On Nonconvex Quadratic Programming with Ball Constraints

Sam Burer

University of Iowa

Kurt-Fest!
SIOPT Optimization Conference
Seattle, WA
June 2, 2023

Outline

Happy Retirement, Kurt!

Happy Retirement, Kurt!

Happy Retirement, Kurt!

What do you get the person who has. . . ?


ANY
TIME
PORSCHE
IOWA
OPTIMAL
JOHNSON
MAY


ANY
TIME
PORSCHE
IOWA
OPTIMAL
JOHNSON
MAY


IOWA
OPTIMAL
JOHNSON
MAY

Outline

Quadratically Constrained Quadratic Programs

$$\min_{x \in \mathbb{R}^n} \left\{ x^T Q x + 2\, q^T x : \begin{array}{c} x \in \mathcal{C} \\ x^T H_i x + 2\, g_i^T x + f_i \le 0 \quad \forall\ i \end{array} \right\}$$

Quadratically Constrained Quadratic Programs

$$\min_{x \in \mathbb{R}^n} \left\{ x^T Q x + 2\, q^T x : \begin{array}{c} x \in \mathcal{C} \\ x^T H_i x + 2\, g_i^T x + f_i \leq 0 \;\; \forall\, i \end{array} \right\}$$

$$\downarrow$$

$$\text{CH} := \text{conv} \left\{ (x, X) : \begin{array}{c} x \text{ feasible} \\ X = x x^T \end{array} \right\}$$

The Shor Relaxation

$$\textsc{Shor} := \left\{ (x, X) : \begin{array}{c} x \in \mathcal{C} \\ H_i \bullet X + 2\, g_i^T x + f_i \le 0 \quad \forall\, i \\ X \succeq x x^T \end{array} \right\}$$

Notes: Introduced by Shor (1987). May need additional constraints to ensure $\textsc{Shor}$ is bounded, even when original feasible set is already bounded.

The RLT Relaxation

Given explicit $a^T x \le \beta$ and $c^T x \le \delta$ in $\mathcal{C}$, we have:

$$
\begin{aligned}
(\beta - a^T x)(\delta - c^T x) \ge 0 \quad &\Longleftrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T x x^T c \ge 0 \\
&\Longrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T X c \ge 0
\end{aligned}
$$

The RLT Relaxation

Given explicit $a^T x \le \beta$ and $c^T x \le \delta$ in $\mathcal{C}$, we have:

$$
\begin{aligned}
(\beta - a^T x)(\delta - c^T x) \ge 0 \quad &\Longleftrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T x x^T c \ge 0 \\
&\Longrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T X c \ge 0
\end{aligned}
$$

$$\text{RLT} := \{(x, X) : \text{ [all such linear inequalities] } \}$$

The RLT Relaxation

Given explicit $a^T x \le \beta$ and $c^T x \le \delta$ in $\mathcal{C}$, we have:

$$(\beta - a^T x)(\delta - c^T x) \ge 0 \quad \Longleftrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T x x^T c \ge 0$$
$$\Longrightarrow \quad \beta\delta - (\beta c + \delta a)^T x + a^T X c \ge 0$$

$$\text{RLT} := \{(x, X) : \text{ [all such linear inequalities] } \}$$

Note: *RLT* stands for *reformulation linearization technique* and was popularized by McCormick (1976) and Sherali-Adams (1999).

The SocRlt Relaxation

Let $\mathcal{L}$ be the second-order cone. Given explicit $a^T x \le \beta$ and $b - Ax \in \mathcal{L}$ in $\mathcal{C}$, we have:

$$\begin{aligned}
(\beta - a^T x)(b - Ax) \in \mathcal{L} \quad &\Longleftrightarrow \quad \beta b - \beta A x - (a^T x) b + (a^T x) A x \in \mathcal{L} \\
&\Longrightarrow \quad \beta b - (\beta A + b a^T) x + A X a \in \mathcal{L}
\end{aligned}$$

The $\textsc{SocRlt}$ Relaxation

Let $\mathcal{L}$ be the second-order cone. Given explicit $a^T x \le \beta$ and $b - Ax \in \mathcal{L}$ in $\mathcal{C}$, we have:

$$
\begin{aligned}
(\beta - a^T x)(b - Ax) \in \mathcal{L} \quad &\Longleftrightarrow \quad \beta b - \beta A x - (a^T x) b + (a^T x) A x \in \mathcal{L} \\
&\Longrightarrow \quad \beta b - (\beta A + b a^T) x + A X a \in \mathcal{L}
\end{aligned}
$$

$$\textsc{SocRlt} := \{(x, X) : \ [\text{ all such SOC constraints }] \ \}$$

The $\textsc{SocRlt}$ Relaxation

Let $\mathcal{L}$ be the second-order cone. Given explicit $a^T x \le \beta$ and $b - Ax \in \mathcal{L}$ in $\mathcal{C}$, we have:

$$\begin{aligned} (\beta - a^T x)(b - Ax) \in \mathcal{L} \quad &\Longleftrightarrow \quad \beta b - \beta A x - (a^T x) b + (a^T x) A x \in \mathcal{L} \\ &\Longrightarrow \quad \beta b - (\beta A + b a^T) x + A X a \in \mathcal{L} \end{aligned}$$

$$\textsc{SocRlt} := \{(x, X) : \text{ [all such SOC constraints] } \}$$

Note: First introduced by Sturm-Zhang (2003).

The Kronecker Relaxation

An example...

$\|x\| \leq 1$ $\qquad\qquad$ $\|x - c\| \leq \rho$

The Kronecker Relaxation

An example...

$$\|x\| \le 1 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & x^T \\ x & I \end{pmatrix} \succeq 0, \qquad \|x - c\| \le \rho$$

The Kronecker Relaxation

An example. . .

$$\|x\| \leq 1 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & x^T \\ x & I \end{pmatrix} \succeq 0, \qquad \|x - c\| \leq \rho \quad \Leftrightarrow \quad \begin{pmatrix} \rho & (x-c)^T \\ x - c & \rho I \end{pmatrix} \succeq 0$$

The Kronecker Relaxation

An example. . .

$$\|x\| \le 1 \quad \Leftrightarrow \quad \begin{pmatrix} 1 & x^T \\ x & I \end{pmatrix} \succeq 0, \qquad \|x - c\| \le \rho \quad \Leftrightarrow \quad \begin{pmatrix} \rho & (x-c)^T \\ x - c & \rho I \end{pmatrix} \succeq 0$$

$$\Downarrow$$

$$\begin{pmatrix} 1 & x^T \\ x & I \end{pmatrix} \otimes \begin{pmatrix} \rho & (x-c)^T \\ x - c & \rho I \end{pmatrix} \succeq 0$$

The Kronecker Relaxation

$$\textsc{Kron} := \{(x, X) : \text{ [all such PSD constraints] } \}$$

The Kronecker Relaxation

$$\textsc{Kron} := \{(x, X) : \text{ [all such PSD constraints] } \}$$

Notes:

- Introduced by Anstreicher (2017).
- Kronecker matrix is big, e.g., size $n^2 \times n^2$. But can reduce the computational burden in some important cases.

A Combined Relaxation

$$\mathrm{CH} \quad \subseteq \quad \textsc{Shor} \cap \textsc{Rlt} \cap \textsc{SocRlt} \cap \textsc{Kron}$$

A Combined Relaxation

$$\text{CH} \quad \subseteq \quad \textsc{Shor} \cap \textsc{Rlt} \cap \textsc{SocRlt} \cap \textsc{Kron}$$

Note: Zhen et al (2022) consider other ways to "multiply" conic constraints to create strong relaxations. See Dick den Hertog's talk later today (MS240).

Outline

Intersection of Euclidean Balls

$$\min_{x \in \mathbb{R}^n} \left\{ x^T Q x + 2\, q^T x : \begin{array}{l} \|x\| \le 1 \\ \|x - c_i\| \le \rho_i \;\; \forall\, i = 2, \ldots, m \end{array} \right\}$$

Intersection of Euclidean Balls

$$\min_{x \in \mathbb{R}^n} \left\{ x^T Q x + 2\, q^T x : \begin{array}{l} \|x\| \le 1 \\ \|x - c_i\| \le \rho_i \ \ \forall\, i = 2, \ldots, m \end{array} \right\}$$

$$\downarrow$$

$$\text{CH} := \text{conv} \left\{ (x, X) : \begin{array}{l} \|x\| \le 1 \\[1em] \|x - c_i\| \le \rho_i \ \ \forall\, i = 2, \ldots, m \\[6em] X = xx^T \end{array} \right\}$$

Intersection of Euclidean Balls

$$\min_{x \in \mathbb{R}^n} \left\{ x^T Q x + 2\, q^T x : \begin{array}{l} \|x\| \le 1 \\ \|x - c_i\| \le \rho_i \;\; \forall\, i = 2, \ldots, m \end{array} \right\}$$

$$\downarrow$$

$$\mathrm{CH} := \mathrm{conv} \left\{ (x, X) : \begin{array}{l} \|x\| \le 1 \\ \|x - c_i\| \le \rho_i \;\; \forall\, i = 2, \ldots, m \\ x^T x \le 1 \\ x^T x \le \rho_i^2 + 2 c_i^T x - c_i^T c_i \;\; \forall\, i = 2, \ldots, m \\ X = x x^T \end{array} \right\}$$

Intersection of Euclidean Balls

- Why do we care?
  - Trust-region subproblem and related variants
  - In particular, the CDT problem, also known as *TTRS* (Celis et al. 1985)
  - Similar problems such as optimal power flow (Chen et al. 2017)
  - Sparse source localization (Beck-Pan 2017)
  - Substructure in more general MINLPs

Intersection of Euclidean Balls

- Why do we care?
  - Trust-region subproblem and related variants
  - In particular, the CDT problem, also known as *TTRS* (Celis et al. 1985)
  - Similar problems such as optimal power flow (Chen et al. 2017)
  - Sparse source localization (Beck-Pan 2017)
  - Substructure in more general MINLPs
- Polynomial-time solvable to ϵ-accuracy for fixed m (Bienstock 2016)

Intersection of Euclidean Balls

- Why do we care?
  - Trust-region subproblem and related variants
  - In particular, the CDT problem, also known as *TTRS* (Celis et al. 1985)
  - Similar problems such as optimal power flow (Chen et al. 2017)
  - Sparse source localization (Beck-Pan 2017)
  - Substructure in more general MINLPs
- Polynomial-time solvable to ϵ-accuracy for fixed m (Bienstock 2016)
- Then what do we care about?
  - Want tight conic relaxations using relaxation machinery

Intersection of Euclidean Balls

- For $m = 1$, Rendl-Wolkowicz (1997) proved $\mathrm{CH} = \textsc{Shor}$
- For $m = 2$, Kelly et al. (2022) gave a disjunctive formulation of CH using two "copies" of X
- Questions
  - Is there a non-disjunctive formulation for $m = 2$?
  - How about when $m \geq 3$?

Case $m = 2$

$$\mathrm{CH} := \mathrm{conv}\left\{ (x, X) : \begin{array}{c} \|x\| \le 1 \\ \|x - c\| \le \rho \\ X = xx^T \end{array} \right\}$$

Case $m = 2$

$$\text{CH} := \text{conv}\left\{ (x, X) : \begin{array}{c} \|x\| \le 1 \\ \|x - c\| \le \rho \\ X = xx^T \end{array} \right\}$$

$$\downarrow$$

$$\text{CH} \subseteq \textsc{Shor} \cap \textsc{Kron}$$

Case $m = 2$

$$\mathrm{CH} := \mathrm{conv}\left\{ (x, X) : \begin{array}{c} \|x\| \le 1 \\ \|x - c\| \le \rho \\ X = xx^T \end{array} \right\}$$

$$\downarrow$$

$$\mathrm{CH} \subseteq \textsc{Shor} \cap \textsc{Kron}$$

Note: The containment is strict.

Case $m = 2$: A Slight Lifting

$$\mathrm{CH}^+ := \mathrm{conv}\left\{ (w, W) : \begin{array}{l} x^T x \le \beta \\ \beta \le 1 \\ \beta \le \rho^2 + 2c^T x - c^T c \\ \\ w = \binom{x}{\beta} \\ W = ww^T \end{array} \right\}$$

Case $m = 2$: A Slight Lifting

$$\mathrm{CH}^+ := \mathrm{conv}\left\{ (w, W) : \begin{array}{l} x^T x \le \beta \\ \beta \le 1 \\ \beta \le \rho^2 + 2c^T x - c^T c \\ \\ w = \binom{x}{\beta} \\ W = ww^T \end{array} \right\}$$

$$\downarrow$$

$$\mathrm{CH}^+ \subseteq \textsc{Shor}^+ \cap \textsc{Rlt}^+ \cap \textsc{SocRlt}^+$$

Case $m = 2$: A Slight Lifting

$$\mathrm{CH}^+ := \mathrm{conv}\left\{ (w, W) : \begin{array}{l} x^T x \le \beta \\ \beta \le 1 \\ \beta \le \rho^2 + 2c^T x - c^T c \\ (1-\beta)(\rho^2 + 2c^T x - c^T c - \beta) = 0 \\ w = \binom{x}{\beta} \\ W = ww^T \end{array} \right\}$$

$$\downarrow$$

$$\mathrm{CH}^+ \subseteq \textsc{Shor}^+ \cap \textsc{Rlt}^+ \cap \textsc{SocRlt}^+$$

Case $m = 2$: A Slight Lifting

$$\mathrm{CH}^+ := \mathrm{conv}\left\{ (w, W) : \begin{array}{l} x^T x \le \beta \\ \beta \le 1 \\ \beta \le \rho^2 + 2c^T x - c^T c \\ (1-\beta)(\rho^2 + 2c^T x - c^T c - \beta) = 0 \\ w = \binom{x}{\beta} \\ W = ww^T \end{array} \right\}$$

$$\downarrow$$

$$\mathrm{CH}^+ \subseteq \textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+$$

Case $m = 2$: A Slight Lifting

Non-Convex Problem	Pre-β	Post-β	Comment
# variables	n	$n+1$	
# linear constraints	0	2	
# SOC constraints	2	1	each size $\approx n$
# quadratic constraints	0	1	

Case $m = 2$: A Slight Lifting

Relaxation (not including $\textsc{Shor}$)	Pre-β	Post-β	Comment
# RLT_0 constraints	0	1	
# $\textsc{SocRlt}$ constraints	0	2	each size $\approx n$
# Kronecker SDP constraints	1	0	size $\approx n^2 \times n^2$

General Case $m \geq 2$

Relaxation (not incl Shor)	Pre-β	Post-β	Comment
# RLT	0	m^2	Regular RLT, not RLT$_0$
# SocRlt	0	m	size n
# Kronecker	m^2	0	each size $n^2 \times n^2$

Outline

Case $m = 2$

Theorem (B 2023 — but gifted to Kurt)

Case $m = 2$

Theorem (B 2023 — but gifted to Kurt)

For $m = 2$, *it holds that* $\mathrm{CH}^+ = \textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+$. *Hence,*

Case $m = 2$

Theorem (B 2023 — but gifted to Kurt)

For $m = 2$*, it holds that* $\mathrm{CH}^+ = \textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+$*. Hence,*

$$\mathrm{CH} = \mathrm{proj}(\textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+).$$

Case $m = 2$

Theorem (B 2023 — but gifted to Kurt)

For $m = 2$, it holds that $\mathrm{CH}^+ = \textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+$. *Hence,*

$$\mathrm{CH} = \mathrm{proj}(\textsc{Shor}^+ \cap \textsc{Rlt}_0^+ \cap \textsc{SocRlt}^+).$$

Proof.

Leverages key idea of Ye-Zhang (2003) with a key assist from Kurt. □

General Case $m \geq 2$

n	m	# Instances	# Solved		Total Time (s)		
			Kron	Beta	Shor	Kron	Beta
2	5	1,000	211	976	0.3	14.9	1.0
2	9	1,000	429	978	0.3	194.8	1.9
4	9	1,000	0	891	0.7	2335.1	3.8

General Case $m \geq 2$

n	m	# Instances	# Solved		Total Time (s)		
			Kron	Beta	Shor	Kron	Beta
2	5	1,000	211	976	0.3	14.9	1.0
2	9	1,000	429	978	0.3	194.8	1.9
4	9	1,000	0	891	0.7	2335.1	3.8

Note: In particular, for $m \geq 3$, the containment

$$\mathrm{CH} \subseteq \mathrm{proj}(\textsc{Shor}^+ \cap \textsc{Rlt}^+ \cap \textsc{SocRlt}^+)$$

is strict.

Outline

Conclusions

CONGRATULATIONS, KURT!