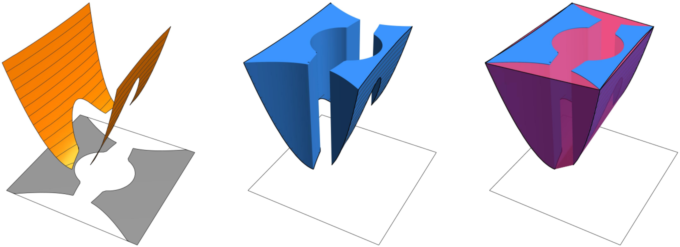
Exactness in SDP relaxations of QCQPs *now with 50% more pictures!

Alex L. Wang , PSE Seminar, Oct. 20



Supported in part by NSF grant CMMI 1454548 and ONR grant N00014-19-1-2321

Quadratically constrained quadratic programs (QCQPs)

- Family of highly expressive optimization problems
- Computer science

  MAX-CUT, MAX-CLIQUE
- Operations research

  Facility location, production planning
- Engineering

  Pooling problem, truss design problem
- More generally

  Binary programming, polynomial programming

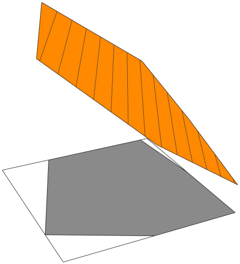
From linear programs to quadratic programs

- Linear programs (LPs)
- $\ell_0, \ell_1, \ldots, \ell_m : \mathbb{R}^n \to \mathbb{R}$ linear functions

$$\ell_i(x) = b_i^\top x + c_i$$

- Want to find

$$\text{Opt} := \min_{x \in \mathbb{R}^n} \left\{ \ell_0(x) : \begin{array}{c} \ell_1(x) \leq 0 \\ \vdots \\ \ell_m(x) \leq 0 \end{array} \right\}$$



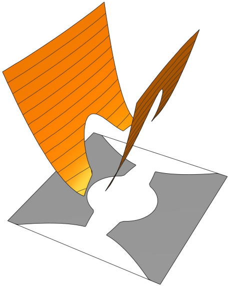
From linear programs to quadratic programs

- Quadratically constrained quadratic programs (QCQPs)
- $q_0, q_1, \ldots, q_m : \mathbb{R}^n \to \mathbb{R}$ quadratic functions

$$q_i(x) = \boxed{x^\top A_i x} + 2b_i^\top x + c_i$$

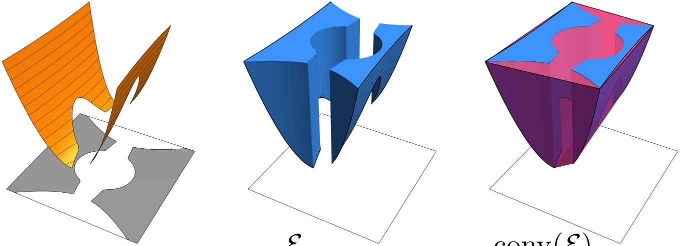
- Want to find

$$\mathrm{Opt} := \min_{x \in \mathbb{R}^n} \left\{ q_0(x) : \begin{array}{c} q_1(x) \leq 0 \\ \vdots \\ q_m(x) \leq 0 \end{array} \right\}$$



The QCQP Epigraph

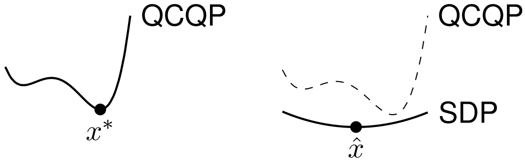
$$\begin{aligned}
\mathrm{Opt} &= \min_{x\in\mathbb{R}^n} \{q_0(x) : \ q_i(x) \le 0, \ \forall i \in [m]\} \\
&= \min_{x,t} \left\{ t : \ \begin{array}{l} q_0(x) \le t \\ q_i(x) \le 0, \ \forall i \in [m] \end{array} \right\} =: \min_{x,t} \left\{ t : \ (x,t) \in \mathcal{E} \right\} \\
&= \min_{x,t} \{ t : \ (x,t) \in \mathrm{conv}(\mathcal{E}) \}
\end{aligned}$$



$\mathcal{E}$ $\qquad$ $\mathrm{conv}(\mathcal{E})$

Convex relaxations

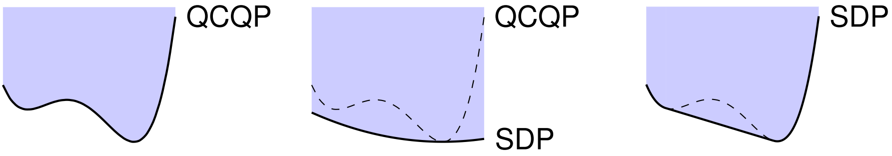
- QCQPs are NP-hard in general ☹
- One issue with QCQPs is nonconvexity!
- Will look for a convex *relaxation*



- SDP relaxation can be solved efficiently

Convex relaxations

- Properties you might want for a convex relaxation
- Types of "exactness"



Objective value

Convex hull

- **Q:** When do these properties hold?

The standard SDP relaxation of QCQP

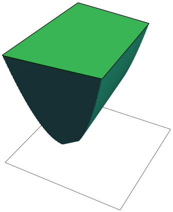
Standard semidefinite program (SDP) relaxation

$$
\begin{aligned}
\mathrm{Opt} &= \min_x \left\{ q_0(x) : \quad q_i(x) \le 0,\ \forall i \ \right\} \\
&= \min_x \left\{ x^\top A_0 x + 2b_0^\top x + c_0 : \quad x^\top A_i x + 2b_i^\top x + c_i \le 0,\ \forall i \ \right\} \\
&= \min_x \left\{ \langle A_0, Y\rangle + 2b_0^\top x + c_0 : \quad \begin{array}{l} \boxed{Y = xx^\top} \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \end{array} \right\} \\
&\ge \min_x \left\{ \langle A_0, Y\rangle + 2b_0^\top x + c_0 : \quad \begin{array}{l} \boxed{\exists\, Y \succeq xx^\top} \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \le 0,\ \forall i \end{array} \right\} \\
&=: \boxed{\mathrm{Opt}_{\mathsf{SDP}}}
\end{aligned}
$$

SDP epigraph

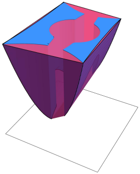
$$
\begin{aligned}
\mathrm{Opt}_{\mathsf{SDP}} &= \min_{x} \left\{ \langle A_0, Y\rangle + 2b_0^\top x + c_0 : \begin{array}{l} \exists Y \succeq xx^\top \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \end{array} \right\} \\
&= \min_{x,\, t} \left\{ t : \begin{array}{l} \exists Y \succeq xx^\top \\ \langle A_0, Y\rangle + 2b_0^\top x + c_0 \leq t \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \end{array} \right\}
\end{aligned}
$$

- Let $\mathcal{E}_{\mathsf{SDP}}$ denote the SDP epigraph

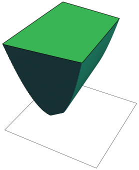


Recap

- QCQPs are highly expressive but NP-hard in general
- Solve SDP relaxation instead
- **Q:** What are sufficient conditions for
  - Convex hull exactness: $\mathrm{conv}(\mathcal{E}) = \mathcal{E}_{\mathsf{SDP}}$?



$=$



?

  - Objective value exactness: $\mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}$?

Outline

- Main result for today:

  If the quadratic forms "interact nicely" and each have "large amounts of symmetry", then convex hull exactness holds

  - The set of convex Lagrange multipliers
  - The quadratic eigenvalue multiplicity
- Example application: robust least squares
- Additional work in this area
- Future directions
- Sneak peek of (very) recent work

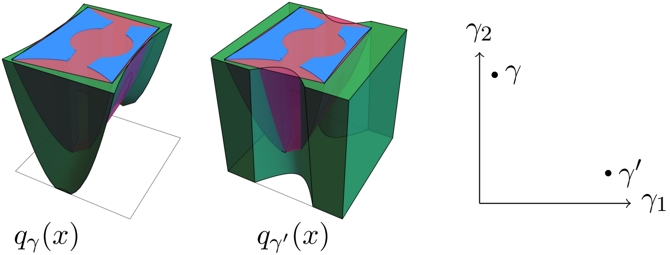
[**W** and Kılınç-Karzan 19]

Aggregation

- For $\gamma \in \mathbb{R}^m_+$, define

$$q_\gamma(x) := q_0(x) + \sum_{i=1}^m \gamma_i q_i(x)$$

- For all $\gamma \in \mathbb{R}^m_+$, $(x,t) \in \mathcal{E} \implies q_\gamma(x) \le t$



Aggregation

- For $\gamma \in \mathbb{R}^m_+$, define

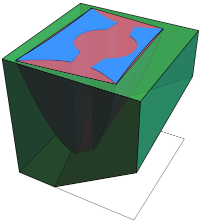
$$q_\gamma(x) \coloneqq q_0(x) + \sum_{i=1}^{m} \gamma_i q_i(x)$$

- Define

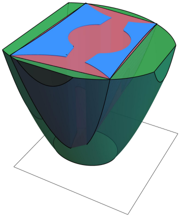
$$\Gamma \coloneqq \left\{\gamma \in \mathbb{R}^m_+ : \ q_\gamma(x) \text{ is convex}\right\}$$

- For all $\gamma \in \Gamma$,

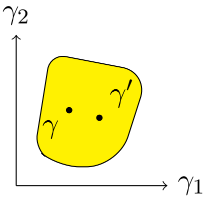
$$(x, t) \in \operatorname{conv}(\mathcal{E}) \implies q_\gamma(x) \le t$$



$q_\gamma(x)$



$q_{\gamma'}(x)$

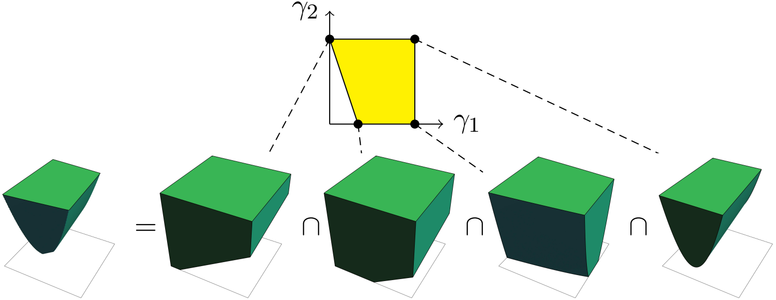


Rewriting the SDP in terms of Γ

Lemma

Suppose primal feasibility and dual strict feasibility, then

$$\mathcal{E}_{\mathsf{SDP}} = \left\{ (x,t) : \max_{\gamma \in \Gamma} q_\gamma(x) \le t \right\}$$

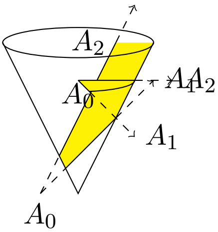


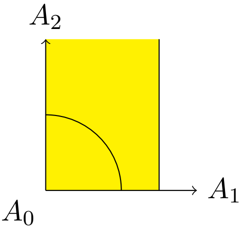
Related: [Fujie and Kojima 97]

What does Γ look like?

- Recall

$$\Gamma = \left\{ \gamma \in \mathbb{R}^m_+ : \; A_0 + \sum_{i=1}^m \gamma_i A_i \succeq 0 \right\}$$





- When A_is diagonal $\implies$ Γ is polyhedral

Quadratic eigenvalue multiplicity

Definition

Let $1 \le k \le n$ be the largest integer such that for each $i = 0, \ldots, m$, the matrix $A_i \in \mathbb{S}^n$ has the following block form

$$\boxed{A_i = \hat{A}_i \otimes I_k} = \begin{pmatrix} \hat{A}_i & & & \\ & \hat{A}_i & & \\ & & \ddots & \\ & & & \hat{A}_i \end{pmatrix}$$

where $\hat{A}_i \in \mathbb{S}^{n/k}$

Quadratic eigenvalue multiplicity

- $A_i = \hat{A}_i \otimes I_k$
- Suppose $n = 4$, i.e. $x \in \mathbb{R}^4$

$x_1^2 + x_2^2 + x_3^2 + x_4^2$	$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$	$k = 4$
$(x_1 - x_2)^2 + (x_3 - x_4)^2$	$\begin{pmatrix} 1 & -1 & & \\ -1 & 1 & & \\ & & 1 & -1 \\ & & -1 & 1 \end{pmatrix}$	$k = 2$

Some results

Corollary

Suppose primal feasibility and dual strict feasibility. If Γ is polyhedral and

$$k \geq \min\left(m,\ \left|\{b_i \neq 0\}_{i=1}^m\right| + 1\right),$$

then $$\operatorname{conv}(\mathcal{E}) = \mathcal{E}_{\mathsf{SDP}} \quad \text{and} \quad \mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}\,.$$

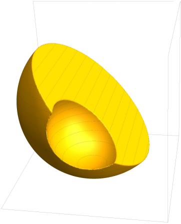
- $m = 1$
- A_is diagonal and $b_1 = b_2 = \cdots = b_m = 0$
- $A_i = \alpha_i I_n$ for all i and $n \geq m$

Related: [Yakubovich 79], [Burer and Ye 19]

Example: Swiss cheese

- Minimizing distance to a piece of Swiss cheese

$$\min_{x\in\mathbb{R}^n} \left\{ \|x\|^2 : \begin{array}{l} \text{inside ball constraints} \\ \text{outside ball constraints} \\ \text{linear constraints} \end{array} \right\}$$



- inside ball $\mapsto I$, outside ball $\mapsto -I$, linear constraints $\mapsto 0$
- If nonempty and $n \geq m$, then the standard SDP relaxation is tight for this QCQP

Some results

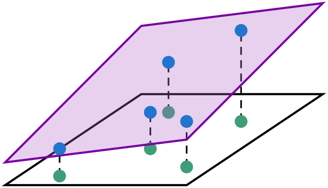
Corollary

Suppose primal feasibility and dual strict feasibility.

- If $k \geq m+2$, then $\mathrm{conv}(\mathcal{E}) = \mathcal{E}_{\mathsf{SDP}}$
- If $k \geq m+1$, then $\mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}$

Related: [Beck 07]

Least squares



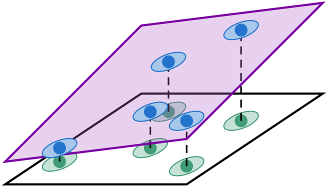
- Input, output pairs

$$Y^* = \begin{pmatrix} (y_1^*)^\top \\ \vdots \\ (y_n^*)^\top \end{pmatrix} \in \mathbb{R}^{n\times k}, \qquad z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{R}^n$$

- Want $Y^*\ell \approx z$

$$\min_{\ell\in\mathbb{R}^k} \|Y^*\ell - z\|_2^2$$

Robust least squares



- Empirical measurement $\hat{Y}$ and uncertainty $\mathcal{U}$, i.e., $Y^* \in \hat{Y} + \mathcal{U}$
- Want to minimize

$$\min_{\ell \in \mathbb{R}^k} \max_{X \in \mathcal{U}} \left\| (\hat{Y} + X)\ell - z \right\|_2^2$$

- When $\mathcal{U}$ is defined by quadratics, can apply our theory!

Robust Least Squares

- Suppose

$$\mathcal{U} = \left\{ X \in \mathbb{R}^{n\times k} :\ \|L_i X\|_F^2 \le c_i,\ \forall i \in [m] \right\}$$

- Consider

$$\max_{X\in\mathbb{R}^{n\times k}} \left\{ \left\| (\hat{Y} + X)\ell - z \right\|_2^2 :\ X \in \mathcal{U} \right\}$$

- Write as a QCQP in the variable $x \in \mathbb{R}^{nk}$
- Quadratic forms

$$L_i^\top L_i \otimes I_k$$

- $k \ge m + 2$ implies convex hull exactness
- $k \ge m + 1$ implies objective value exactness

Recap, additional work, future work

- QCQPs are NP-hard
- Analyzed SDP relaxation
  - Convex Lagrange multipliers,
  - "Amount of symmetry" k
- **Q:** When is the SDP relaxation exact?

	Polyhedral Γ	General Γ
Cvx. hull	$k \geq \min\left(m, \left\lvert\{b_i \neq 0\}_{i=1}^m\right\rvert + 1\right)$	$k \geq m + 2$
Obj. val.	$k \geq \min\left(m, \left\lvert\{b_i \neq 0\}_{i=1}^m\right\rvert + 1\right)$	$k \geq m + 1$

- Current and future work:
  - Is the polyhedral case fundamentally different from the general case?
  - Conditions that only depend on the constraints?
  - Can approximation results be explained in this framework?

Related: [Argue, Kılınç-Karzan, and **W** 20]

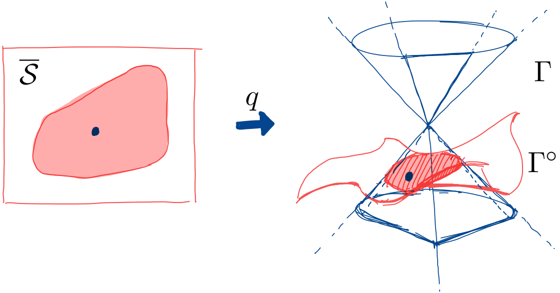
Thank you. Questions?

Slides
cs.cmu.edu/~alw1

Full version
A. L. Wang and F. Kılınç-Karzan. "On the tightness of SDP relaxations of QCQPs". In: *arXiv preprint* arXiv:1911.09195 (2019)

When Γ is non-polyhedral

- The geometry of Γ and Γ°
- Rounding scheme: Given $x \in \overline{\mathcal{S}}$,
  - Either output convex decomposition for $x \in \operatorname{conv}(\mathcal{S})$
  - Or, claim $x' \in \overline{\mathcal{S}} \setminus \operatorname{conv}(\mathcal{S})$.
- Correct when Γ° is facially exposed
- Extends framework to handle complementarity constraints $\longrightarrow$ sparse regression



References I

- C.J. Argue, F. Kılınç-Karzan, and A. L. Wang. “Necessary and sufficient conditions for rank-one generated cones”. In: *arXiv preprint* arXiv:2007.07433 (2020).
- A. Beck. “Quadratic matrix programming”. In: *SIAM J. Optim.* 17.4 (2007), pp. 1224–1238.
- S. Burer and Y. Ye. “Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs”. In: *Math. Program.* 181 (2019), pp. 1–17.
- T. Fujie and M. Kojima. “Semidefinite programming relaxation for nonconvex quadratic programs”. In: *J. Global Optim.* 10.4 (1997), pp. 367–380.
- V. A. Yakubovich. “S-procedure in nolinear control theory”. In: *Vestnik Leningrad Univ. Math.* (1971), pp. 73–93.