An Introduction to Semidefinite Program Relaxations of Quadratically Constrained Quadratic Programs

Fatma Kılınç-Karzan

Carnegie Mellon University
Tepper School of Business

IPCO Summer School

June 19-20, 2023

Recap: QCQP and its SDP relaxation

- $$q_i(x) := x^\top A_i x + 2b_i^\top x + c_i = \begin{pmatrix} x \\ 1 \end{pmatrix}^\top \underbrace{\begin{pmatrix} A_i & b_i \\ b_i^\top & c_i \end{pmatrix}}_{=:\, M_i} \begin{pmatrix} x \\ 1 \end{pmatrix} = \left\langle M_i, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix} \right\rangle$$

- $$\text{Opt} = \inf_{x \in \mathbb{R}^n} \left\{ \left\langle M_{\text{obj}}, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix} \right\rangle : \left\langle M_i, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix} \right\rangle \leq 0,\ \forall i \in [m] \right\}$$
$$\geq \inf_{x \in \mathbb{R}^n,\, X \in \mathbb{S}^n,\, Z \in \mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : \begin{array}{l} \langle M_i, Z \rangle \leq 0,\ \forall i \in [m] \\ Z = \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix} \succeq 0 \end{array} \right\} = \text{Opt}_{\text{SDP}}$$

Recap: forms of exactness

- What does exactness mean?
  - Objective value exactness: $\mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}$
    - Optimizer exactness: $\arg\min \mathrm{Opt} = \arg\min \mathrm{Opt}_{\mathsf{SDP}}$
  - Convex hull exactness: $\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}}$ $\longleftarrow$ convexification of substructures
  - Rank-one generated (ROG) property:

    "SDP exactness that is oblivious to the objective function"

    $\longrightarrow$ exactness in the lifted SDP space

Today's outline

- **Rank-one generated (ROG) property of SDPs**
  - Sufficient (necessary) conditions
  - Examples
- **Exactness in the original space**
  - Convex hull
  - Objective value
- **Efficient algorithms for solving SDPs**

Exactness in the lifted SDP space: ROG property

References:

Argue, C., K.-K., F., and Wang, A. L. (2022+). Necessary and sufficient conditions for rank-one generated cones. *Math. Oper. Res.*, Forthcoming, (arXiv:2007.07433)

K.-K., F. and Wang, A. L. (2021). Exactness in SDP relaxations of QCQPs: Theory and applications. Tut. in Oper. Res. INFORMS

Recap: ROG

- Given $\mathcal{M} \subseteq \mathbb{S}^{n+1}$, define $\mathcal{S}(\mathcal{M}) := \{Z \in \mathbb{S}^{n+1}_{+} : \langle M, Z \rangle \leq 0, \ \forall M \in \mathcal{M}\}$

Definition

A closed cone $\mathcal{S} \subseteq \mathbb{S}^{n+1}_{+}$ is rank-one generated (ROG) if

$$\mathcal{S} = \operatorname{conv}\left(\mathcal{S} \cap \left\{zz^\top : z \in \mathbb{R}^{n+1}\right\}\right).$$

Equivalently, if all extreme rays are generated by rank-one matrices.

- **Analogy:** (Integer programs, integral polyhedra) $\approx$ (QCQPs, ROG)
- ROG implies exactness (objective value and convex hull via the projected SDP)

$$\mathcal{S}(\mathcal{M}) := \{Z \in \mathbb{S}_+^{n+1} : \langle M, Z \rangle \leq 0, \ \forall M \in \mathcal{M}\}$$

Well-known ROG sets:

- Positive semidefinite cone $\mathbb{S}_+^{n+1}$ itself!
- Any single linear matrix inequality (LMI) or equation (LME):

Theorem (S-lemma)

$\mathcal{S}(\{M\})$ for any $M \in \mathbb{S}^{n+1}$ is ROG.

[Fradkov and Yakubovich, 1979, Sturm and Zhang, 2003]

When is $\mathcal{S}(\mathcal{M})$ ROG?

- **Question:** for what $\mathcal{M} \subseteq \mathbb{S}^{n+1}$ is $\mathcal{S}(\mathcal{M})$ ROG?

When is $\mathcal{S}(\mathcal{M})$ ROG?

- **Question:** for what $\mathcal{M} \subseteq \mathbb{S}^{n+1}$ is $\mathcal{S}(\mathcal{M})$ ROG?
- Can we analyze ROG property of $\mathcal{S}(\mathcal{M})$ from ROG property of

$$\mathcal{T}(\mathcal{M}) := \left\{ Z \in \mathbb{S}^{n+1}_{+} : \langle M, Z \rangle = 0,\ \forall M \in \mathcal{M} \right\}?$$

When is $\mathcal{S}(\mathcal{M})$ ROG?

- **Question:** for what $\mathcal{M} \subseteq \mathbb{S}^{n+1}$ is $\mathcal{S}(\mathcal{M})$ ROG?
- Can we analyze ROG property of $\mathcal{S}(\mathcal{M})$ from ROG property of

$$\mathcal{T}(\mathcal{M}) := \left\{ Z \in \mathbb{S}^{n+1}_{+} : \langle M, Z \rangle = 0,\ \forall M \in \mathcal{M} \right\}?$$

- Caveat:

When is $\mathcal{S}(\mathcal{M})$ ROG?

- **Question:** for what $\mathcal{M} \subseteq \mathbb{S}^{n+1}$ is $\mathcal{S}(\mathcal{M})$ ROG?
- Can we analyze ROG property of $\mathcal{S}(\mathcal{M})$ from ROG property of

$$\mathcal{T}(\mathcal{M}) := \left\{ Z \in \mathbb{S}^{n+1}_{+} : \langle M, Z \rangle = 0, \ \forall M \in \mathcal{M} \right\}?$$

- Caveat:
  - When $\mathcal{M}$ is finite, $\mathcal{S}(\mathcal{M})$ can be "lifted" into $\mathcal{T}(\mathcal{M}')$. But, ROG property is not necessarily preserved in such liftings.

Facial structure

$$\mathcal{S}(\mathcal{M}) = \left\{ Z \in \mathbb{S}^{n+1}_+ : \langle M, Z \rangle \leq 0, \ \forall M \in \mathcal{M} \right\}$$
$$\mathcal{T}(\mathcal{M}) = \left\{ Z \in \mathbb{S}^{n+1}_+ : \langle M, Z \rangle = 0, \ \forall M \in \mathcal{M} \right\}$$

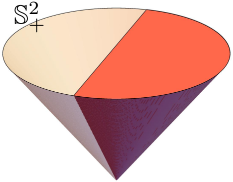
$$\mathcal{S}(\mathcal{M}) = \{Z \in \mathbb{S}_+^{n+1} : \langle M, Z\rangle \le 0,\ \forall M \in \mathcal{M}\}$$
$$\mathcal{T}(\mathcal{M}) = \{Z \in \mathbb{S}_+^{n+1} : \langle M, Z\rangle = 0,\ \forall M \in \mathcal{M}\}$$

Proposition

- $\boxed{\mathcal{S}(\mathcal{M}) \text{ is ROG}} \iff \boxed{\text{every face of } \mathcal{S}(\mathcal{M}) \text{ is ROG}}$
- $\boxed{\mathcal{S}(\mathcal{M}) \text{ is ROG}} \implies \boxed{\mathcal{T}(\mathcal{M}) \text{ is ROG}}$, $\not\Longleftarrow$
- When $\mathcal{M}$ is compact,
  $\boxed{\mathcal{S}(\mathcal{M}) \text{ is ROG}} \iff \boxed{\forall \varnothing \neq \mathcal{M}' \subseteq \mathcal{M},\ \mathcal{S}(\mathcal{M}) \cap \mathcal{T}(\mathcal{M}') \text{ is ROG}}$
- When $\mathcal{M}$ is finite, $\boxed{\forall \mathcal{M}' \subseteq \mathcal{M},\ \mathcal{T}(\mathcal{M}') \text{ is ROG}} \implies \boxed{\mathcal{S}(\mathcal{M}) \text{ is ROG}}$

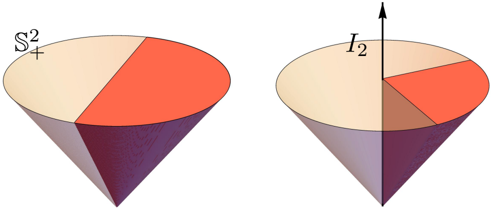
Facial structure → sufficient conditions

- $|\mathcal{M}| = 1$, then both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG (S-lemma)



Facial structure → sufficient conditions

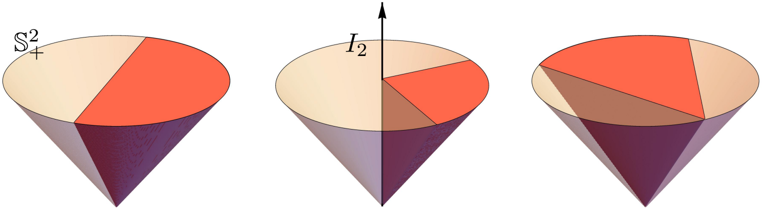
- $|\mathcal{M}| = 1$, then both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG (S-lemma)
- Not necessarily true if $|\mathcal{M}| \geq 2$



Facial structure $\to$ sufficient conditions

- $|\mathcal{M}| = 1$, then both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG (S-lemma)
- Not necessarily true if $|\mathcal{M}| \geq 2$
- Two LMIs $\langle M_1, Z\rangle \leq 0$ and $\langle M_2, Z\rangle \leq 0$ are "non-interacting" when

$$\exists(\alpha_1, \alpha_2) \neq (0,0),\ \alpha_1 M_1 + \alpha_2 M_2 \succeq 0$$

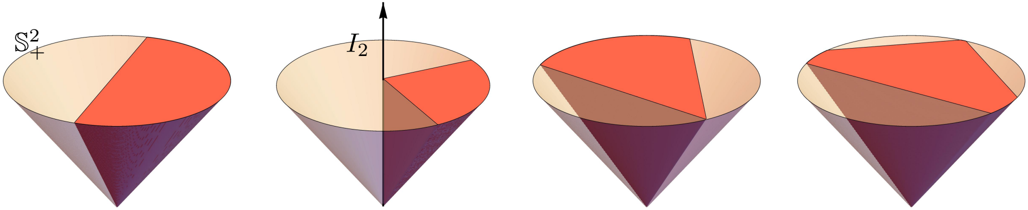


- $|\mathcal{M}| = 1$, then both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG (S-lemma)
- Not necessarily true if $|\mathcal{M}| \geq 2$
- Two LMIs $\langle M_1, Z\rangle \leq 0$ and $\langle M_2, Z\rangle \leq 0$ are "non-interacting" when

$$\exists(\alpha_1, \alpha_2) \neq (0,0),\ \alpha_1 M_1 + \alpha_2 M_2 \succeq 0$$

Lemma

If every pair (M_i, M_j) is "non-interacting" in $\mathcal{M} = \{M_1, \ldots, M_k\}$, then $\mathcal{T}(\mathcal{M})$ and $\mathcal{S}(\mathcal{M})$ are ROG.



The ROG property and solutions to quadratic systems

- Let $\mathcal{E}(Z, \mathcal{M}) := \left\{ z \in \mathbb{R}^{n+1} : \langle M, Z \rangle \leq z^\top M z \leq 0, \ \forall M \in \mathcal{M} \right\}$

The ROG property and solutions to quadratic systems

- Let $\mathcal{E}(Z,\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : \langle M, Z\rangle \le z^\top M z \le 0,\ \forall M \in \mathcal{M}\}$

Proposition

$\mathcal{S}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{S}(\mathcal{M})$, $\operatorname{range}(Z) \cap \mathcal{E}(Z,\mathcal{M}) \neq \{0\}$

- Let $\mathcal{E}(Z, \mathcal{M}) := \{z \in \mathbb{R}^{n+1} : \langle M, Z \rangle \leq z^\top M z \leq 0, \ \forall M \in \mathcal{M}\}$

Proposition

$\mathcal{S}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{S}(\mathcal{M})$, $\mathrm{range}(Z) \cap \mathcal{E}(Z, \mathcal{M}) \neq \{0\}$

- $\mathcal{T}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{T}(\mathcal{M})$, $\mathrm{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$,

  where $\mathcal{N}(\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : z^\top M z = 0, \ \forall M \in \mathcal{M}\}$

The ROG property and solutions to quadratic systems

- Let $\mathcal{E}(Z,\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : \langle M, Z\rangle \le z^\top M z \le 0,\ \forall M \in \mathcal{M}\}$

Proposition

$\mathcal{S}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{S}(\mathcal{M})$, $\operatorname{range}(Z) \cap \mathcal{E}(Z,\mathcal{M}) \ne \{0\}$

- $\mathcal{T}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{T}(\mathcal{M})$, $\operatorname{range}(Z) \cap \mathcal{N}(\mathcal{M}) \ne \{0\}$,

  where $\mathcal{N}(\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : z^\top M z = 0,\ \forall M \in \mathcal{M}\}$
- $\mathcal{N}(\mathcal{M}) \subseteq \mathcal{E}(Z,\mathcal{M})$ for all $Z \in \mathcal{S}(\mathcal{M})$

The ROG property and solutions to quadratic systems

- Let $\mathcal{E}(Z,\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : \langle M, Z\rangle \leq z^\top M z \leq 0,\ \forall M \in \mathcal{M}\}$

Proposition

$\mathcal{S}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{S}(\mathcal{M})$, $\operatorname{range}(Z) \cap \mathcal{E}(Z,\mathcal{M}) \neq \{0\}$

- $\mathcal{T}(\mathcal{M})$ is ROG $\iff$ for all nonzero $Z \in \mathcal{T}(\mathcal{M})$, $\operatorname{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$,

  where $\mathcal{N}(\mathcal{M}) := \{z \in \mathbb{R}^{n+1} : z^\top M z = 0,\ \forall M \in \mathcal{M}\}$
- $\mathcal{N}(\mathcal{M}) \subseteq \mathcal{E}(Z,\mathcal{M})$ for all $Z \in \mathcal{S}(\mathcal{M})$
- Suffices to check these for all Z with $\operatorname{rank}(Z) \geq 2$

Sufficient conditions - II

- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when

Sufficient conditions - II

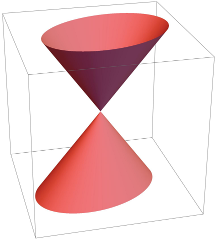
- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\operatorname{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$

Sufficient conditions - II

- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\mathrm{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$
  - E.g., when $\mathcal{N}(\mathcal{M})$ contains a hyperplane $a^{\perp} := \{\xi \in \mathbb{R}^{n+1} : a^{\top}\xi = 0\}$

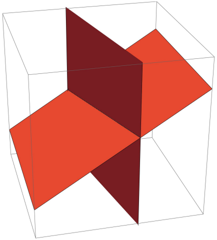
Sufficient conditions - II

- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\text{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$
  - E.g., when $\mathcal{N}(\mathcal{M})$ contains a hyperplane $a^\perp := \{\xi \in \mathbb{R}^{n+1} : a^\top \xi = 0\}$
- $\mathcal{N}(\{M\})$ contains $a^\perp$?

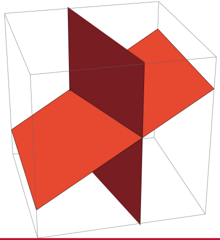


Sufficient conditions - II

- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\text{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$
  - E.g., when $\mathcal{N}(\mathcal{M})$ contains a hyperplane $a^\perp := \{\xi \in \mathbb{R}^{n+1} : a^\top \xi = 0\}$
- $\mathcal{N}(\{M\})$ contains $a^\perp \iff M = ab^\top + ba^\top$ for some $b \in \mathbb{R}^{n+1}$



- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\operatorname{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$
  - E.g., when $\mathcal{N}(\mathcal{M})$ contains a hyperplane $a^\perp := \{\xi \in \mathbb{R}^{n+1} : a^\top \xi = 0\}$
- $\mathcal{N}(\{M\})$ contains $a^\perp \iff M = ab^\top + ba^\top$ for some $b \in \mathbb{R}^{n+1}$

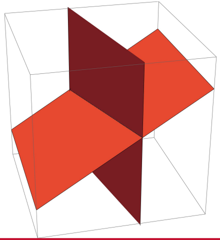


Proposition

Let $a \in \mathbb{R}^{n+1}$, $\mathcal{B} \subseteq \mathbb{R}^{n+1}$ and $\mathcal{M} := \left\{ab^\top + ba^\top : b \in \mathcal{B}\right\}$. Then, both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG.

Sufficient conditions - II

- Both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG when
  - For all $Z \in \mathcal{T}(\mathcal{M})$ with rank ≥ 2, $\operatorname{range}(Z) \cap \mathcal{N}(\mathcal{M}) \neq \{0\}$
  - E.g., when $\mathcal{N}(\mathcal{M})$ contains a hyperplane $a^\perp := \{\xi \in \mathbb{R}^{n+1} : a^\top \xi = 0\}$
- $\mathcal{N}(\{M\})$ contains $a^\perp \iff M = ab^\top + ba^\top$ for some $b \in \mathbb{R}^{n+1}$



Proposition

Let $a \in \mathbb{R}^{n+1}$, $\mathcal{B} \subseteq \mathbb{R}^{n+1}$ and $\mathcal{M} := \left\{ab^\top + ba^\top : b \in \mathcal{B}\right\}$. Then, both $\mathcal{S}(\mathcal{M})$ and $\mathcal{T}(\mathcal{M})$ are ROG.

- For any closed convex cone $\mathbb{K} \subseteq \mathbb{R}^{n+1} \implies \left\{Z \in \mathbb{S}^{n+1}_+ : Za \in \mathbb{K}\right\}$ is ROG.

Summary for two LMIs

$$\mathcal{S}(\{M_1, M_2\}) = \left\{Z \in \mathbb{S}^{n+1}_{+} : \langle M_1, Z\rangle \le 0,\ \langle M_2, Z\rangle \le 0\right\}$$

Summary for two LMIs

$$\mathcal{S}(\{M_1, M_2\}) = \left\{Z \in \mathbb{S}_+^{n+1} : \langle M_1, Z\rangle \leq 0,\ \langle M_2, Z\rangle \leq 0\right\}$$

Theorem

$\mathcal{S}(\{M_1, M_2\})$ is ROG if at least one of the following holds

1. $\exists(\alpha_1, \alpha_2) \neq (0,0)$ s.t. $\alpha_1 M_1 + \alpha_2 M_2 \succeq 0$,
2. $\exists a, b_1, b_2 \in \mathbb{R}^{n+1}$ s.t. $M_1 = ab_1^\top + b_1 a^\top$ and $M_2 = ab_2^\top + b_2 a^\top$.

Summary for two LMIs

$$\mathcal{S}(\{M_1, M_2\}) = \left\{Z \in \mathbb{S}^{n+1}_+ : \langle M_1, Z\rangle \le 0,\ \langle M_2, Z\rangle \le 0\right\}$$

Theorem

$\mathcal{S}(\{M_1, M_2\})$ is ROG if and only if at least one of the following holds

1. $\exists(\alpha_1, \alpha_2) \neq (0,0)$ s.t. $\alpha_1 M_1 + \alpha_2 M_2 \succeq 0$,
2. $\exists a, b_1, b_2 \in \mathbb{R}^{n+1}$ s.t. $M_1 = ab_1^\top + b_1 a^\top$ and $M_2 = ab_2^\top + b_2 a^\top$.

Summary for two LMIs

$$\mathcal{S}(\{M_1, M_2\}) = \{Z \in \mathbb{S}_+^{n+1} : \langle M_1, Z\rangle \le 0, \langle M_2, Z\rangle \le 0\}$$

Theorem

$\mathcal{S}(\{M_1, M_2\})$ is ROG if and only if at least one of the following holds

1. $\exists(\alpha_1, \alpha_2) \neq (0,0)$ s.t. $\alpha_1 M_1 + \alpha_2 M_2 \succeq 0$,
2. $\exists a, b_1, b_2 \in \mathbb{R}^{n+1}$ s.t. $M_1 = ab_1^\top + b_1 a^\top$ and $M_2 = ab_2^\top + b_2 a^\top$.

- Complete characterization of ROG property in the case of two LMIs (c.f., S-lemma).

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

$$\text{(TRS):} \inf_{x \in \mathbb{R}^n} \left\{ x^\top A_{\mathsf{obj}} x + 2 b_{\mathsf{obj}}^\top x : \|x\|_2 \leq 1 \right\}$$

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

(TRS): $\inf_{x \in \mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2 b_{\text{obj}}^\top x : \|x\|_2 \leq 1 \right\}$

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

$$\text{(TRS):} \quad \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \leq 1 \right\}$$

$$= \inf_{x\in\mathbb{R}^n} \left\{ \left\langle M_{\text{obj}}, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix} \right\rangle : \operatorname{tr}(xx^\top) - 1 \leq 0 \right\}$$

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

(TRS): $\inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \leq 1\right\}$

$$= \inf_{x\in\mathbb{R}^n} \left\{\left\langle M_{\text{obj}}, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix}\right\rangle : \operatorname{tr}(xx^\top) - 1 \leq 0\right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}} \left\{\langle M_{\text{obj}}, zz^\top\rangle : z_{n+1} = 1, \ \langle L, zz^\top\rangle \leq 0\right\}, \quad \text{where } L := \operatorname{Diag}(1,\ldots,1,-1)$$

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

$$
\begin{aligned}
\text{(TRS):} \ \inf_{x\in\mathbb{R}^n} &\left\{x^\top A_{\mathsf{obj}} x + 2 b_{\mathsf{obj}}^\top x : \ \|x\|_2 \le 1\right\} \\
= \inf_{x\in\mathbb{R}^n} &\left\{\left\langle M_{\mathsf{obj}}, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix}\right\rangle : \ \mathrm{tr}(xx^\top) - 1 \le 0\right\} \\
= \inf_{z\in\mathbb{R}^{n+1}} &\left\{\langle M_{\mathsf{obj}}, zz^\top\rangle : \ z_{n+1} = 1, \ \langle L, zz^\top\rangle \le 0\right\}, \quad \text{where } L := \mathrm{Diag}(1,\dots,1,-1) \\
= \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} &\left\{\langle M_{\mathsf{obj}}, Z\rangle : \ Z_{n+1,n+1} = 1, \ \langle L, Z\rangle \le 0, \ Z = zz^\top\right\}
\end{aligned}
$$

What other sets are ROG?

- ROG property is extensively studied in the context of **Trust-region subproblem (TRS)** and its variants
  [Sturm and Zhang, 2003], [Burer, 2015] and references therein, [Yang et al., 2018]

$$
\begin{aligned}
\text{(TRS):}\ & \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\mathsf{obj}} x + 2b_{\mathsf{obj}}^\top x : \|x\|_2 \le 1 \right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{ \left\langle M_{\mathsf{obj}}, \begin{pmatrix} xx^\top & x \\ x^\top & 1 \end{pmatrix} \right\rangle : \operatorname{tr}(xx^\top) - 1 \le 0 \right\} \\
&= \inf_{z\in\mathbb{R}^{n+1}} \left\{ \langle M_{\mathsf{obj}}, zz^\top \rangle : z_{n+1} = 1,\ \langle L, zz^\top \rangle \le 0 \right\}, \quad \text{where } L := \operatorname{Diag}(1,\dots,1,-1) \\
&= \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\mathsf{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ Z = zz^\top \right\} \\
&\ge \inf_{Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\mathsf{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ Z \succeq 0 \right\}
\end{aligned}
$$

Trust region subproblem

$$\text{(TRS):} \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{\langle M_{\text{obj}}, Z\rangle : Z_{n+1,n+1} = 1,\ \langle L, Z\rangle \le 0,\ Z = zz^\top\right\}$$

$$\ge \inf_{Z\in\mathbb{S}^{n+1}} \left\{\langle M_{\text{obj}}, Z\rangle : Z_{n+1,n+1} = 1,\ \langle L, Z\rangle \le 0,\ Z \succeq 0\right\},$$

where $L := \text{Diag}(1, \dots, 1, -1)$.

Trust region subproblem

$$\text{(TRS):} \quad \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z\rangle : Z_{n+1,n+1} = 1,\ \langle L, Z\rangle \le 0,\ Z = zz^\top \right\}$$
$$\ge \inf_{Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z\rangle : Z_{n+1,n+1} = 1,\ \langle L, Z\rangle \le 0,\ Z \succeq 0 \right\},$$

where $L := \mathrm{Diag}(1, \ldots, 1, -1)$.

Theorem ([Sturm and Zhang, 2003])

$$\mathrm{cl\,conv}\left\{ zz^\top : \left\langle L, zz^\top \right\rangle \le 0 \right\} = \left\{ Z \in \mathbb{S}^{n+1}_+ : \langle L, Z\rangle \le 0 \right\} = \mathcal{S}(\{L\}).$$

(recall $\mathcal{S}(\{L\})$ is ROG).
Furthermore,

$$\mathrm{cl\,conv}\left\{ (x, xx^\top) : \|x\|_2 \le 1 \right\}$$
$$= \left\{ (x, X) : \exists Z = \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix},\ Z_{n+1,n+1} = 1,\ \langle L, Z\rangle \le 0,\ Z \succeq 0 \right\}.$$

Extended TRS:

$$\text{(e-TRS):}\ \inf_{x\in\mathbb{R}^n}\left\{ x^\top A_{\mathsf{obj}}x + 2b_{\mathsf{obj}}^\top x :\ \|x\|_2 \le 1,\ c_i^\top \underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{:= z} \ge 0,\ \forall i \in [m] \right\}$$

Extended TRS:

(e-TRS): $$\inf_{x\in\mathbb{R}^n}\left\{ x^\top A_{\text{obj}}x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ c_i^\top \underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{:=z} \ge 0, \forall i\in[m] \right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}}\left\{ \langle M_{\text{obj}}, zz^\top\rangle : z_{n+1}=1,\ \|\overbrace{(z_1,\ldots,z_n)}^{=x}\|_2 \le 1,\ c_i^\top z \ge 0, \forall i\in[m] \right\}$$

Extended TRS:

$$\text{(e-TRS):} \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ c_i^\top \underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{:= z} \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1,\ \| \overbrace{(z_1, \ldots, z_n)}^{=x} \|_2 \le 1,\ c_i^\top z \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ c_i^\top z \ge 0,\ \forall i \in [m],\ Z = zz^\top \right\}$$

Extended TRS:

$$\text{(e-TRS):} \quad \inf_{x \in \mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2 b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ c_i^\top \underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{:= z} \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z \in \mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1,\ \| \overbrace{(z_1, \ldots, z_n)}^{=x} \|_2 \le 1,\ c_i^\top z \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z \in \mathbb{R}^{n+1}, Z \in \mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ c_i^\top z \ge 0,\ \forall i \in [m],\ Z = zz^\top \right\}$$

$$?=? \inf_{z \in \mathbb{R}^{n+1}, Z \in \mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ c_i^\top z \ge 0,\ \forall i \in [m],\ Z \succeq zz^\top \right\}$$

Extended TRS:

$$\text{(e-TRS):} \quad \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ c_i^\top \underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{:= z} \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1,\ \| \overbrace{(z_1,\ldots,z_n)}^{=x} \|_2 \le 1,\ c_i^\top z \ge 0,\ \forall i \in [m] \right\}$$

$$= \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ c_i^\top z \ge 0,\ \forall i \in [m],\ Z = zz^\top \right\}$$

$$? =? \inf_{z\in\mathbb{R}^{n+1}, Z\in\mathbb{S}^{n+1}} \left\{ \langle M_{\text{obj}}, Z \rangle : Z_{n+1,n+1} = 1,\ \langle L, Z \rangle \le 0,\ c_i^\top z \ge 0,\ \forall i \in [m],\ Z \succeq zz^\top \right\}$$

Not really! We need a stronger relaxation.

How to strengthen the standard SDP relaxation?

$$\text{(e-TRS)} = \inf_{z \in \mathbb{R}^{n+1}} \left\{ \left\langle M_{\text{obj}}, zz^\top \right\rangle : \; z_{n+1} = 1, \; \|(z_1, \ldots, z_n)\|_2 \leq 1, \; c_i^\top z \geq 0, \; \forall i \in [m] \right\}$$

How to strengthen the standard SDP relaxation?

$$\text{(e-TRS)} = \inf_{z \in \mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1, \ \|(z_1, \ldots, z_n)\|_2 \le 1, \ c_i^\top z \ge 0, \forall i \in [m] \right\}$$

- **Linear RLT:** $\left. \begin{array}{l} c_1^\top z \ge 0 \\ c_2^\top z \ge 0 \end{array} \right\} \implies c_1^\top z z^\top c_2 \ge 0 \implies c_1^\top Z c_2 \ge 0$

How to strengthen the standard SDP relaxation?

$$\text{(e-TRS)} = \inf_{z \in \mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1, \ \|(z_1, \ldots, z_n)\|_2 \le 1, \ c_i^\top z \ge 0, \ \forall i \in [m] \right\}$$

- **Linear RLT:** $\left. \begin{array}{l} c_1^\top z \ge 0 \\ c_2^\top z \ge 0 \end{array} \right\} \implies c_1^\top zz^\top c_2 \ge 0 \implies c_1^\top Z c_2 \ge 0$

- $\mathbb{L}^{n+1} := \left\{ z \in \mathbb{R}^{n+1} : z_{n+1} \ge \|(z_1, \ldots, z_n)\|_2 \right\}$ denote the SOC in $\mathbb{R}^{n+1}$. Then,

$$\text{(e-TRS)} = \inf_{z \in \mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1, \ z \in \mathbb{L}^{n+1}, \ c_i^\top z \ge 0, \ \forall i \in [m] \right\}$$

How to strengthen the standard SDP relaxation?

$$\text{(e-TRS)} = \inf_{z\in\mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1,\ \|(z_1,\ldots,z_n)\|_2 \le 1,\ c_i^\top z \ge 0,\ \forall i \in [m] \right\}$$

- **Linear RLT:** $\left.\begin{array}{l} c_1^\top z \ge 0 \\ c_2^\top z \ge 0 \end{array}\right\} \implies c_1^\top zz^\top c_2 \ge 0 \implies c_1^\top Z c_2 \ge 0$
- $\mathbb{L}^{n+1} := \left\{ z \in \mathbb{R}^{n+1} : z_{n+1} \ge \|(z_1,\ldots,z_n)\|_2 \right\}$ denote the SOC in $\mathbb{R}^{n+1}$. Then,

$$\text{(e-TRS)} = \inf_{z\in\mathbb{R}^{n+1}} \left\{ \langle M_{\text{obj}}, zz^\top \rangle : z_{n+1} = 1,\ z \in \mathbb{L}^{n+1},\ c_i^\top z \ge 0,\ \forall i \in [m] \right\}$$

- **SOC RLT:** $\left.\begin{array}{l} c_1^\top z \ge 0 \\ z \in \mathbb{L}^{n+1} \end{array}\right\} \implies zz^\top c_1 \in \mathbb{L}^{n+1} \implies Zc_1 \in \mathbb{L}^{n+1}$

ROG characterization of (e-TRS)

$$\text{(e-TRS)} = \inf_{z\in\mathbb{R}^{n+1}} \left\{ \left\langle M_{\text{obj}}, zz^\top \right\rangle : z_{n+1} = 1,\ z \in \mathbb{L}^{n+1},\ c_i^\top z \geq 0,\ \forall i \in [m] \right\}$$

Theorem ([Burer and Anstreicher, 2013, Burer, 2015] and references therein)

Suppose $c_i^\top z \geq 0$ for $i \in [m]$ are s.t. whenever $\bar{z}$ is feasible to (e-TRS) and $c_\ell^\top \bar{z} = 0$ for some $\ell \in [m]$, then $c_j^\top \bar{z} \geq 0$ for all $j \in [m]$. Then, the set

$$\left\{ Z \in \mathbb{S}_+^{n+1} :\ \langle L, Z\rangle \leq 0,\ Zc_i \in \mathbb{L}^{n+1}, \forall i \in [m],\ c_i^\top Z c_j \geq 0, \forall i,j \in [m] \right\}$$

is ROG and it is equal to $\operatorname{conv}\left\{ zz^\top :\ z \in \mathbb{L}^{n+1},\ c_i^\top z \geq 0, \forall i \in [m] \right\}$.

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$\text{(tb-TRS)} \;=\; \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\mathsf{obj}} x + 2b_{\mathsf{obj}}^\top x :\ \|x\|_2 \le 1,\ \|x-c\|_2 \le \tilde{r} \right\}$$

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$\text{(tb-TRS)} \;=\; \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\mathsf{obj}} x + 2 b_{\mathsf{obj}}^\top x :\; \|x\|_2 \le 1,\; \|x-c\|_2 \le \tilde{r} \right\}$$

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$
\begin{aligned}
\text{(tb-TRS)} &= \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ \|x-c\|_2 \le \tilde{r} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le 1,\ \|x\|_2^2 - 2c^\top x + \|c\|_2^2 \le \tilde{r} \right\}
\end{aligned}
$$

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$
\begin{aligned}
\text{(tb-TRS)} &= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ \|x-c\|_2 \le \tilde{r}\right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le 1,\ \|x\|_2^2 - 2c^\top x + \|c\|_2^2 \le \tilde{r}\right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le \min\left\{1,\ 2c^\top x \underbrace{-\|c\|_2^2 + \tilde{r}}_{:=r}\right\}\right\}
\end{aligned}
$$

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$
\begin{aligned}
\text{(tb-TRS)} &= \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ \|x-c\|_2 \le \tilde{r} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le 1,\ \|x\|_2^2 - 2c^\top x + \|c\|_2^2 \le \tilde{r} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le \min\left\{ 1,\ 2c^\top x \underbrace{-\|c\|_2^2 + \tilde{r}}_{:=r} \right\} \right\} \\
&= \inf_{x\in\mathbb{R}^n, t\in\mathbb{R}} \left\{ x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le t,\ t = \min\{1,\ 2c^\top x + r\} \right\}
\end{aligned}
$$

Intersection of two Euclidean balls

- ROG characterization of the intersection of two Euclidean balls is studied in [Kelly et al., 2022, Burer, 2023]

$$
\begin{aligned}
\text{(tb-TRS)} &= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2 \le 1,\ \|x-c\|_2 \le \tilde{r}\right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le 1,\ \|x\|_2^2 - 2c^\top x + \|c\|_2^2 \le \tilde{r}\right\} \\
&= \inf_{x\in\mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le \min\left\{1,\ 2c^\top x \underbrace{-\|c\|_2^2 + \tilde{r}}_{:=r}\right\}\right\} \\
&= \inf_{x\in\mathbb{R}^n, t\in\mathbb{R}} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x : \|x\|_2^2 \le t,\ t = \min\left\{1,\ 2c^\top x + r\right\}\right\}
\end{aligned}
$$

Theorem ([Burer, 2023], informal)

Consider (tb-TRS) in the (x, t) space. Then, its strengthened $\mathcal{S}(\mathcal{M})$ set which contains 1 LMI from the norm constraint, 2 SOC-RLT constraints, and 1 LME from the linear RLT, is ROG.

What about nonconvex quadratics?

Theorem ([Yang et al., 2018], informal)

Consider the intersection of

- "ball": $\|x\|_2 \leq 1$
- "cuts": $Cx \geq d$
- "holes": $x^\top A_i x + 2b_i^\top x + c_i \geq 0$, where each $A_i \succ 0$, for all $i \in [k]$.

If none of the cuts and holes touch each other, then the strengthened $\mathcal{S}(\mathcal{M})$ set which contains

- 1 LMI from the norm constraint,
- all SOC-RLT and linear RLT constraints from the cuts, and
- all LMIs $\langle A_i, X \rangle + 2b_i^\top x + c_i \geq 0$ from the holes,

is ROG.

Open questions

Here is a deceivingly simple looking open question:

Open questions

Here is a deceivingly simple looking open question:

Open question

Given $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$, what is the set $\mathcal{S}(\mathcal{M})$ that gives the ROG characterization of

$$\{x \in \mathbb{R}^n : \|x\|_2 \leq 1,\ \|Ax - c\|_2 \leq \tilde{r}\}?$$

- Kronecker RLT constraints? [Anstreicher, 2017]

Open questions

Here is a deceivingly simple looking open question:

Open question

Given $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$, what is the set $\mathcal{S}(\mathcal{M})$ that gives the ROG characterization of

$$\{x \in \mathbb{R}^n : \|x\|_2 \leq 1,\ \|Ax - c\|_2 \leq \tilde{r}\}?$$

- Kronecker RLT constraints? [Anstreicher, 2017]
- What about cuts? holes?

Open questions

Here is a deceivingly simple looking open question:

Open question

Given $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$, what is the set $\mathcal{S}(\mathcal{M})$ that gives the ROG characterization of

$$\{x \in \mathbb{R}^n : \|x\|_2 \le 1,\ \|Ax - c\|_2 \le \tilde{r}\}?$$

- Kronecker RLT constraints? [Anstreicher, 2017]
- What about cuts? holes?
- Fejes-Tóth conjecture (1964) (one of Kurt Anstreicher's favorite open problems that can significantly simplify the proof of Kepler conjecture)

Lemma

Suppose

- $\mathcal{M} = \bigcup_{\alpha \in \mathcal{F}} \mathcal{M}_\alpha$ for some family of matrices $\{\mathcal{M}_\alpha\}_{\alpha \in \mathcal{F}}$, and
- $\mathcal{S}(\mathcal{M}_\alpha)$ is ROG for every $\alpha \in \mathcal{F}$.

Then, $\mathcal{S}(\mathcal{M})$ is ROG iff $\operatorname{extr}(\mathcal{S}(\mathcal{M})) \subseteq \bigcap_{\alpha \in \mathcal{F}} \operatorname{extr}(\mathcal{S}(\mathcal{M}_\alpha))$.

Simple ROG preserving operations

Lemma

Suppose

- $\mathcal{M} = \bigcup_{\alpha\in\mathcal{F}} \mathcal{M}_\alpha$ for some family of matrices $\{\mathcal{M}_\alpha\}_{\alpha\in\mathcal{F}}$, and
- $\mathcal{S}(\mathcal{M}_\alpha)$ is ROG for every $\alpha \in \mathcal{F}$.

Then, $\mathcal{S}(\mathcal{M})$ is ROG iff $\operatorname{extr}(\mathcal{S}(\mathcal{M})) \subseteq \bigcap_{\alpha\in\mathcal{F}} \operatorname{extr}(\mathcal{S}(\mathcal{M}_\alpha))$.

Lemma

Suppose

- $\mathcal{M} = \bigcup_{i=1}^{k} \mathcal{M}_i$, i.e., a *finite* union of compact sets, and
- the following "non-interacting" assumption holds:

  for all $0 \neq Z \in \mathbb{S}^{n+1}_{+}$ *and* $i \in [k]$, *if* $\langle M_i, Z\rangle = 0$ *for some* $M_i \in \mathcal{M}_i$, *then* $\langle M, Z\rangle < 0$ *for all* $M \in \mathcal{M} \setminus \mathcal{M}_i$.

Then, $\mathcal{S}(\mathcal{M})$ is ROG iff $\mathcal{S}(\mathcal{M}_i)$ is ROG for all $i \in [k]$.

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, . . .)

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, . . .)
- Sufficient conditions based on “non-interacting” or “solutions to quadratic systems”

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, . . .)
- Sufficient conditions based on “non-interacting” or “solutions to quadratic systems”
- Complete characterization of when $\mathcal{S}(\{M_1, M_2\})$ is ROG

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, ...)
- Sufficient conditions based on "non-interacting" or "solutions to quadratic systems"
- Complete characterization of when $\mathcal{S}(\{M_1, M_2\})$ is ROG
- Toolkit for ROG property

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, ...)
- Sufficient conditions based on “non-interacting” or “solutions to quadratic systems”
- Complete characterization of when $\mathcal{S}(\{M_1, M_2\})$ is ROG
- Toolkit for ROG property
- Many ROG sets arising from variants of TRS...

Summary of the ROG property

- ROG property implies (closed) convex hull exactness for any* objective function
- Other applications (e.g., minimizing ratios of quadratic functions, PSD matrix completion, . . .)
- Sufficient conditions based on “non-interacting” or “solutions to quadratic systems”
- Complete characterization of when $\mathcal{S}(\{M_1, M_2\})$ is ROG
- Toolkit for ROG property
- Many ROG sets arising from variants of TRS. . .
- Many more open questions about ROG characterizations of sets defined by quadratics. . .

Exactness in the original space

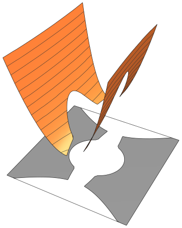
References:

Wang, A. L. and K.-K., F. (2022c). On the tightness of SDP relaxations of QCQPs. *Math. Program.*, 193:33–73

Wang, A. L. and K.-K., F. (2020). A geometric view of SDP exactness in QCQPs and its applications. *arXiv preprint*, 2011.07155

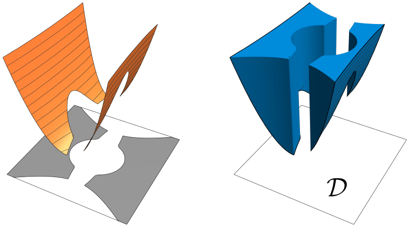
The QCQP epigraph

- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$



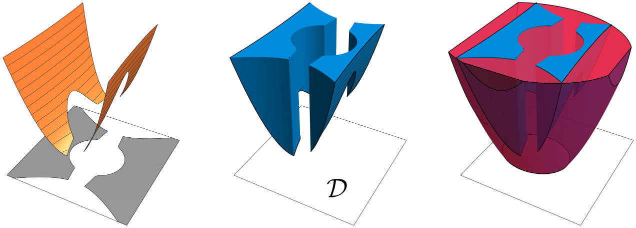
The QCQP epigraph

- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$



The QCQP epigraph

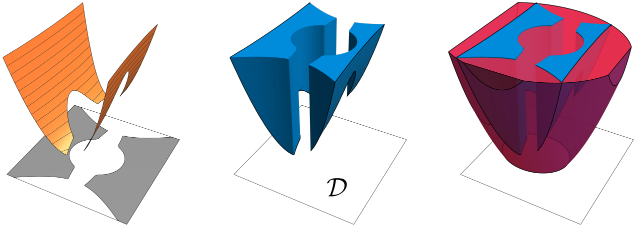
- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \leq t \\ q_i(x) \leq 0,\ \forall i \in [m] \end{array} \right\}$



- How can we derive convex relaxations of $\mathcal{D}$?

The QCQP epigraph

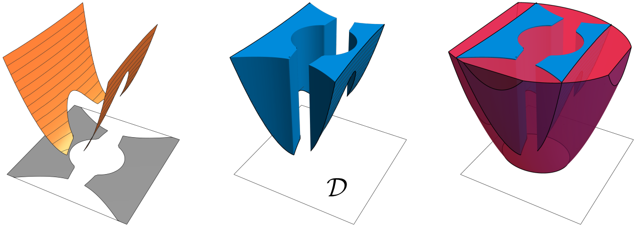
- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$



- How can we derive ~~convex~~ relaxations of $\mathcal{D}$? **Lagrangian aggregation!**

The QCQP epigraph

- QCQP epigraph $\quad \mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$

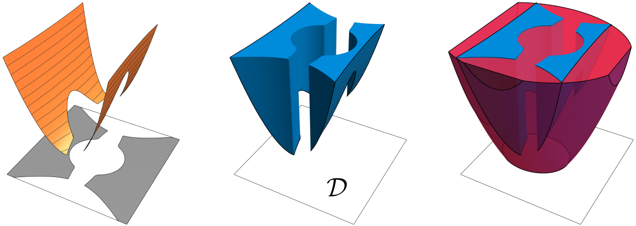


- How can we derive ~~convex~~ relaxations of $\mathcal{D}$? **Lagrangian aggregation!**
- For any $\gamma \in \mathbb{R}^m_+$, the aggregated inequality $\quad q_{\text{obj}}(x) + \sum_{i=1}^m \gamma_i q_i(x) \le t$

  is valid for all $(x,t) \in \mathcal{D}$.

The QCQP epigraph

- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\mathsf{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$



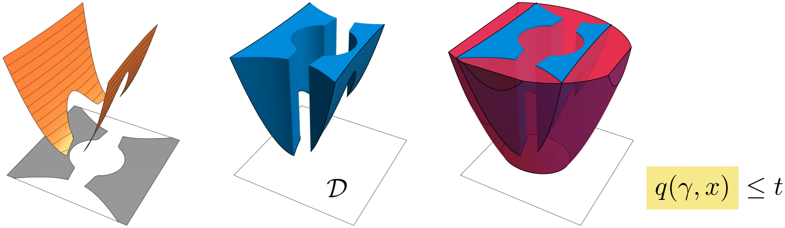
- How can we derive ~~convex~~ relaxations of $\mathcal{D}$? **Lagrangian aggregation!**
- For any $\gamma \in \mathbb{R}^m_+$, the aggregated inequality

$$\underbrace{q_{\mathsf{obj}}(x) + \sum_{i=1}^m \gamma_i q_i(x)}_{=:\, q(\gamma,x)} \le t$$

is valid for all $(x,t) \in \mathcal{D}$.

The QCQP epigraph

- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \le t \\ q_i(x) \le 0,\ \forall i \in [m] \end{array} \right\}$



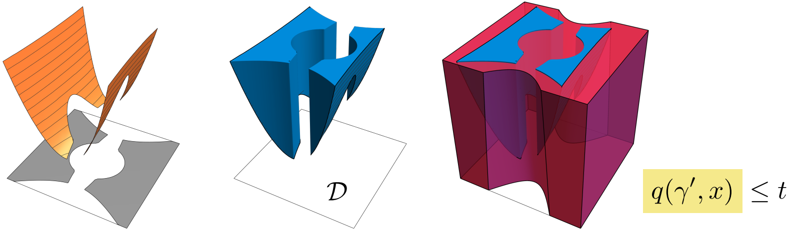
- How can we derive ~~convex~~ relaxations of $\mathcal{D}$? **Lagrangian aggregation!**
- For any $\gamma \in \mathbb{R}^m_+$, the aggregated inequality

$$\underbrace{q_{\text{obj}}(x) + \sum_{i=1}^{m} \gamma_i q_i(x)}_{=: q(\gamma, x)} \le t$$

is valid for all $(x,t) \in \mathcal{D}$.

The QCQP epigraph

- QCQP epigraph $\mathcal{D} := \left\{ (x,t) \in \mathbb{R}^{n+1} : \begin{array}{l} q_{\text{obj}}(x) \leq t \\ q_i(x) \leq 0,\ \forall i \in [m] \end{array} \right\}$



- How can we derive ~~convex~~ relaxations of $\mathcal{D}$? **Lagrangian aggregation!**
- For any $\gamma \in \mathbb{R}^m_+$, the aggregated inequality

$$\underbrace{q_{\text{obj}}(x) + \sum_{i=1}^{m} \gamma_i q_i(x)}_{=: q(\gamma, x)} \leq t$$

  is valid for all $(x,t) \in \mathcal{D}$.

SDP relaxation is Lagrangian aggregation

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i \in [m]} \gamma_i^* A_i \succ 0$.

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i\in[m]} \gamma_i^* A_i \succ 0$.

$$\begin{aligned}
\text{Opt} &= \inf_{x\in\mathbb{R}^n} \{q_{\text{obj}}(x) : q_i(x) \le 0, \ \forall i \in [m]\} \\
&= \inf_{x\in\mathbb{R}^n} \{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : x^\top A_i x + 2b_i^\top x + c_i \le 0, \forall i \in [m]\} \\
&\ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\text{obj}}, X\rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0, \ \forall i \in [m] \end{array} \right\}
\end{aligned}$$

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i\in[m]} \gamma_i^* A_i \succ 0$.

$$\begin{aligned}
\text{Opt} &= \inf_{x\in\mathbb{R}^n} \{q_{\text{obj}}(x) : q_i(x) \le 0,\ \forall i \in [m]\} \\
&= \inf_{x\in\mathbb{R}^n} \{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : x^\top A_i x + 2b_i^\top x + c_i \le 0, \forall i \in [m]\} \\
&\ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\text{obj}}, X\rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \in [m] \end{array} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \inf_{X\in\mathbb{S}^n} \dots
\end{aligned}$$

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i\in[m]} \gamma_i^* A_i \succ 0$.

$$\begin{aligned}
\text{Opt} &= \inf_{x\in\mathbb{R}^n} \{q_{\text{obj}}(x) : \; q_i(x) \le 0, \; \forall i \in [m]\} \\
&= \inf_{x\in\mathbb{R}^n} \{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \; x^\top A_i x + 2b_i^\top x + c_i \le 0, \forall i \in [m]\} \\
&\ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\text{obj}}, X\rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0, \; \forall i \in [m] \end{array} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \inf_{X\in\mathbb{S}^n} \dots \\
&= \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) = \text{Opt}_{\text{SDP}}
\end{aligned}$$

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i\in[m]} \gamma_i^* A_i \succ 0$.

$$\begin{aligned}
\text{Opt} &= \inf_{x\in\mathbb{R}^n} \{q_{\text{obj}}(x) : q_i(x) \le 0,\ \forall i \in [m]\} \\
&= \inf_{x\in\mathbb{R}^n} \{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : x^\top A_i x + 2b_i^\top x + c_i \le 0, \forall i \in [m]\} \\
&\ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\text{obj}}, X\rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \in [m] \end{array} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \inf_{X\in\mathbb{S}^n} \dots \\
&= \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) = \text{Opt}_{\text{SDP}}
\end{aligned}$$

where $\Gamma_1 := \left\{ \gamma \in \mathbb{R}^m_+ : A_{\text{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\}$

Related: Fujie and Kojima [1997]

SDP relaxation is Lagrangian aggregation

Assumption

Dual strict feasibility holds, i.e., $\exists \gamma^* \in \mathbb{R}^m_+$ s.t. $A_{\text{obj}} + \sum_{i\in[m]} \gamma_i^* A_i \succ 0$.

$$\begin{aligned}
\text{Opt} &= \inf_{x\in\mathbb{R}^n} \{q_{\text{obj}}(x) : q_i(x) \le 0,\ \forall i \in [m]\} \\
&= \inf_{x\in\mathbb{R}^n} \{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : x^\top A_i x + 2b_i^\top x + c_i \le 0, \forall i \in [m]\} \\
&\ge \inf_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \left\{ \langle A_{\text{obj}}, X\rangle + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \begin{array}{l} X - xx^\top \succeq 0 \\ \langle A_i, X\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \in [m] \end{array} \right\} \\
&= \inf_{x\in\mathbb{R}^n} \inf_{X\in\mathbb{S}^n} \dots \\
&= \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) = \text{Opt}_{\text{SDP}}
\end{aligned}$$

where $\Gamma_1 := \left\{ \gamma \in \mathbb{R}^m_+ : A_{\text{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\} = \{\gamma \in \mathbb{R}^m_+ : q(\gamma, x) \text{ is convex in } x\}$

Related: Fujie and Kojima [1997]

Revisiting the SDP relaxation

- $\boxed{\mathrm{Opt}_{\mathsf{SDP}}} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_1} q(\gamma, x)$ where $\boxed{\Gamma_1} = \left\{ \gamma \in \mathbb{R}^m_+ : \; A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\}$

Revisiting the SDP relaxation

- $$\text{Opt}_{\text{SDP}} = \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) \text{ where } \Gamma_1 = \left\{\gamma \in \mathbb{R}^m_+ :\ A_{\text{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$$
- Define $\Gamma := \left\{(\gamma_{\text{obj}}, \gamma) \in \mathbb{R}^{1+m}_+ :\ \gamma_{\text{obj}} A_{\text{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$, i.e., $\Gamma = \text{cl}\,\text{cone}\,\{(1,\gamma) : \gamma \in \Gamma_1\}$

Revisiting the SDP relaxation

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma,x) \text{ where } \Gamma_1 = \left\{\gamma \in \mathbb{R}^m_+ :\ A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$$
- Define $\Gamma := \left\{(\gamma_{\mathsf{obj}},\gamma) \in \mathbb{R}^{1+m}_+ :\ \gamma_{\mathsf{obj}} A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$, i.e., $\Gamma = \mathrm{cl\,cone}\,\{(1,\gamma) : \gamma \in \Gamma_1\}$

  Then, $\mathcal{D}_{\mathsf{SDP}} := \bigcap_{(\gamma_{\mathsf{obj}},\gamma)\in\Gamma} \left\{(x,t) :\ \gamma_{\mathsf{obj}}(q_{\mathsf{obj}}(x) - t) + \sum_{i=1}^m \gamma_i q_i(x) \le 0\right\}$

Revisiting the SDP relaxation

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_1} q(\gamma, x) \text{ where } \Gamma_1 = \left\{ \gamma \in \mathbb{R}^m_+ : A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\}$$

- Define $\Gamma := \left\{ (\gamma_{\mathsf{obj}}, \gamma) \in \mathbb{R}^{1+m}_+ : \gamma_{\mathsf{obj}} A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\}$, i.e., $\Gamma = \mathrm{cl\,cone}\,\{(1, \gamma) : \gamma \in \Gamma_1\}$

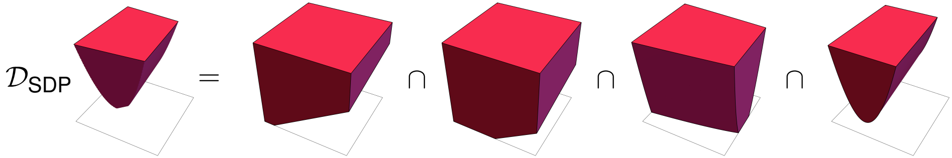
  Then, $\mathcal{D}_{\mathsf{SDP}} := \bigcap_{(\gamma_{\mathsf{obj}}, \gamma) \in \Gamma} \left\{ (x, t) : \gamma_{\mathsf{obj}} (q_{\mathsf{obj}}(x) - t) + \sum_{i=1}^m \gamma_i q_i(x) \leq 0 \right\}$

- Projected SDP relaxation = impose all *convex* aggregated inequalities!

Revisiting the SDP relaxation

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma,x) \text{ where } \Gamma_1 = \left\{\gamma \in \mathbb{R}^m_+ : A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$$
- Define $\Gamma := \left\{(\gamma_{\mathsf{obj}},\gamma) \in \mathbb{R}^{1+m}_+ : \gamma_{\mathsf{obj}} A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i A_i \succeq 0\right\}$, i.e., $\Gamma = \mathrm{cl\,cone}\left\{(1,\gamma) : \gamma \in \Gamma_1\right\}$

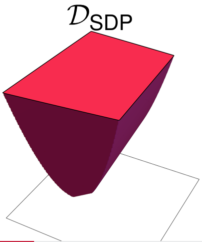
  Then, $$\mathcal{D}_{\mathsf{SDP}} := \bigcap_{(\gamma_{\mathsf{obj}},\gamma)\in\Gamma} \left\{(x,t) : \gamma_{\mathsf{obj}}(q_{\mathsf{obj}}(x) - t) + \sum_{i=1}^m \gamma_i q_i(x) \le 0\right\}$$
- Projected SDP relaxation = impose all *convex* aggregated inequalities!



Rewriting SDP in terms of Γ

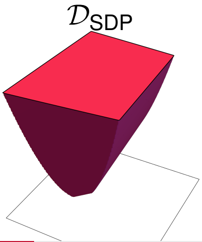
Lemma

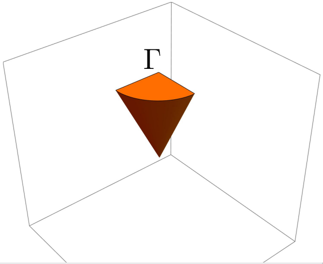
$$\mathcal{D}_{\mathsf{SDP}} = \left\{ (x,t) : \left\langle \begin{pmatrix} \gamma_{\mathsf{obj}} \\ \gamma \end{pmatrix}, \begin{pmatrix} q_{\mathsf{obj}}(x) - t \\ q(x) \end{pmatrix} \right\rangle \le 0, \ \ \forall \begin{pmatrix} \gamma_{\mathsf{obj}} \\ \gamma \end{pmatrix} \in \Gamma \right\}$$



Lemma

$$\mathcal{D}_{\mathsf{SDP}} = \left\{ (x,t) : \left\langle \begin{pmatrix} \gamma_{\mathsf{obj}} \\ \gamma \end{pmatrix}, \begin{pmatrix} q_{\mathsf{obj}}(x) - t \\ q(x) \end{pmatrix} \right\rangle \leq 0, \;\; \forall \begin{pmatrix} \gamma_{\mathsf{obj}} \\ \gamma \end{pmatrix} \in \Gamma \right\}$$



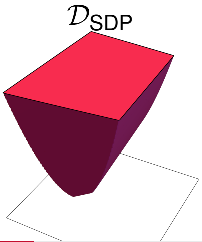


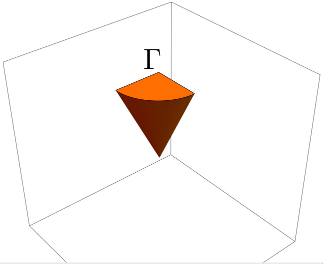
Rewriting SDP in terms of Γ

Lemma

$$\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \left\langle \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix}, \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \right\rangle \le 0, \ \forall \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix} \in \Gamma \right\}$$

For a cone K, the polar cone $K^\circ := \{\xi : \langle \xi, \zeta \rangle \le 0, \forall \zeta \in K\}$



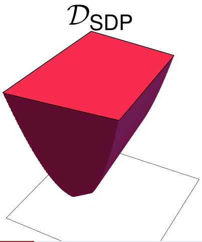


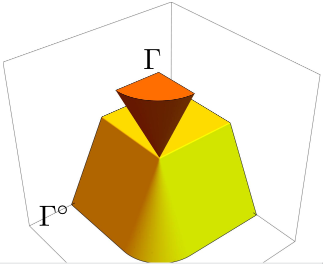
Lemma

$$\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \left\langle \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix}, \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \right\rangle \leq 0, \ \forall \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix} \in \Gamma \right\}$$
$$= \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\},$$

where Γ° is the polar cone of Γ.

For a cone K, the polar cone $K^\circ := \{\xi : \langle \xi, \zeta \rangle \leq 0, \forall \zeta \in K\}$



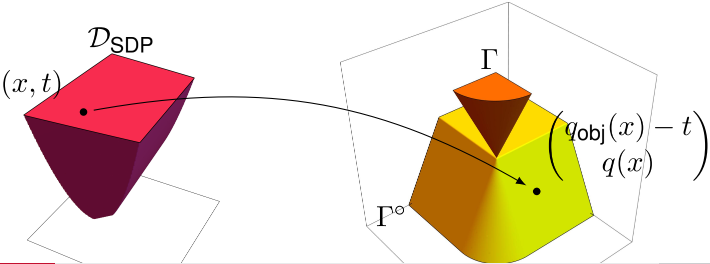


Lemma

$$\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \left\langle \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix}, \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \right\rangle \le 0, \ \forall \begin{pmatrix} \gamma_{\text{obj}} \\ \gamma \end{pmatrix} \in \Gamma \right\}$$
$$= \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\},$$

where Γ° is the polar cone of Γ.

For a cone K, the polar cone $K^\circ := \{\xi : \langle \xi, \zeta \rangle \le 0, \forall \zeta \in K\}$

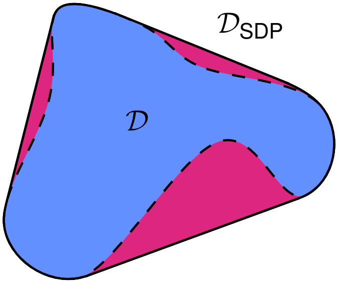


Convex hull exactness

- $\mathrm{conv}(\mathcal{D}) \stackrel{?}{=} \mathcal{D}_{\mathsf{SDP}}$

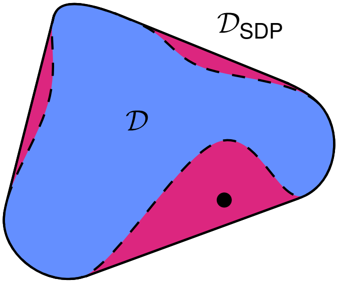
Convex hull exactness

- $\text{conv}(\mathcal{D}) \stackrel{?}{=} \mathcal{D}_{\mathsf{SDP}}$



Convex hull exactness

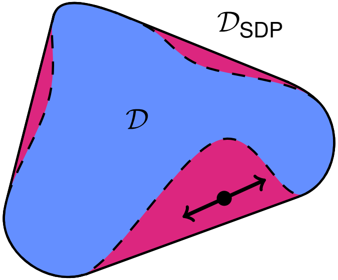
- $\operatorname{conv}(\mathcal{D}) \stackrel{?}{=} \mathcal{D}_{\mathsf{SDP}}$



$\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}} \quad \Longleftrightarrow$ "Given any point in $\mathcal{D}_{\mathsf{SDP}} \setminus \mathcal{D}$, there exists a direction such that we can move forward and backward inside $\mathcal{D}_{\mathsf{SDP}}$"

Convex hull exactness

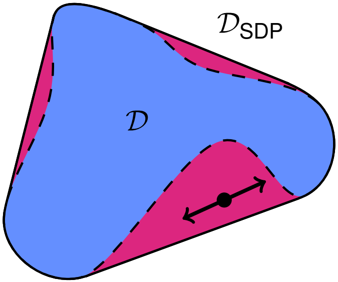
- $\mathrm{conv}(\mathcal{D}) \overset{?}{=} \mathcal{D}_{\mathsf{SDP}}$



$\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}}$ $\Longleftrightarrow$ "Given any point in $\mathcal{D}_{\mathsf{SDP}} \setminus \mathcal{D}$, there exists a direction such that we can move forward and backward inside $\mathcal{D}_{\mathsf{SDP}}$"

Convex hull exactness

- $\text{conv}(\mathcal{D}) \overset{?}{=} \mathcal{D}_{\text{SDP}}$

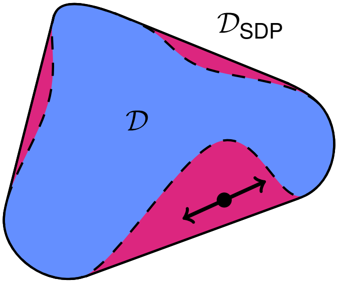


$\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}} \quad \Longleftrightarrow$ "Given any point in $\mathcal{D}_{\text{SDP}} \setminus \mathcal{D}$, there exists a direction such that we can move forward and backward inside $\mathcal{D}_{\text{SDP}}$"

- When do these "rounding" directions exist?

Convex hull exactness

- $\mathrm{conv}(\mathcal{D}) \stackrel{?}{=} \mathcal{D}_{\mathsf{SDP}}$

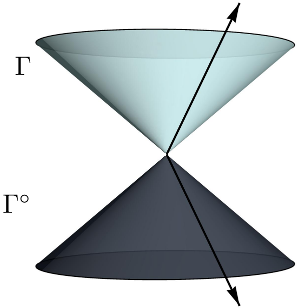


$\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}} \iff$ "Given any point in $\mathcal{D}_{\mathsf{SDP}} \setminus \mathcal{D}$, there exists a direction such that we can move forward and backward inside $\mathcal{D}_{\mathsf{SDP}}$"

- When do these "rounding" directions exist? $\longleftarrow$ Can carry out this idea for QCQPs!

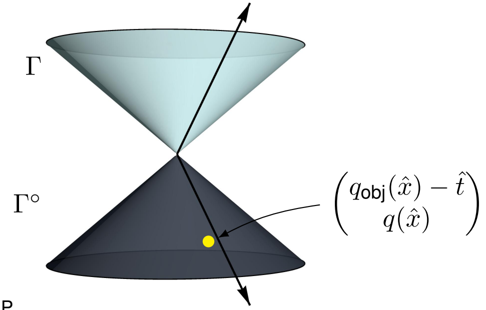
Faces of Γ and Γ°

- Recall $\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\}.$



Faces of Γ and Γ°

- Recall $\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\}$.

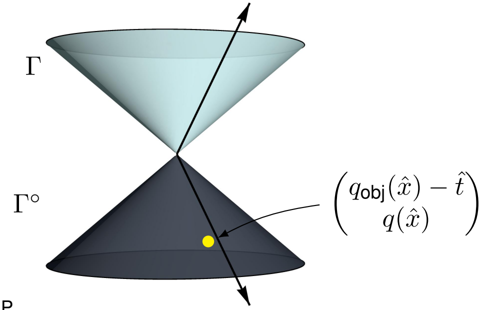


- $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$

$\iff \forall (\hat{x},\hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}, \exists (x',t') \neq 0$ s.t. $\begin{pmatrix} q_{\text{obj}}(\hat{x} \pm \alpha x') - (\hat{t} \pm \alpha t') \\ q(\hat{x} \pm \alpha x') \end{pmatrix} \in \Gamma^\circ$ for some $\alpha > 0$.

Faces of Γ and Γ°

- Recall $\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\}.$

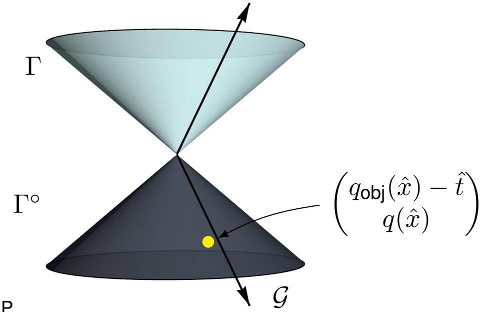


- $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$

  $\iff \forall (\hat{x},\hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}, \exists (x',t') \neq 0$ s.t. $\begin{pmatrix} q_{\text{obj}}(\hat{x} \pm \alpha x') - (\hat{t} \pm \alpha t') \\ q(\hat{x} \pm \alpha x') \end{pmatrix} \in \Gamma^\circ$ for some $\alpha > 0$.

Faces of Γ and Γ°

- Recall $\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\}$.



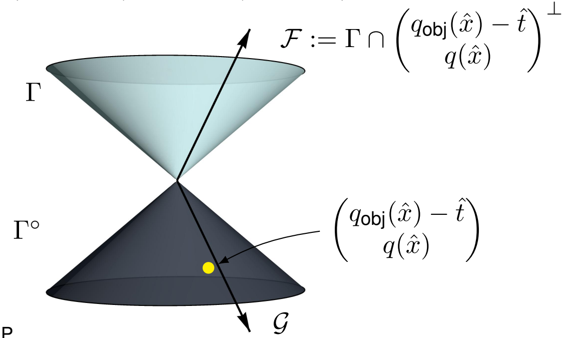
- $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$

  $\iff \forall (\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}, \exists (x', t') \neq 0$ s.t. $\begin{pmatrix} q_{\text{obj}}(\hat{x} \pm \alpha x') - (\hat{t} \pm \alpha t') \\ q(\hat{x} \pm \alpha x') \end{pmatrix} \in \Gamma^\circ$ for some $\alpha > 0$.

- Let $\mathcal{G}(\hat{x}, \hat{t})$ denote the minimal face of Γ° containing $\begin{pmatrix} q_{\text{obj}}(\hat{x}) - \hat{t} \\ q(\hat{x}) \end{pmatrix}$.

Faces of Γ and Γ°

- Recall $\mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \begin{pmatrix} q_{\text{obj}}(x) - t \\ q(x) \end{pmatrix} \in \Gamma^\circ \right\}$.



- $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$

  $\Longleftrightarrow \forall (\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}, \exists (x', t') \neq 0$ s.t. $\begin{pmatrix} q_{\text{obj}}(\hat{x} \pm \alpha x') - (\hat{t} \pm \alpha t') \\ q(\hat{x} \pm \alpha x') \end{pmatrix} \in \Gamma^\circ$ for some $\alpha > 0$.

- Let $\mathcal{G}(\hat{x}, \hat{t})$ denote the minimal face of Γ° containing $\begin{pmatrix} q_{\text{obj}}(\hat{x}) - \hat{t} \\ q(\hat{x}) \end{pmatrix}$.

Sufficient condition for convex hull exactness

- Given $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}}$, look for a subset of directions (x', t') s.t. $\left[(\hat{x}, \hat{t}) \pm \alpha(x', t')\right] \in \mathcal{D}_{\text{SDP}}$ for some $\alpha > 0$

Sufficient condition for convex hull exactness

- Given $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}}$, look for a subset of directions (x', t') s.t. $\left[(\hat{x}, \hat{t}) \pm \alpha(x', t')\right] \in \mathcal{D}_{\text{SDP}}$ for some $\alpha > 0$

Theorem

If for every $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}$, the set

$$\mathcal{R}'(\hat{x}, \hat{t}) := \left\{ (x', t') \in \mathbb{R}^{n+1} : \begin{pmatrix} q_{\text{obj}}(\hat{x} + \alpha x') - (\hat{t} + \alpha t') \\ q(\hat{x} + \alpha x') \end{pmatrix} \in \operatorname{span}(\mathcal{G}(\hat{x}, \hat{t})), \forall \alpha \in \mathbb{R} \right\}.$$

is nontrivial, then $\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$.

Sufficient condition for convex hull exactness

- Given $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}}$, look for a subset of directions (x', t') s.t. $\left[(\hat{x}, \hat{t}) \pm \alpha(x', t')\right] \in \mathcal{D}_{\text{SDP}}$ for some $\alpha > 0$

Theorem

If for every $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}$, the set

$$\mathcal{R}'(\hat{x}, \hat{t}) := \left\{ (x', t') \in \mathbb{R}^{n+1} : \begin{pmatrix} q_{\text{obj}}(\hat{x} + \alpha x') - (\hat{t} + \alpha t') \\ q(\hat{x} + \alpha x') \end{pmatrix} \in \operatorname{span}(\mathcal{G}(\hat{x}, \hat{t})), \ \forall \alpha \in \mathbb{R} \right\}.$$

is nontrivial, then $\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$.

- Whenever Γ° is facially exposed (e.g., whenever Γ° is polyhedral), this condition identifies all rounding directions:

Sufficient condition for convex hull exactness

- Given $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}}$, look for a subset of directions (x', t') s.t. $\big[(\hat{x}, \hat{t}) \pm \alpha(x', t')\big] \in \mathcal{D}_{\text{SDP}}$ for some $\alpha > 0$

Theorem

If for every $(\hat{x}, \hat{t}) \in \mathcal{D}_{\text{SDP}} \setminus \mathcal{D}$, the set

$$\mathcal{R}'(\hat{x}, \hat{t}) := \left\{ (x', t') \in \mathbb{R}^{n+1} : \begin{pmatrix} q_{\text{obj}}(\hat{x} + \alpha x') - (\hat{t} + \alpha t') \\ q(\hat{x} + \alpha x') \end{pmatrix} \in \operatorname{span}(\mathcal{G}(\hat{x}, \hat{t})), \ \forall \alpha \in \mathbb{R} \right\}.$$

is nontrivial, then $\operatorname{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}}$.

- Whenever Γ° is facially exposed (e.g., whenever Γ° is polyhedral), this condition identifies all rounding directions:

  $\Longrightarrow$ This sufficient condition becomes also necessary.

Sufficient condition for convex hull exactness

- When Γ° is facially exposed, $\mathcal{R}'(x,t)$ admits further simplification.

Sufficient condition for convex hull exactness

- When Γ° is facially exposed, $\mathcal{R}'(x,t)$ admits further simplification.
- For example,

Proposition

Suppose Γ is strictly feasible. Consider any $(x,t) \in \mathcal{D}_{\text{SDP}}$ with $t = \sup_{\gamma \in \Gamma_1} q(\gamma, x)$, and let $(1,f) \in \operatorname{rint}(\mathcal{F}(x,t))$.
If Γ is polyhedral, then

$$\mathcal{R}'(x,t) = \left\{ (x',t') \in \mathbb{R}^{n+1} : \begin{array}{l} x' \in \ker(A(f)), \\ \langle b(\gamma), x' \rangle - t' = 0, \ \forall (1,\gamma) \in \mathcal{F}(x,t) \end{array} \right\}.$$

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \leq 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.

Related: Yıldıran [2009]

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \leq 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.
- Then, $\mathcal{D} = \mathcal{X} \times \mathbb{R}_+$ and $\mathcal{D}_{\text{SDP}} = \mathcal{X}_{\text{SDP}} \times \mathbb{R}_+$.

Related: Yıldıran [2009]

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \leq 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.
- Then, $\mathcal{D} = \mathcal{X} \times \mathbb{R}_+$ and $\mathcal{D}_{\text{SDP}} = \mathcal{X}_{\text{SDP}} \times \mathbb{R}_+$.

Proposition

Related: Yıldıran [2009]

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \leq 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.
- Then, $\mathcal{D} = \mathcal{X} \times \mathbb{R}_+$ and $\mathcal{D}_{\text{SDP}} = \mathcal{X}_{\text{SDP}} \times \mathbb{R}_+$.

Proposition

- Suppose $\exists \gamma^* \in \mathbb{R}^2_+$ s.t. $\gamma_1^* A_1 + \gamma_2^* A_2 \succ 0$, and let $\gamma^{(1)}, \gamma^{(2)} \in \mathbb{R}^2_+$ be generators of Γ_1.

Related: Yıldıran [2009]

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \le 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.
- Then, $\mathcal{D} = \mathcal{X} \times \mathbb{R}_+$ and $\mathcal{D}_{\text{SDP}} = \mathcal{X}_{\text{SDP}} \times \mathbb{R}_+$.

Proposition

- Suppose $\exists \gamma^* \in \mathbb{R}^2_+$ s.t. $\gamma^*_1 A_1 + \gamma^*_2 A_2 \succ 0$, and let $\gamma^{(1)}, \gamma^{(2)} \in \mathbb{R}^2_+$ be generators of Γ_1.
- Suppose $\mathcal{X}$ is strictly feasible and q_1, q_2 are both nonconvex.

Related: Yıldıran [2009]

Example: SDP convex hull exactness for $m = 2$

- Consider $\mathcal{X} = \{x : q_i(x) \le 0, \forall i \in [2]\}$, i.e., $q_{\text{obj}} = 0$ and $m = 2$.
- Then, $\mathcal{D} = \mathcal{X} \times \mathbb{R}_+$ and $\mathcal{D}_{\text{SDP}} = \mathcal{X}_{\text{SDP}} \times \mathbb{R}_+$.

Proposition

- Suppose $\exists \gamma^* \in \mathbb{R}^2_+$ s.t. $\gamma_1^* A_1 + \gamma_2^* A_2 \succ 0$, and let $\gamma^{(1)}, \gamma^{(2)} \in \mathbb{R}^2_+$ be generators of Γ_1.
- Suppose $\mathcal{X}$ is strictly feasible and q_1, q_2 are both nonconvex.
- Then, $\operatorname{cl}\operatorname{conv}(\mathcal{X}) = \mathcal{X}_{\text{SDP}}$ if and only if for both $i = 1, 2$, we have that

$$\ker(A(\gamma^{(i)})) \cap b(\gamma^{(i)})^\perp \text{ is nontrivial.}$$

Related: Yıldıran [2009]

Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs, a.k.a., **quadratic matrix programs** (QMPs):

$$x \in \mathbb{R}^n \quad \longrightarrow \quad X \in \mathbb{R}^{n\times k} \text{ and}$$

$$x^\top A x + 2b^\top x + c \quad \longrightarrow \quad \operatorname{tr}\left(X^\top \mathbb{A} X\right) + 2\left\langle B, X\right\rangle + c$$

Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs, a.k.a., **quadratic matrix programs** (QMPs):

$$x \in \mathbb{R}^n \quad \longrightarrow \quad X \in \mathbb{R}^{n\times k} \text{ and}$$

$$x^\top A x + 2b^\top x + c \quad \longrightarrow \quad \operatorname{tr}\left(X^\top \mathbb{A} X\right) + 2\langle B, X\rangle + c =: q_i(X)$$

- General QMP: $$\min_{X\in\mathbb{R}^{n\times k}} \{q_{\text{obj}}(X) : \; q_i(X) \le 0, \; \forall i \in [m]\}$$

Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs, a.k.a., **quadratic matrix programs** (QMPs):

$$x \in \mathbb{R}^n \longrightarrow X \in \mathbb{R}^{n\times k} \text{ and}$$

$$x^\top Ax + 2b^\top x + c \longrightarrow \operatorname{tr}\left(X^\top \mathbb{A} X\right) + 2\langle B, X\rangle + c =: q_i(X)$$

- General QMP: $$\min_{X\in\mathbb{R}^{n\times k}} \{q_{\text{obj}}(X) : \; q_i(X) \le 0, \; \forall i \in [m]\}$$

- **Applications**:
  - Robust least squares, sphere packing problems, QCQPs with spherical constraints, orthogonal Procrustes problem

Related: Beck [2007], Beck et al. [2012]

Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs, a.k.a., **quadratic matrix programs** (QMPs):

$$x \in \mathbb{R}^n \quad \longrightarrow \quad X \in \mathbb{R}^{n\times k} \text{ and}$$

$$x^\top A x + 2b^\top x + c \quad \longrightarrow \quad \operatorname{tr}\left(X^\top \mathbb{A} X\right) + 2\langle B, X\rangle + c =: q_i(X)$$

- General QMP: $$\min_{X\in\mathbb{R}^{n\times k}} \left\{ q_{\text{obj}}(X) : \; q_i(X) \le 0, \; \forall i \in [m] \right\}$$

- Can be written as a QCQP by defining $A_{\text{obj}} = I_k \otimes \mathbb{A}_{\text{obj}}$, $A_i = I_k \otimes \mathbb{A}_i$ $\forall i \in [m]$

$$A = I_k \otimes \mathbb{A} = \begin{pmatrix} \mathbb{A} & & \\ & \ddots & \\ & & \mathbb{A} \end{pmatrix}$$

- Convex hull exactness holds whenever $k \ge m$

Related: Beck [2007], Beck et al. [2012]

Sufficient condition for objective value exactness

Objective value exactness has been studied a lot:

- TRS and S-lemma

  [Yakubovich, 1971]

- Extended TRS

  [Jeyakumar and Li, 2014, Ben-Tal and den Hertog, 2014, Locatelli, 2016, Ho-Nguyen and K.-K., 2017, Bomze et al., 2018]

- Sign-definite SDPs

  [Sojoudi and Lavaei, 2014]

- SDPs with simultaneously diagonalizable matrices

  [Burer and Ye, 2019, Locatelli, 2022]

- SDPs with certain sparsity patterns (forest, bipartite)

  [Azuma et al., 2022b,a]

- ...

Sufficient condition for objective value exactness

- Give primal and also dual sufficient conditions for optimizer exactness, i.e.,

$$\arg\min_{(x,t)\in\mathcal{D}} t = \arg\min_{(x,t)\in\mathcal{D}_{\mathsf{SDP}}} t.$$

Related: [Burer and Ye, 2019, Locatelli, 2022]

Sufficient condition for objective value exactness

- Give primal and also dual sufficient conditions for optimizer exactness, i.e.,

$$\arg\min_{(x,t)\in\mathcal{D}} t = \arg\min_{(x,t)\in\mathcal{D}_{\mathsf{SDP}}} t.$$

- Dual sufficient condition originates from

$$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} \{q(\gamma, x)\}$$

Related: [Burer and Ye, 2019, Locatelli, 2022]

Sufficient condition for objective value exactness

- Give primal and also dual sufficient conditions for optimizer exactness, i.e.,

$$\arg\min_{(x,t)\in\mathcal{D}} t = \arg\min_{(x,t)\in\mathcal{D}_{\mathsf{SDP}}} t.$$

- Dual sufficient condition originates from

$$\begin{aligned}\mathrm{Opt}_{\mathsf{SDP}} &= \inf_{x\in\mathbb{R}^n}\sup_{\gamma\in\Gamma_1}\{q(\gamma,x)\}\\ &= \sup_{\gamma\in\Gamma_1}\underbrace{\inf_{x\in\mathbb{R}^n}\{q(\gamma,x)\}}_{:=\mathbf{d}(\gamma)} = \sup_{\gamma\in\Gamma_1}\mathbf{d}(\gamma).\end{aligned}$$

(by coercivity [Ekeland and Temam, 1999])

Related: [Burer and Ye, 2019, Locatelli, 2022]

Sufficient condition for objective value exactness

- Give primal and also dual sufficient conditions for optimizer exactness, i.e.,

$$\arg\min_{(x,t)\in\mathcal{D}} t = \arg\min_{(x,t)\in\mathcal{D}_{\text{SDP}}} t.$$

- Dual sufficient condition originates from

$$\begin{aligned}\text{Opt}_{\text{SDP}} &= \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} \{q(\gamma,x)\} \\ &= \sup_{\gamma\in\Gamma_1} \underbrace{\inf_{x\in\mathbb{R}^n} \{q(\gamma,x)\}}_{:= \mathbf{d}(\gamma)} = \sup_{\gamma\in\Gamma_1} \mathbf{d}(\gamma).\end{aligned}$$

(by coercivity [Ekeland and Temam, 1999])

Theorem

Suppose $\sup_{\gamma\in\Gamma_1} \mathbf{d}(\gamma)$ is achieved at some γ^* for which $A_{\text{obj}} + A(\gamma^*) \succ 0$. Then, $\arg\min_{(x,t)\in\mathcal{D}} t = \arg\min_{(x,t)\in\mathcal{D}_{\text{SDP}}} t$.

Related: [Burer and Ye, 2019, Locatelli, 2022]

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- **Further applications**:

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- **Further applications**:
  - Diagonal QCQPs with sign-definite linear terms, QCQPs with centered constraints and polyhedral Γ, QCQPs with spherical constraints, random and semi-random QCQPs, ratios of quadratic functions

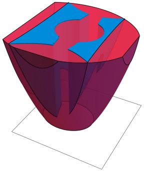
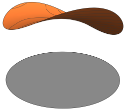
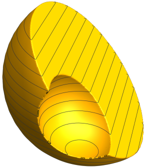
Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- **Further applications**:
  - Diagonal QCQPs with sign-definite linear terms, QCQPs with centered constraints and polyhedral Γ, QCQPs with spherical constraints, random and semi-random QCQPs, ratios of quadratic functions

Summary of exactness in the original space

- Sufficient conditions for convex hull exactness
  - **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- **Further applications**:
  - Diagonal QCQPs with sign-definite linear terms, QCQPs with centered constraints and polyhedral Γ, QCQPs with spherical constraints, random and semi-random QCQPs, ratios of quadratic functions

SDPs provide exact reformulations for broad classes of QCQPs!

Efficient algorithms for exact SDPs

References:

Wang, A. L., Lu, Y., and K.-K., F. (2023+). Implicit regularity and linear convergence rates for the generalized trust-region subproblem. *SIAM J. Optim.*, Forthcoming, (arXiv:2112.13821)

Wang, A. L. and K.-K., F. (2022a). Accelerated first-order methods for a class of semidefinite programs. *arXiv preprint*, 2206.00224

Solving the SDP relaxation

- SDPs can be solved in polynomial time

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n$

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n$ $\longrightarrow$ $O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:
  $\longrightarrow$ **Burer-Monteiro method:** [Burer and Monteiro, 2003], extremely popular in ML...

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:
  $\longrightarrow$ **Burer-Monteiro method:** [Burer and Monteiro, 2003], extremely popular in ML. . .

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n$ $\longrightarrow$ $O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:
  $\longrightarrow$ **Burer-Monteiro method:** [Burer and Monteiro, 2003], extremely popular in ML...
  - *Motivation:* For a primal SDP (P) with m LMEs, ∃an optimal solution Z^* with $\text{rank}(Z^*) \leq \lceil \sqrt{2m} \rceil$ [Pataki, 1998, Barvinok, 2001]

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:
  $\longrightarrow$ **Burer-Monteiro method:** [Burer and Monteiro, 2003], extremely popular in ML. . .
  - *Motivation:* For a primal SDP (P) with m LMEs, ∃an optimal solution Z^* with $\mathrm{rank}(Z^*) \leq \lceil \sqrt{2m} \rceil$ [Pataki, 1998, Barvinok, 2001]
  - *Main Idea:* Solve (P) as an NLP by replacing Z with $VV^\top$ where V has at least $\lceil \sqrt{2m} \rceil$ columns

Solving the SDP relaxation

- SDPs can be solved in polynomial time $\longleftarrow$ too expensive in modern applications
- Usual SDP relaxation in $x \in \mathbb{R}^n$ and $X \in \mathbb{S}^n \longrightarrow O(n^2)$ variables
  - Classical interior point methods:
    $\longrightarrow$ iterations are too expensive requiring $O(mn^3 + m^2n^2 + m^3)$ operations (time) and $O(n^2 + m^2)$ storage
- High storage requirements and expensive iterations led to alternative approaches:
  $\longrightarrow$ **Burer-Monteiro method:** [Burer and Monteiro, 2003], extremely popular in ML. . .
  - *Motivation:* For a primal SDP (P) with m LMEs, ∃an optimal solution Z^* with $\mathrm{rank}(Z^*) \leq \lceil\sqrt{2m}\rceil$ [Pataki, 1998, Barvinok, 2001]
  - *Main Idea:* Solve (P) as an NLP by replacing Z with $VV^\top$ where V has at least $\lceil\sqrt{2m}\rceil$ columns
  - Recent theory showing that under some regularity conditions, for almost all objective functions, B-M method finds the global optimum. [Boumal et al., 2016, 2020]

Solving the SDP relaxation

We can solve an SDP more efficiently and with $O(kn + m)$ storage if it is rank-k-exact (regular)!

Solving the SDP relaxation

We can solve an SDP more efficiently and with $O(kn+m)$ storage if it is rank-k-exact (regular)!

- Recall $$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x)$$

We can solve an SDP more efficiently and with $O(kn+m)$ storage if it is rank-k-exact (regular)!

- Recall
$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x\in\mathbb{R}^n} \left(\sup_{\gamma\in\Gamma_1} q(\gamma, x) \right)$$

Solving the SDP relaxation

We can solve an SDP more efficiently and with $O(kn+m)$ storage if it is rank-k-exact (regular)!

- Recall
$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x \in \mathbb{R}^n} \left(\sup_{\gamma \in \Gamma_1} q(\gamma, x) \right)$$

  This is a minimization problem in the original space $\longrightarrow$ n variables

We can solve an SDP more efficiently and with $O(kn+m)$ storage if it is rank-k-exact (regular)!

- Recall
$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x\in\mathbb{R}^n} \left(\sup_{\gamma\in\Gamma_1} q(\gamma, x) \right)$$
This is a minimization problem in the original space $\longrightarrow$ n variables
- Exactness (regularity) will allow us to efficiently deal with max-type obj. structure

Linear-time algorithm for the Generalized TRS

- Generalized TRS: $$\mathrm{Opt} = \inf_{x\in\mathbb{R}^n} \left\{ q_{\mathsf{obj}}(x) : q_1(x) \le 0 \right\}$$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020]

Linear-time algorithm for the Generalized TRS

- Generalized TRS: $\mathrm{Opt} = \inf_{x\in\mathbb{R}^n} \left\{ q_{\mathsf{obj}}(x) : q_1(x) \le 0 \right\}$
- Recall convex hull exactness holds: $\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}} = \left\{ (x,t) : \sup_{\gamma\in\Gamma_1} q(\gamma,x) \le t \right\}$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020]

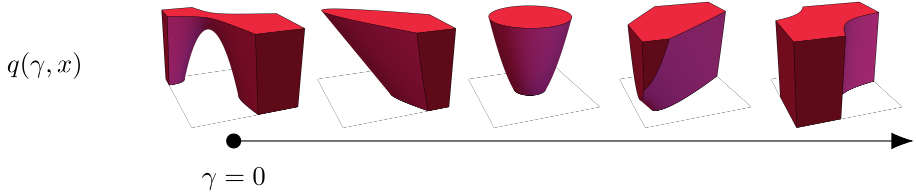
Linear-time algorithm for the Generalized TRS

- Generalized TRS: $\mathrm{Opt} = \inf_{x \in \mathbb{R}^n} \left\{ q_{\mathsf{obj}}(x) : q_1(x) \leq 0 \right\}$
- Recall convex hull exactness holds: $\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}} = \left\{ (x,t) : \sup_{\gamma \in \Gamma_1} q(\gamma, x) \leq t \right\}$
- where $\Gamma_1 = \{\gamma \in \mathbb{R}_+ : q(\gamma, x) \text{ is convex in } x\}$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020]

Linear-time algorithm for the Generalized TRS

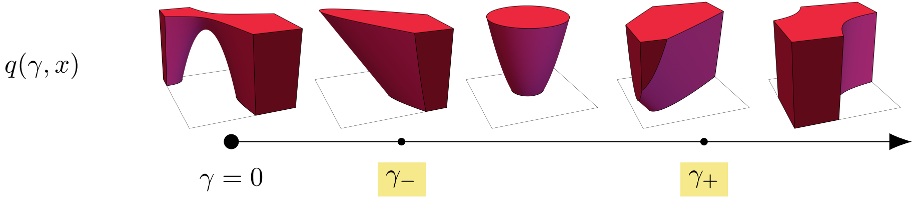
- Generalized TRS: $\text{Opt} = \inf_{x \in \mathbb{R}^n} \{q_{\text{obj}}(x) : q_1(x) \le 0\}$
- Recall convex hull exactness holds: $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}} = \left\{(x,t) : \sup_{\gamma \in \Gamma_1} q(\gamma, x) \le t\right\}$
- where $\Gamma_1 = \{\gamma \in \mathbb{R}_+ : q(\gamma, x) \text{ is convex in } x\}$



Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020]

Linear-time algorithm for the Generalized TRS

- Generalized TRS: $\text{Opt} = \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : q_1(x) \leq 0 \right\}$
- Recall convex hull exactness holds: $\text{conv}(\mathcal{D}) = \mathcal{D}_{\text{SDP}} = \left\{ (x,t) : \sup_{\gamma \in \Gamma_1} q(\gamma, x) \leq t \right\}$
- where $\Gamma_1 = \{\gamma \in \mathbb{R}_+ : q(\gamma, x) \text{ is convex in } x\}$



Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020]

Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+]$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in [\gamma_-, \gamma_+]} q(\gamma, x)$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

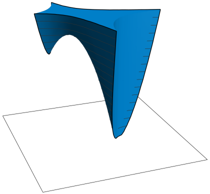
Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

Linear-time algorithm for the Generalized TRS

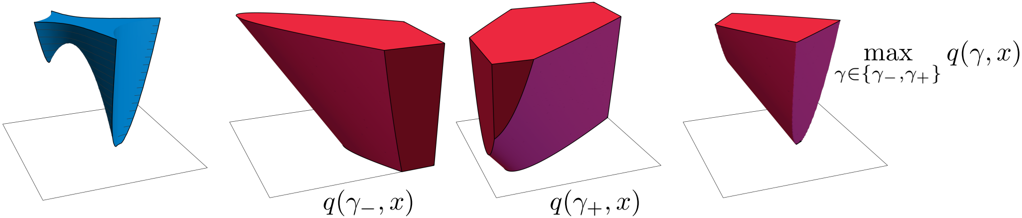
- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

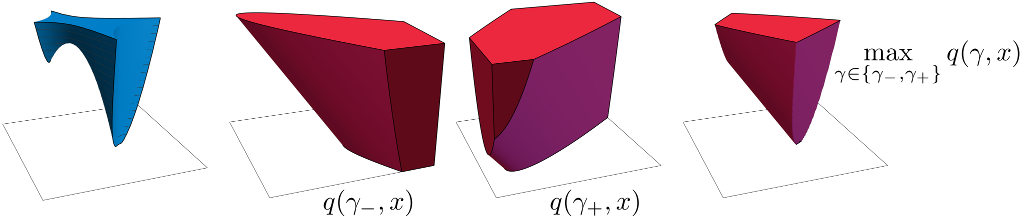
Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

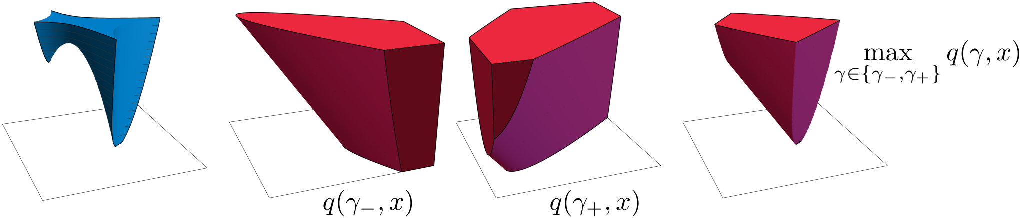
- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



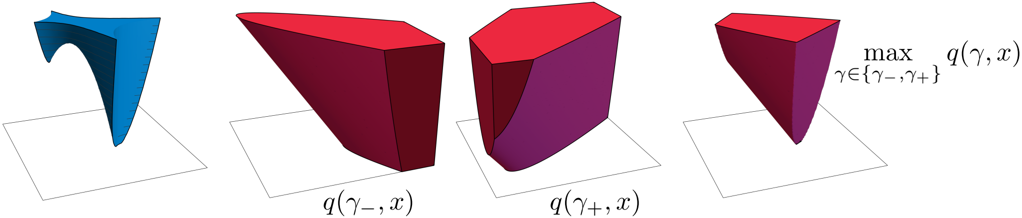
- **Algorithmic idea**
  - Compute γ_- and γ_+ to some accuracy
  - Apply Accelerated Gradient Descent for smooth minimax problems

---

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



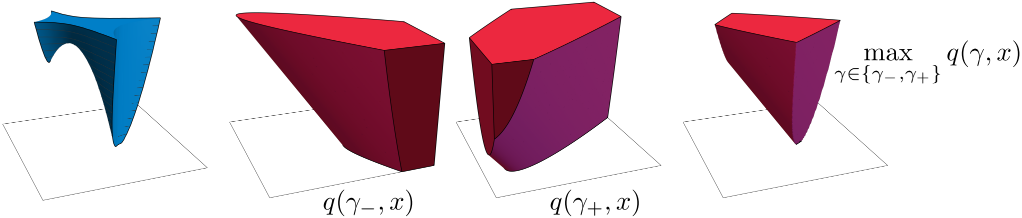
- **Algorithmic idea**
  - Compute γ_- and γ_+ to some accuracy
  - Apply Accelerated Gradient Descent for smooth minimax problems

$\implies \tilde{O}\left(\frac{N}{\sqrt{\epsilon}} \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right)$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

Linear-time algorithm for the Generalized TRS

- $\Gamma_1 = [\gamma_-, \gamma_+] \implies \text{Opt} = \text{Opt}_{\mathsf{SDP}} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\gamma_-, \gamma_+\}} q(\gamma, x)$



- **Algorithmic idea**
  - Compute γ_- and γ_+ to some accuracy
  - Apply Accelerated Gradient Descent for smooth minimax problems

$$\implies \tilde{O}\left(\frac{N}{\sqrt{\epsilon}} \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right) \approx \frac{1}{\sqrt{\epsilon}}$$

Related: Hazan and Koren [2016], Ho-Nguyen and K.-K. [2017], Jiang and Li [2019, 2020], Wang and K.-K. [2022b]

- Dual problem:

$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) = \sup_{\gamma\in\Gamma_1} \inf_{x\in\mathbb{R}^n} q(\gamma, x)$$

Regularity

- Dual problem:

$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x \in \mathbb{R}^n} \sup_{\gamma \in \Gamma_1} q(\gamma, x) = \sup_{\gamma \in \Gamma_1} \inf_{x \in \mathbb{R}^n} q(\gamma, x)$$

Definition

Let γ^* be a dual optimizer. Define $\mu^* := \lambda_{\min}\left(A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i^* A_i\right)$. Note $\mu^* \geq 0$ by definition of Γ_1. QCQP instance is regular if $\mu^* > 0$.

Regularity

- Dual problem:

$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x) = \sup_{\gamma\in\Gamma_1} \inf_{x\in\mathbb{R}^n} q(\gamma, x)$$

Definition

Let γ^* be a dual optimizer. Define $\mu^* := \lambda_{\min}\left(A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i^* A_i\right)$. Note $\mu^* \geq 0$ by definition of Γ_1. QCQP instance is regular if $\mu^* > 0$.

- Regularity $\implies$ optimizer exactness

$$\mu^* > 0 \implies \operatorname*{arg\,min}_{x\in\mathbb{R}^n} \{q_{\mathsf{obj}}(x) : \; q_i(x) \leq 0, \forall i \in [m]\} = \operatorname*{arg\,min}_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma_1} q(\gamma, x)$$

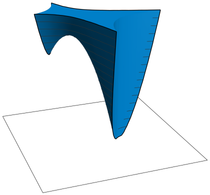
Linear convergence for regular GTRS

- Suppose $\mu^* > 0$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

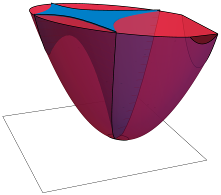
- Suppose $\mu^* > 0$



Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

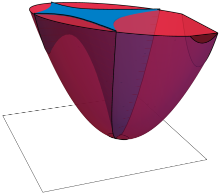
- Suppose $\mu^* > 0$



Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$

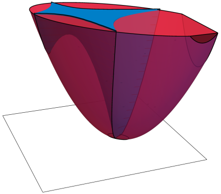


- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$

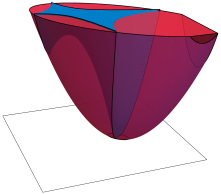


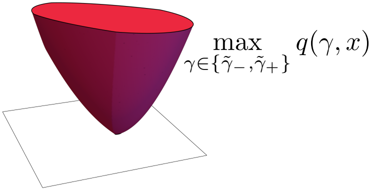
- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$



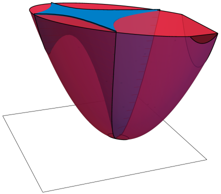


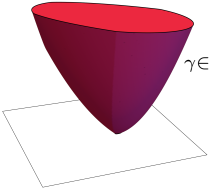
- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





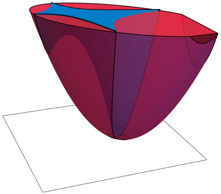
$$\max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$$

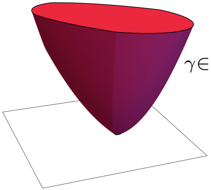
- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$
- Suffices to estimate γ^* to low accuracy

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





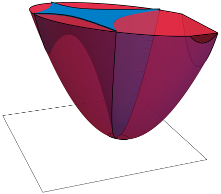
$$\max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$$

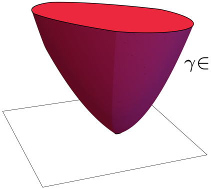
- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \mathrm{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$
- Suffices to estimate γ^* to low accuracy and then we can exploit strong convexity

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





$$\max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$$

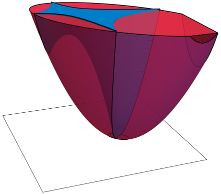
- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x \in \mathbb{R}^n} \max_{\gamma \in \{\tilde{\gamma}_-, \tilde{\gamma}_+\}} q(\gamma, x)$
- Suffices to estimate γ^* to low accuracy and then we can exploit strong convexity

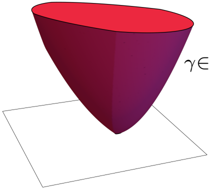
$$\tilde{O}\left(\frac{N}{\sqrt{\mu^*}} \log\left(\frac{1}{\mu^*}\right) \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right)$$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





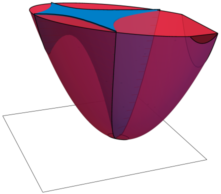
- Suppose $\gamma^* \in [\tilde\gamma_-, \tilde\gamma_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x\in\mathbb{R}^n} \max_{\gamma\in\{\tilde\gamma_-,\tilde\gamma_+\}} q(\gamma, x)$
- Suffices to estimate γ^* to low accuracy and then we can exploit strong convexity

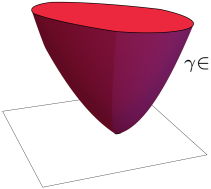
$$\tilde{O}\left(\frac{N}{\sqrt{\mu^*}} \log\left(\frac{1}{\mu^*}\right) \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right) \approx \log\left(\frac{1}{\epsilon}\right)$$

Related: Carmon and Duchi [2018]

Linear convergence for regular GTRS

- Suppose $\mu^* > 0$





$$\max_{\gamma\in\{\tilde{\gamma}_-,\tilde{\gamma}_+\}} q(\gamma, x)$$

- Suppose $\gamma^* \in [\tilde{\gamma}_-, \tilde{\gamma}_+] \subseteq \Gamma_1 \implies \text{Opt} = \inf_{x\in\mathbb{R}^n} \max_{\gamma\in\{\tilde{\gamma}_-,\tilde{\gamma}_+\}} q(\gamma, x)$
- Suffices to estimate γ^* to low accuracy and then we can exploit strong convexity

$$\tilde{O}\left(\frac{N}{\sqrt{\mu^*}} \log\left(\frac{1}{\mu^*}\right) \log\left(\frac{n}{p}\right) \log\left(\frac{1}{\epsilon}\right)\right) \approx \log\left(\frac{1}{\epsilon}\right)$$

$\implies$ This rate is linear in both N and $\log(1/\epsilon)$

Related: Carmon and Duchi [2018]

How to generalize?

- How to handle SDP relaxations of general QCQPs with multiple constraints?

How to generalize?

- How to handle SDP relaxations of general QCQPs with multiple constraints?
- For QCQPs, we desire rank-1 solutions in the SDP relaxations.
  What about SDPs in which we seek rank-k solutions?

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \ M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i \succeq 0 \right\}$$

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$
- **k-exact SDPs:**

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

- k**-exact SDPs**:
  - Strong duality holds, both problems are solvable

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

- k**-exact SDPs**:
  - Strong duality holds, both problems are solvable
  - W, subspace of dimension $n - k$ such that $Y^*_{W^\perp} \succ 0$ is known

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

- k**-exact SDPs**:
  - Strong duality holds, both problems are solvable
  - W, subspace of dimension $n-k$ such that $Y^*_{W^\perp} \succ 0$ is known
    - This talk: $W = \mathbb{R}^{n-k}$ and $Y^*_{W^\perp} = I_k$

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

- k**-exact SDPs**:
  - Strong duality holds, both problems are solvable
  - W, subspace of dimension $n - k$ such that $Y^*_{W^\perp} \succ 0$ is known
    - This talk: $W = \mathbb{R}^{n-k}$ and $Y^*_{W^\perp} = I_k$
  - Strict complementarity + exactness

General setup

- $$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y \succeq 0 \end{array} \right\}$$
$$\geq \sup_{\gamma \in \mathbb{R}^m} \left\{ d^\top \gamma : \underbrace{M_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i M_i}_{:= M(\gamma)} \succeq 0 \right\}$$

- k**-exact SDPs**:
  - Strong duality holds, both problems are solvable
  - W, subspace of dimension $n-k$ such that $Y^*_{W^\perp} \succ 0$ is known
    - This talk: $W = \mathbb{R}^{n-k}$ and $Y^*_{W^\perp} = I_k$
  - Strict complementarity + exactness: there exists Y^*, γ^* such that $\mathrm{rank}(Y^*) = k$ and $\mathrm{rank}(M(\gamma^*)) = n-k$

Motivation

- SDP relaxation of QMP in $X \in \mathbb{R}^{(n-k)\times k}$

Motivation

- SDP relaxation of QMP in $X \in \mathbb{R}^{(n-k)\times k}$, $\begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \mapsto \begin{pmatrix} * & * \\ * & I_k \end{pmatrix}$

Motivation

- SDP relaxation of QMP in $X \in \mathbb{R}^{(n-k)\times k}$, $\begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \mapsto \begin{pmatrix} * & * \\ * & I_k \end{pmatrix}$

$$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0,\ \forall i \in [m] \\ Y = \begin{pmatrix} * & * \\ * & I_k \end{pmatrix} \succeq 0 \end{array} \right\}$$

Motivation

- SDP relaxation of QMP in $X \in \mathbb{R}^{(n-k)\times k}$, $\begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \mapsto \begin{pmatrix} * & * \\ * & I_k \end{pmatrix}$

$$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0, \ \forall i \in [m] \\ Y = \begin{pmatrix} * & * \\ * & I_k \end{pmatrix} \succeq 0 \end{array} \right\}$$

Taking $W = \mathbb{R}^{n-k}$, we know $Y^*_{W^\perp} = I_k \succ 0$

Motivation

- SDP relaxation of QMP in $X \in \mathbb{R}^{(n-k)\times k}$, $\begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \mapsto \begin{pmatrix} * & * \\ * & I_k \end{pmatrix}$

$$\mathrm{Opt}_{\mathsf{SDP}} = \inf_{Y \in \mathbb{S}^n} \left\{ \langle M_{\mathsf{obj}}, Y \rangle : \begin{array}{l} \langle M_i, Y \rangle + d_i = 0,\ \forall i \in [m] \\ Y = \begin{pmatrix} * & * \\ * & I_k \end{pmatrix} \succeq 0 \end{array} \right\}$$

  Taking $W = \mathbb{R}^{n-k}$, we know $Y^*_{W^\perp} = I_k \succ 0$

- Equivalently, k-exact SDPs originate from QCQPs and QMPs that admit exact SDP relaxations

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} X X^\top & X \\ X^\top & I_k \end{pmatrix}$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} X X^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \;= \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \; = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \ = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

- Let $q_i(X) \coloneqq \operatorname{tr}(X^\top A_i X) + 2\left\langle \tilde{B}_i, X \right\rangle + \tilde{c}_i$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \; = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

- Let $q_i(X) := \operatorname{tr}(X^\top A_i X) + 2\left\langle \tilde{B}_i, X \right\rangle + \tilde{c}_i$
- We have reduced SDP to QMP

$$X^* = \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} \{q_{\text{obj}}(X) : \; q_i(X) = 0, \; \forall i \in [m]\}$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \; = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

- Let $q_i(X) := \operatorname{tr}(X^\top A_i X) + 2\left\langle \tilde{B}_i, X \right\rangle + \tilde{c}_i$
- We have reduced SDP to QMP

$$X^* = \operatorname*{arg\,min}_{X\in\mathbb{R}^{(n-k)\times k}} \{q_{\text{obj}}(X) : q_i(X) = 0,\ \forall i \in [m]\}$$

$$= \operatorname*{arg\,min}_{X\in\mathbb{R}^{(n-k)\times k}} \sup_{\gamma\in\mathbb{R}^m} q_{\text{obj}}(X) + \sum_{i=1}^m \gamma_i q_i(X)$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto \; = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

- Let $q_i(X) := \operatorname{tr}(X^\top A_i X) + 2\left\langle \tilde{B}_i, X \right\rangle + \tilde{c}_i$
- We have reduced SDP to QMP

$$X^* = \operatorname*{arg\,min}_{X\in\mathbb{R}^{(n-k)\times k}} \{q_{\text{obj}}(X) : q_i(X) = 0, \ \forall i \in [m]\}$$

$$= \operatorname*{arg\,min}_{X\in\mathbb{R}^{(n-k)\times k}} \sup_{\gamma\in\mathbb{R}^m} \underbrace{q_{\text{obj}}(X) + \sum_{i=1}^m \gamma_i q_i(X)}_{=: \, q(\gamma, X)}$$

Deriving a strongly convex minimax problem

- Suppose k-exact and $Y^*_{W^\perp} = I_k$
- Know $Y^* = \begin{pmatrix} X^* X^{*\top} & X^* \\ X^{*\top} & I_k \end{pmatrix}$ for $X^* \in \mathbb{R}^{(n-k)\times k}$ $\longrightarrow$ replace $Y \mapsto \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix}$

$$\langle M_i, Y\rangle + d_i \mapsto = \left\langle \begin{pmatrix} A_i & B_i \\ B_i^\top & C_i \end{pmatrix}, \begin{pmatrix} XX^\top & X \\ X^\top & I_k \end{pmatrix} \right\rangle + d_i =: q_i(X)$$

- Let $q_i(X) := \operatorname{tr}(X^\top A_i X) + 2\left\langle \tilde{B}_i, X \right\rangle + \tilde{c}_i$
- We have reduced SDP to QMP

$$X^* = \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} \{q_{\text{obj}}(X) : q_i(X) = 0,\ \forall i \in [m]\}$$

$$= \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} \underbrace{q_{\text{obj}}(X) + \sum_{i=1}^m \gamma_i q_i(X)}_{=:\ q(\gamma, X)\ =\ \operatorname{tr}(X^\top A(\gamma) X) + \dots}$$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \operatorname*{arg\,min}_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$

$$= \operatorname*{arg\,min}_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$
$$= \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \operatorname*{arg\,min}_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$
$$= \operatorname*{arg\,min}_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\implies$ $A(\gamma^*) \succ 0$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$
$$= \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\implies A(\gamma^*) \succ 0 \implies q(\gamma^*, X)$ is strongly convex

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$

$$= \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\implies A(\gamma^*) \succ 0 \implies q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$

$$= \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\implies A(\gamma^*) \succ 0 \implies q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X)$$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X\in\mathbb{R}^{(n-k)\times k}} \sup_{\gamma\in\mathbb{R}^m} q(\gamma, X)$$

$$= \arg\min_{X\in\mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\implies A(\gamma^*) \succ 0 \implies q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

$$X^* = \arg\min_{X\in\mathbb{R}^{(n-k)\times k}} \max_{\gamma\in\mathcal{U}} q(\gamma, X) \qquad \text{(QMMP)}$$

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$\begin{aligned} X^* &= \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X) \\ &= \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X) \end{aligned}$$

- Strict complementarity $\Longrightarrow$ $A(\gamma^*) \succ 0$ $\Longrightarrow$ $q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

$$X^* = \mathop{\arg\min}_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X) \qquad \text{(QMMP)}$$

- **Questions left**:

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \underset{X \in \mathbb{R}^{(n-k)\times k}}{\arg\min} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$
$$= \underset{X \in \mathbb{R}^{(n-k)\times k}}{\arg\min} q(\gamma^*, X)$$

- Strict complementarity $\Longrightarrow$ $A(\gamma^*) \succ 0$ $\Longrightarrow$ $q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

$$X^* = \underset{X \in \mathbb{R}^{(n-k)\times k}}{\arg\min} \max_{\gamma \in \mathcal{U}} q(\gamma, X) \qquad \text{(QMMP)}$$

- **Questions left**:
  - How to construct the certificate $\mathcal{U}$?

Deriving a strongly convex minimax problem

- Note strong duality + strict complementarity implies

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \sup_{\gamma \in \mathbb{R}^m} q(\gamma, X)$$

$$= \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} q(\gamma^*, X)$$

- Strict complementarity $\Longrightarrow$ $A(\gamma^*) \succ 0$ $\Longrightarrow$ $q(\gamma^*, X)$ is strongly convex
- **Cert. of strict complementarity**: $\gamma^* \in \mathcal{U} \subseteq \mathbb{R}^m$ and $A(\gamma) \succ 0$ for all $\gamma \in \mathcal{U}$

$$X^* = \arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X) \qquad \text{(QMMP)}$$

- **Questions left**:
  - How to construct the certificate $\mathcal{U}$?
  - How to solve the strongly convex quadratic matrix minimax problem (QMMP)?

Algorithms

- Given $\mathcal{U}$, how to solve strongly convex QMMP

$$\operatorname*{arg\,min}_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X)?$$

Related: Nesterov [2005], Devolder et al. [2013, 2014], Nesterov [2018]

Algorithms

- Given $\mathcal{U}$, how to solve strongly convex QMMP

$$\arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X)?$$

  - Develop an **inexact** variant of Nesterov's accelerated gradient descent (AGD) method for minimax functions (each "prox-map" is a saddle point problem of its own)

$\longrightarrow$ CautiousAGD: $O\left(\epsilon^{-1/2}\log(\epsilon^{-1})\right)$

Related: Nesterov [2005], Devolder et al. [2013, 2014], Nesterov [2018]

Algorithms

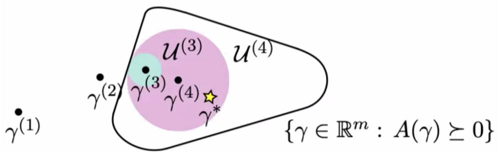
- Given $\mathcal{U}$, how to solve strongly convex QMMP

$$\arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X)?$$

  - Develop an **inexact** variant of Nesterov's accelerated gradient descent (AGD) method for minimax functions (each "prox-map" is a saddle point problem of its own)

  $\longrightarrow$ CautiousAGD: $O\left(\epsilon^{-1/2}\log(\epsilon^{-1})\right)$

- How to find the certificate $\mathcal{U}$?



Related: Nesterov [2005], Devolder et al. [2013, 2014], Nesterov [2018]

Algorithms

- Given $\mathcal{U}$, how to solve strongly convex QMMP

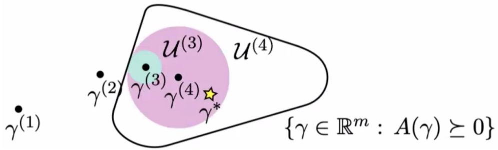
$$\arg\min_{X \in \mathbb{R}^{(n-k)\times k}} \max_{\gamma \in \mathcal{U}} q(\gamma, X)?$$

  - Develop an **inexact** variant of Nesterov's accelerated gradient descent (AGD) method for minimax functions (each "prox-map" is a saddle point problem of its own)

  $\longrightarrow$ CautiousAGD: $\quad O\left(\epsilon^{-1/2}\log(\epsilon^{-1})\right)$

- How to find the certificate $\mathcal{U}$?
  - Generate $\gamma^{(i)} \to \gamma^*$ and neighborhoods $\mathcal{U}^{(i)} \subseteq \{\gamma : A(\gamma) \succeq 0\}$ and monitor convergence of CautiousAGD for QMMP with $\mathcal{U}^{(i)}$.

  $\longrightarrow$ CertSDP



Related: Nesterov [2005], Devolder et al. [2013, 2014], Nesterov [2018]

CertSDP guarantees

Theorem

Given $\epsilon > 0$, CertSDP produces iterates X_t such that

$$\left\langle M_{\text{obj}}, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle \leq \mathrm{Opt}_{\text{SDP}} + \epsilon \qquad \text{and} \qquad \left\| \left(\left\langle M_i, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle + d_i \right)_i \right\|_2 \leq \epsilon.$$

after completing

- **iteration count**: $t \approx O(1) + O\left(\log(\epsilon^{-1})\right)$

CertSDP guarantees

Theorem

Given $\epsilon > 0$, CertSDP produces iterates X_t such that

$$\left\langle M_{\text{obj}}, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle \leq \mathrm{Opt}_{\text{SDP}} + \epsilon \quad \text{and} \quad \left\| \left(\left\langle M_i, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle + d_i \right)_i \right\|_2 \leq \epsilon.$$

after completing

- **iteration count**: $t \approx O(1) + O\left(\log(\epsilon^{-1})\right)$
- **iteration complexity**: $O(m\epsilon^{-1})$ matrix-vector products per iteration

CertSDP guarantees

Theorem

Given $\epsilon > 0$, CertSDP produces iterates X_t such that

$$\left\langle M_{\mathsf{obj}}, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle \leq \mathrm{Opt}_{\mathsf{SDP}} + \epsilon \qquad \text{and} \qquad \left\| \left(\left\langle M_i, \begin{pmatrix} X_t X_t^\top & X_t \\ X_t^\top & I_k \end{pmatrix} \right\rangle + d_i \right)_i \right\|_2 \leq \epsilon.$$

after completing

- **iteration count**: $t \approx O(1) + O\left(\log(\epsilon^{-1})\right)$
- **iteration complexity**: $O(m\epsilon^{-1})$ matrix-vector products per iteration
- **storage**: $O(m + nk)$ entries

A glimpse on numerical results

- Random instances of k-exact **distance-minimization QMP**

$$\inf_{X \in \mathbb{R}^{(n-k)\times k}} \left\{ \|X\|_F^2 : \ q_i(X) = 0, \ \forall i \in [m] \right\}$$

with $k = m = 10$, $(n-k) = 10^3, 10^4, 10^5$ (10 instances per setting)

Related: Ding et al. [2021], Yurtsever et al. [2021], Souto et al. [2020], O'Donoghue et al. [2016]

A glimpse on numerical results

- Random instances of k-exact **distance-minimization QMP**

$$\inf_{X \in \mathbb{R}^{(n-k)\times k}} \left\{ \|X\|_F^2 : \ q_i(X) = 0, \ \forall i \in [m] \right\}$$

with $k = m = 10$, $(n-k) = 10^3, \ 10^4, \ 10^5$ (10 instances per setting)

Algorithm	time (s)	std.	$\|X - X^*\|_F^2$	std.	memory (MB)	std.
CertSDP	1.3×10^3	7.6×10^2	1.9×10^{-22}	4.2×10^{-23}	0.0	0.0
CSSDP	3.0×10^3	5.8×10^{-1}	7.3×10^{-2}	3.4×10^{-2}	0.0	0.0
SketchyCGAL	3.0×10^3	8.5	1.1	6.6×10^{-1}	1.0×10^1	1.0×10^1
ProxSDP	2.1×10^2	1.1×10^1	1.2×10^{-19}	3.2×10^{-19}	4.8×10^1	1.9×10^1
SCS	3.1×10^3	2.5×10^1	5.1×10^{-5}	9.5×10^{-5}	5.3×10^2	4.3×10^1

$n - k = 10^3$, time limit 3×10^3 seconds

Related: Ding et al. [2021], Yurtsever et al. [2021], Souto et al. [2020], O'Donoghue et al. [2016]

A glimpse on numerical results

- Random instances of k-exact **distance-minimization QMP**

$$\inf_{X \in \mathbb{R}^{(n-k)\times k}} \left\{ \|X\|_F^2 : \; q_i(X) = 0, \; \forall i \in [m] \right\}$$

with $k = m = 10$, $(n-k) = 10^3, \, 10^4, \, 10^5$ (10 instances per setting)

Algorithm	time (s)	std.	$\|X - X^*\|_F^2$	std.	memory (MB)	std.
CertSDP	4.5×10^3	7.0×10^2	1.9×10^{-22}	5.2×10^{-23}	8.5	1.2×10^1
CSSDP	1.0×10^4	6.6×10^{-1}	2.7	9.4×10^{-1}	6.2	1.5×10^1
SketchyCGAL	9.7×10^3	1.8×10^2	4.0	1.4	2.7×10^1	2.2×10^1
ProxSDP	1.2×10^4	1.1×10^2	2.9	9.9×10^{-1}	1.9×10^4	1.2×10^2

$n - k = 10^4$, time limit 10^4 seconds

Related: Ding et al. [2021], Yurtsever et al. [2021], Souto et al. [2020], O'Donoghue et al. [2016]

A glimpse on numerical results

- Random instances of k-exact **distance-minimization QMP**

$$\inf_{X \in \mathbb{R}^{(n-k)\times k}} \left\{ \|X\|_F^2 : \; q_i(X) = 0, \; \forall i \in [m] \right\}$$

with $k = m = 10$, $(n-k) = 10^3, \, 10^4, \, 10^5$ (10 instances per setting)

Algorithm	time (s)	std.	$\|X - X^*\|_F^2$	std.	memory (MB)	std.
CertSDP	5.0×10^4	6.2×10^2	2.5×10^{-2}	6.5×10^{-2}	2.3×10^2	2.0×10^2
CSSDP	5.0×10^4	4.7	2.8	5.1×10^{-1}	2.0×10^2	2.5×10^2
SketchyCGAL	4.7×10^4	3.3×10^3	4.0	2.1	3.7×10^2	2.0×10^2

$n - k = 10^5$, time limit 5×10^4 seconds

Related: Ding et al. [2021], Yurtsever et al. [2021], Souto et al. [2020], O'Donoghue et al. [2016]

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs
  (especially when constraints interact nicely and there are large amounts of symmetry)

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:
  - Can we approach **approximation quality** similarly?

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:
  - Can we approach **approximation quality** similarly?
  - Can these tools for **proving exactness** guide us to **design** better convex relaxations?

Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:
  - Can we approach **approximation quality** similarly?
  - Can these tools for **proving exactness** guide us to **design** better convex relaxations?

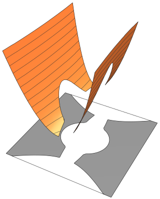
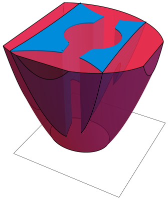
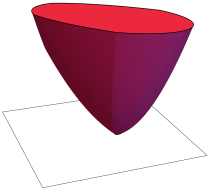
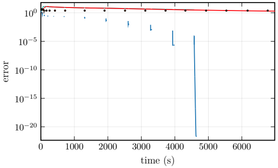
Summary

- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:
  - Can we approach **approximation quality** similarly?
  - Can these tools for **proving exactness** guide us to **design** better convex relaxations?
  - More generally, **exactness** $\approx$ **efficiency?**

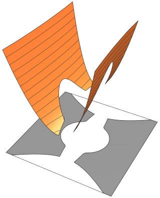
Summary

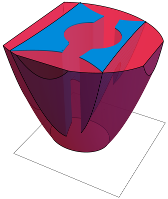
- SDPs provide exact reformulations for broad classes of QCQPs and QMPs (especially when constraints interact nicely and there are large amounts of symmetry)
- Rank-k exact SDPs can be solved very efficiently via first-order methods
- **Future directions**:
  - Can we approach **approximation quality** similarly?
  - Can these tools for **proving exactness** guide us to **design** better convex relaxations?
  - More generally, **exactness** $\approx$ **efficiency?**
  - Can we develop efficient algorithms for SDPs admitting **approximately** low-rank solutions?

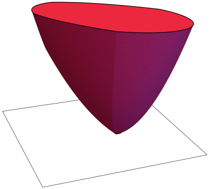
Thank you!

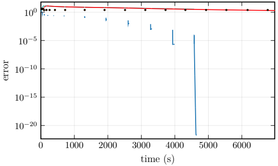




10^0
10^{-5}
10^{-10}
10^{-15}
10^{-20}
error
0
1000
2000
3000
4000
5000
6000
time (s)

Questions?









10^0
10^{-5}
10^{-10}
10^{-15}
10^{-20}
error
0
1000
2000
3000
4000
5000
6000
time (s)

References I

Anstreicher, K. M. (2017). Kronecker product constraints with an application to the two-trust-region subproblem. *SIAM Journal on Optimization*, 27(1):368–378.

Argue, C., K.-K., F., and Wang, A. L. (2022+). Necessary and sufficient conditions for rank-one generated cones. *Math. Oper. Res.*, Forthcoming, (arXiv:2007.07433).

Azuma, G., Fukuda, M., Kim, S., and Yamashita, M. (2022a). Exact sdp relaxations for quadratic programs with bipartite graph structures. *Journal of Global Optimization*, pages 1–21.

Azuma, G., Fukuda, M., Kim, S., and Yamashita, M. (2022b). Exact sdp relaxations of quadratically constrained quadratic programs with forest structures. *Journal of Global Optimization*, pages 1–20.

Barvinok, A. I. (2001). A remark on the rank of positive semidefinite matrices subject to affine constraints.

Beck, A. (2007). Quadratic matrix programming. *SIAM J. Optim.*, 17(4):1224–1238.

Beck, A., Drori, Y., and Teboulle, M. (2012). A new semidefinite programming relaxation scheme for a class of quadratic matrix problems. *Oper. Res. Lett.*, 40(4):298–302.

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

References II

Bomze, I. M., Jeyakumar, V., and Li, G. (2018). Extended trust-region problems with one or two balls: exact copositive and lagrangian relaxations. *J. Global Optim.*, 71(3):551–569.

Boumal, N., Voroninski, V., and Bandeira, A. (2016). The non-convex Burer–Monteiro approach works on smooth semidefinite programs. volume 29.

Boumal, N., Voroninski, V., and Bandeira, A. S. (2020). Deterministic guarantees for burer-monteiro factorizations of smooth semidefinite programs. *Communications on Pure and Applied Mathematics*, 73(3):581–608.

Burer, S. (2015). A gentle, geometric introduction to copositive optimization. *Math. Program.*, 151:89–116.

Burer, S. (2023). A slightly lifted convex relaxation for nonconvex quadratic programming with ball constraints. *arXiv preprint arXiv:2303.01624*.

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 Monteiro, R. D. (2003). A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Math. Program.*, 95:329–357.

Burer, S. and Ye, Y. (2019). Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs. *Math. Program.*, 181:1–17.

Carmon, Y. and Duchi, J. C. (2018). Analysis of Krylov subspace solutions of regularized nonconvex quadratic problems. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pages 10728–10738.

Devolder, O., Glineur, F., and Nesterov, Y. (2013). First-order methods with inexact oracle: the strongly convex case. Technical Report 2013016.

Devolder, O., Glineur, F., and Nesterov, Y. (2014). First-order methods of smooth convex optimization with inexact oracle. *Math. Program.*, 146(1):37–75.

Ding, L., Yurtsever, A., Cevher, V., Tropp, J. A., and Udell, M. (2021). An optimal-storage approach to semidefinite programming using approximate complementarity. *SIAM J. Optim.*, 31(4):2695–2725.

Ekeland, I. and Temam, R. (1999). *Convex Analysis and Variational Problems*, volume 28 of *Classics Appl. Math.* SIAM.

Fradkov, A. L. and Yakubovich, V. A. (1979). The S-procedure and duality relations in nonconvex problems of quadratic programming. *Vestnik Leningrad Univ. Math.*, 6:101–109.

References IV

Fujie, T. and Kojima, M. (1997). Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4):367–380.

Hazan, E. and Koren, T. (2016). A linear-time algorithm for trust region problems. *Math. Program.*, 158:363–381.

Ho-Nguyen, N. and K.-K., F. (2017). A second-order cone based approach for solving the Trust Region Subproblem and its variants. *SIAM J. Optim.*, 27(3):1485–1512.

Jeyakumar, V. and Li, G. Y. (2014). Trust-region problems with linear inequality constraints: Exact SDP relaxation, global optimality and robust optimization. *Math. Program.*, 147:171–206.

Jiang, R. and Li, D. (2019). Novel reformulations and efficient algorithms for the Generalized Trust Region Subproblem. *SIAM J. Optim.*, 29(2):1603–1633.

Jiang, R. and Li, D. (2020). A linear-time algorithm for generalized trust region problems. *SIAM J. Optim.*, 30(1):915–932.

K.-K., F. and Wang, A. L. (2021). Exactness in SDP relaxations of QCQPs: Theory and applications. Tut. in Oper. Res. INFORMS.

Kelly, S., Ouyang, Y., and Yang, B. (2022). A note on semidefinite representable reformulations for two variants of the trust-region subproblem. *Manuscript, School of Mathematical and Statistical Sciences, Clemson University, Clemson, South Carolina, USA*.

References V

Locatelli, M. (2016). Exactness conditions for an SDP relaxation of the extended trust region problem. *Oper. Res. Lett.*, 10(6):1141–1151.

Locatelli, M. (2022). Kkt-based primal-dual exactness conditions for the shor relaxation. *Journal of Global Optimization*, pages 1–17.

Nesterov, Y. (2005). Excessive gap technique in nonsmooth convex minimization. *SIAM J. Optim.*, 16(1):235–249.

Nesterov, Y. (2018). *Lectures on convex optimization*, volume 137 of *Springer Optimization and Its Applications*. Springer.

O'Donoghue, B., Chu, E., Parikh, N., and Boyd, S. (2016). Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068.

Pataki, G. (1998). On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358.

Sojoudi, S. and Lavaei, J. (2014). Exactness of semidefinite relaxations for nonlinear optimization problems with underlying graph structure. *SIAM J. Optim.*, 24(4):1746–1778.

References VI

Souto, M., Garcia, J. D., and Veiga, A. (2020). Exploiting low-rank structure in semidefinite programming by approximate operator splitting. *Optimization*, pages 1–28.

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

Wang, A. L. and K.-K., F. (2020). A geometric view of SDP exactness in QCQPs and its applications. *arXiv preprint*, 2011.07155.

Wang, A. L. and K.-K., F. (2022a). Accelerated first-order methods for a class of semidefinite programs. *arXiv preprint*, 2206.00224.

Wang, A. L. and K.-K., F. (2022b). The generalized trust region subproblem: solution complexity and convex hull results. *Math. Program.*, 191:445–486.

Wang, A. L. and K.-K., F. (2022c). On the tightness of SDP relaxations of QCQPs. *Math. Program.*, 193:33–73.

Wang, A. L., Lu, Y., and K.-K., F. (2023+). Implicit regularity and linear convergence rates for the generalized trust-region subproblem. *SIAM J. Optim.*, Forthcoming, (arXiv:2112.13821).

Yakubovich, V. A. (1971). S-procedure in nonlinear control theory. *Vestnik Leningrad Univ. Math.*, pages 62–77.

Yang, B., Anstreicher, K., and Burer, S. (2018). Quadratic programs with hollows. *Math. Program.*, 170:541–553.

Yıldıran, U. (2009). Convex hull of two quadratic constraints is an LMI set. *IMA J. Math. Control Inform.*, 26(4):417–450.

Yurtsever, A., Tropp, J. A., Fercoq, O., Udell, M., and Cevher, V. (2021). Scalable semidefinite programming. *SIAM Journal on Mathematics of Data Science*, 3(1):171–200.