

Semidefinite optimization and convex algebraic geometry

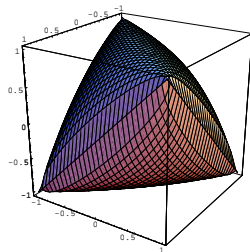
Pablo A. Parrilo

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

NIPS - December 2012

This talk

- Convex sets with algebraic descriptions
- The role of semidefinite programming and sums of squares
- Unifying idea: convex hull of algebraic varieties
- Examples and applications throughout
- Discuss results, but also open questions
- Computational considerations
- Connections with other areas of mathematics



Convex sets: geometry vs. algebra

The geometric theory of convex sets (e.g., Minkowski, Carathéodory, Fenchel) is very rich and well-understood.

Enormous importance in applied mathematics and engineering, in particular in optimization.

But, what if we are concerned with the *representation* of these geometric objects? For instance, basic semialgebraic sets?

How do the *algebraic*, *geometric*, and *computational* aspects interact?

Ex: Convex optimization is not always “easy”.

The polyhedral case

Consider first the case of *polyhedra*, which are described by finitely many *linear* inequalities $\{x \in \mathbb{R}^n : a_i^T x \leq b_i\}$.

- Behave well under projections (Fourier-Motzkin)
- Farkas' lemma (or duality) gives emptiness certificates
- Good associated computational techniques
- Optimization over polyhedra is linear programming (LP)

Great. But how to move away from linearity? For instance, if we want convex sets described by polynomial inequalities?

Claim: semidefinite programming is an essential tool.

The polyhedral case

Consider first the case of *polyhedra*, which are described by finitely many *linear* inequalities $\{x \in \mathbb{R}^n : a_i^T x \leq b_i\}$.

- Behave well under projections (Fourier-Motzkin)
- Farkas' lemma (or duality) gives emptiness certificates
- Good associated computational techniques
- Optimization over polyhedra is linear programming (LP)

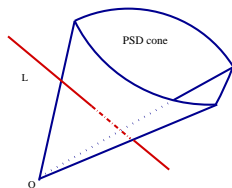
Great. But how to move away from linearity? For instance, if we want convex sets described by polynomial inequalities?

Claim: semidefinite programming is an essential tool.

Semidefinite programming (SDP, LMIs)

A broad generalization of LP to symmetric matrices

$$\min \text{Tr } CX \quad \text{s.t.} \quad X \in \mathcal{L} \cap \mathcal{S}_+^n$$

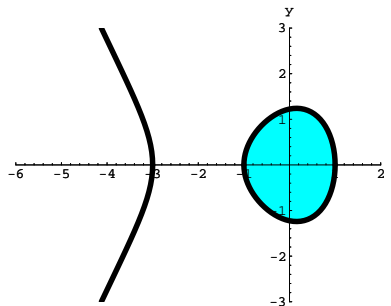


- Intersection of an affine subspace \mathcal{L} and the cone of positive semidefinite matrices.
- Feasible set is called *spectrahedron*
- Lots of applications. A true “revolution” in computational methods for engineering applications
- Convex finite dimensional optimization. Nice duality theory.
- Essentially, solvable in **polynomial time** (interior point, etc.)

Example

Consider the feasible set of the SDP:

$$\begin{bmatrix} x & 0 & y \\ 0 & 1 & -x \\ y & -x & 1 \end{bmatrix} \succeq 0.$$



- Convex, but not necessarily polyhedral
- In general, piecewise-smooth
- Determinant vanishes on the boundary

In this case, the determinant is the elliptic curve $x - x^3 = y^2$.

Semidefinite representations

A natural question in convex optimization:

What sets can be represented using semidefinite programming?

In the LP case, well-understood question: finite number of extreme points/rays (polyhedral sets)

Are there “obstructions” to SDP representability?

Semidefinite representations

A natural question in convex optimization:

What sets can be represented using semidefinite programming?

In the LP case, well-understood question: finite number of extreme points/rays (polyhedral sets)

Are there “obstructions” to SDP representability?

Semidefinite representations

A natural question in convex optimization:

What sets can be represented using semidefinite programming?

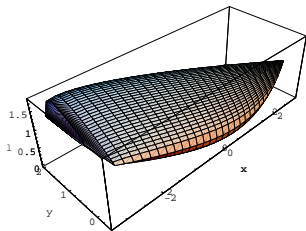
In the LP case, well-understood question: finite number of extreme points/rays (polyhedral sets)

Are there “obstructions” to SDP representability?

Known SDP-representable sets

- Many interesting sets are known to be SDP-representable (e.g., polyhedra, convex quadratics, matrix norms, etc.)
- Preserved by “natural” properties: affine transformations, convex hull, polarity, etc.
- Several known structural results (e.g., facial exposedness)

Work of Nesterov-Nemirovski, Ramana, Tunçel, Güler, Renegar, Chua, etc.



Existing results

Obvious necessary conditions: \mathcal{S} must be convex and semialgebraic.

Several versions of the problem:

- *Exact vs. approximate* representations.
- “Direct” (non-lifted) representations: no additional variables.

$$x \in \mathcal{S} \iff A_0 + \sum_i x_i A_i \succeq 0$$

- “Lifted” representations: can use extra variables (projection)

$$x \in \mathcal{S} \iff \exists y \text{ s.t. } A_0 + \sum_i x_i A_i + \sum_j y_j B_j \succeq 0$$

Projection helps a lot!

Existing results

Obvious necessary conditions: \mathcal{S} must be convex and semialgebraic.

Several versions of the problem:

- *Exact* vs. *approximate* representations.
- “Direct” (non-lifted) representations: no additional variables.

$$x \in \mathcal{S} \quad \Leftrightarrow \quad A_0 + \sum_i x_i A_i \succeq 0$$

- “Lifted” representations: can use extra variables (projection)

$$x \in \mathcal{S} \quad \Leftrightarrow \quad \exists y \text{ s.t. } A_0 + \sum_i x_i A_i + \sum_j y_j B_j \succeq 0$$

Projection helps a lot!

Liftings and projections

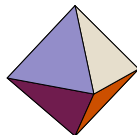
Often, “simpler” descriptions of convex sets from higher-dimensions.

Ex: The n -dimensional crosspolytope (ℓ_1 unit ball). Requires 2^n linear inequalities, of the form

$$\pm x_1 \pm x_2 \pm \cdots \pm x_n \leq 1.$$

However, can efficiently represent it as a *projection*:

$$\{(x, y) \in \mathbb{R}^{2n}, \quad \sum_{i=1}^n y_i = 1, \quad -y_i \leq x_i \leq y_i \quad i = 1, \dots, n\}$$



Only $2n$ variables, and $2n + 1$ constraints!

In convexity, elimination is *not* always a good idea.

Quite the opposite, it is often advantageous to go to higher-dimensional spaces, where descriptions (can) become simpler.

Liftings and projections

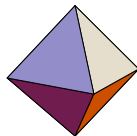
Often, “simpler” descriptions of convex sets from higher-dimensions.

Ex: The n -dimensional crosspolytope (ℓ_1 unit ball). Requires 2^n linear inequalities, of the form

$$\pm x_1 \pm x_2 \pm \cdots \pm x_n \leq 1.$$

However, can efficiently represent it as a *projection*:

$$\{(x, y) \in \mathbb{R}^{2n}, \quad \sum_{i=1}^n y_i = 1, \quad -y_i \leq x_i \leq y_i \quad i = 1, \dots, n\}$$



Only $2n$ variables, and $2n + 1$ constraints!

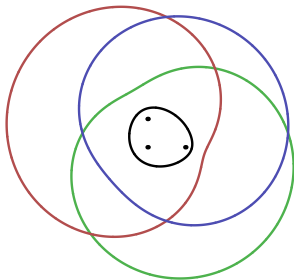
In convexity, elimination is *not* always a good idea.

Quite the opposite, it is often advantageous to go to higher-dimensional spaces, where descriptions (can) become simpler.

Example: k -ellipse

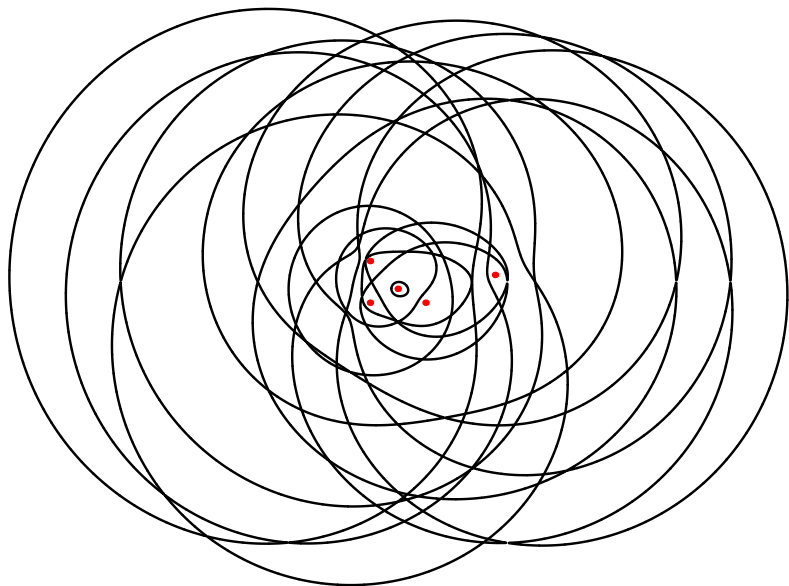
Fix a positive real number d and fix k distinct points (u_i, v_i) in \mathbb{R}^2 . The k -ellipse with foci (u_i, v_i) and radius d is the following curve in the plane:

$$\left\{ (x, y) \in \mathbb{R}^2 : \sum_{i=1}^k \sqrt{(x - u_i)^2 + (y - v_i)^2} = d \right\}.$$



Thm:(Nie-P.-Sturmfels 07) The k -ellipse has degree 2^k if k is odd and degree $2^k - \binom{k}{k/2}$ if k is even. It has an explicit $2^k \times 2^k$ SDP representation.

5-ellipse



Results on exact SDP representations

- Direct representations:
 - Necessary condition: **rigid convexity**. Helton & Vinnikov (2004) showed that in \mathbb{R}^2 , rigid convexity is also sufficient.
 - Related to hyperbolic polynomials and the Lax conjecture (Güler, Renegar, Lewis-P.-Ramana 2005)
 - For higher dimensions the problem is open.
- Lifted representations:
 - No known nontrivial obstructions.
 - Does every convex basic SA set have a lifted exact SDP representation?
 - (Helton & Nie 2007): Under strict positive curvature assumptions on the boundary, this is true.

Results on exact SDP representations

- Direct representations:
 - Necessary condition: **rigid convexity**. Helton & Vinnikov (2004) showed that in \mathbb{R}^2 , rigid convexity is also sufficient.
 - Related to hyperbolic polynomials and the Lax conjecture (Güler, Renegar, Lewis-P.-Ramana 2005)
 - For higher dimensions the problem is open.
- Lifted representations:
 - No known nontrivial obstructions.
 - Does every convex basic SA set have a lifted exact SDP representation?
 - (Helton & Nie 2007): Under strict positive curvature assumptions on the boundary, this is true.

Sum of squares

A multivariate polynomial $p(x)$ is a sum of squares (SOS) if

$$p(x) = \sum_i q_i^2(x), \quad q_i(x) \in \mathbb{R}[x].$$

- If $p(x)$ is SOS, then clearly $p(x) \geq 0 \forall x \in \mathbb{R}^n$.
- Converse not true, in general (Hilbert). Counterexamples exist.
- For univariate or quadratics, nonnegativity is equivalent to SOS.
- Convex condition, can be reduced to SDP.

Checking the SOS condition

Basic “Gram matrix” method (Shor 87, Choi-Lam-Reznick 95, Powers-Wörmann 98, Nesterov, Lasserre, P., etc.)

A polynomial $F(x)$ is SOS if and only if

$$F(x) = w(x)^T Q w(x),$$

where $w(x)$ is a vector of monomials, and $Q \succeq 0$.

Checking the SOS condition

Let $F(x) = \sum f_\alpha x^\alpha$. Index rows and columns of Q by monomials. Then,

$$F(x) = w(x)^T Q w(x) \quad \Leftrightarrow \quad f_\alpha = \sum_{\beta+\gamma=\alpha} Q_{\beta\gamma}$$

Thus, we have the SDP feasibility problem

$$f_\alpha = \sum_{\beta+\gamma=\alpha} Q_{\beta\gamma}, \quad Q \succeq 0$$

- Factorize $Q = L^T L$. The SOS is given by $f = Lz$.

SOS Example

$$\begin{aligned} F(x, y) &= 2x^4 + 5y^4 - x^2y^2 + 2x^3y \\ &= \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix} \\ &= q_{11}x^4 + q_{22}y^4 + (q_{33} + 2q_{12})x^2y^2 + 2q_{13}x^3y + 2q_{23}xy^3 \end{aligned}$$

An SDP with equality constraints. Solving, we obtain:

$$Q = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = L^T L, \quad L = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

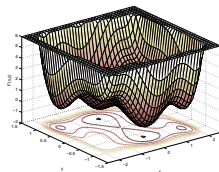
And therefore $F(x, y) = \frac{1}{2}(2x^2 - 3y^2 + xy)^2 + \frac{1}{2}(y^2 + 3xy)^2$

From feasibility to optimization

SOS directly yield lower bounds for optimization!

$$F(x) - \gamma \text{ is SOS} \quad \Rightarrow \quad F(x) \geq \gamma \text{ for all } x$$

- Finding the best such γ is also an SDP
- Typically, very high-quality bounds
- If exact, can recover exact solution
- Natural extensions to constrained case

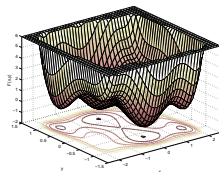


From feasibility to optimization

SOS directly yield lower bounds for optimization!

$$F(x) - \gamma \text{ is SOS} \quad \Rightarrow \quad F(x) \geq \gamma \text{ for all } x$$

- Finding the best such γ is also an SDP
- Typically, very high-quality bounds
- If exact, can recover exact solution
- Natural extensions to constrained case



Convex hulls of algebraic varieties

Back to SDP representations...

Focus here on a specific, but very important case.

Given a set $S \subset \mathbb{R}^n$, we can define its *convex hull*

$$\operatorname{conv} S := \left\{ \sum_i \lambda_i x_i : x_i \in S, \sum_i \lambda_i = 1, \lambda_i \geq 0 \right\}$$

We are interested in the case where S is a real algebraic variety.

Convex hulls of algebraic varieties

Back to SDP representations...

Focus here on a specific, but very important case.

Given a set $S \subset \mathbb{R}^n$, we can define its *convex hull*

$$\operatorname{conv} S := \left\{ \sum_i \lambda_i x_i : x_i \in S, \sum_i \lambda_i = 1, \lambda_i \geq 0 \right\}$$

We are interested in the case where S is a real algebraic variety.

Why?

Many interesting problems require or boil down *exactly* to understanding and describing convex hulls of (toric) algebraic varieties.

- Nonnegative polynomials and optimization
- Polynomial games
- Convex relaxations for minimum-rank

We discuss these next.

Polynomial optimization

Consider the unconstrained minimization of a multivariate polynomial

$$p(x) = \sum_{\alpha \in S} p_{\alpha} x^{\alpha},$$

where $x \in \mathbb{R}^n$ and S is a given set of monomials (e.g., all monomials of total degree less than or equal to $2d$, in the dense case).

Define the (real, toric) algebraic variety $V_S \subset \mathbb{R}^{|S|}$:

$$V_S := \{(x^{\alpha_1}, \dots, x^{\alpha_{|S|}}) : x \in \mathbb{R}^n\}.$$

This is the image of \mathbb{R}^n under the monomial map (e.g., in the homogeneous case, the Veronese embedding).

Want to study the *convex hull* of V_S . Extends to the constrained case.

Polynomial optimization

Consider the unconstrained minimization of a multivariate polynomial

$$p(x) = \sum_{\alpha \in S} p_{\alpha} x^{\alpha},$$

where $x \in \mathbb{R}^n$ and S is a given set of monomials (e.g., all monomials of total degree less than or equal to $2d$, in the dense case).

Define the (real, toric) algebraic variety $V_S \subset \mathbb{R}^{|S|}$:

$$V_S := \{(x^{\alpha_1}, \dots, x^{\alpha_{|S|}}) : x \in \mathbb{R}^n\}.$$

This is the image of \mathbb{R}^n under the monomial map (e.g., in the homogeneous case, the Veronese embedding).

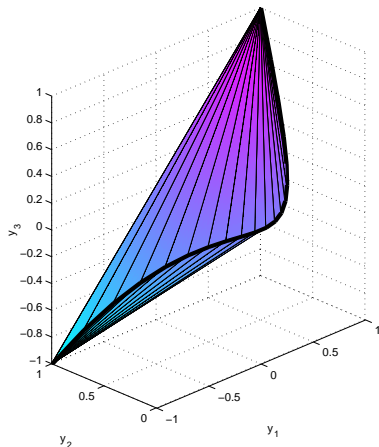
Want to study the *convex hull* of V_S . Extends to the constrained case.

Univariate case

Convex hull of the rational normal curve
 $(1, t, \dots, t^d)$.

Not polyhedral.

Known geometry (Karlin-Shapley)



“Simplicial”: every supporting hyperplane yields a simplex.
Related to cyclic polytopes.

Polynomial optimization

We have then (almost trivially):

$$\inf_{x \in \mathbb{R}^n} p(x) = \inf\{p^T y : y \in \text{conv } V_S\}$$

Optimizing a nonconvex polynomial is equivalent to linear optimization over a convex set (!)

Unfortunately, in general, it is NP-hard to check membership in $\text{conv } V_S$. Nevertheless, we can turn this around, and use SOS relaxations to obtain “good” approximate SDP descriptions of the convex hull V_S .

Polynomial optimization

We have then (almost trivially):

$$\inf_{x \in \mathbb{R}^n} p(x) = \inf \{ p^T y : y \in \text{conv } V_S \}$$

Optimizing a nonconvex polynomial is equivalent to linear optimization over a convex set (!)

Unfortunately, in general, it is NP-hard to check membership in $\text{conv } V_S$. Nevertheless, we can turn this around, and use SOS relaxations to obtain “good” approximate SDP descriptions of the convex hull V_S .

A geometric interlude

How is this possible? Convex optimization for solving nonconvex problems?

Convexity is *relative*. Every problem can be trivially “lifted” to a convex setting (in general, infinite dimensional).

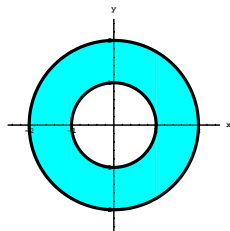
Ex: mixed strategies in games, “relaxed” controls, Fokker-Planck, etc. Interestingly, however, often a finite (and small) dimension is enough.

Consider the set defined by

$$1 \leq x^2 + y^2 \leq 2$$

Clearly non-convex.

Can we use convex optimization?



A geometric interlude

How is this possible? Convex optimization for solving nonconvex problems?

Convexity is *relative*. Every problem can be trivially “lifted” to a convex setting (in general, infinite dimensional).

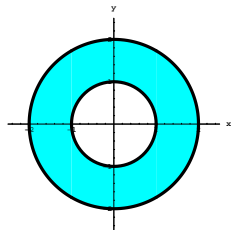
Ex: mixed strategies in games, “relaxed” controls, Fokker-Planck, etc. Interestingly, however, often a finite (and small) dimension is enough.

Consider the set defined by

$$1 \leq x^2 + y^2 \leq 2$$

Clearly non-convex.

Can we use convex optimization?



A geometric interlude

How is this possible? Convex optimization for solving nonconvex problems?

Convexity is *relative*. Every problem can be trivially “lifted” to a convex setting (in general, infinite dimensional).

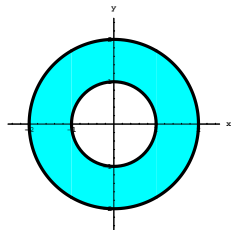
Ex: mixed strategies in games, “relaxed” controls, Fokker-Planck, etc. Interestingly, however, often a finite (and small) dimension is enough.

Consider the set defined by

$$1 \leq x^2 + y^2 \leq 2$$

Clearly non-convex.

Can we use convex optimization?



Geometric interpretation

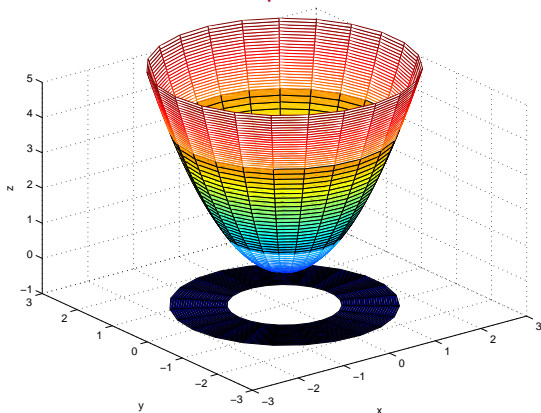
A polynomial “lifting” to a higher dimensional space:

$$(x, y) \mapsto (x, y, x^2 + y^2)$$

The nonconvex set is the **projection** of the **extreme points** of a convex set.

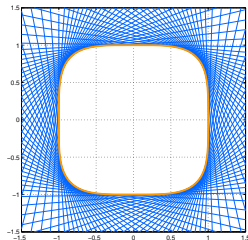
In particular, the convex set defined by

$$\begin{aligned} x^2 + y^2 &\leq z \\ 1 &\leq z \leq 4 \end{aligned}$$



A “polar” viewpoint

Any convex set \mathcal{S} is uniquely defined by its supporting hyperplanes.



Thus, if we can optimize a *linear function* over a set using SDP, we effectively have an SDP representation.

Need to solve (or approximate)

$$\min c^T x \quad \text{s.t. } x \in \mathcal{S}$$

If \mathcal{S} is defined by polynomial equations/inequalities, can use SOS.

Example: orthogonal matrices

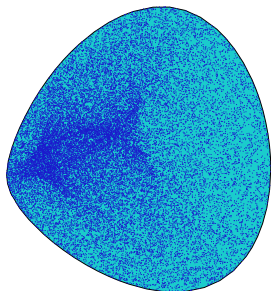
Consider $O(3)$, the group of 3×3 orthogonal matrices of determinant one. It has two connected components (sign of determinant).

We can use the double-cover of $SO(3)$ with $SU(2)$ to provide an exact SDP representation of the convex hull of $SO(3)$:

$$\begin{bmatrix} Z_{11} + Z_{22} - Z_{33} - Z_{44} & 2Z_{23} - 2Z_{14} & 2Z_{24} + 2Z_{13} \\ 2Z_{23} + 2Z_{14} & Z_{11} - Z_{22} + Z_{33} - Z_{44} & 2Z_{34} - 2Z_{12} \\ 2Z_{24} - 2Z_{13} & 2Z_{34} + 2Z_{12} & Z_{11} - Z_{22} - Z_{33} + Z_{44} \end{bmatrix}, \quad Z \succeq 0, \quad \text{Tr } Z = 1.$$

This is a convex set in \mathbb{R}^9 .

Here is a two-dimensional projection.



Minimum rank and convex relaxations

Consider the rank minimization problem

$$\text{minimize rank } X \quad \text{subject to } \mathcal{A}(X) = b,$$

where $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$ is a linear map.

Find the minimum-rank matrix in a given subspace. In general, NP-hard.

Since rank is hard, let's use instead its *convex envelope*, the nuclear norm. The nuclear norm of a matrix (alternatively, Schatten 1-norm, Ky Fan r -norm, or trace class norm) is the sum of its singular values, i.e.,

$$\|X\|_* := \sum_{i=1}^r \sigma_i(X).$$

Minimum rank and convex relaxations

Consider the rank minimization problem

$$\text{minimize rank } X \quad \text{subject to } \mathcal{A}(X) = b,$$

where $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$ is a linear map.

Find the minimum-rank matrix in a given subspace. In general, NP-hard.

Since rank is hard, let's use instead its *convex envelope*, the nuclear norm. The nuclear norm of a matrix (alternatively, Schatten 1-norm, Ky Fan r -norm, or trace class norm) is the sum of its singular values, i.e.,

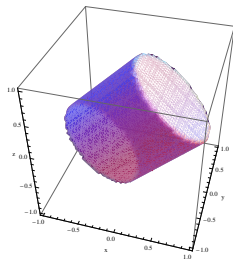
$$\|X\|_* := \sum_{i=1}^r \sigma_i(X).$$

Convex hulls and nuclear norm

Nuclear norm ball is convex hull of rank one matrices!

$$B = \text{conv}\{uv^T : u \in \mathbb{R}^m, v \in \mathbb{R}^n, \|u\|^2 = 1, \|v\|^2 = 1\}$$

Exactly SDP-characterizable.



Under certain conditions (e.g., if \mathcal{A} is “random”), optimizing the nuclear norm yields the true minimum rank solution.

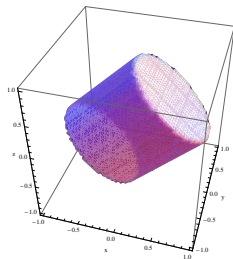
For details, see Recht-Fazel-P., “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,” *SIAM Review*, 2010.

Convex hulls and nuclear norm

Nuclear norm ball is convex hull of rank one matrices!

$$B = \text{conv}\{uv^T : u \in \mathbb{R}^m, v \in \mathbb{R}^n, \|u\|^2 = 1, \|v\|^2 = 1\}$$

Exactly SDP-characterizable.



Under certain conditions (e.g., if \mathcal{A} is “random”), optimizing the nuclear norm yields the true minimum rank solution.

For details, see Recht-Fazel-P., “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,” *SIAM Review*, 2010.

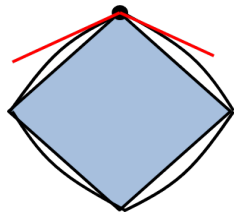
Rank, sparsity, and beyond: atomic norms

Exactly the same constructions can be applied to more general situations:
atomic norms.

Structure-inducing regularizer is convex hull of atom set, e.g., low-rank matrices/tensors, permutation matrices, cut matrices, etc.

Generally NP-hard to compute, but good SDP approximations.

Statistical guarantees for recovery based on *Gaussian width of tangent cones*. Interesting interplay between computational and sample complexities.



For details, see Chandrasekaran-Recht-P.-Willsky, "The convex geometry of linear inverse problems," *Found. Comp. Math.*, 2012.

Connections

Many fascinating links to other areas of mathematics:

- Probability (moments, exchangeability and de Finetti, etc)
- Operator theory (via Gelfand-Neimark-Segal)
- Harmonic analysis on semigroups
- Noncommutative probability (i.e., quantum mechanics)
- Complexity and proof theory (degrees of certificates)
- Graph theory (perfect graphs)
- Tropical geometry (SDP over more general fields)

Algebraic structure

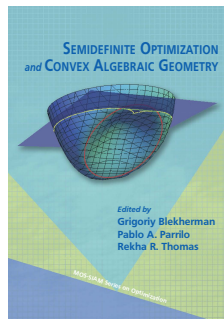
- **Algebraic sparsity:** polynomials with few nonzero coefficients.
 - Newton polytopes techniques.
- **Ideal structure:** equality constraints.
 - SOS on *quotient rings*.
 - Compute in the coordinate ring. Quotient bases.
- **Graph structure:**
 - Dependency graph among the variables.
- **Symmetries:** invariance under a group (w/ **K. Gatermann**)
 - SOS on *invariant rings*
 - Representation theory and invariant-theoretic methods.
 - Enabling factor in applications (e.g., Markov chains)

Numerical structure

- Rank one SDPs.
 - Dual coordinate change makes all constraints rank one
 - Efficient computation of Hessians and gradients
- Representations
 - Interpolation representation
 - Orthogonalization
- Displacement rank
 - Fast solvers for search direction

Summary

- A very rich class of optimization problems
- Methods have enabled many new applications
- Interplay of many branches of mathematics
- Structure must be exploited for reliability and efficiency
- Combination of numerical and algebraic techniques.



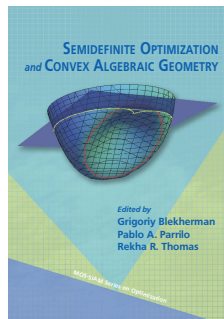
If you want to know more:

- Papers, slides, lecture notes, software, etc.: www.mit.edu/~parrilo
- NSF FRG project "SDP and convex algebraic geometry" website www.math.washington.edu/~thomas/frg/frg.html (Helton/P./Nie/Sturmfels/Thomas), and **new SIAM book!**

Thanks for your attention!

Summary

- A very rich class of optimization problems
- Methods have enabled many new applications
- Interplay of many branches of mathematics
- Structure must be exploited for reliability and efficiency
- Combination of numerical and algebraic techniques.



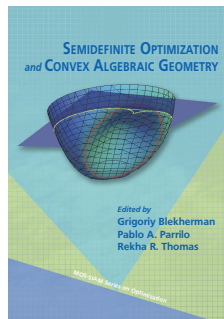
If you want to know more:

- Papers, slides, lecture notes, software, etc.: www.mit.edu/~parrilo
- NSF FRG project “SDP and convex algebraic geometry” website www.math.washington.edu/~thomas/frg/frg.html (Helton/P./Nie/Sturmfels/Thomas), and **new SIAM book!**

Thanks for your attention!

Summary

- A very rich class of optimization problems
- Methods have enabled many new applications
- Interplay of many branches of mathematics
- Structure must be exploited for reliability and efficiency
- Combination of numerical and algebraic techniques.



If you want to know more:

- Papers, slides, lecture notes, software, etc.: www.mit.edu/~parrilo
- NSF FRG project “SDP and convex algebraic geometry” website www.math.washington.edu/~thomas/frg/frg.html (Helton/P./Nie/Sturmfels/Thomas), and **new SIAM book!**

Thanks for your attention!