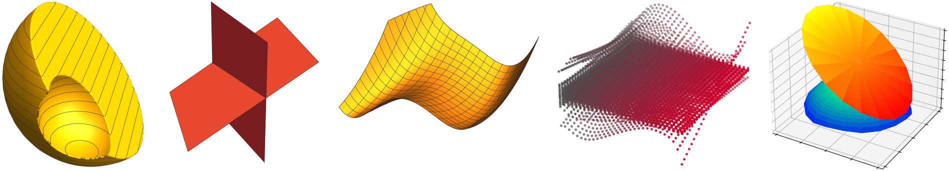
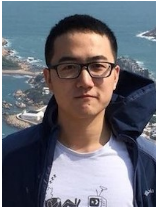
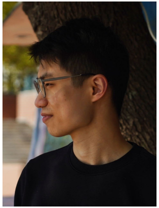
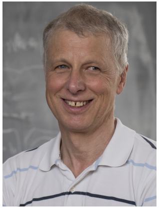
Accurately and efficiently solving structured nonconvex optimization problems

Alex L. Wang
Carnegie Mellon University



These slides are publicly available at `cs.cmu.edu/~alw1`

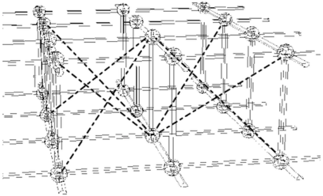




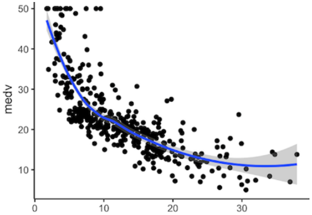



Carnegie Mellon University (OR, Math, CS), Northwestern University, Fudan University, Peking University

Convex optimization

- Convex optimization is influential in many different fields





Engineering
Controller stability, power allocation, truss design, +

Statistics
(Linear) Regression, parameter estimation, +

Finance
Portfolio optimization, risk analysis, +

- Convex optimization is accurate and efficient

Convex optimization, meet nonconvex problems

- Unfortunately, many practical optimization problems are nonconvex
- Example: Low-rank matrix completion (**Netflix problem**)

Movies

Users $\begin{pmatrix} 5 & 4 & ? & ? & 4 \\ 3 & ? & ? & 3 & ? \\ ? & 2 & 4 & 1 & 1 \\ ? & 3 & ? & ? & 4 \end{pmatrix}$

$$\min_{X \in \mathbb{R}^{n \times k}} \{\mathrm{rank}(X) : X \text{ agrees with revealed entries}\}$$

- Rank constraints, binary constraints, sparsity constraints $\longleftarrow$ generally hard
- **But not always!**
  Some nonconvex problems can be solved using convex optimization

Research goal

Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Completed work**:
  - Nonconvex problems: quadratically constrained quadratic programs (QCQPs)
  - Convex relaxations: semidefinite programs (SDPs)

Today's questions

Understand **structures within QCQPs** that enable us to solve them exactly and efficiently **using SDPs**

Today's outline

- **Preliminaries**

  QCQPs and their applications, the SDP relaxation

- **Understand structures within QCQPs that enable us to solve them...**

  - **exactly** [IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*]

    Objective value, convex hull exactness, applications

  - **efficiently** [Math. Prog. 20], [SIAM J. Optim. *under review*], [Ongoing]

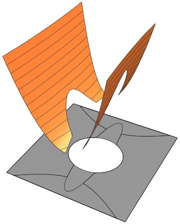
    The generalized trust-region subproblem and regular QCQPs

- **Conclusion and future directions**

1. Preliminaries
2. Objective value exactness, convex hull exactness, applications
3. Efficient algorithms for regular QCQPs
4. Conclusion and future directions

Quadratically constrained quadratic programs (QCQPs)

- $q_{\mathsf{obj}}, q_1, \dots, q_m : \mathbb{R}^n \to \mathbb{R}$ quadratic (possibly nonconvex!)



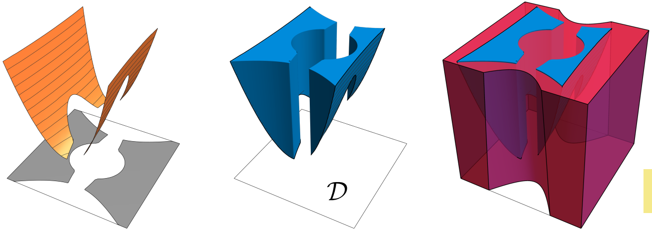
$$\mathrm{Opt} := \inf_{x \in \mathbb{R}^n} \left\{ q_{\mathsf{obj}}(x) : \ q_i(x) \le 0, \ \forall i \in [m] \right\}$$

$$q_i(x) = x^\intercal A_i x + 2 b_i^\intercal x + c_i$$

- Highly expressive:
  - MAX-CUT, MAX-CLIQUE, pooling, truss design, facility location, production planning
  - binary programs $x_1(1 - x_1) = 0$
  - polynomial optimization problems $x_1 x_2 = z_{12}$
- NP-hard in general

The QCQP epigraph

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

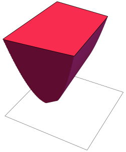
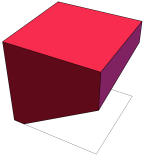
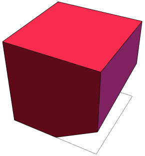
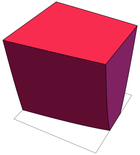
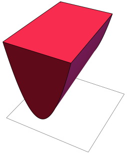


$q(\gamma', x) \leq t$

- How can we derive ~~convex~~ relaxations of $\mathcal{D}$?
- If $\gamma \in \mathbb{R}^m_+$, then $\forall (x,t) \in \mathcal{D}, \quad \underbrace{q_{\text{obj}}(x) + \sum_{i=1}^m \gamma_i q_i(x)}_{=:\, q(\gamma, x)} \leq t$

The SDP relaxation

- SDP relaxation = impose all convex aggregated inequalities!

$\mathcal{D}_{\text{SDP}}$  $=$  $\cap$  $\cap$  $\cap$ 

- Formally,

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

$$\mathcal{D}_{\text{SDP}} := \bigcap_{\gamma\in\Gamma} \{(x,t) : q(\gamma, x) \le t\} = \left\{(x,t) \in \mathbb{R}^{n+1} : \sup_{\gamma\in\Gamma} q(\gamma, x) \le t\right\}$$

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

The usual SDP relaxation

$$\inf_{x\in\mathbb{R}^n} \{q_{\mathrm{obj}}(x) : q_i(x) \leq 0,\ \forall i \in [m]\}$$

$$= \inf_{x\in\mathbb{R}^n, Y\in\mathbb{S}^n} \left\{ \langle A_{\mathrm{obj}}, Y\rangle + 2b_{\mathrm{obj}}^\top x + c_{\mathrm{obj}} : \begin{array}{l} Y = xx^\top \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \in [m] \end{array} \right\}$$

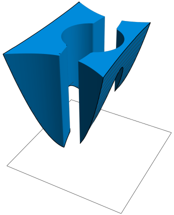
$$\geq \inf_{x\in\mathbb{R}^n, Y\in\mathbb{S}^n} \left\{ \langle A_{\mathrm{obj}}, Y\rangle + 2b_{\mathrm{obj}}^\top x + c_{\mathrm{obj}} : \begin{array}{l} Y - xx^\top \succeq 0 \\ \langle A_i, Y\rangle + 2b_i^\top x + c_i \leq 0,\ \forall i \in [m] \end{array} \right\}$$

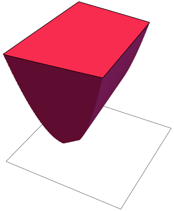
$$= \inf_{x\in\mathbb{R}^n} \inf_{Y\in\mathbb{S}^n} \dots$$

$$= \inf_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma} q(\gamma, x)$$

- Main objects of interest

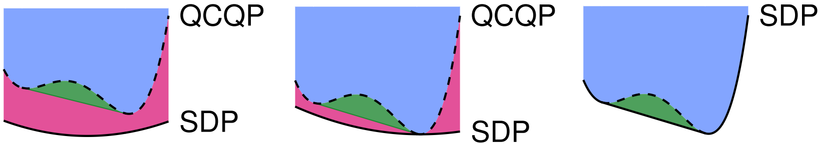
	nonconvex QCQP	convex SDP
Optimum value	Opt	$\mathrm{Opt}_{\mathsf{SDP}}$
Epigraph	$\mathcal{D}$	$\mathcal{D}_{\mathsf{SDP}}$





- Useful for analysis:
  - $q(\gamma, x) =$ Lagrangian function
  - $\Gamma =$ aggregation weights such that $q(\gamma, x)$ is convex

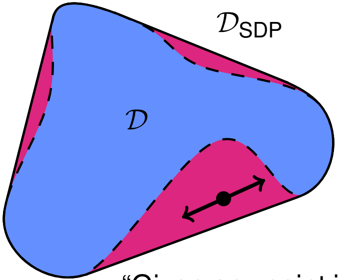
- What does exactness mean?
    - Objective value exactness: $\mathrm{Opt} = \mathrm{Opt}_{\mathsf{SDP}}$
    - Convex hull exactness: $\mathrm{conv}(\mathcal{D}) = \mathcal{D}_{\mathsf{SDP}}$ ⟵ convexification of substructures



Obj. val. ex. ✗	Obj. val. ex. ✓	Obj. val. ex. ✓
Conv. hull ex. ✗	Conv. hull ex. ✗	Conv. hull ex. ✓

Convex hull exactness

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



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

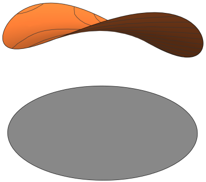
- When do these directions exist? $\longleftarrow$ Can carry out this idea for QCQPs!
- Sufficient conditions based on abstract properties $\longrightarrow$ concrete conditions

Based on: [IPCO 19], [Math. Prog. 21], [Math. Prog. *under review*]

Example: the trust-region subproblem

- Convex hull exactness in the case of single ball constraint

$$\text{Opt} = \inf_{x \in \mathbb{R}^n} \left\{ q_{\text{obj}}(x) : \|x\|^2 \leq 1 \right\}$$



- **Applications**:
  - Nonlinear minimization (trust-region methods), combinatorial optimization, robust optimization

Based on: **[IPCO 19]**, **[Math. Prog. 20]**
Related: Yakubovich [1971], Yıldıran [2009], Ho-Nguyen and Kılınç-Karzan [2017]

Example: QCQPs with symmetry

- Convex hull exactness in the case of "highly symmetric" QCQPs
- Suppose $A_{\text{obj}} = I_k \otimes \mathbb{A}_{\text{obj}}$, $A_i = I_k \otimes \mathbb{A}_i$ for all $i \in [m]$

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

  and $k \geq m$
- **Applications**:
  - Robust least squares, sphere packing, QCQPs with spherical constraints, orthogonal Procrustes problem

Based on: **[Math. Prog. *under review*]**
Related: Beck [2007], Beck et al. [2012]

Example: QCQPs with sign-definite linear terms

- Objective value exactness in the case of diagonal, sign-definite QCQPs
- Suppose $A_{\text{obj}}, \ldots, A_m$ diagonal and

$$\forall j \in [n], \left\{(b_{\text{obj}})_j, (b_1)_j, \ldots, (b_m)_j\right\} \text{ have the same sign}$$

- Example:

$$\min_{x \in \mathbb{R}^n} \left\{x^\top A_{\text{obj}} x + 2b_{\text{obj}}^\top x + c_{\text{obj}} : \|x\|_2 \leq 1, \|x\|_\infty \leq \alpha\right\}$$

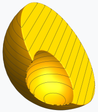
Based on: **[Tut. Oper. Res. 21]**

Related: Burer and Ye [2019], Sojoudi and Lavaei [2014]

Summary of Part 1

- Sufficient conditions for convex hull exactness
- **Necessary and sufficient** if Γ is polyhedral (dual facially exposed)
- Sufficient conditions for objective value exactness
- Rank-one-generated (ROG) cones: QCQP-SDP analogue of integrality
- **Applications**:
  - Random, semi-random QCQPs, ratios of quadratic functions

Exactness	ROG Cones
[IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*]	[Tut. Oper. Res. 21], [Math. Oper. Res. 21]

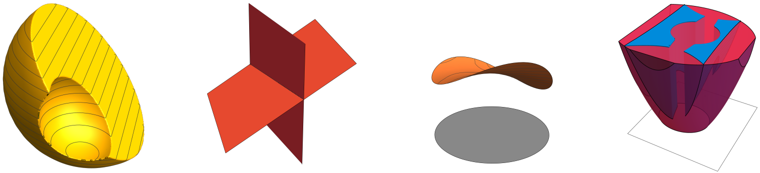


Take-home message

Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

SDPs provide exact reformulations for broad classes of QCQPs!



Revisiting the SDP relaxation

- SDPs polynomial time $\longleftarrow$ too expensive in modern machine learning regimes
- Usual SDP relaxation
  - Interior point method $\longrightarrow$ iterations expensive $O(mn^3 + m^2n^2 + m^3)$ time

We can solve an SDP more efficiently if it is exact (regular)!

- Our view:
$$\mathrm{Opt}_{\mathsf{SDP}} := \inf_{x \in \mathbb{R}^n} \left(\sup_{\gamma \in \Gamma} q(\gamma, x) \right)$$
  is a minimization problem in the original space
- Regularity will allow us to deal with max-type structure

Based on: [Math. Prog. 20], [SIAM J. Optim. *under review*], [Ongoing]

- Dual problem

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

Definition

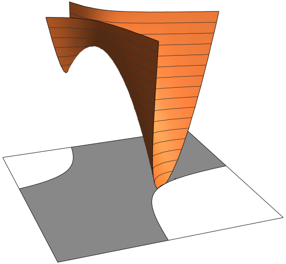
Let γ^* be dual optimizer. Define $\mu^* := \lambda_{\min}\left(A_{\mathsf{obj}} + \sum_{i=1}^m \gamma_i^* A_i\right)$.
QCQP is regular if $\mu^* > 0$.

- Regularity $\Longrightarrow$ optimizer exactness

$$\mu^* > 0 \quad \Longrightarrow \quad \arg\min_{x\in\mathbb{R}^n} \{q_{\mathsf{obj}}(x) : q_i(x) \leq 0,\ \forall i \in [m]\} = \arg\min_{x\in\mathbb{R}^n} \sup_{\gamma\in\Gamma} q(\gamma, x)$$

Based on: [Ongoing]

The generalized trust-region subproblem (GTRS)



- Special setting with single constraint ($\leq$ or $=$)

$$\text{Opt} := \inf_{x \in \mathbb{R}^n} \{q_{\text{obj}}(x) : q_1(x) \leq 0\}$$

- **TRS applications**:
  - nonlinear programming (trust-region methods), combinatorial optimization, robust optimization

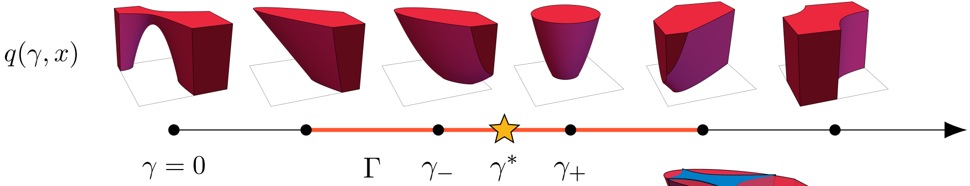
- **GTRS applications**:
  - minimizing quartics of the form $q(x, p(x))$

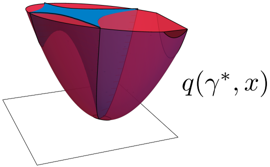
$$\inf_{x \in \mathbb{R}^n, \alpha} \{q(x, \alpha) : \alpha = p(x)\}$$

    (source localization, constrained rank-one approximation), regression with adversarial data, iterative QCQP solvers
- Assume $\mu^* > 0 \longleftarrow$ most GTRS

Based on: **[Math. Prog. 20]**, **[SIAM J. Optim. *under review*]**, **[*under review*]**

- $\Gamma = \{\gamma \in \mathbb{R}_+ : q(\gamma, x) \text{ is convex in } x\}$



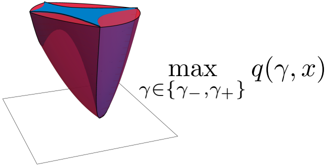


- **Thought experiment**: If γ^* known

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

- **Key observation**: If $\gamma^* \in [\gamma_-, \gamma_+] \subseteq \Gamma$, then

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

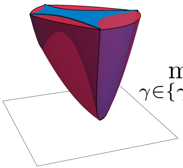


Based on: [SIAM J. Optim. *under review*]

Efficient algorithms for regular GTRS

- **Key observation**: If $\gamma^* \in [\gamma_-, \gamma_+] \subseteq \Gamma$, then

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



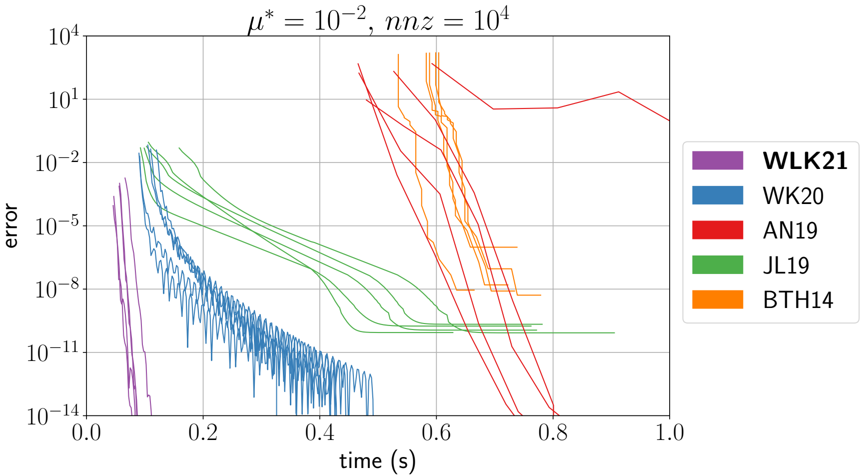
- **Algorithmic idea**:
  - Solve for γ^* to low accuracy, $\gamma^* \in [\gamma_-, \gamma_+] \subseteq \mathrm{int}(\Gamma)$
  - Apply Accelerated Gradient Descent
    $\longleftarrow$ for strongly convex nonsmooth function
- Putting pieces together:

$$\Longrightarrow \quad \tilde{O}\left(\frac{T}{\sqrt{\mu^*}} \log\left(\frac{1}{\epsilon}\right)\right)$$

where T is time for matrix vector product $\longleftarrow$ think $O(n)$

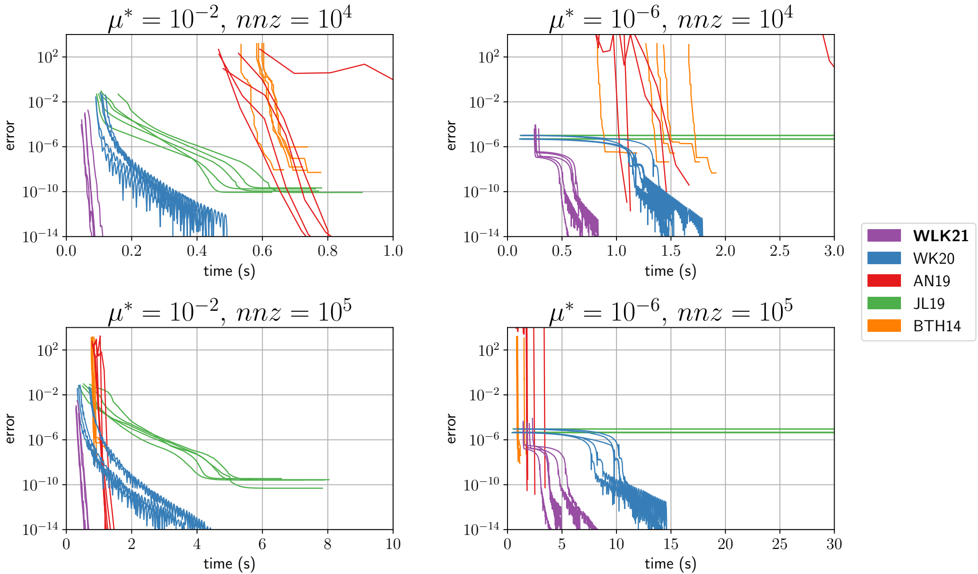
Based on: **[SIAM J. Optim. *under review*]**
Related: Carmon and Duchi [2018], Jiang and Li [2019], Adachi and Nakatsukasa [2019]

Numerical experiments, n = 1,000

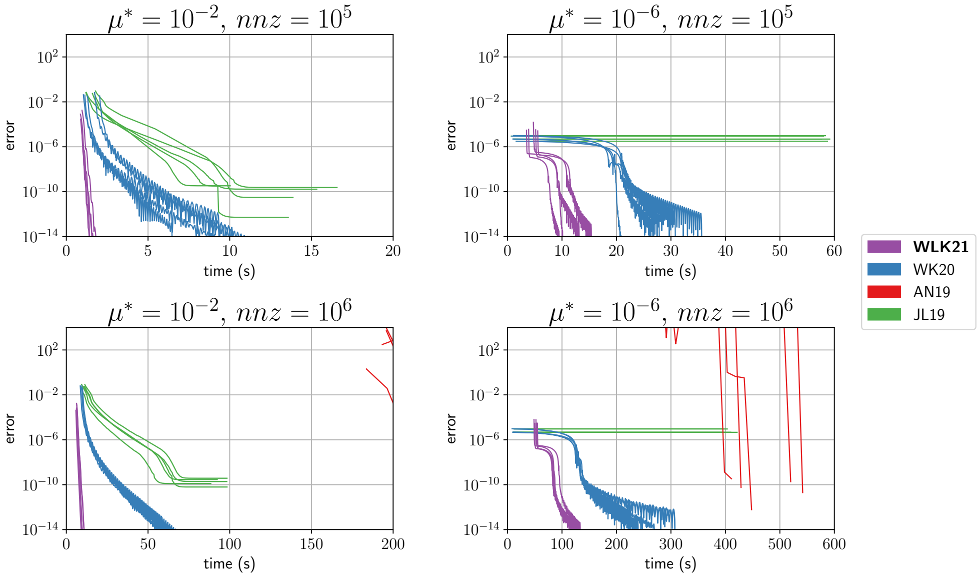


Based on: **[Math. Prog. 20]**, **[SIAM J. Optim. *under review*]**
Related: Ben-Tal and den Hertog [2014], Jiang and Li [2019], Adachi and Nakatsukasa [2019]

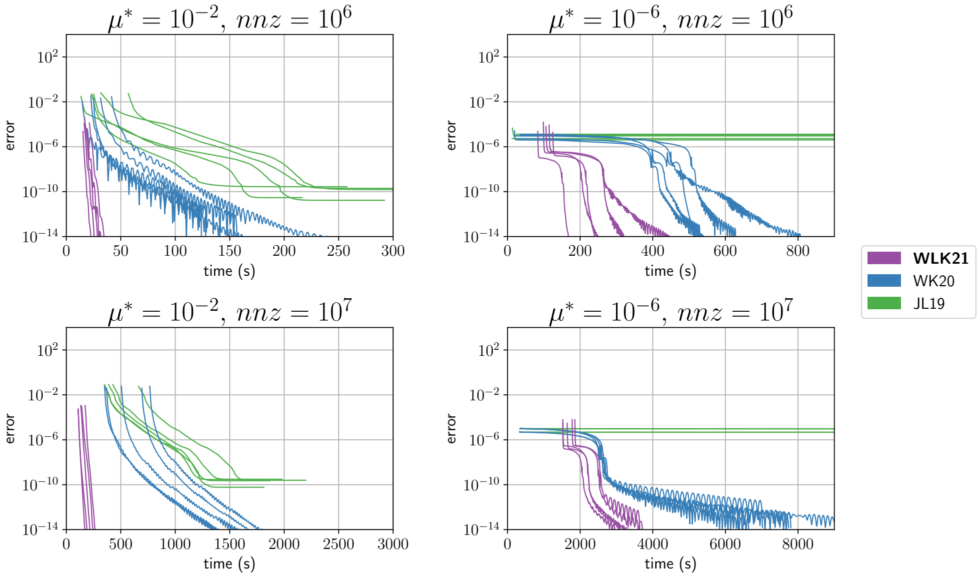
Numerical experiments, n = 1,000



Numerical experiments, n = 10,000



Numerical experiments, n = 100,000



Efficient algorithms for regular QCQPs: Algorithm

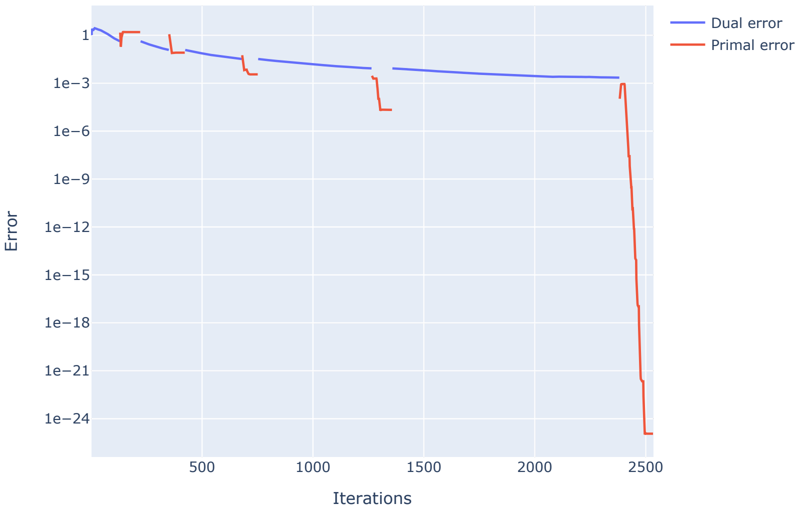
- Regularity $\Longrightarrow$ optimizer exactness
- Regularity holds in a number of statistical recovery problems: phase-retrieval, clustering
- Will leverage regularity to design efficient algorithms
- **Key observation**: If $\gamma^* \in \mathcal{U} \subseteq \Gamma$, then

$$\mathrm{Opt}_{\mathsf{SDP}} = \min_{x\in\mathbb{R}^n} \max_{\gamma\in\mathcal{U}} q(\gamma, x)$$

- **Algorithm sketch**:
  - Construct $\mathcal{U}$ $\longrightarrow$ $O(1)$ iterations, $O\left(\frac{mT}{\sqrt{\epsilon}}\right)$ / iter.
  - Solve min-max problem $\longrightarrow$ $O\left(\frac{1}{\sqrt{\mu^*}}\log\left(\frac{1}{\epsilon}\right)\right)$ iterations, $O\left(\frac{mT}{\epsilon}\right)$ / iter.

Based on: [Ongoing]

Convergence behavior



Preliminary numerical experiments

- 10 synthetic instances: $n = 5000$, $m = 50$, density $= 0.001$, $\mu^* = 0.1$
- Primal-dual solver

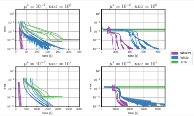
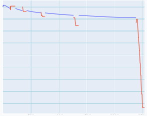
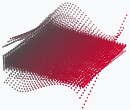
Section	ncalls	avg time	%tot	avg error
Total	10	1740s	100%	
dual_solve	10	1691s	97.2%	1.15e-02
eigsolve	33.1k	506ms	96.0%	
primal_solve	10	48.7s	2.80%	3.61e-12

- Splitting Conic Solver (SCS): `avg time 18150s`

Based on: [Ongoing]

Summary of Part 2

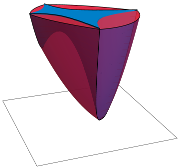
- Efficient algorithms for the GTRS
  $\longrightarrow$ nonlinear programming, iterative QCQP solvers, regression
- Efficient algorithms for regular QCQPs (low-rank SDPs)
  $\longrightarrow$ statistical recovery problems
- Algorithms for diagonalizing QCQPs

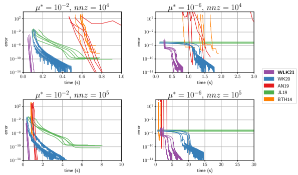
GTRS	Regular QCQPs	Diagonalizing QCQPs
[Math. Prog. 20], [SIAM J. Optim. *under review*], [*under review*]	[Ongoing]	[Math. Prog. *under review*]
		

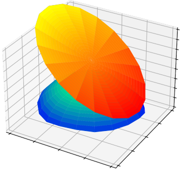
Long-term research goal

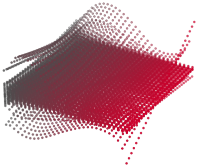
Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

Some nonconvex problems can be solved efficiently via first-order methods!









Future work

Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Solving nonconvex problems accurately**
  - **Completed**: SDPs provide exact reformulations for broad classes of QCQPs!
  - **Future**:
    - Can we understand approximation quality systematically within general framework?
    - Can we understand exactness/approximation for other convex relaxations?

Future work

Long-term research goal

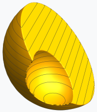
Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

- **Solving nonconvex problems efficiently**
  - **Completed**: Some nonconvex problems can be solved efficiently via first-order methods!
  - **Future**:
    - Exactness $\approx$ efficiency?
    - Can we develop efficient algorithms for semidefinite programs with low-rank solutions
    - Can we approximate "expensive" tools (e.g., SDPs) with cheap tools (e.g., linear programs, second-order cone programs)

Summary of my research

Exactness

[IPCO 20], [Math. Prog. 21], [Math. Prog. *under review*]

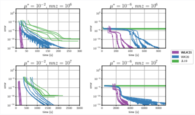


ROG Cones

[Tut. Oper. Res. 21], [Math. Oper. Res. 21]

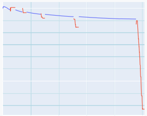
GTRS

[Math. Prog. 20], [SIAM J. Optim. *under review*], [*under review*]



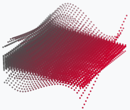
Regular QCQPs

[Ongoing]



Diagonalizing QCQPs

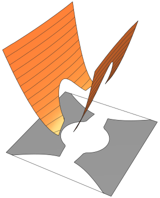
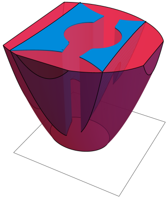
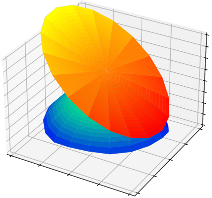
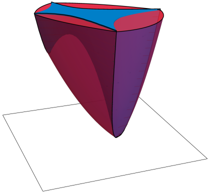
[Math. Prog. *under review*]



Long-term research goal

Understand **structures within nonconvex problems** that enable us to solve them "well" using **convex optimization**

Thank you! Questions?

References I

Adachi, S. and Nakatsukasa, Y. (2019). Eigenvalue-based algorithm and analysis for nonconvex QCQP with one constraint. *Math. Program.*, 173:79–116.

Argue, C., Kılınç-Karzan, F., and Wang, A. L. (2020). Necessary and sufficient conditions for rank-one generated cones. *arXiv preprint*, 2007.07433.

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.

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. pages 10728–10738.

Dutta, A., Vijayaraghavan, A., and Wang, A. L. (2017). Clustering stable instances of Euclidean k-means. In *Advances in Neural Information Processing Systems*, pages 6500–6509.

References II

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

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

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

Miller, G. L., Walkington, N. J., and Wang, A. L. (2019). Hardy-Muckenhoupt bounds for Laplacian eigenvalues. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, pages 8:1–8:19.

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.

Wang, A. L. and Jiang, R. (2021). New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs. *arXiv preprint*, 2101.12141.

Wang, A. L. and Kılınç-Karzan, F. (2020a). The generalized trust region subproblem: solution complexity and convex hull results. *Math. Program.* Forthcoming.

References III

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

Wang, A. L. and Kılınç-Karzan, F. (2020c). On convex hulls of epigraphs of QCQPs. In *Integer Programming and Combinatorial Optimization (IPCO 2020)*, pages 419–432. Springer.

Wang, A. L. and Kılınç-Karzan, F. (2021). On the tightness of SDP relaxations of QCQPs. *Math. Program.* Forthcoming.

Wang, J., Huang, W., Jiang, R., Li, X., and Wang, A. L. (2022). Solving stackelberg prediction games with least squares loss via spherically constrained least squares.

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

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

Rank-one-generated cones

Definition

Cone $\mathcal{S} \subseteq \mathbb{S}^n_+$ is rank-one-generated (ROG) if $\mathcal{S} = \operatorname{conv}\left(\mathcal{S} \cap \left\{xx^\top\right\}\right)$.

Compare: $P \subseteq [0,1]^n$ is integral if $P = \operatorname{conv}(P \cap \{0,1\}^n)$

- Given QCQP, if constraints correspond to ROG cone, then objective value exactness and convex hull exactness regardless of objective function
- Suppose $\mathcal{S} = \left\{X \in \mathbb{S}^n_+ : \langle M, X\rangle \leq 0,\ \forall M \in \mathcal{M}\right\}$

Goal

What properties of $\mathcal{M} = \{M_1, \dots, M_k\}$ imply $\mathcal{S}$ is ROG?

Based on: [Math. Oper. Res. 21], [Tut. Oper. Res. 21]

Rank-one-generated cones

Theorem (Sufficient conditions)

$\mathcal{S}$ is ROG if

- for all $i \neq j$, there exists $(\alpha, \beta) \neq (0,0)$ such that $\alpha M_i + \beta M_j \succeq 0$, or
- there exists $a \in \mathbb{R}^n$ such that $M_i = ab_i^\top + b_i a^\top$.

Theorem (Characterization of ROG for $|\mathcal{M}| = 2$)

Suppose $\mathcal{M} = \{M_1, M_2\}$. Then sufficient condition above is also necessary.

Based on: [Math. Oper. Res. 21], [Tut. Oper. Res. 21]

ROG and ratios of quadratic functions

- $$\begin{aligned} \inf_{z\in\mathbb{R}^{n+1}} &\left\{ \frac{z^\top M_{\mathsf{obj}} z}{z^\top B z} : \begin{array}{l} z^\top M_i z \le 0,\ \forall i\in[m] \\ z^\top B z > 0 \\ z_{n+1}^2 = 1 \end{array} \right\} \\ = \inf_{\tilde z\in\mathbb{R}^{n+1}} &\left\{ \tilde z^\top M_{\mathsf{obj}} \tilde z : \begin{array}{l} \tilde z^\top M_i \tilde z \le 0,\ \forall i\in[m] \\ \tilde z^\top B \tilde z = 1 \\ \tilde z_{n+1}^2 > 0 \end{array} \right\} \\ \ge \inf_{Z\in\mathbb{S}^{n+1}_+} &\left\{ \langle M_{\mathsf{obj}}, Z\rangle : \begin{array}{l} \langle M_i, Z\rangle \le 0,\ \forall i\in[m] \\ \langle B, Z\rangle = 1 \end{array} \right\} \end{aligned}$$
- Equality holds if $\mathcal{S}(\{M_1,\dots,M_m\})$ is ROG (+ minor assumptions)
- Example: Regularized total least squares

Based on: [Math. Oper. Res. 21], [Tut. Oper. Res. 21]

Stackelberg prediction games with least squares losses

- True data $(x_i, \alpha_i, \beta_i)_{i=1}^m$
- Leader (learner) chooses $w \in \mathbb{R}^n$
- Follower (data provider) modifies $x_i \to \tilde{x}_i$ so that $\langle w, \tilde{x}_i \rangle \approx \beta_i$
- Leader has loss $(\langle w, \tilde{x}_i \rangle - \alpha_i)^2$
- $$\min_{w \in \mathbb{R}^n} \left\{ \sum_{i=1}^m (\langle w, \tilde{x}_i \rangle - \alpha_i)^2 : \tilde{x}_i \in \arg\min_{x \in \mathbb{R}^n} \gamma \left\| x - x_i \right\|^2 + (\langle w, x \rangle - \beta_i)^2 \right\}$$

Based on: ***[under review]***

Stackelberg prediction games with least squares losses

- $$\begin{aligned}
&\min_{w\in\mathbb{R}^n} \left\{ \sum_{i=1}^m (\langle w, \tilde{x}_i\rangle - \alpha_i)^2 :\ \tilde{x}_i \in \arg\min_{x\in\mathbb{R}^n} \gamma \left\| x - x_i \right\|^2 + (\langle w, x\rangle - \beta_i)^2 \right\} \\
&= \min_{w\in\mathbb{R}^n} \left\{ \left\| \tilde{X}^\top w - \alpha \right\|^2 :\ \tilde{X} = (\gamma I + ww^\top)^{-1}(\gamma X + w\beta^\top) \right\} \\
&= \min_{w\in\mathbb{R}^n} \left\| \frac{\left\|w\right\|^2 \beta + \gamma X w}{\left\|w\right\|^2 + \gamma} - \alpha \right\|^2 \\
&= \min_{w\in\mathbb{R}^n, t\in\mathbb{R}} \left\{ \left\| \frac{t\beta + \gamma X w}{t+\gamma} - \alpha \right\|^2 :\ t = \left\|w\right\|^2 \right\} \\
&= \min_{\tilde{w}\in\mathbb{R}^n, \tilde{t}\in\mathbb{R}} \left\{ \left\| \frac{\tilde{t}}{2}\beta + \frac{\sqrt{\gamma}}{2} X\tilde{w} - \left(\alpha - \frac{\beta}{2}\right) \right\|^2 :\ \left\|\tilde{w}\right\|^2 + \tilde{t}^2 = 1 \right\}
\end{aligned}$$

Based on: [*under review*]

Diagonalizing QCQPs

Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is simultaneously diagonalizable via congruence (SDC) if there exists invertible $P \in \mathbb{R}^{n\times n}$ such that $P^\top A_i P$ is diagonal $\forall i$.

- Nice property because: SDP relaxation of diagonal QCQP is SOCP (faster), Γ is polyhedral (better understanding of exactness)

Goal

Most sets of matrices are not SDC, can we find other computationally variants of SDC and understand such properties?

Based on: [Math. Prog. *under review*]

Diagonalizing QCQPs: variants of SDC

Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is almost SDC (ASDC) if for all $\epsilon > 0$, there exists $\|A_i' - A_i\| \leq \epsilon$ such that $\{A_i'\}$ is SDC.

- "Limit of SDC sets"

Definition

$\{A_i\} \subseteq \mathbb{S}^n$ is d-restricted SDC (d-RSDC) if there exists $A_i' = \begin{pmatrix} A_i & * \\ * & * \end{pmatrix} \in \mathbb{S}^{n+d}$ such that $\{A_i'\}$ is SDC.

- "Restriction of SDC sets"

Based on: [Math. Prog. *under review*]

Diagonalizing QCQPs: ASDC

Theorem

Let $\{A, B\} \subseteq \mathbb{S}^n$ and suppose A invertible. Then $\{A, B\}$ is ASDC if and only if $A^{-1}B$ has real spectrum. (+ construction)

Theorem

Let $\{A, B\} \subseteq \mathbb{S}^n$. If $\{A, B\}$ is singular, then it is ASDC. (+ construction)

Theorem

Let $\{A, B, C\} \subseteq \mathbb{S}^n$ and suppose A invertible. Then $\{A, B, C\}$ is ASDC if and only if $\{A^{-1}B, A^{-1}C\}$ commute and have real spectrum. (+ construction)

Based on: **[Math. Prog. *under review*]**

Theorem

Let $\{A, B\} \subseteq \mathbb{S}^n$. If A is invertible and $A^{-1}B$ has simple eigenvalues, then $\{A, B\}$ is 1-RSDC. (+ construction)

- Condition holds generically

Based on: **[Math. Prog. *under review*]**

"Fuzzy" spectral partitioning

- Connected graph $G = (V, E)$
- Vertex masses $\mu : V \to \mathbb{R}_{++}$ and edge weights $\kappa : E \to \mathbb{R}_{++}$
- Laplacian $L = D - A$ w.r.t. κ

Theorem (Cheeger's inequality)

If $\mu_v = d_v$, then
$$\frac{\Phi^2}{2} \leq \lambda_2(L, M) \leq 2\Phi$$

- $\lambda_2(L, M)$ is first nontrivial generalized eigenvalue
- Φ is sparsest cut

Based on: [APPROX 19]

"Fuzzy" spectral partitioning

- We define "Fuzzy cuts"

Definition

$$\Psi \approx \min_{A,B} \left\{ \frac{\kappa_{\text{eff}}(A,B)}{\min\left(\mu(A),\mu(B)\right)},\ A,B \neq \varnothing,\ A \cap B = \varnothing \right\}$$

- Φ must partition, Ψ may leave out. $\Psi = \Phi$ if A, B is a partition.

Theorem

$$\frac{\Psi}{4} \leq \lambda_2(L,M) \leq \Psi$$

Based on: [APPROX 19]

- k-means clustering: $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$
- Suppose there exist true clustering that is unique optimum even if for all i, $x_i \mapsto x_i' \in B(x_i, \epsilon)$

Theorem

Two clusters. There exists $c \geq 1$ such that for any fixed $\epsilon > 0$, we can recover true clustering in time $d \cdot n^{O(\epsilon^{-c})}$.

Additional results for ≥ 3 clusters given an additional "separation" assumption

---

Based on: **[NeurIPS 17]**