad (5) \\ & \qquad\qquad\qquad i = 1, \ldots, k, j = 1, \ldots, m. \end{aligned}$$

Convex relaxations

Adding RLT constraints to the basic SDP relaxation (SDP) yields the following strengthened reformulation:

$$
\begin{array}{lll}
(\mathrm{SDP}_{\mathrm{RLT}}) & \min & Q_0 \cdot X + c_0^T x \\
& \text{s.t.} & (1), (2), (3), (4).
\end{array}
$$

Similarly, by adding SOC-RLT constraints to $(\mathrm{SDP}_{\mathrm{RLT}})$ yields an even better strengthened reformulation:

$$
\begin{array}{lll}
(\mathrm{SDP}_{\mathrm{SOC\text{-}RLT}}) & \min & Q_0 \cdot X + c_0^T x \\
& \text{s.t.} & (1), (2), (3), (4), (5).
\end{array}
$$

Different RLT-like constraints

We investigate the following several extensions of the RLT-like techniques,

$$
\begin{aligned}
\mathrm{L} \times \mathrm{L} &\implies \mathrm{RLT} \text{ ([Sherali and Adams, 2013])} \\
\mathrm{SOC(convex)} \times \mathrm{L} &\implies \text{SOC-RLT ([Sturm and Zhang, 2003])} \\
\mathrm{SOC(nonconvex)} \times \mathrm{L} &\implies \mathrm{GSRT} \\
\mathrm{M}(\succeq 0) \circ \mathrm{M}(\succeq 0) &\implies \mathrm{HSOC} \text{ ([Zheng et al., 2011])} \\
\mathrm{SOC} \times \mathrm{SOC} &\implies \mathrm{SST} \\
\mathrm{M}(\succeq 0) \otimes \mathrm{M}(\succeq 0) &\implies \mathrm{KSOC}(\succeq 0) \text{ ([Anstreicher, 2016])}
\end{aligned}
$$

Motivation

W.l.o.g., we assume that Q_i is positive semidefinite for $i \in \mathcal{C} = \{1, \ldots, k\}$ and Q_i is not positive semidefinite for $i \in \mathcal{N} = \{k+1, \ldots, l\}$, where $0 \leq k < l$.

All the existing methods in the literature lose their effectiveness when dealing with the nonconvex quadratic constraints.

- The only existing method to deal with nonconvex QCQP is to lift the nonconvex quadratic constraints directly (as the basic SDP relaxation does)

$$Q_i \cdot X + c_i^T x + d_i \leq 0, i = k+1, \ldots, l.$$

This motivates our first kind of valid constraints, generalized SOC-RLT (GSRT) constraints.

GSRT-A constraints

- Decompose each non-positive-semidefinite matrix Q_i as $Q_i = L_i^T L_i - M_i^T M_i$, $i = k+1, \cdots, l$. Then $x^T Q_i x + c_i^T x + d_i \leq 0$ is equivalent to the following two constraints,

$$\left\| L_i x, \frac{1}{2}(c_i^T x + d_i + 1) \right\| \leq z_{i-k}, \tag{6}$$

$$\left\| M_i x, \frac{1}{2}(c_i^T x + d_i - 1) \right\| = z_{i-k}. \tag{7}$$

- Relaxing the intractable nonconvex equality in (7) to an inequality yields a second order cone constraint.
- Lift $\begin{pmatrix} x \\ z \end{pmatrix}$ to $\begin{pmatrix} X & S \\ S^T & Z \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} (x \ z)$ and relax it to $\begin{pmatrix} X & S \\ S^T & Z \end{pmatrix} \succeq \begin{pmatrix} x \\ z \end{pmatrix} (x \ z)$, i.e., $\begin{pmatrix} 1 & x^T & z^T \\ x & X & S \\ z & S^T & Z \end{pmatrix} \succeq 0.$

GSRT-A constraints

- Afterwards, multiplying the both sides by a linear constraint $a_j^T x - b_j$, and linearizing them give rise to

$$\left\| L_i b_j x - L_i X a_j, \frac{1}{2}(c_i^T(b_j x - X a_j) + (d_i + 1)(b_j - a_j^T x)) \right\| \le z_{i-k} b_j - S_{i-k} a_j,$$

$$\left\| M_i b_j x - M_i X a_j, \frac{1}{2}(c_i^T(b_j x - X a_j) + (d_i - 1)(b_j - a_j^T x)) \right\| \le z_{i-k} b_j - S_{i-k} a_j,$$

  where S_{i-k} is the $(i-k)$th column of the matrix S.

- To further strengthen these valid inequalities, construct the following constraint from linearization of the square of (7),

$$Z_{i-k,i-k} = X \cdot M_i^T M_i + \frac{1}{4}(c_i c_i^T X + (d_i - 1)^2 + 2c_i^T x(d_i - 1)).$$

These five types of the new constraints make up the new class of constraints, termed Type A generalized SOC-RLT (GSRT-A) constraints.

GSRT-B constraints

Alternative decomposition: Define $x_0 = \frac{1}{2} Q_i^{\dagger} c_i$.

The main difference of GSRT-B from GSRT-A is that we use an alternative expression of (6) and (7):

- i) If $\frac{1}{4}(c_i^T Q_i^{\dagger} c_i) - d_i > 0$, we set $\|L_i(x + x_0), \Delta\| \leq z_{i-k}$ and $\|M_i(x + x_0)\| = z_{i-k}$, where $\Delta = \sqrt{\frac{1}{4}(c_i^T Q_i^{\dagger} c_i) - d_i}$;
- ii) Otherwise, we set $\|L_i(x + x_0)\| \leq z_{i-k}$ and $\|M_i(x + x_0), \Delta\| = z_{i-k}$ with $\Delta = \sqrt{d_i - \frac{1}{4}(c_i^T Q_i^{\dagger} c_i)}$.

GSRT-B constraints

If $\frac{1}{4}(c_i^T Q_i^\dagger c_i) - d_i > 0$, $\Delta = \sqrt{\frac{1}{4}(c_i^T Q_i^\dagger c_i) - d_i}$,

$$\left\| L_i(x + \frac{1}{2} Q_i^\dagger c_i) \right\| \le z_i, i = k+1, \cdots, l, \tag{8}$$

$$\left\| M_i(x + \frac{1}{2} Q_i^\dagger c_i), \Delta \right\| \le z_i, i = k+1, \cdots, l, \tag{9}$$

$$Z_{i-k,i-k} = M_i^T M_i \cdot (X + \frac{1}{4} Q_i^\dagger c_i c_i^T Q_i^\dagger + Q_i^\dagger c_i x^T) + \Delta^2, i = k+1, \ldots, l,$$

$$\left\| L_i(b_j x - X a_j + \frac{1}{2} Q_i^\dagger c_i (b_j - a_j^T x)) \right\| \le z_i b_j - a_j^T S_{i-k},$$

$$\left\| M_i(b_j x - X a_j + \frac{1}{2} Q_i^\dagger c_i (b_j - a_j^T x)), \Delta(b_j - a_j^T x) \right\| \le z_i b_j - a_j^T S_{i-k},$$

$$i = k+1, \cdots, l, j = 1, \cdots, m;$$

GSRT-B constraints

Otherwise $\Delta = \sqrt{d_i - \frac{1}{4}(c_i^T Q_i^\dagger c_i)}$,

$$\left\| L_i(x + \frac{1}{2}Q_i^\dagger c_i), \Delta \right\| \leq z_i, i = k+1, \cdots, l, \tag{10}$$

$$\left\| M_i(x + \frac{1}{2}Q_i^\dagger c_i) \right\| \leq z_i, i = k+1, \cdots, l, \tag{11}$$

$$Z_{i-k,i-k} = M_i^T M_i \cdot (X + \frac{1}{4}Q_i^\dagger c_i c_i^T Q_i^\dagger + Q_i^\dagger c_i x^T), i = k+1, \ldots, l,$$

$$\left\| L_i(b_j x - X a_j + \frac{1}{2}Q_i^\dagger c_i(b_j - a_j^T x)), \Delta(b_j - a_j^T x) \right\| \leq z_i b_j - a_j^T S_{i-k},$$

$$\left\| M_i(b_j x - X a_j + \frac{1}{2}Q_i^\dagger c_i(b_j - a_j^T x)) \right\| \leq z_i b_j - a_j^T S_{i-k},$$

$$i = k+1, \cdots, l, j = 1, \cdots, m.$$

Convex relaxation with GSRT constraints

With GSRT constraints, we get the following convex relaxations,

$$
\begin{aligned}
(\mathrm{SDP}_{\mathrm{GSRT-A}}) \quad & \min \quad && Q_0 \cdot X + c_0^T x \\
& \text{s.t.} && (1), (2), (4), (5), \\
& && \textit{GSRT-A constraints}, \\
& && \begin{pmatrix} 1 & x^T & z^T \\ x & X & S \\ z & S^T & Z \end{pmatrix} \succeq 0.
\end{aligned}
$$

$$
\begin{aligned}
(\mathrm{SDP}_{\mathrm{GSRT-B}}) \quad & \min \quad && Q_0 \cdot X + c_0^T x \\
& \text{s.t.} && (1), (2), (4), (5) \\
& && \textit{GSRT-B constraints}, \\
& && \begin{pmatrix} 1 & x^T & z^T \\ x & X & S \\ z & S^T & Z \end{pmatrix} \succeq 0.
\end{aligned}
$$

Dominance relationship

Theorem 1

$v(\mathrm{P}) \geq v(\mathrm{SDP}_{\mathrm{GSRT}}) \geq v(\mathrm{SDP}_{\mathrm{SOC\text{-}RLT}}) \geq v(\mathrm{SDP}_{\mathrm{RLT}}) \geq v(\mathrm{SDP})$,

where GSRT represents either GSRT-A or GSRT-B.

Numerical experiments showed GSRT-B constraints are always tighter than GSRT-A constraints, i.e., $v(\mathrm{SDP}_{\mathrm{GSRT\text{-}B}}) \geq \mathrm{v}(\mathrm{SDP}_{\mathrm{GSRT\text{-}A}})$.

Example 1

$$
\begin{array}{llll}
(\mathrm{P}) & \min & x^T Q_0 x + c_0^T x & \\
& \text{s.t.} & x^T Q_1 x + c_1^T x + d_1 \leq 0, & \\
& & a_1^T x \leq b_1, &
\end{array}
$$

$$
Q_0 = \begin{pmatrix} 0.3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2.4 \end{pmatrix};\ Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix};\ a_1 = \begin{pmatrix} -0.6 \\ -2 \\ 0.8 \end{pmatrix};
$$

$$
b_1 = -0.5;\ c_0 = \begin{pmatrix} -0.2 \\ 0.8 \\ 0.2 \end{pmatrix};\ c_1 = 0;\ d_1 = -1.
$$

- $v(\mathrm{P}) = -1.21788$,
  $v(\mathrm{SDP}_{\text{GSRT-A}}) = -1.2249$,
  $v(\mathrm{SDP}_{\text{RLT}}) = -1.9900$.

Example 2

Q_0, Q_1, c_0, c_1, d_1, a_1, b_1 remain the same with Example 1 but there is another linear constraint with $a_2 = (0.3, 0.2, 0.6)^T$ and $b_2 = -0.3$.

- The optimal solution is $v(\mathrm{P}) = -0.7449$ with

  $x^* = (-0.1264, 1.3250, -0.8785)^T$.

- $v(\mathrm{SDP}_{\mathrm{GSRT\text{-}A}}) = -0.7449$,

  $v(\mathrm{SDP}_{\mathrm{RLT}}) = -1.9252$,

  $v(\mathrm{SDP}) = -1.9900$.

Note that the above two examples only have nonconvex quadratic constraints, so the SOC-RLT constraints don't exist in these two cases.

Some special cases

- A case with no need to introduce augmented variable z_i:
  When $\frac{1}{4}(c_i^T Q_i^\dagger c_i) - d_i < 0$, $M_i(x + \frac{1}{2}Q_i^\dagger c_i)$ is a scalar and $M_i(x + \frac{1}{2}Q_i^\dagger c_i) \geq 0$, the corresponding GSRT-B constraint reduces to

$$\left\| L_i(x + \frac{1}{2}Q_i^\dagger c_i),\ \Delta \right\| \leq M_i(x + \frac{1}{2}Q_i^\dagger c_i).$$

- [Jin et al., 2013] proved that the SDP relaxation with GSRT constraints admits no gap for the problem of minimizing a quadratic objective subject to
  $x^T x \leq (a_1 + a_2^T x + a_3^T x)^2,\ a_1 + a_2^T x + a_3^T x \geq a_4 \geq 0.$ or
  $x^T x \leq (a_1 + a_2^T x + a_3^T x)^2,\ a_5 \geq a_1 + a_2^T x + a_3^T x \geq a_4 \geq 0.$

Valid inequalities with extra redundant constraint $\alpha_u \geq u^T x$

[Zheng et al., 2011] introduced an artificial inequality, $\alpha_u = \max\{u^T x \mid x \in \Omega\} > 0$, with a chosen $u \in \Re^n_{++} = \{y \in \Re^n \mid y_i > 0,\ i = 1, \ldots, n\}$, where Ω is some suitable set that contains the feasible region.
Using the following fact,

$$\begin{pmatrix} \mathrm{diag}(u)\mathrm{diag}(x) & \mathrm{diag}(u)x \\ x^T \mathrm{diag}(u) & \alpha_u \end{pmatrix} \succeq 0 \Leftrightarrow \alpha_u \geq u^T x,$$

they derived the valid linear matrix inequality

$$X \preceq \alpha_u \mathrm{diag}(u)^{-1} \mathrm{diag}(x). \tag{12}$$

Improvement of the decomposition-approximation method

Each convex quadratic constraint, $x^T Q_i x + c_i^T x + d_i \le 0$, is equivalent to the following linear matrix inequality (by representing Q_i as $B_i B_i^T$),

$$0 \succeq \begin{pmatrix} -I_n & B_i x \\ x^T B_i^T & c_i^T x + d_i \end{pmatrix}.$$

Then applying Hadamard product yields [Zheng et al., 2011]

$$\begin{aligned} 0 \;&\succeq\; \begin{pmatrix} -I_n & B_i x \\ x^T B_i^T & c_i^T x + d_i \end{pmatrix} \circ \begin{pmatrix} \mathrm{diag}(u)\mathrm{diag}(x) & \mathrm{diag}(u)x \\ x^T \mathrm{diag}(u) & \alpha_u \end{pmatrix} && (13) \\ &= \begin{pmatrix} -\mathrm{diag}(u)\mathrm{diag}(x) & \mathrm{diag}(u)\mathrm{Diag}(B_i x x^T) \\ (\mathrm{Diag}(B_i x x^T))^T \mathrm{diag}(u) & \alpha_u (c_i^T x + d_i) \end{pmatrix}. && (14) \end{aligned}$$

Linearizing (14) yields another valid inequality, termed HSOC,

$$\begin{pmatrix} -\mathrm{diag}(u)\mathrm{diag}(x) & \mathrm{diag}(u)\mathrm{Diag}(B_i X) \\ (\mathrm{Diag}(B_i X))^T \mathrm{diag}(u) & \alpha_u (c_i^T x + d_i) \end{pmatrix} \preceq 0. \qquad (15)$$

Dominance of the valid inequalities

Theorem 2

The valid inequality (12) is dominated by the the RLT constraints generated by $x \geq 0$ and $\alpha_u \geq u^T x$, i.e, $\alpha_u x_i \geq u^T X_{.i}$, $i = 1, \ldots, n$.

Theorem 3

The HSOC valid inequality (15) is dominated by the SOC-RLT constraints generated by $x \geq 0$, $\alpha_u \geq u^T x$ and $\|B_i x\|^2 \leq -c_i^T x - d_i$, i.e.,

$$\left\| \begin{pmatrix} B_i X_{\cdot,j} \\ \frac{1}{2}(x_j + c_i^T X_{\cdot,j} + d_i x_j) \end{pmatrix} \right\| \leq \frac{1}{2}(x_j - c_i^T X_{\cdot,j} - d_i x_j) \qquad (16)$$

and

$$\begin{aligned} & \left\| \begin{pmatrix} \alpha_u B_i x - B_i X u \\ \frac{1}{2}(\alpha_u(1 + c_i^T x + d_i) - (1 + d_i) u^T x - u^T X c_i) \end{pmatrix} \right\| \\ \leq \quad & \frac{1}{2}(\alpha_u(1 - c_i^T x - d_i) - (1 - d_i) u^T x + u^T X c_i). \end{aligned} \qquad (17)$$

Valid inequalities with extra linear constraint

Theorem 4

Assume relaxation $(\mathrm{SDP}_{\alpha\mathrm{GSRT}})$ *is* $(\mathrm{SDP}_{\mathrm{GSRT}})$ *enhanced with RLT, SOC-RLT, and GSRT constraints corresponding to the extra linear constraint* $u^T x \leq \alpha_u$*, then we have* $v(\mathrm{SDP}_{\alpha\mathrm{GSRT}}) \geq v(\mathrm{SDP}_{\mathrm{GSRT}})$.

Similar technique can also be found in [Burer and Anstreicher, 2013].

- In their paper, they add extra redundant linear constraints and add extra SOC-RLT constraints of quadratic constraints and the extra redundant linear constraints. By adding these extra SOC-RLT constraints they successfully improve the relaxation gap.

Example 3

$$\begin{aligned} \min \quad & 21x_1^2 + 34x_1x_2 - 24x_2^2 + 2x_1 - 14x_2 \\ \text{s.t} \quad & x_1^2 + 4x_1x_2 + 2x_2^2 + 8x_1 + 6x_2 - 9 \le 0, \\ & -5x_1^2 - 8x_1x_2 - 5x_2^2 - 4x_1 + 4x_2 + 4 \le 0, \\ & x_1 + 2x_2 \le 2, \\ & x \in [0,1]^2. \end{aligned}$$

- The optimal solution is $v^* = -3.327$ with $x^* = (0.427, 0.588)^T$.
- In [Zheng et al., 2011], they set $u = (1,2)^T$ and $\alpha_u = 1.8029$, and got a tighter bound -10.86.
  Regarding $u^T x \le \alpha_u$ as an extra linear constraint in Example 3, we get tighter bounds in the following table.

Example 3

Table: SDP bounds for Example 3

SDP relaxation	Lower bound	Extra linear constraint	Lower bound
(SDP)	-20.28	—	—
$(\mathrm{SDP}_{\mathrm{RLT}})$	-16.23	$(\mathrm{SDP}_{\alpha\mathrm{RLT}})$	-11.66
$(\mathrm{SDP}_{\mathrm{SOC\text{-}RLT}})$	-13.99	$(\mathrm{SDP}_{\alpha\mathrm{SOC\text{-}RLT}})$	-8.445
$(\mathrm{SDP}_{\alpha_{\mathrm{u}}})$	-10.86	—	—
$(\mathrm{SDP}_{\mathrm{GSRT\text{-}A}})$	-6.011	$(\mathrm{SDP}_{\alpha\mathrm{GSRT\text{-}A}})$	-4.887
$(\mathrm{SDP}_{\mathrm{GSRT\text{-}B}})$	-3.331	$(\mathrm{SDP}_{\alpha\mathrm{GSRT\text{-}B}})$	-3.327

Singly constrained QCQP

Now let us consider a special QCQP problem with one quadratic constraint, also known as the generalized trust region subproblem (GTRS),

$$\begin{aligned} (\mathrm{G}) \quad \min \quad & f_1(x) := \frac{1}{2}x^T Q_1 x + b_1^T x \\ \text{s.t.} \quad & f_2(x) := \frac{1}{2}x^T Q_2 x + b_2^T x \le c, \end{aligned}$$

- The classical trust region subproblem (TRS) first arose as a subproblem in trust region methods for unconstrained nonlinear programming.
  - TRS: $Q_2 = I$, $b_2 = 0$, $c = 1$,
  - TRS is a quadratic approximation in the trust region around the current point.
- Other applications: least square, robust optimization.

GTRS in Consensus ADMM

- A subproblem in Consensus ADMM algorithms for nonconvex QCQP with numerous applications in signal processing, machine learning, and wireless communications [Huang and Sidiropoulos, 2016].
- Problem reformulation:

$$\begin{array}{ll} \min & x^H A_0 x - 2\Re\{b_0^H x\} \\ \text{s.t.} & x^H A_i x - 2\Re\{b_i^H x\} \le c_i, \forall i = 1, \ldots, m. \end{array}$$

- Consensus form, $z_i^H A_i z_i - 2\Re\{b_i^H z_i\} \le c_i,\ z_i = x,\ \forall i = 1, \ldots, m$
- Consensus ADMM updates:

$$\begin{aligned} x &\leftarrow (A_0 + m\rho I)^{-1} \left(b_0 + \rho \sum_{i=1}^{m} (z_i + u_i) \right), \\ z_i &\leftarrow \arg\min_{z_i} \|z_i - x + u_i\|^2, \\ &\qquad\qquad \text{s.t. } z_i^H A_i z_i - 2\Re\{b_i^H z_i\} \le c_i \\ u_i &\leftarrow u_i + z_i - x. \end{aligned}$$

Methodology

- Iterative algorithms for TRS
  [Moré and Sorensen, 1983, Martínez, 1994, Ye, 1992]
- SDP relaxations [Sturm and Zhang, 2003, Ye and Zhang, 2003, Burer and Anstreicher, 2013, Burer and Yang, 2015]
- A linear-time algorithm for the TRS with respect to the nonzero entries in the input [Hazan and Koren, 2015]
- Iterative algorithms for GTRS
  [Moré, 1993, Ben-Tal and Teboulle, 1996, Sturm and Zhang, 2003, Feng et al., 2012]
- GTRS with an interval constraint ($c_1 \leq \frac{1}{2}x^T Ax + b^T x \leq c_2$)
  [Stern and Wolkowicz, 1995, Pong and Wolkowicz, 2014]

Literature review: Simultaneous Diagonalization

[Ben-Tal and den Hertog, 2014] If the two matrices in the quadratic forms are *simultaneously diagonalizable* (SD), then problem (GTRS) can be transformed into a second order cone programming (SOCP) problem, which can be solved much faster than the SDP algorithm and thus applicable to large scale problems.

Then one may ask:

- When the two matrices in the objective and constraint functions are SD? (Answered in [Jiang and Li, 2016])
- Can this result be extended to a general pair of matrices that are not SD? (Answered in [Jiang et al., 2017])

Main contributions for solving GTRS

Our contributions are summarized as follows:

- Show that all GTRS with optimal value bounded from below are SOCP representable.
- Derive conditions for the attainableness of the optimal value.
- Derive a closed-form solution when $I_{PD} = \{\lambda : Q_1 + \lambda Q_2 \succeq 0\}$ is a singleton, which includes the case that the two matrices are not SD.
- Extend our method with slight modification to the equality and interval constraint variants of GTRS.

To obtain the SOCP representation, we invoke and extend the simultaneous block diagonalization canonical form in [Uhlig, 1976], the transformation methods in [Ben-Tal and den Hertog, 2014] and the S-lemma.

Simultaneous diagonalization

[Ben-Tal and den Hertog, 2014] Assume that A and D can be *simultaneously diagonalized* by a nonsingular S. Using a one to one change of variables $z = Sx$, problem (GTRS) can be transformed as:

$$
\begin{aligned}
(\mathrm{Q}_1) \quad \min \quad & \sum_i (\frac{1}{2}\delta_i x_i^2 + e_i x_i) \\
\text{s.t.} \quad & \sum_i (\frac{1}{2}\alpha_i x_i^2 + b_i x_i) + c \leq 0.
\end{aligned}
$$

By setting $y_i = \frac{1}{2}x_i^2$ and relaxing it to $\frac{1}{2}x_i^2 \leq y_i$, (Q_1) can be relaxed to

$$
\begin{aligned}
(\mathrm{Q}_2) \quad \min \quad & \delta^T y + e^T x \\
\text{s.t.} \quad & \alpha^T y + b^T x + c \leq 0, \\
& \frac{1}{2}x_i^2 \leq y_i, i = 1, \ldots, n.
\end{aligned}
$$

The above relaxation is exact when Slater condition holds.

Notations

Define $\mathrm{diag}(A_1, \ldots, A_k)$ as matrix

$$\begin{pmatrix} A_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & A_k \end{pmatrix}.$$

Define E as the the block diagonal anti-diagonal matrix

$$\begin{pmatrix} 0 & & & 1 \\ & & \cdot & \\ & \cdot & & \\ \cdot & & & \\ 1 & & & 0 \end{pmatrix}.$$

Define F as the lower striped matrix.

$$\begin{pmatrix} & & & & 0 \\ & & & 0 & 1 \\ & & \cdots & & \\ & 0 & 1 & & \\ 0 & 1 & & & \end{pmatrix}.$$

Define $J(\lambda, m)$ as an $m \times m$ Jordan block

$$\begin{pmatrix} \lambda & e & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \lambda & e \\ & & & & \lambda \end{pmatrix}.$$

Canonical form of two symmetric matrices

Theorem 5 (Canonical form)

For any two $n \times n$ real symmetric matrices A and D, there exists an $n \times n$ real invertible matrix S such that

$$\begin{aligned} S^T A S = \quad & \mathrm{diag}(\tau_1 E_1, \ldots, \tau_k E_k, \tau_{k+1} E_{k+1}, \ldots, \tau_p E_p, \\ & \tau_{p+1} F_{p+1}, \ldots, \tau_m F_m, 0, \ldots, 0) \end{aligned} \tag{18}$$

and

$$\begin{aligned} S^T D S = \quad & \mathrm{diag}(\tau_1 E_1 J(\kappa_1, n_1), \ldots, \tau_k E_k J(\kappa_k, n_k), \tau_{k+1} F_{k+1}, \\ & \ldots, \tau_p F_p, \tau_{p+1} E_{p+1}, \ldots, \tau_m E_m, 0 \ldots, 0) \end{aligned} \tag{19}$$

where $\dim E_i = \dim F_i = n_i,\ i = k+1, \ldots, m$, and $\tau_i = \pm 1,\ i = 1, \ldots, m$. The signs of τ_i are uniquely (up to permutations) determined by the associated Jordan blocks $J(\kappa_i, n_i),\ E_i$ or F_i. The values of κ_i are uniquely (up to permutations) determined by the associated Jordan blocks $J(\kappa_i, n_i)$.

Canonical form of two symmetric matrices

Three types of block pairs in (18) and (20):

- Type A block pairs: $(\tau_i E_i, \tau_i E_i J(\kappa_i, n_i))$ or $(\tau_i E_i, \tau_i F_i)$;
- Type B block pairs: $(\tau_i F_i, \tau_i E_i)$;
- Type C block pairs: $(0, 0)$.

The following assumption avoids naive cases.

Assumption 1

i) There is at least one feasible solution in problem (GTRS)*; ii) The following three conditions do not hold true at the same time:* $Q_2 \succeq 0$, $b_2 \in \mathrm{Range}(Q_2)$ *and* $c = \frac{1}{2} b_2^T Q_2^\dagger b_2$.

Assumption 1 is equivalent to the Slater condition, i.e., there exists an $\overline{x}$ such that $f_2(\overline{x}) < 0$.

Unbounded cases

Using the canonical form of two matrices, we have the following theorem.

Theorem 6 (Unbounded cases)

The objective value of (GTRS) *is unbounded from below, if any of the following conditions hold,*

1. *there is a type A block pair and the size of the associated Jordan block is greater than 2 and the associated eigenvalue is real;*
2. *there is a type A block pair and the eigenvalues of the associated Jordan block form a complex pair;*
3. *there is a type B block pair* $(\tau_i F_i, \tau_i E_i)$ *and* $\dim F_i \geq 2$.

Enlighten by Theorem 6, we can now focus on a subproblem associated a type A block with size 2 and a real eigenvalue, by fixing all other variables.

A subproblem with a single block

Theorem 7 (Subproblem)

Consider the case where there exists a type A block pair $(\tau_i E_i, \tau_i E_i J(\kappa_i, n_i))$ in problem (GTRS) and the eigenvalue of the associated Jordan block $J_i(\lambda, 2)$ is real. Assume there is a feasible solution $\bar{x} = (\bar{z}^T, \bar{y}^T)^T$ and let $\pi = \tau_1 \bar{z}_1 \bar{z}_2$. Let $\rho = \inf\{\ \tau_i \lambda z_1 z_2 + \frac{1}{2}\tau_i z_2^2 + e_1 z_1 + e_2 z_2 \mid \tau_i z_1 z_2 \leq \pi\}$. We have the following three cases:

1. *When $\tau_i = 1$. If $(\lambda \leq 0,\ e_1 = 0,\ e_2 \neq 0)$ or $(\lambda = 0,\ e_1 = 0,\ e_2 = 0,\ \pi \geq 0)$ or $(\lambda < 0,\ e_1 = 0,\ e_2 = 0,\ \pi = 0)$, then $\rho = \lambda\pi - \frac{1}{2}e_2^2$ and the infimum is attainable;*
2. *When $\tau_i = 1$. If $(\lambda = 0,\ e_1 = 0,\ e_2 = 0,\ \pi < 0)$ or $(\lambda < 0,\ e_1 = 0,\ e_2 = 0,\ \pi \neq 0)$, then $\rho = \lambda\pi - \frac{1}{2}e_2^2$ and the infimum is unattainable;*
3. *Otherwise, $\rho = -\infty$ and thus problem (GTRS) is unbounded from below.*

Necessary conditions

Now we are ready to provide necessary conditions for problem (GTRS) being bounded from below.

Theorem 8 (Necessary conditions)

If problem (GTRS) *has an optimal value bounded from below, then:*

1. $\dim E_i \leq 2,\ i = 1, \ldots, p$, $\dim E_i = 1,\ i = p+1, \ldots, m$, *and there is no complex eigenvalue pair in* $J(\kappa_i, n_i)$*;*
2. *If for some index* i, $\dim E_i = 2$, *then the* i*th block satisfies the conditions in either case 1 or case 2 in Theorem 7.*

We next provide an SOCP reformulation for problem (GTRS) under the two necessary conditions in Theorem 8.

Reformulation via canonical form

Now we can rearrange the block pairs and express A and D as

$$A = \mathrm{diag}(\alpha_1, \ldots, \alpha_l, E_1, \ldots, E_{\frac{n-l}{2}}), \tag{20}$$

$$D = \mathrm{diag}(\delta_1, \ldots, \delta_l, E_1 J(\zeta_1, 2), \ldots, E_{\frac{n-l}{2}} J(\zeta_{\frac{n-l}{2}}, 2)), \tag{21}$$

where l is the number of block pairs with size one.

W.o.l.g, assume $b_i = 0,\ i = l+1, \ldots, n$ and $e_{l+2j-1} = 0,\ j = 1, \ldots, \frac{n-l}{2}$ (case 1 and 2 in Theorem 7).

Then problem (GTRS) can be reduced to the following problem, termed (P_1):

$$\begin{aligned} \min \quad & \sum_{i=1}^{l} (\delta_i x_i^2 + e_i x_i) + \sum_{j=1,\ldots,\frac{n-l}{2}} (\zeta_j x_{l+2j-1} x_{l+2j} + \frac{1}{2} x_{l+2j}^2 + e_{l+2j} x_{l+2j}) \\ \mathrm{s.t.} \quad & \sum_{i=1}^{l} (\alpha_i x_i^2 + b_i x_i) + \sum_{j=1,\ldots,\frac{n-l}{2}} (x_{l+2j-1} x_{l+2j}) + c \le 0. \end{aligned}$$

SOCP reformulation

SOCP reformulation for problem (GTRS):

$$
\begin{aligned}
(\mathrm{P}_2) \qquad \min \quad & \sum_{i=1}^{l}(\delta_i y_i + e_i x_i) + \sum_{j=1}^{\frac{n-l}{2}} \zeta_j z_j + c_0 \\
\text{s.t.} \quad & \sum_{i=1}^{l}(\alpha_i y_i + b_i x_i) + \sum_{j=1}^{\frac{n-l}{2}} z_j + c \le 0, \\
& \frac{1}{2}x_i^2 - y_i \le 0, \ \forall i = 1, 2, \ldots, l, \\
& x, y \in \Re^l, \ z \in \Re^{\frac{n-l}{2}},
\end{aligned}
$$

where $c_0 = -\sum_{j=1,\ldots,\frac{n-l}{2}} \frac{1}{2} e_{l+2j}^2$.

We next fully characterize the equivalence of (P_2) and (GTRS) and the attainableness of problem (GTRS) in Theorem 10.

Theorem 9 (Reformulation and attainableness of GTRS)

Assume that items 1 and 2 in Theorem 8 are satisfied, then $v(\mathrm{GTRS}) = v(\mathrm{P}_1) = v(\mathrm{P}_2)$*, More specifically, if* (P_2) *admits an optimal solution, then there exists an optimal solution* $(\bar{x}, \bar{y}, \bar{z})$ *to* (P_2) *with* $\frac{1}{2}\bar{x}_i^2 = \bar{y}_i,\ i = 1, 2, \ldots, l$*. Moreover, we can find an optimal solution (or an* ϵ *optimal solution)* $\tilde{x}$ *to* (P_2) *with*

$$
\begin{aligned}
&\tilde{x}_i = \bar{x}_i, i = 1, \ldots, l,\\
&\tilde{x}_{l+2j} = \begin{cases} 1/M & \textit{if} \begin{cases} \zeta_j = 0, e_{l+2j} = 0, \bar{z}_j < 0, \\ \textit{or } \zeta_j < 0, e_{l+2j} = 0, \bar{z}_j \neq 0, \end{cases} \\ -e_{l+2j} & \textit{otherwise}, \end{cases}\\
&\qquad\qquad j = 1, \ldots, \tfrac{n-l}{2},\\
&\tilde{x}_{l+2j-1} = \tfrac{\bar{z}_j}{\tilde{x}_{l+2j}},\ j = 1, \ldots, \tfrac{n-l}{2}.
\end{aligned}
$$

Particularly, if (P_2) *is bounded from below, then the optimal value of* (P_2) *is unattainable if and only if* $\zeta_j = 0,\ e_{l+2j} = 0,\ \bar{z}_j < 0$ *or* $\zeta_j < 0,\ e_{l+2j} = 0,\ \bar{z}_j \neq 0$*. In this case, for any* $\epsilon > 0$*, there exists an* ϵ *optimal solution* $\tilde{x}$ *such that* $f(\tilde{x}) - v(\mathrm{P}_1) < \epsilon$ *with a sufficiently large* $M > 0$*.*

Remarks

- It seems that there is no stable algorithm to compute the canonical form. But the canonical form still provides rich information for solving the GTRS efficiently.
- When the set $\{\lambda : Q_1 + \lambda Q_2 \succeq 0\}$ is an interval, the two matrices are SD and the case reduces to [Ben-Tal and den Hertog, 2014].
- When the set $\{\lambda : Q_1 + \lambda Q_2 \succeq 0\}$ is a singleton, the SOCP reformulation implies a closed form solution. Further investigation tells us that an optimal solution can be computed without calculating the canonical form.

Convex quadratic reformulation for TRS

Next we will derive a new reformulation for the GTRS. Before that, let us review a convex reformulation for the TRS
[Wang and Xia, 2016, Ho-Nguyen and Kilinc-Karzan, 2016].

$$\begin{array}{ll} \min & f_1(x) := \frac{1}{2}x^T(Q_1 - \lambda_{min(Q_1)}I)x + b_1^T x + \frac{1}{2}\lambda_{min(Q_1)} \\ \text{s.t.} & x^T x \leq 1. \end{array}$$

The above problem can be solved efficiently by Nesterov's accelerated gradient projection method.

Remark: This reformulation can be easily extended to the GTRS with positive definite Q_2.

Quadratic convex reformulation

Define $I_{PD} = \{\lambda : Q_1 + \lambda Q_2 \succeq 0\} \cap \mathbb{R}_+$, which is an interval [Moré, 1993].

Consider the case that the set $I_{PD} = [\lambda_1, \lambda_2]$ with $\lambda_1 < \lambda_2$ is nonempty.

We have the epigraph Reformulation:

$$(\mathrm{G}_1)\ \ \min_{x,t}\{t : \ f_1(x) \le t,\ f_2(x) \le 0\}.$$

Theorem 10

Suppose the set $I_{PD} = [\lambda_1, \lambda_2]$ with $\lambda_1 < \lambda_2$ is nonempty. By defining $h_i(x) = f_1(x) + \lambda_i f_2(x), i = 1, 2$, problem (G) is equivalent to

$$(\mathrm{G}_2)\ \ \min_{x,t}\{t : \ h_1(x) \le t,\ h_2(x) \le t\}.$$

Minmax Formulation

(G_2) is further equivalent to the minmax problem

$$\min h(x) := \max\{h_1(x), h_2(x)\}.$$

We developed two steepest descent algorithms based on two line search rules for the step size β_k,

$$x_{k+1} = x_k + \beta_k d_k.$$

Kurdyka-Łojasiewicz (KL) property

Definition 1

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a proper lower semicontinuous function satisfying that the restriction of* f *to its domain is a continuous function. The function* f *is said to have the KL property if for any* $\forall x^* \in \{x : 0 \in \partial f(x)\}$, *there exist* $C, \epsilon > 0$ *and* $\theta \in [0, 1)$ *such that*

$$C \|y\| \geq |f(x) - f^*(x)|^\theta, \quad \forall x \in B(x^*, \epsilon),\ \forall y \in \partial f(x),$$

where θ *is known as the KL exponent.*

Theorem 11

Assume that $\min h_1(x) < \min H(x)$ *and* $\min h_2(x) < \min H(x)$. *Then* $h(x)$ *has the KL property with exponent* $\theta = 1/2$.

Algorithm 1

Assume that h_i is the active function and h_j is the other function. Let d_k be the steepest descent direction and the associated step size be chosen as follows.

- When $h_1(x_k) = h_2(x_k)$, $\beta_k = 1/L$.
- When $h_1(x_k) \neq h_2(x_k)$ and the following quadratic equation for γ,

$$\mathbf{a}\gamma^2 + \mathbf{b}\gamma + \mathbf{c} = 0, \tag{22}$$

  where $\mathbf{a} = \frac{1}{2}\nabla h_i(x_k)^\top (A_i - A_j)\nabla h_i(x_k)$, $\mathbf{b} = (\nabla h_i(x_k)^\top - \nabla h_j(x_k)^\top)\nabla h_i(x_k)$ and $\mathbf{c} = h_i(x_k) - h_j(x_k)$, has no positive solution or any positive solution $\gamma \geq 1/L$, set $d_k = -\nabla h_i(x_k)$ with and $\beta_k = 1/L$;
- When $h_1(x_k) \neq h_2(x_k)$ and the quadratic equation (22) has a positive solution $\gamma < 1/L$, set $\beta_k = \gamma$ and $d_k = -\nabla h_i(x_k)$.

Then the sequence $\{f(x_k)\}$ has a local convergence rate $O(\log(1/\epsilon))$ under assumptions in Theorem 11 and a global convergence rate $O(1/\epsilon)$

Algorithm 2

Modified Armijo rule

Choose the smallest nonnegative integer k such that the following inequality holds for the step size $\beta_k = \xi s^k$ with $0 < \xi \leq 1$ and $0 < s < 1$,

$$f(x_k + \beta_k p_k) \leq f(x_k) - \sigma \beta_k p_k^\top d,$$

where $0 \leq \sigma \leq 0.5$ and d is the steepest descent direction.

This algorithm converges to an approximate stationary point.

Preliminary numerical results

- Data setting: Q_1 is positive definite and Q_2 is indefinite
  - Use Matlab package eigifp to calculate the minimum generalized eigenvalue of the matrix pair (A, B) for positive definite B
- If the null space of the Hessian matrix, $Q_1 + \lambda^* Q_2$, with λ^* being the optimal Lagrangian multiplier of problem (P), is orthogonal to $b_1 + \lambda^* b_2$, we are in the hard case; otherwise we are in the easy case.
- Preliminary numerical results showed that our algorithms outperform the state-of-the-art methods in [Pong and Wolkowicz, 2014] (the ERW algorithm in the tables on the following pages)
- Algorithm 2 is more efficient than Algorithm 1 in most cases. This may be because of the aggressiveness of Armijo line search rule.

Table: Numerical results for positive definite A and indefinite B

cond	n	Easy Case							
		Alg1		Alg2		$\mathrm{time}_{\mathrm{eig}}$	ERW		
		iter	time	iter	time		iter	time	fail
10	10000	90	**1.03**	109.3	1.24	1.45	5.9	4.89	0
10	20000	52	**2.82**	72.2	3.91	9.20	6.8	25.1	0
10	30000	60.9	**9.81**	83.2	13.4	25.2	6.6	75.0	0
10	40000	58.3	**17.1**	95.2	27.8	49.7	6.8	153	0
100	10000	417.7	**4.26**	424.9	4.34	3.99	5.9	11.4	0
100	20000	474.3	24.6	342.4	**17.8**	18.4	6.1	69.4	0
100	30000	196.9	28.0	162.1	**23.1**	51.8	6.2	147	0
100	40000	135.8	40.1	114.7	**33.9**	153.6	6.3	309	0
1000	10000	4245	44.7	1706.7	**17.8**	14.2	5.3	56.7	0
1000	20000	4177.3	216	1182.7	**61.2**	70.8	6.10	368	0
1000	30000	2023.8	289	813.7	**116**	189	5.9	1220	0
1000	40000	2519.8	652	1003	**301**	640.9	6.8	2960	0

Table: Numerical results for positive definite A and indefinite B

cond	n	Hard Case 1							
		Alg1		Alg2		$\mathrm{time}_{\mathrm{eig}}$	ERW		
		iter	time	iter	time		iter	time	fail
10	10000	1490	16.7	609.6	**6.81**	1.19	6	11.1	1
10	20000	530.3	27.9	313.9	**16.7**	7.56	6.5	53.9	0
10	30000	1014.6	157	270.6	**41.0**	30.1	7.3	170	1
10	40000	1866.4	520	782.7	**219**	54.0	7.1	356	1
100	10000	3328.2	33.9	1131.6	**13.6**	3.63	5.7	24.6	3
100	20000	6494.9	350	1410	**76.8**	42.2	6.4	123	5
100	30000	2836.6	420	1197.9	**176**	44.2	5.2	388	0
100	40000	906.7	257	506.1	**143**	173.5	6.5	639	0
1000	10000	25982.6	261	5090.7	**51.3**	24.0	5.75	81.1	6
1000	20000	26214.8	1360	2726.8	**139**	98.1	5.8	346	5
1000	30000	15311.4	2190	2591.9	**385**	195	5.8	1530	3
1000	40000	8735.8	3060	1343	**1020**	853	6.25	3280	2

Conclusions

- Derive the GSRT valid inequalities for nonconvex quadratic constraints using RLT-like techniques.
- Deduce an SOCP reformulation for singly constrained QCQP without the simultaneous diagonalization condition.
- Propose a convex quadratic reformulation for singly constrained QCQP and further show its equivalence to a minimax problem. Also develop an efficient algorithms that outperform the state-of-the-art algorithms in the literature.

Anstreicher, K. M. (2016).
Kronecker product constraints for semidefinite optimization.
http://www.optimization-online.org/DB_HTML/2016/06/5477.html.

Ben-Tal, A. and den Hertog, D. (2014).
Hidden conic quadratic representation of some nonconvex quadratic optimization problems.
Math. Program., 143(1-2):1–29.

Ben-Tal, A. and Teboulle, M. (1996).
Hidden convexity in some nonconvex quadratically constrained quadratic programming.
Math. Program., 72(1):51–63.

Burer, S. and Anstreicher, K. M. (2013).
Second-order-cone constraints for extended trust-region subproblems.
SIAM J. Optim., 23(1):432–451.

Burer, S. and Saxena, A. (2012).
The MILP road to MIQCP.
In *Mixed Integer Nonlinear Programming*, pages 373–405. Springer.

Burer, S. and Yang, B. (2015).
The trust region subproblem with non-intersecting linear constraints.
Math. Program., 149(1-2):253–264.

Feng, J.-M., Lin, G.-X., Sheu, R.-L., and Xia, Y. (2012).
Duality and solutions for quadratic programming over single non-homogeneous quadratic constraint.
J. Global Optim., 54(2):275–293.

Hazan, E. and Koren, T. (2015).
A linear-time algorithm for trust region problems.
Math. Program., pages 1–19.

Ho-Nguyen, N. and Kilinc-Karzan, F. (2016).
A second-order cone based approach for solving the trust region subproblem and its variants.
arXiv preprint arXiv:1603.03366.

Huang, K. and Sidiropoulos, N. D. (2016).
Consensus-admm for general quadratically constrained quadratic programming.
IEEE Transactions on Signal Processing, 64(20):5297–5310.

Jiang, R. and Li, D. (2016).
Simultaneous diagonalization of matrices and its application in quadratic constrained quadratic programming.
SIAM J. Optim., page to appear.

Jiang, R., Li, D., and Wu, B. (2017).
Socp reformulation for the generalized trust region subproblem via a canonical form of two symmetric matrices.
Mathematical Programming, pages 1–33.

Jin, Q., Tian, Y., Deng, Z., Fang, S.-C., and Xing, W. (2013).
Exact computable representation of some second-order cone constrained quadratic programming problems.
J. Oper. Res. Soc. China, 1(1):107–134.

Martínez, J. M. (1994).
Local minimizers of quadratic functions on euclidean balls and spheres.
SIAM J. Optim., 4(1):159–176.

Moré, J. J. (1993).
Generalizations of the trust region problem.
Optimization Methods and Software, 2(3-4):189–209.

Moré, J. J. and Sorensen, D. C. (1983).

Computing a trust region step.
SIAM Journal on Scientific and Statistical Computing, 4(3):553–572.

Pong, T. K. and Wolkowicz, H. (2014).
The generalized trust region subproblem.
Computational Optimization and Applications, 58(2):273–322.

Sherali, H. D. and Adams, W. P. (2013).
A reformulation-linearization technique for solving discrete and continuous nonconvex problems, volume 31.
Springer Science & Business Media.

Stern, R. J. and Wolkowicz, H. (1995).
Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations.
SIAM J. Optim., 5(2):286–313.

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

Uhlig, F. (1976).
A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil.
Linear Algebra and its Applications, 14(3):189–209.

Wang, J. and Xia, Y. (2016).
A linear-time algorithm for the trust region subproblem based on hidden convexity.
Optimization Letters, pages 1–8.

Ye, Y. (1992).
A new complexity result on minimization of a quadratic function with a sphere constraint.
In *Recent Advances in Global Optimization*, pages 19–31. Princeton University Press.

Ye, Y. and Zhang, S. (2003).

New results on quadratic minimization.
SIAM J. Optim., 14(1):245–267.

Zheng, X. J., Sun, X. L., and Li, D. (2011).
Convex relaxations for nonconvex quadratically constrained quadratic programming: matrix cone decomposition and polyhedral approximation.
Math. Program., 129(2):301–329.