Nonconvex Quadratic Optimization over Simple Ground Sets

Kurt M. Anstreicher

Department of Management Sciences
University of Iowa

Lunteren Conference
January, 2012

Outline

Outline

Outline

1. Nonconvex Quadratic Optimization
2. Standard QP
3. Box QP

Outline

Outline

Outline

Outline

We consider a problem of the form

$$\text{(QP)} \qquad \min \quad x^T Q x + c^T x \qquad \text{s.t.} \quad x \in \mathcal{F}$$

where $\mathcal{F} \in \Re^n$ is a feasible region or ground set of a simple form. The symmetric matrix Q is *not* assumed to be positive semidefinite. The cases of greatest interest are:

We consider a problem of the form

$$\text{(QP)} \qquad \min \quad x^T Qx + c^T x \qquad \text{s.t.} \quad x \in \mathcal{F}$$

where $\mathcal{F} \in \Re^n$ is a feasible region or ground set of a simple form. The symmetric matrix Q is *not* assumed to be positive semidefinite. The cases of greatest interest are:

- The simplex, $\mathcal{F} = \{x \geq 0 \,|\, e^T x = 1\}$. In this case QP is often referred to as the “standard QP” problem (QPS).

We consider a problem of the form

$$\text{(QP)} \qquad \min \quad x^T Qx + c^T x$$
$$\text{s.t.} \quad x \in \mathcal{F}$$

where $\mathcal{F} \in \Re^n$ is a feasible region or ground set of a simple form. The symmetric matrix Q is *not* assumed to be positive semidefinite. The cases of greatest interest are:

- The simplex, $\mathcal{F} = \{x \geq 0 \,|\, e^T x = 1\}$. In this case QP is often referred to as the "standard QP" problem (QPS).
- The hypercube, $\mathcal{F} = \{x \,|\, 0 \leq x \leq e\}$. In this case QP is often referred to as the "box QP" problem (QPB).

We consider a problem of the form

$$\text{(QP)} \qquad \min \quad x^T Qx + c^T x \qquad \text{s.t.} \quad x \in \mathcal{F}$$

where $\mathcal{F} \in \Re^n$ is a feasible region or ground set of a simple form. The symmetric matrix Q is *not* assumed to be positive semidefinite. The cases of greatest interest are:

- The simplex, $\mathcal{F} = \{x \geq 0 \,|\, e^T x = 1\}$. In this case QP is often referred to as the "standard QP" problem (QPS).
- The hypercube, $\mathcal{F} = \{x \,|\, 0 \leq x \leq e\}$. In this case QP is often referred to as the "box QP" problem (QPB).
- The hypersphere, $\mathcal{F} = \{x \,|\, \|x\| \leq 1\}$. In this case QP is often referred to as the "trust-region subproblem" (TRS).

We assume throughout that $\mathcal{F}$ is a compact convex set. Our approach to QP is to consider the convex hull of quadratic forms on $\mathcal{F}$,

$$Q[\mathcal{F}] := \mathrm{Co}\left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T \mid x \in \mathcal{F} \right\}.$$

We assume throughout that $\mathcal{F}$ is a compact convex set. Our approach to QP is to consider the convex hull of quadratic forms on $\mathcal{F}$,

$$Q[\mathcal{F}] := \mathrm{Co}\left\{ \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T \mid x \in \mathcal{F} \right\}.$$

We often write an element of $Q[\mathcal{F}]$ as

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

Since the extreme points of $Q[\mathcal{F}]$ are points $Y(x, xx^T)$ where $x \in \mathcal{F}$, the problem QP can be written equivalently as

$$\text{(QP)} \qquad \min \quad Q \bullet X + c^T x$$
$$\text{s.t.} \quad Y(x, X) \in Q[\mathcal{F}],$$

which is a linear optimization problem over $Q[\mathcal{F}]$. Written in this form, tractability of QP depends on the ability to efficiently characterize $Q[\mathcal{F}]$.

Since the extreme points of $Q[\mathcal{F}]$ are points $Y(x, xx^T)$ where $x \in \mathcal{F}$, the problem QP can be written equivalently as

$$\text{(QP)} \qquad \min \quad Q \bullet X + c^T x \qquad \text{s.t.} \quad Y(x, X) \in Q[\mathcal{F}],$$

which is a linear optimization problem over $Q[\mathcal{F}]$. Written in this form, tractability of QP depends on the ability to efficiently characterize $Q[\mathcal{F}]$.

An interesting issue that we will *not* consider is the complexity of approximation results for continuous optimization problems on these ground sets: see de Klerk (2008).

For polyhedral $\mathcal{F}$, result of Burer (2009) provides a general mechanism for expressing $Q[\mathcal{F}]$ using completely positive (CP) matrices. Let $\mathcal{C}_n$ denote the cone of $n \times n$ CP matrices.

For polyhedral $\mathcal{F}$, result of Burer (2009) provides a general mechanism for expressing $Q[\mathcal{F}]$ using completely positive (CP) matrices. Let $\mathcal{C}_n$ denote the cone of $n \times n$ CP matrices.

Theorem (CP representation of $Q[\mathcal{F}]$)

Let $\mathcal{F} = \{x \geq 0 \,|\, Ax = b\}$, *where A is* $m \times n$ *and* $\mathcal{F}$ *is bounded. Then* $Q[\mathcal{F}] = \{Y(x, X) \in \mathcal{C}_{n+1} \,|\, a_i x = b_i,\ a_i^T X a_i = b_i^2,\ i = 1, \ldots, m\}$.

Outline

For the case of Standard QP have $\mathcal{F} = \{x \geq 0 \mid e^T x = 1\}$. Note that if $e^T x = 1$ and $X = xx^T$, then $Xe = x$. For X with $E \bullet X = 1$, consider

$$Y(Xe, X) = \begin{pmatrix} e^T \\ I \end{pmatrix} X \begin{pmatrix} e & I \end{pmatrix} = \begin{pmatrix} 1 & e^T X \\ Xe & X \end{pmatrix}.$$

For the case of Standard QP have $\mathcal{F} = \{x \geq 0 \,|\, e^T x = 1\}$. Note that if $e^T x = 1$ and $X = xx^T$, then $Xe = x$. For X with $E \bullet X = 1$, consider

$$Y(Xe, X) = \begin{pmatrix} e^T \\ I \end{pmatrix} X \begin{pmatrix} e & I \end{pmatrix} = \begin{pmatrix} 1 & e^T X \\ Xe & X \end{pmatrix}.$$

Then $e^T(Xe) = 1$, and $X \in \mathcal{C}_n \implies Y(Xe, X) \in \mathcal{C}_{n+1}$. Burer's CP representation then immediately implies that

$$Q[\mathcal{F}] = \{Y(Xe, X) \,|\, X \in \mathcal{C}_n,\ E \bullet X = 1\}.$$

For the case of Standard QP have $\mathcal{F} = \{x \geq 0 \,|\, e^T x = 1\}$. Note that if $e^T x = 1$ and $X = xx^T$, then $Xe = x$. For X with $E \bullet X = 1$, consider

$$Y(Xe, X) = \begin{pmatrix} e^T \\ I \end{pmatrix} X \begin{pmatrix} e & I \end{pmatrix} = \begin{pmatrix} 1 & e^T X \\ Xe & X \end{pmatrix}.$$

Then $e^T(Xe) = 1$, and $X \in \mathcal{C}_n \implies Y(Xe, X) \in \mathcal{C}_{n+1}$. Burer's CP representation then immediately implies that

$$Q[\mathcal{F}] = \{Y(Xe, X) \,|\, X \in \mathcal{C}_n,\ E \bullet X = 1\}.$$

Let $\mathcal{D}_n$ be the cone of $n \times n$ doubly nonnegative (DNN) matrices. Replacing $\mathcal{C}_n$ with $\mathcal{D}_n$ gives a tractable DNN relaxation of QPS, which is exact for $n \leq 4$.

- CP representation for QPS was first obtained by Bomze et al. (2000).

- CP representation for QPS was first obtained by Bomze et al. (2000).
- Applications of QPS include formulation of max stable set problem. In this case DNN relaxation corresponds to Lovasz-Schrijver bound ϑ'.

- CP representation for QPS was first obtained by Bomze et al. (2000).
- Applications of QPS include formulation of max stable set problem. In this case DNN relaxation corresponds to Lovasz-Schrijver bound ϑ'.
- Representation of $Q[\mathcal{F}]$ where $\mathcal{F}$ is the standard simplex can also be used to derive representations for some other sets of interest.

Let $\mathcal{T}$ denote the convex hull of $n+1$ affinely independent points in $\Re^n$, $\mathcal{T} = \{Ax \mid x \in \mathcal{S} \subset \Re^{n+1}\}$, where $\mathcal{S}$ is a standard simplex and the columns of A are the extreme points of $\mathcal{T}$. (So $\mathcal{T}$ is a triangle in $\Re^2$ or a tetrahedron in $\Re^3$).

Let $\mathcal{T}$ denote the convex hull of $n+1$ affinely independent points in $\Re^n$, $\mathcal{T} = \{Ax \,|\, x \in \mathcal{S} \subset \Re^{n+1}\}$, where $\mathcal{S}$ is a standard simplex and the columns of A are the extreme points of $\mathcal{T}$. (So $\mathcal{T}$ is a triangle in $\Re^2$ or a tetrahedron in $\Re^3$).
Since there is an invertible affine mapping from $\mathcal{T} \in \Re^n$ to $\mathcal{S} \in \Re^{n+1}$, above result also gives a representation for $Q[\mathcal{T}]$;

$$Q[\mathcal{T}] = \left\{ \begin{pmatrix} 1 & e^T X A^T \\ AXe & AXA^T \end{pmatrix} \,|\, X \in \mathcal{C}_{n+1},\; E \bullet X = 1 \right\}.$$

Let $\mathcal{T}$ denote the convex hull of $n+1$ affinely independent points in $\Re^n$, $\mathcal{T} = \{Ax \,|\, x \in \mathcal{S} \subset \Re^{n+1}\}$, where $\mathcal{S}$ is a standard simplex and the columns of A are the extreme points of $\mathcal{T}$. (So $\mathcal{T}$ is a triangle in $\Re^2$ or a tetrahedron in $\Re^3$).
Since there is an invertible affine mapping from $\mathcal{T} \in \Re^n$ to $\mathcal{S} \in \Re^{n+1}$, above result also gives a representation for $Q[\mathcal{T}]$;

$$Q[\mathcal{T}] = \left\{ \begin{pmatrix} 1 & e^T X A^T \\ AXe & AXA^T \end{pmatrix} \,|\, X \in \mathcal{C}_{n+1},\; E \bullet X = 1 \right\}.$$

Replacing $\mathcal{C}_{n+1}$ with $\mathcal{D}_{n+1}$ gives a tractable DNN relaxation, which is exact for $n \leq 3$.

Next consider the case where $\mathcal{F} \subset \Re^n$ is a triangulated polytope

$$\mathcal{F} = \mathcal{P} = \cup_{i=1}^{k} \mathcal{T}_i,$$

where each $\mathcal{T}_i$ is the convex hull of $n+1$ affinely independent points, $\mathcal{T}_i = \{A_i x \mid x \in \mathcal{S} \subset \Re^{n+1}\}$. We are primarily interested in cases where $\mathcal{P}$ has a simple enough structure so that a triangulation can be explicitly given.

Next consider the case where $\mathcal{F} \subset \Re^n$ is a triangulated polytope

$$\mathcal{F} = \mathcal{P} = \cup_{i=1}^{k} \mathcal{T}_i,$$

where each $\mathcal{T}_i$ is the convex hull of $n+1$ affinely independent points, $\mathcal{T}_i = \{A_i x \mid x \in \mathcal{S} \subset \Re^{n+1}\}$. We are primarily interested in cases where $\mathcal{P}$ has a simple enough structure so that a triangulation can be explicitly given. From previous result we immediately obtain

$$Q[\mathcal{P}] = \left\{ \sum_{i=1}^{k} \begin{pmatrix} e^T X_i e & e^T X_i A_i^T \\ A_i X_i e & A_i X_i A_i^T \end{pmatrix} \mid \sum_{i=1}^{k} E \bullet X_i = 1,\ X_i \in \mathcal{C}_{n+1}\ \forall i \right\}.$$

Next consider the case where $\mathcal{F} \subset \Re^n$ is a triangulated polytope

$$\mathcal{F} = \mathcal{P} = \cup_{i=1}^k \mathcal{T}_i,$$

where each $\mathcal{T}_i$ is the convex hull of $n+1$ affinely independent points, $\mathcal{T}_i = \{A_i x \mid x \in \mathcal{S} \subset \Re^{n+1}\}$. We are primarily interested in cases where $\mathcal{P}$ has a simple enough structure so that a triangulation can be explicitly given. From previous result we immediately obtain

$$Q[\mathcal{P}] = \left\{ \sum_{i=1}^k \begin{pmatrix} e^T X_i e & e^T X_i A_i^T \\ A_i X_i e & A_i X_i A_i^T \end{pmatrix} \mid \sum_{i=1}^k E \bullet X_i = 1,\ X_i \in \mathcal{C}_{n+1}\ \forall i \right\}.$$

Replacing $\mathcal{C}_{n+1}$ with $\mathcal{D}_{n+1}$ gives a tractable DNN relaxation, which is again exact for $n \leq 3$.

Outline

Box QP problem QPB corresponds to $\mathcal{F} = \{x \mid 0 \le x \le e\}$. To use Burer's CP representation, write $\mathcal{F}$ in form

$$\{(x, s) \ge 0 \mid x_i + s_i = 1,\ i = 1, \ldots, n\}$$

and consider matrix

$$Y^+ = \begin{pmatrix} 1 & x^T & s^T \\ x & X & Z \\ s & Z^T & S \end{pmatrix}.$$

Box QP problem QPB corresponds to $\mathcal{F} = \{x \,|\, 0 \le x \le e\}$. To use Burer's CP representation, write $\mathcal{F}$ in form

$$\{(x,s) \ge 0 \,|\, x_i + s_i = 1,\ i = 1,\ldots,n\}$$

and consider matrix

$$Y^+ = \begin{pmatrix} 1 & x^T & s^T \\ x & X & Z \\ s & Z^T & S \end{pmatrix}.$$

By Burer's result, QPB is equivalent to the problem

$$\begin{array}{ll} \min & Q \bullet X + c^T x \\ \text{s.t.} & x + s = e, \quad \mathrm{diag}(X + 2Z + S) = e, \\ & Y^+ \in \mathcal{C}_{2n+1}. \end{array}$$

- Replacing $\mathcal{C}_{2n+1}$ with $\mathcal{D}_{2n+1}$ gives tractable doubly nonnegative (DNN) relaxation.

- Replacing $\mathcal{C}_{2n+1}$ with $\mathcal{D}_{2n+1}$ gives tractable doubly nonnegative (DNN) relaxation.
- Can easily show that DNN relaxation is equivalent to "SDP+RLT" relaxation that imposes $Y(x, X) \succeq 0$ and the RLT constraints

$$x_{ij} \geq x_i + x_j - 1, \; x_{ij} \geq 0, \; x_{ij} \leq x_i, \; x_{ij} \leq x_j.$$

- Replacing $\mathcal{C}_{2n+1}$ with $\mathcal{D}_{2n+1}$ gives tractable doubly nonnegative (DNN) relaxation.
- Can easily show that DNN relaxation is equivalent to "SDP+RLT" relaxation that imposes $Y(x, X) \succeq 0$ and the RLT constraints

$$x_{ij} \geq x_i + x_j - 1, \; x_{ij} \geq 0, \; x_{ij} \leq x_i, \; x_{ij} \leq x_j.$$

- For $n = 2$, DNN "relaxation" is equivalent to QPB (A. and Burer, 2010). Result can be viewed as strengthening of well-know fact that for two variables, RLT constraints generate convex hull of $\{x_1 x_2 \,|\, 0 \leq x_i \leq 1, \; i = 1, 2\}$.

- Constraints from Boolean Quadric Polytope (BQP) are valid for off-diagonal components of X (Burer and Letchford, 2009). For example can impose triangle (TRI) inequalities

$$
\begin{aligned}
x_1 + x_2 + x_3 &\le x_{12} + x_{13} + x_{23} + 1, \\
x_{12} + x_{13} &\le x_1 + x_{23}, \\
x_{12} + x_{23} &\le x_2 + x_{13}, \\
x_{13} + x_{23} &\le x_3 + x_{12}.
\end{aligned}
$$

- Constraints from Boolean Quadric Polytope (BQP) are valid for off-diagonal components of X (Burer and Letchford, 2009). For example can impose triangle (TRI) inequalities

$$\begin{aligned} x_1 + x_2 + x_3 &\leq x_{12} + x_{13} + x_{23} + 1, \\ x_{12} + x_{13} &\leq x_1 + x_{23}, \\ x_{12} + x_{23} &\leq x_2 + x_{13}, \\ x_{13} + x_{23} &\leq x_3 + x_{12}. \end{aligned}$$

- For $n = 3$, BQP is completely determined by triangle inequalities and RLT constraints. However, can still be a gap when using SDP+RLT+TRI relaxation; example from Burer and Letchford ($n = 3$) has solution value 1.0, bound value 1.093.

- Constraints from Boolean Quadric Polytope (BQP) are valid for off-diagonal components of X (Burer and Letchford, 2009). For example can impose triangle (TRI) inequalities

$$\begin{aligned} x_1 + x_2 + x_3 &\leq x_{12} + x_{13} + x_{23} + 1, \\ x_{12} + x_{13} &\leq x_1 + x_{23}, \\ x_{12} + x_{23} &\leq x_2 + x_{13}, \\ x_{13} + x_{23} &\leq x_3 + x_{12}. \end{aligned}$$

- For $n = 3$, BQP is completely determined by triangle inequalities and RLT constraints. However, can still be a gap when using SDP+RLT+TRI relaxation; example from Burer and Letchford ($n = 3$) has solution value 1.0, bound value 1.093.
- For Burer-Letchford example (n=3), solution matrix Y^+ has 5×5 principal submatrix that is *not* strictly positive, and is *not* CP. Can obtain copositive cut, re-solve problem, and repeat.

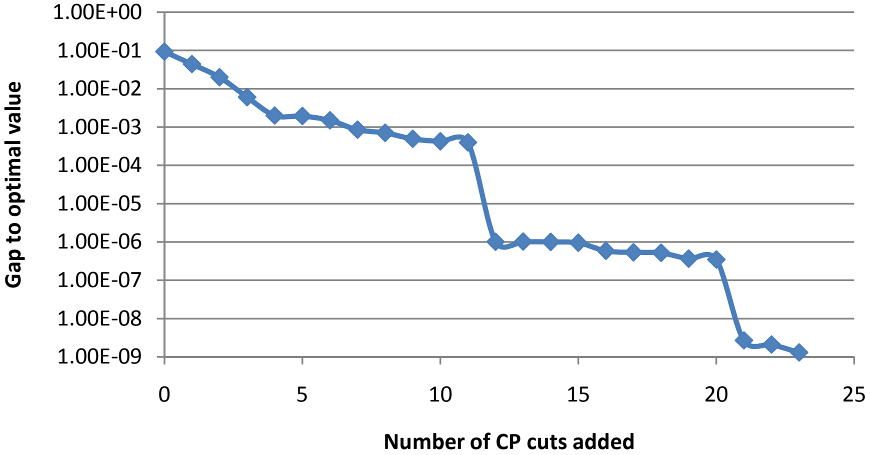

Figure: Gap to optimal value for Burer-Letchford QPB problem ($n = 3$)

- Can obtain better bounds using tighter approximations of $\mathcal{C}_n$ or $\mathcal{C}_n^*$. For Burer-Letchford example, using $\mathcal{K}_7^1 = \mathcal{Q}_7^1$ in place of $\mathcal{D}_7^*$ obtains exact solution value.

- Can obtain better bounds using tighter approximations of $\mathcal{C}_n$ or $\mathcal{C}_n^*$. For Burer-Letchford example, using $\mathcal{K}_7^1 = \mathcal{Q}_7^1$ in place of $\mathcal{D}_7^*$ obtains exact solution value.
- For $n = 3$ can obtain exact representation using result for QPS and triangulation of the cube; this representation uses 5 or 6 matrices in $\mathcal{D}_4$.

- Can obtain better bounds using tighter approximations of $\mathcal{C}_n$ or $\mathcal{C}_n^*$. For Burer-Letchford example, using $\mathcal{K}_7^1 = \mathcal{Q}_7^1$ in place of $\mathcal{D}_7^*$ obtains exact solution value.
- For $n = 3$ can obtain exact representation using result for QPS and triangulation of the cube; this representation uses 5 or 6 matrices in $\mathcal{D}_4$.
- For larger n, methodology based on imposing SDP+RLT+TRI constraints often gives excellent bounds. This approach was first considered by Yajima and Fujie (1998).

Consider 54 QPB maximization problems with $n = 20, 30, 40, 50, 60$ from Vandenbussche and Nemhauser (2003). Density of (c, Q) varies from 30% to 100%. Compare bounds using SDP ($Y(x, X) \succeq 0$ with added bounds $x_{ii} \leq x_i$ on diagonal components), SDP+RLT and SDP+RLT+TRI. When using TRI inequalities, generate RLT and TRI inequalities in several rounds, with decreasing infeasibility tolerance.

Consider 54 QPB maximization problems with $n = 20, 30, 40, 50, 60$ from Vandenbussche and Nemhauser (2003). Density of (c, Q) varies from 30% to 100%. Compare bounds using SDP ($Y(x, X) \succeq 0$ with added bounds $x_{ii} \leq x_i$ on diagonal components), SDP+RLT and SDP+RLT+TRI. When using TRI inequalities, generate RLT and TRI inequalities in several rounds, with decreasing infeasibility tolerance.

- Exact solution of 50 problems accomplished using branch and cut with polyhedral bounds by Vandenbussche and Nemhauser, using up to $28,000$ LPs and $500,000$ cuts per problem.

Consider 54 QPB maximization problems with $n = 20, 30, 40, 50, 60$ from Vandenbussche and Nemhauser (2003). Density of (c, Q) varies from 30% to 100%. Compare bounds using SDP ($Y(x, X) \succeq 0$ with added bounds $x_{ii} \le x_i$ on diagonal components), SDP+RLT and SDP+RLT+TRI. When using TRI inequalities, generate RLT and TRI inequalities in several rounds, with decreasing infeasibility tolerance.

- Exact solution of 50 problems accomplished using branch and cut with polyhedral bounds by Vandenbussche and Nemhauser, using up to $28,000$ LPs and $500,000$ cuts per problem.
- For 15 problems with $n = 30$, root gap for polyhedral bound averages 71%; root gap using RLT averages 77%; root gap for BARON after bound-tightening averages 74%.

Problem	Objective Value OPT	SDP	+RLT	+RLT+TRI	Cuts Added RLT	TRI	% Gaps to OPT SDP	+RLT	+RLT+TRI
20-100-1	706.50	739.39	706.52	706.50	197	55	4.655%	0.002%	0.000%
20-100-2	856.50	900.20	857.97	856.50	184	172	5.102%	0.171%	0.000%
20-100-3	772.00	785.51	772.00		168		1.750%	0.000%	
30-060-1	706.00	768.12	714.68	706.00	371	777	8.799%	1.229%	0.000%
30-060-2	1377.17	1426.94	1377.17		381		3.614%	0.000%	
30-060-3	1293.50	1370.13	1298.26	1293.50	394	288	5.924%	0.368%	0.000%
30-070-1	654.00	746.43	674.00	654.00	369	784	14.133%	3.058%	0.000%
30-070-2	1313.00	1375.07	1313.00		449		4.727%	0.000%	
30-070-3	1657.40	1719.77	1657.57	1657.40	452	442	3.763%	0.010%	0.000%
30-080-1	952.73	1050.76	965.25	952.73	365	718	10.290%	1.315%	0.000%
30-080-2	1597.00	1622.81	1597.00		376		1.616%	0.000%	
30-080-3	1809.78	1836.79	1809.78		317		1.492%	0.000%	
30-090-1	1296.50	1348.48	1296.50		370		4.009%	0.000%	
30-090-2	1466.84	1527.87	1466.84		344		4.160%	0.000%	
30-090-3	1494.00	1516.81	1494.00		420		1.527%	0.000%	
30-100-1	1227.13	1285.74	1227.13		356		4.777%	0.000%	
30-100-2	1260.50	1365.32	1261.11	1260.50	427	465	8.316%	0.048%	0.000%
30-100-3	1511.05	1611.11	1513.15	1511.05	377	265	6.622%	0.139%	0.000%
40-030-1	839.50	876.60	839.50		656		4.419%	0.000%	
40-030-2	1429.00	1496.83	1429.00		889		4.747%	0.000%	
40-030-3	1086.00	1156.52	1086.00		705		6.494%	0.000%	
40-040-1	837.00	956.09	863.09	837.00	710	1966	14.228%	3.117%	0.000%
40-040-2	1428.00	1452.53	1428.00		600		1.718%	0.000%	
40-040-3	1173.50	1269.83	1180.85	1173.50	745	1427	8.209%	0.626%	0.000%
40-050-1	1154.50	1276.79	1160.44	1154.50	797	1608	10.592%	0.515%	0.000%
40-050-2	1430.98	1517.51	1436.05	1430.98	788	961	6.047%	0.354%	0.000%
40-050-3	1653.63	1747.31	1653.63		680		5.665%	0.000%	

Problem	Objective Value OPT	SDP	+RLT	+RLT+TRI	Cuts Added RLT	TRI	% Gaps to OPT SDP	+RLT	+RLT+TRI
40-060-1	1322.67	1481.96	1352.92	1322.67	696	1722	12.043%	2.287%	0.000%
40-060-2	2004.23	2099.58	2004.23		739		4.758%	0.000%	
40-060-3	2454.50	2508.68	2454.50		701		2.207%	0.000%	
40-070-1	1605.00	1663.98	1605.00		584		3.675%	0.000%	
40-070-2	1867.50	1931.34	1867.50		650		3.418%	0.000%	
40-070-3	2436.50	2522.71	2436.50		828		3.538%	0.000%	
40-080-1	1838.50	1936.17	1838.50		615		5.312%	0.000%	
40-080-2	1952.50	2012.92	1952.50		639		3.094%	0.000%	
40-080-3	2545.50	2638.34	2545.89	2545.50	755	742	3.647%	0.015%	0.000%
40-090-1	2135.50	2262.51	2135.50		763		5.948%	0.000%	
40-090-2	2113.00	2268.86	2113.75	2113.00	731	336	7.376%	0.035%	0.000%
40-090-3	2535.00	2594.26	2535.00		598		2.338%	0.000%	
40-100-1	2476.38	2557.23	2476.38		673		3.265%	0.000%	
40-100-2	2102.50	2216.62	2106.37	2102.50	707	1251	5.428%	0.184%	0.000%
40-100-3	1866.07	2037.31	1908.19	1866.07	664	1732	9.176%	2.257%	0.000%
50-030-1	1324.50	1389.09	1324.50		903		4.877%	0.000%	
50-030-2	1668.00	1755.68	1671.33	1668.00	831	233	5.257%	0.200%	0.000%
50-030-3	1453.61	1565.76	1454.88	1453.61	830	180	7.715%	0.087%	0.000%
50-040-1	1411.00	1483.01	1411.00		1017		5.103%	0.000%	
50-040-2	1745.76	1881.33	1749.46	1745.76	868	509	7.766%	0.212%	0.000%
50-040-3	2094.50	2176.98	2094.50		1081		3.938%	0.000%	
50-050-1	1198.41	1417.77	1302.24	1200.14	723	1531	18.304%	8.664%	0.144%
50-050-2	1776.00	1942.53	1789.58	1776.00	867	667	9.377%	0.765%	0.000%
50-050-3	2106.10	2268.04	2121.93	2106.10	937	933	7.689%	0.752%	0.000%
60-020-1	1212.00	1297.42	1212.00		1199		7.048%	0.000%	
60-020-2	1925.50	2010.57	1925.50		1319		4.418%	0.000%	
60-020-3	1483.00	1604.60	1491.06	1483.00	1040	735	8.200%	0.543%	0.000%
						Average:	5.969%	0.499%	

Outline

For TRS interested in $\mathcal{F} = \{x \mid \|x\| \leq 1\}$. SDP representation for $Q[\mathcal{F}]$ first given by Rendl and Wolkowocz (1997):

$$Q[\mathcal{F}] = \{\, Y(x, X) \succeq 0 \mid \operatorname{tr}(X) \leq 1 \}.$$

For TRS interested in $\mathcal{F} = \{x \mid \|x\| \leq 1\}$. SDP representation for $Q[\mathcal{F}]$ first given by Rendl and Wolkowocz (1997):

$$Q[\mathcal{F}] = \{ Y(x, X) \succeq 0 \mid \mathrm{tr}(X) \leq 1 \}.$$

Since $\mathcal{F}$ is not polyhedral, Burer's CP representation cannot be applied. Can instead use Pataki's (1998) rank result for extreme points of SDP constraint systems.

Proposition (Pataki's rank result)

Consider an SDP feasible set in block standard form: $F := \{X^j \succeq 0,\ j = 1,\dots,p : \sum_{j=1}^{p} A_i^j \bullet X^j = b_i,\ i = 1,\dots,m\}$. *Let* $(X^1,\dots,X^p)$ *be an extreme point of* F*, and define* $r_j := \mathrm{rank}(X^j)$. *Then* $\sum_{j=1}^{p} r_j(r_j+1) \le 2m$.

Proposition (Pataki's rank result)

Consider an SDP feasible set in block standard form:
$F := \{X^j \succeq 0,\ j = 1, \ldots, p : \sum_{j=1}^{p} A_i^j \bullet X^j = b_i,\ i = 1, \ldots, m\}$. *Let* $(X^1, \ldots, X^p)$ *be an extreme point of* F*, and define* $r_j := \operatorname{rank}(X^j)$. *Then* $\sum_{j=1}^{p} r_j(r_j + 1) \leq 2m$.

Note that standard form LP problem corresponds to diagonal matrix, with $p = n$ 1×1"blocks." Rank of block for x_i is zero if $x_i = 0$, and one otherwise, so result is that for an extreme point must have $\sum_{j=1}^{n} r_j(r_j + 1) = \sum_{j=1}^{n} 2r_j \leq 2m$, meaning that at most m variables are positive.

Lemma (SDP representation of TRS)

Suppose that $Y(x, X)$ *is an extreme point of the convex set* $\{Y(x, X) \succeq 0 \mid \mathrm{tr}(X) \leq 1\}$. *Then* $X = xx^T$, *where* $\|x\| \leq 1$.

Lemma (SDP representation of TRS)

Suppose that $Y(x, X)$ is an extreme point of the convex set $\{Y(x, X) \succeq 0 \mid \mathrm{tr}(X) \leq 1\}$. Then $X = xx^T$, where $\|x\| \leq 1$.

Proof.

The given convex set can be expressed in the form of Pataki's result as

$$F := \left\{ Y = \begin{pmatrix} x_0 & x^T \\ x & X \end{pmatrix} \succeq 0,\ s \geq 0 \,\middle|\, x_0 = 1,\ \mathrm{tr}(X) + s = 1 \right\}.$$

Lemma (SDP representation of TRS)

Suppose that $Y(x, X)$ *is an extreme point of the convex set* $\{Y(x, X) \succeq 0 \mid \operatorname{tr}(X) \leq 1\}$. *Then* $X = xx^T$, *where* $\|x\| \leq 1$.

Proof.

The given convex set can be expressed in the form of Pataki's result as

$$F := \left\{ Y = \begin{pmatrix} x_0 & x^T \\ x & X \end{pmatrix} \succeq 0,\ s \geq 0 \,\middle|\, x_0 = 1,\ \operatorname{tr}(X) + s = 1 \right\}.$$

Pataki's result then implies that if (Y, s) is an extreme point of F, $r_Y(r_Y + 1) + r_s(r_s + 1) \leq 4$, where $r_Y = \operatorname{rank}(Y)$ and $r_s = \operatorname{rank}(s)$. Then $r_Y \leq 1$, and since $Y \neq 0$ it must be that $r_Y = 1$, implying $X = xx^T$. The fact that $\|x\| \leq 1$ follows from $Y(x, X) \succeq 0$ and $\operatorname{tr}(X) \leq 1$. □

Outline

Outline

Next consider a problem TRS1 that adds one linear constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ a^T x \le u\}$.

Next consider a problem TRS1 that adds one linear constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \leq 1,\ a^T x \leq u\}$. Sturm and Zhang (2003) proved that in this case $Q[\mathcal{F}]$ has a representation as a mixed SDP/SOCP system

$$Q[\mathcal{F}] = \{Y(x, X) \succeq 0 \mid \mathrm{tr}(X) \leq 1,\ \|ux - Xa\| \leq u - a^T x\}.$$

Next consider a problem TRS1 that adds one linear constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ a^T x \le u\}$. Sturm and Zhang (2003) proved that in this case $Q[\mathcal{F}]$ has a representation as a mixed SDP/SOCP system

$$Q[\mathcal{F}] = \{Y(x, X) \succeq 0 \mid \operatorname{tr}(X) \le 1,\ \|ux - Xa\| \le u - a^T x\}.$$

Derivation of the constraint $\|ux - Xa\| \le u - a^T x$ can be viewed as an extension of the well-known RLT procedure to SOC constraints, since

$$\|(u - a^T x)x\| = (u - a^T x)\|x\| \le u - a^T x$$

for a feasible x, and $(a^T x) = xx^T a$.

Next consider a problem TRS1 that adds one linear constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ a^T x \le u\}$. Sturm and Zhang (2003) proved that in this case $Q[\mathcal{F}]$ has a representation as a mixed SDP/SOCP system

$$Q[\mathcal{F}] = \{Y(x, X) \succeq 0 \mid \operatorname{tr}(X) \le 1,\ \|ux - Xa\| \le u - a^T x\}.$$

Derivation of the constraint $\|ux - Xa\| \le u - a^T x$ can be viewed as an extension of the well-known RLT procedure to SOC constraints, since

$$\|(u - a^T x)x\| = (u - a^T x)\|x\| \le u - a^T x$$

for a feasible x, and $(a^T x) = xx^T a$. Replacing xx^T with X then gives the SOC-RLT constraint $\|ux - Xa\| \le u - a^T x$.

Outline

Problem TRS2p adds two parallel linear constraints to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ l \le a^T x \le u\}$.

Problem TRS2p adds two parallel linear constraints to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \leq 1,\ l \leq a^T x \leq u\}$. Ye and Zhang (2003) showed that TRS2p could be solved via a case analysis, but could not obtain a single convex programming representation.

Problem TRS2p adds two parallel linear constraints to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \leq 1,\ l \leq a^T x \leq u\}$. Ye and Zhang (2003) showed that TRS2p could be solved via a case analysis, but could not obtain a single convex programming representation.

Burer and A. (2011) show that for TRS2p,

$$Q[\mathcal{F}] = \left\{ Y(x, X) \succeq 0 \mid \mathrm{tr}(X) \leq 1,\ \begin{array}{c} \|ux - Xa\| \leq u - a^T x \\ \|lx - Xa\| \leq a^T x - l \\ (l + u)a^T x - a^T Xa \geq lu \end{array} \right\}.$$

Problem TRS2p adds two parallel linear constraints to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ l \le a^T x \le u\}$. Ye and Zhang (2003) showed that TRS2p could be solved via a case analysis, but could not obtain a single convex programming representation.

Burer and A. (2011) show that for TRS2p,

$$Q[\mathcal{F}] = \left\{ Y(x,X) \succeq 0 \mid \mathrm{tr}(X) \le 1,\ \begin{array}{r} \|ux - Xa\| \le u - a^T x \\ \|lx - Xa\| \le a^T x - l \\ (l+u)a^T x - a^T X a \ge lu \end{array} \right\}.$$

The constraints in the above representation include two SOC-RLT constraints, each obtained from one linear inequality and the SOC constraint $\|x\| \le 1$, as well as the ordinary RLT constraint obtained from the two linear inequalities together.

Outline

The two-trust-region subproblem (TTRS) is the problem obtained by adding a second full-dimensional ellipsoidal constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ \|H^{1/2}(x-h)\| \le 1\}$ where $H \succ 0$ and $h \in \Re^n$ is the center of the second ellipsoid. TTRS has been heavily studied in the NLP literature.

The two-trust-region subproblem (TTRS) is the problem obtained by adding a second full-dimensional ellipsoidal constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ \|H^{1/2}(x-h)\| \le 1\}$ where $H \succ 0$ and $h \in \Re^n$ is the center of the second ellipsoid. TTRS has been heavily studied in the NLP literature.

Standard SDP relaxation for TTRS is

$$\min \left\{ Q \bullet X + c^T x \mid \begin{array}{c} \mathrm{tr}(X) \le 1,\ Y(x,X) \succeq 0 \\ H \bullet X - 2h^T Hx + h^T Hh \le 1 \end{array} \right\}.$$

The two-trust-region subproblem (TTRS) is the problem obtained by adding a second full-dimensional ellipsoidal constraint to TRS, corresponding to $\mathcal{F} = \{x \mid \|x\| \le 1,\ \|H^{1/2}(x-h)\| \le 1\}$ where $H \succ 0$ and $h \in \Re^n$ is the center of the second ellipsoid. TTRS has been heavily studied in the NLP literature.

Standard SDP relaxation for TTRS is

$$\min \left\{ Q \bullet X + c^T x \mid \begin{array}{c} \mathrm{tr}(X) \le 1,\ \ Y(x,X) \succeq 0 \\ H \bullet X - 2h^T Hx + h^T Hh \le 1 \end{array} \right\}.$$

Known that this relaxation may have a nonzero gap. Not immediately clear how to strengthen it; for example, no explicit linear inequality constraints from which to derive SOC-RLT constraints.

Consider SOC-RLT constraints derived from supporting hyperplanes of the ball $B = \{x : \|x\| \leq 1\}$. Given any vector a with $\|a\| = 1$, the inequality $a^T x \leq 1$ supports B at a, so the SOC-RLT constraint $\|H^{1/2}(x - Xa - (1 - a^T x)h)\| \leq 1 - a^T x$ can be used to strengthen the SDP relaxation.

Consider SOC-RLT constraints derived from supporting hyperplanes of the ball $B = \{x : \|x\| \le 1\}$. Given any vector a with $\|a\| = 1$, the inequality $a^T x \le 1$ supports B at a, so the SOC-RLT constraint $\|H^{1/2}(x - Xa - (1 - a^T x)h)\| \le 1 - a^T x$ can be used to strengthen the SDP relaxation.

- Can show that it is equivalent to generate SOC-RLT cuts using supporting hyperplanes from second ellipsoid instead of B.

Consider SOC-RLT constraints derived from supporting hyperplanes of the ball $B = \{x : \|x\| \leq 1\}$. Given any vector a with $\|a\| = 1$, the inequality $a^T x \leq 1$ supports B at a, so the SOC-RLT constraint $\|H^{1/2}(x - Xa - (1 - a^T x)h)\| \leq 1 - a^T x$ can be used to strengthen the SDP relaxation.

- Can show that it is equivalent to generate SOC-RLT cuts using supporting hyperplanes from second ellipsoid instead of B.
- Separation problem of finding a vector a that generates an SOC-RLT cut for which a given point is infeasible can be formulated as an ordinary TRS problem.

Consider SOC-RLT constraints derived from supporting hyperplanes of the ball $B = \{x : \|x\| \le 1\}$. Given any vector a with $\|a\| = 1$, the inequality $a^T x \le 1$ supports B at a, so the SOC-RLT constraint $\|H^{1/2}(x - Xa - (1 - a^T x)h)\| \le 1 - a^T x$ can be used to strengthen the SDP relaxation.

- Can show that it is equivalent to generate SOC-RLT cuts using supporting hyperplanes from second ellipsoid instead of B.
- Separation problem of finding a vector a that generates an SOC-RLT cut for which a given point is infeasible can be formulated as an ordinary TRS problem.
- Consider 4 examples from literature for which SDP relaxation of TTRS known to have gap. By using SOC-RLT cuts, can get gap to zero in all cases using at most 5 cuts.

However cannot always drive gap to zero using these SOC-RLT cuts. Consider TTRS of form

$$\min\{x^TQx + c^Tx : \|x\| \leq 1, \|H^{1/2}x\| \leq 1\},$$

where second ellipsoid is also centered at origin.

However cannot always drive gap to zero using these SOC-RLT cuts. Consider TTRS of form

$$\min\{x^T Qx + c^T x : \|x\| \le 1, \|H^{1/2}x\| \le 1\},$$

where second ellipsoid is also centered at origin.

- Ye and Zhang (2003) show that SDP relaxation is tight for such a problem if $c = 0$. (Can easily be proved using Pataki rank result.)

However cannot always drive gap to zero using these SOC-RLT cuts. Consider TTRS of form

$$\min\{x^T Qx + c^T x : \|x\| \le 1, \|H^{1/2}x\| \le 1\},$$

where second ellipsoid is also centered at origin.

- Ye and Zhang (2003) show that SDP relaxation is tight for such a problem if $c = 0$. (Can easily be proved using Pataki rank result.)
- However, for instance with $n = 2$ and

$$H = \frac{1}{2}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} -4 & 1 \\ 1 & -2 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

  exact solution value is -4 at $x^* = (\pm 1, \mp 1)^T/\sqrt{2}$, and SDP relaxation has value of -4.25. Optimal value using SOC-RLT cuts is ≈ -4.0360, leaving a 0.9% gap to the true solution value.

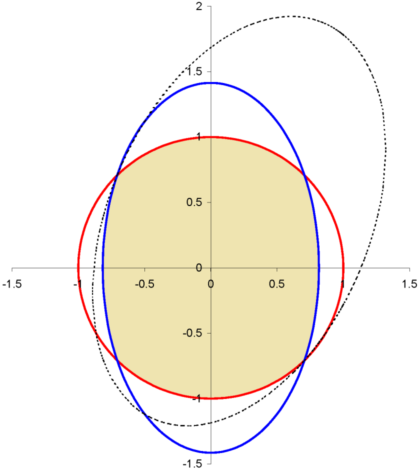

Figure: TTRS with nonzero gap using SOC-RLT constraints

To further investigate use of SOC-RLT constraints numerically, use theorem of Martinez (1994) to generate instances of TRS having one global solution and another local, nonglobal minimum. Add second ellipsoidal constraint that cuts off global solution. Resulting problems are good candidates for "difficult" TTRS.

To further investigate use of SOC-RLT constraints numerically, use theorem of Martinez (1994) to generate instances of TRS having one global solution and another local, nonglobal minimum. Add second ellipsoidal constraint that cuts off global solution. Resulting problems are good candidates for "difficult" TTRS.

Consider 1000 instances each for $n = 5, 10, 20$. First solve SDP relaxation. If solution is *not* rank-one, add up to 25 SOC-RLT cuts. Consider rank measure λ_n/λ_{n-1}, applied to $Y(x, X)$. Consider solution to be numerically rank-one if rank measure $> 10^4$.

To further investigate use of SOC-RLT constraints numerically, use theorem of Martinez (1994) to generate instances of TRS having one global solution and another local, nonglobal minimum. Add second ellipsoidal constraint that cuts off global solution. Resulting problems are good candidates for "difficult" TTRS.

Consider 1000 instances each for $n = 5, 10, 20$. First solve SDP relaxation. If solution is *not* rank-one, add up to 25 SOC-RLT cuts. Consider rank measure λ_n/λ_{n-1}, applied to $Y(x, X)$. Consider solution to be numerically rank-one if rank measure $> 10^4$.

	Solved		Unsolved
n	SDP	+SOC-RLT	
5	92.1%	4.9 %	3.0%
10	17.5%	74.7%	7.8%
20	7.7%	84.5%	7.7%

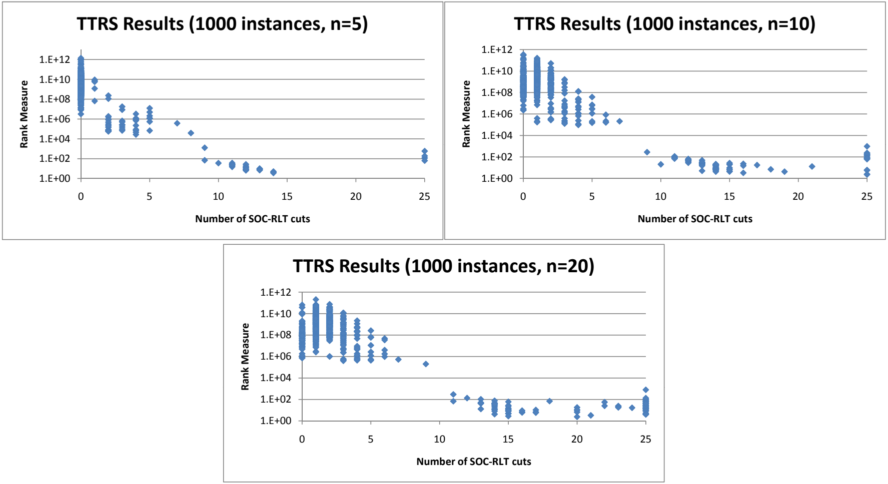

Figure: TTRS results based on Martínez (1994) with $n = 5, 10$ and 20.

Outline

Open Problems

- Complete description of $Q[\mathcal{F}]$ for box constraints $\mathcal{F} = \{x \mid 0 \leq x \leq e\}$ for $n = 3$, using only original (x, X) variables.

Open Problems

- Complete description of $Q[\mathcal{F}]$ for box constraints $\mathcal{F} = \{x \,|\, 0 \le x \le e\}$ for $n = 3$, using only original (x, X) variables.
- TRS2 - TRS with two additional linear inequalities (not parallel). Same construction with RLT and SOC-RLT constraints as used for TRS2p applies, but current proof does not. May be distinction between cases where constraints do and do not intersect in ball.

Open Problems

- Complete description of $Q[\mathcal{F}]$ for box constraints $\mathcal{F} = \{x \mid 0 \le x \le e\}$ for $n = 3$, using only original (x, X) variables.
- TRS2 - TRS with two additional linear inequalities (not parallel). Same construction with RLT and SOC-RLT constraints as used for TRS2p applies, but current proof does not. May be distinction between cases where constraints do and do not intersect in ball.
- TTRS - devise polynomial-time solution procedure, or demonstrate that problem is NP-hard.

Thank You