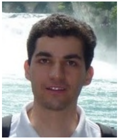
PSD factorizations of nonnegative matrices and lower bounds on semidefinite representations

Pablo A. Parrilo

Laboratory for Information and Decision Systems
Electrical Engineering and Computer Science
Massachusetts Institute of Technology



Based on joint work with **Hamza Fawzi** (MIT),
João Gouveia (U. Coimbra) and **Rekha Thomas** (U. Washington).

Nonnegative factorizations

Given a nonnegative matrix $A \in \mathbb{R}^{n \times m}$, a factorization

$$A = UV$$

where $U \in \mathbb{R}^{n \times k}$, $V \in \mathbb{R}^{k \times m}$ are also nonnegative.

- The smallest such k is the *nonnegative rank* of the matrix A.
- Many applications: statistics, factor models, machine learning, ...
- Very difficult problem, many heuristics exist.

Factorizations and hidden variables

Let X, Y be discrete random variables, with joint distribution

$$\mathbf{P}[X = i, Y = j] = P_{ij}.$$

The nonnegative rank of P is the smallest support of a random variable Z, such that X and Y are *conditionally independent* given Z (i.e., $X - Z - Y$ is Markov):

$$\mathbf{P}[X = i, Y = j] = \sum_{s=1,\ldots,k} \mathbf{P}[Z = s] \cdot \mathbf{P}[X = i | Z = s] \cdot \mathbf{P}[Y = j | Z = s].$$

- Relations with information theory, “correlation generation,” communication complexity, etc.
- Quantum versions are also of interest.

As we’ll see, fundamental in optimization …

Factorizations and hidden variables

Let X, Y be discrete random variables, with joint distribution

$$\mathbf{P}[X = i, Y = j] = P_{ij}.$$

The nonnegative rank of P is the smallest support of a random variable Z, such that X and Y are *conditionally independent* given Z (i.e., $X - Z - Y$ is Markov):

$$\mathbf{P}[X = i, Y = j] = \sum_{s=1,\dots,k} \mathbf{P}[Z = s] \cdot \mathbf{P}[X = i | Z = s] \cdot \mathbf{P}[Y = j | Z = s].$$

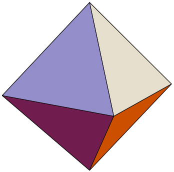
- Relations with information theory, "correlation generation," communication complexity, etc.
- Quantum versions are also of interest.

As we'll see, fundamental in optimization ...

Motivating example

The *crosspolytope* $\mathcal{C}_n$ is the unit ball of the ℓ_1 ball:

$$\mathcal{C}_n := \{x \in \mathbb{R}^n \,:\, \sum_{i=1}^{n} |x_i| \leq 1\}.$$



It is a polytope defined by 2^n linear inequalities:

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

The "obvious" linear program is exponentially large!

A better representation

By introducing *slack* or *auxiliary* variables, the set $\mathcal{C}_n$ can be represented more conveniently:

$$\mathcal{C}_n := \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n, \quad -y_i \leq x_i \leq y_i, \quad \sum_{i=1}^{n} y_i = 1\}.$$

This has only $2n$ variables $(x_1, y_1, \ldots, x_n, y_n)$ and $2n + 1$ constraints. A "small" linear program. Much better!

What is going on in here?

A better representation

By introducing *slack* or *auxiliary* variables, the set $\mathcal{C}_n$ can be represented more conveniently:

$$\mathcal{C}_n := \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n, \quad -y_i \leq x_i \leq y_i, \quad \sum_{i=1}^{n} y_i = 1\}.$$

This has only $2n$ variables $(x_1, y_1, \ldots, x_n, y_n)$ and $2n + 1$ constraints. A "small" linear program. Much better!

What is going on in here?

Geometric viewpoint

Geometrically, we are representing our polytope as a *projection* of a higher-dimensional polytope.

The number of *vertices* does not increase, but the number of *facets* can grow exponentially!

“Complicated” objects are sometimes easily described as “projections” of “simpler” ones.

A general theme: algebraic varieties, graphical models, . . .

Geometric viewpoint

Geometrically, we are representing our polytope as a *projection* of a higher-dimensional polytope.

The number of *vertices* does not increase, but the number of *facets* can grow exponentially!

“Complicated” objects are sometimes easily described as “projections” of “simpler” ones.

A general theme: algebraic varieties, graphical models, ...

Extended formulations

These representations are usually called *extended formulations*.
Particularly relevant in combinatorial optimization (e.g., TSP).

Seminal work by Yannakakis (1991), who used them to disprove the existence of a "symmetric" LP formulation for the TSP polytope.
Nice recent survey by Conforti-Cornuéjols-Zambelli (2010).

Our goal: to understand this phenomenon for convex optimization, not just LP.

Extended formulations

These representations are usually called *extended formulations*.
Particularly relevant in combinatorial optimization (e.g., TSP).

Seminal work by Yannakakis (1991), who used them to disprove the existence of a "symmetric" LP formulation for the TSP polytope.
Nice recent survey by Conforti-Cornuéjols-Zambelli (2010).

Our goal: to understand this phenomenon for convex optimization, not just LP.

Extended formulations

These representations are usually called *extended formulations*.
Particularly relevant in combinatorial optimization (e.g., TSP).

Seminal work by Yannakakis (1991), who used them to disprove the existence of a “symmetric” LP formulation for the TSP polytope.
Nice recent survey by Conforti-Cornuéjols-Zambelli (2010).

Our goal: to understand this phenomenon for convex optimization, not just LP.

“Extended formulations” in SDP

Many convex sets and functions can be modeled by SDP or SOCP in nontrivial ways. Among others:

- Sums of eigenvalues of symmetric matrices
- Convex envelope of univariate polynomials
- Multivariate polynomials that are sums of squares
- Unit ball of matrix operator and nuclear norms
- Geometric and harmonic means
- (Some) orbitopes – convex hulls of group orbits

E.g., Nesterov/Nemirovski, Boyd/Vandenberghe, Ben-Tal/Nemirovski, Sanyal/Sottile/Sturmfels, etc.

Often, clever and non-obvious reformulations.

“Extended formulations” in SDP

Many convex sets and functions can be modeled by SDP or SOCP in nontrivial ways. Among others:

- Sums of eigenvalues of symmetric matrices
- Convex envelope of univariate polynomials
- Multivariate polynomials that are sums of squares
- Unit ball of matrix operator and nuclear norms
- Geometric and harmonic means
- (Some) orbitopes – convex hulls of group orbits

E.g., Nesterov/Nemirovski, Boyd/Vandenberghe, Ben-Tal/Nemirovski, Sanyal/Sottile/Sturmfels, etc.

Often, clever and non-obvious reformulations.

Our questions

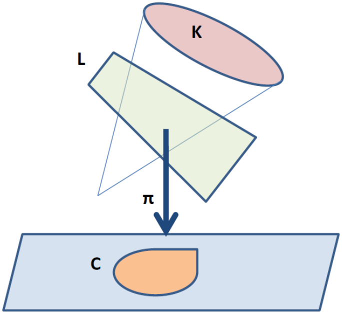
Existence and efficiency:

- When is a convex set representable by conic optimization?
- How to quantify the number of additional variables that are needed?

Given a convex set C, is it possible to represent it as

$$C = \pi(K \cap L)$$

where K is a cone, L is an affine subspace, and π is a linear map?



Cone lifts of convex bodies

When do such representations exist?

Even ignoring complexity aspects, this question is not well understood.

- Why a sphere is not a polytope?
- Can every basic closed semialgebraic set be represented using semidefinite programming?

What are “obstructions” for cone representability?

This talk: polytopes

What happens in the case of polytopes?

$$P = \{x \in \mathbb{R}^n \; : \; f_i^T x \leq 1\}$$

(WLOG, compact with $0 \in \operatorname{int} P$).

Polytopes have a finite number of facets f_i and vertices v_j.
Define a nonnegative matrix, called the *slack matrix* of the polytope:

$$[S_P]_{ij} = 1 - f_i^T v_j, \qquad i = 1, \ldots, |F| \quad j = 1, \ldots, |V|$$

This talk: polytopes

What happens in the case of polytopes?

$$P = \{x \in \mathbb{R}^n \,:\, f_i^T x \leq 1\}$$

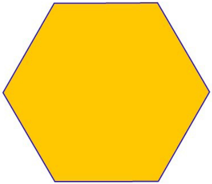
(WLOG, compact with $0 \in \operatorname{int} P$).

Polytopes have a finite number of facets f_i and vertices v_j.
Define a nonnegative matrix, called the *slack matrix* of the polytope:

$$[S_P]_{ij} = 1 - f_i^T v_j, \qquad i = 1, \ldots, |F| \quad j = 1, \ldots, |V|$$

Example: hexagon (I)

Consider a regular hexagon in the plane.



It has 6 vertices, and 6 facets. Its slack matrix is

$$S_H = \begin{pmatrix} 0 & 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \end{pmatrix}.$$

“Trivial” representation requires 6 facets. Can we do better?

Cone factorizations and representability

"Geometric" LP formulations exactly correspond to "algebraic" factorizations of the slack matrix.

For polytopes, this amounts to a *nonnegative factorization* of the slack matrix:

$$S_{ij} = \langle a_i, b_j \rangle, \qquad i = 1, \ldots, v, \qquad j = 1, \ldots, f$$

where a_i, b_i are nonnegative vectors.

Yannakakis (1991) showed that the minimal lifting dimension is equal to the nonnegative rank of the slack matrix.

Cone factorizations and representability

“Geometric” LP formulations exactly correspond to “algebraic” factorizations of the slack matrix.

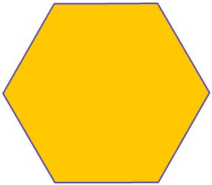
For polytopes, this amounts to a *nonnegative factorization* of the slack matrix:

$$S_{ij} = \langle a_i, b_j \rangle, \qquad i = 1, \ldots, v, \qquad j = 1, \ldots, f$$

where a_i, b_i are nonnegative vectors.

Yannakakis (1991) showed that the minimal lifting dimension is equal to the nonnegative rank of the slack matrix.

Example: hexagon (II)



Regular hexagon in the plane.

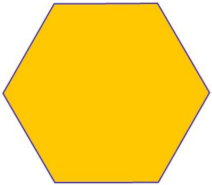
Its slack matrix is

$$S_H = \begin{pmatrix} 0 & 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \end{pmatrix}.$$

Nonnegative rank is 5.

Example: hexagon (II)

Regular hexagon in the plane.



Its slack matrix is

$$S_H = \begin{pmatrix} 0 & 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \end{pmatrix}.$$

Nonnegative rank is 5.

Beyond LPs and nonnegative factorizations

LPs are nice, but what about broader representability questions?

In [GPT11], a generalization of Yannakakis' theorem to the general convex case. General theme:

"Geometric" extended formulations exactly correspond to "algebraic" factorizations of a slack operator.

polytopes/LP	convex sets/convex cones
slack matrix	slack operators
vertices	extreme points of C
facets	extreme points of polar C°
nonnegative factorizations	conic factorizations
$S_{ij} = \langle a_i, b_j \rangle, \quad a_i \geq, b_j \geq 0$	$S_{ij} = \langle a_i, b_j \rangle, \quad a_i \in K, b_j \in K^*$

Polytopes and PSD factorizations

Even for polytopes, PSD factorizations can be interesting.

Well-known example: the *stable set* or *independent set* polytope.

For perfect graphs, we have efficient SDP representations, but no known subexponential LP.

Natural notion: *positive semidefinite rank* ([GPT 11]).

Exactly captures the complexity of SDP-representability.

Polytopes and PSD factorizations

Even for polytopes, PSD factorizations can be interesting.

Well-known example: the *stable set* or *independent set* polytope.

For perfect graphs, we have efficient SDP representations, but no known subexponential LP.

Natural notion: *positive semidefinite rank* ([GPT 11]).
Exactly captures the complexity of SDP-representability.

PSD rank of a nonnegative matrix

Let $M \in \mathbb{R}^{m \times n}$ be a nonnegative matrix.

The *PSD rank* of M, denoted rank_{psd}, is the smallest r for which there exists $r \times r$ PSD matrices $\{A_1, \ldots, A_m\}$ and $\{B_1, \ldots, B_n\}$ such that

$$M_{ij} = \text{trace}\, A_i B_j, \qquad i = 1, \ldots, m \quad j = 1, \ldots, n.$$

Natural generalization of nonnegative rank.

The PSD rank determines the “best” semidefinite lifting.

PSD rank of a nonnegative matrix

Let $M \in \mathbb{R}^{m \times n}$ be a nonnegative matrix.

The *PSD rank* of M, denoted rank_{psd}, is the smallest r for which there exists $r \times r$ PSD matrices $\{A_1, \ldots, A_m\}$ and $\{B_1, \ldots, B_n\}$ such that

$$M_{ij} = \text{trace}\, A_i B_j, \qquad i = 1, \ldots, m \quad j = 1, \ldots, n.$$

Natural generalization of nonnegative rank.

The PSD rank determines the "best" semidefinite lifting.

Some inequalities

- For any nonnegative matrix M

$$\frac{1}{2}\sqrt{1 + 8\,\mathrm{rank}(M)} - \frac{1}{2} \leq \mathrm{rank}_{psd}(M) \leq \mathrm{rank}_{+}(M).$$

- Gap between $\mathrm{rank}_{+}(M)$ and $\mathrm{rank}_{psd}(M)$ can be arbitrarily large:

$$M_{ij} = (i-j)^2 = \left\langle \begin{pmatrix} i^2 & -i \\ -i & 1 \end{pmatrix}, \begin{pmatrix} 1 & j \\ j & j^2 \end{pmatrix} \right\rangle$$

has $\mathrm{rank}_{psd}(M) = 2$, but $\mathrm{rank}_{+}(M) = \Omega(\log n)$.

Arbitrarily large gaps between all pairs of ranks (rank, rank_{+} and $\mathrm{rank}_{\mathrm{psd}}$). For slack matrices of polytopes, arbitrarily large gaps between rank and rank_{+}, and rank and $\mathrm{rank}_{\mathrm{psd}}$.

Bounding nonnegative rank

Want techniques to *lower bound* the nonnegative rank of a matrix.

In applications, these bounds may yield:

- Minimal size of latent variables
- Complexity lower bounds on extended representations

Known bounds exist (e.g. rank bound, combinatorial bounds, etc.).

Good lower bounds using "nonnegative nuclear norm" (Fawzi-P. 2012).

Improved SOS/SDP techniques (Fawzi-P. 2013), also extend to other "product cone" ranks (e.g., NN tensor rank, CP-rank, etc).

Bounding nonnegative rank

Want techniques to *lower bound* the nonnegative rank of a matrix.

In applications, these bounds may yield:

- Minimal size of latent variables
- Complexity lower bounds on extended representations

Known bounds exist (e.g. rank bound, combinatorial bounds, etc.).

Good lower bounds using "nonnegative nuclear norm" (Fawzi-P. 2012).
Improved SOS/SDP techniques (Fawzi-P. 2013), also extend to other "product cone" ranks (e.g., NN tensor rank, CP-rank, etc).

Lower bounding PSD rank?

Currently extending our bounds to PSD rank, since combinatorial methods (based on sparsity patterns) don't quite work.

But, a few unexpected difficulties...

- In the PSD case, the underlying norm is non-atomic, and the corresponding "obvious" inequalities do not hold...
- "Noncommutative" trace positivity, quite complicated structure...

Nice links between rank_{psd} and quantum communication complexity, mirroring the situation between rank_+ and classical communication complexity (e.g., Fiorini *et al.* (2011), Jain *et al.* (2011), Zhang (2012)).

Nevertheless, in some cases one can get nice strong results...

Lower bounding PSD rank?

Currently extending our bounds to PSD rank, since combinatorial methods (based on sparsity patterns) don't quite work.

But, a few unexpected difficulties...

- In the PSD case, the underlying norm is non-atomic, and the corresponding "obvious" inequalities do not hold...
- "Noncommutative" trace positivity, quite complicated structure...

Nice links between rank_{psd} and quantum communication complexity, mirroring the situation between rank_+ and classical communication complexity (e.g., Fiorini *et al.* (2011), Jain *et al.* (2011), Zhang (2012)).

Nevertheless, in some cases one can get nice strong results...

Lifts with product cones

- Let d fixed. We are interested in lifts over the cone:

$$\underbrace{\mathcal{S}_+^d \times \cdots \times \mathcal{S}_+^d}_{r \text{ copies}} = (\mathcal{S}_+^d)^r$$

- Why?
  - $d = 1 \rightarrow$ LP lifts
  - $d = 2 \rightarrow$ SOCP lifts: ice-cream cone is affinely isomorphic to $\mathcal{S}_+^2$:

$$\sqrt{x^2 + y^2} \leq t \quad \Leftrightarrow \quad \begin{bmatrix} t + x & y \\ y & t - x \end{bmatrix} \succeq 0.$$

  - In practice, these are SDPs we can solve efficiently (d small)

Factorization theorem

Factorization theorem [GPT] applied to $(\mathcal{S}_+^d)^r$ lifts:

Theorem

Polytope P has a $(\mathcal{S}_+^d)^r$ lift iff slack matrix S_P can be written as the sum of r matrices of psd rank $\leq d$.

- Define $\text{rank}_{\mathcal{S}_+^d}(A)$ as:

$$\text{rank}_{\mathcal{S}_+^d}(A) = \min r \text{ s.t. } A \text{ can be written as the sum of } r \text{ matrices of psd rank} \leq d$$

- $\text{rank}_{\mathcal{S}_+^d}$ is atomic rank
- Interesting in quantum information: *classical-quantum* states (a.k.a., cq-states) where the dimension of the quantum part is $\leq d$.

Correlation polytope

- Correlation polytope

$$COR(n) = \text{conv}\left(bb^T \ : \ b \in \{0,1\}^n\right).$$

- $COR(n)$ is affinely isomorphic to cut polytope of K_{n+1}.

- Fiorini et al. 2012, and subsequent papers, showed that any LP lift of $COR(n)$ has exponential size.
  $\rightarrow$ Proof based on analysis of a submatrix of slack matrix of $COR(n)$, known as *unique-disjointness* matrix.

- What about $(\mathcal{S}_+^d)^r$ lifts?

$(\mathcal{S}_+^d)^r$ lifts of $COR(n)$

Theorem (Fawzi-P. 2013)

If $COR(n)$ has a $(\mathcal{S}_+^d)^r$-lift then necessarily $r \geq \kappa(d) \cdot c(d)^n$ where $c(d) > 1$ and $\kappa(d) > 0$ explicit constants:

$$c(d) = (1 - 3^{-d})^{-1/d} > 1 \qquad \kappa(d) = (3^d - 1)^{-(1-1/d)}$$

Implications:

- No small LP representations (case $d = 1$, c.f. Fiorini et. al)
- No small SOCP representations (case $d = 2$)
- For any constant d, exponentially many LMIs...

Results are *unconditional* (no P/NP assumptions, not about specific constructions)

$(\mathcal{S}_+^d)^r$ lifts of $COR(n)$

Theorem (Fawzi-P. 2013)

If $COR(n)$ has a $(\mathcal{S}_+^d)^r$-lift then necessarily $r \geq \kappa(d) \cdot c(d)^n$ where $c(d) > 1$ and $\kappa(d) > 0$ explicit constants:

$$c(d) = (1 - 3^{-d})^{-1/d} > 1 \qquad \kappa(d) = (3^d - 1)^{-(1-1/d)}$$

Proof by showing that $\text{rank}_{\mathcal{S}_+^d}(UDISJ(n))$ is exponentially large.

Main ingredients:

- Analysis of sparsity pattern of certain matrices of small psd rank
- + induction argument inspired from Kaibel and Weltge [KW13].

Unique disjointness matrix

- $2^n \times 2^n$ matrix indexed by n-bit strings:

$$UDISJ(n)_{a,b} = (1 - a^T b)^2 \quad \forall a \in \{0,1\}^n, b \in \{0,1\}^n$$

- Submatrix of slack matrix of $COR(n)$
- Two key facts about $UDISJ$:
  - if $a^T b = 0$ then $UDISJ_{a,b} = 1$
  - if $a^T b = 1$ then $UDISJ_{a,b} = 0$
- $UDISJ$ for $n = 1$ and $n = 2$:

$$\begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array} \qquad \begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Atoms

$$\mathcal{A}_{\mathcal{S}_+^d}(n) = \{M \in \mathbb{R}_+^{2^n \times 2^n} \;:\; \text{rank}_{psd}(M) \leq d \text{ and } M_{a,b} = 0 \;\text{ for } a^T b = 1\}.$$

- In any $(\mathcal{S}_+^d)^r$-factorization of *UDISJ*, all terms must be in $\mathcal{A}_{\mathcal{S}_+^d}$
- To prove lower bound on $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$, we show that matrices in $\mathcal{A}_{\mathcal{S}_+^d}(n)$ must be very sparse, and thus $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$ has to be large to be able to fill up the > 0 entries of *UDISJ*

Precisely, let

$$\text{val}(M) = \text{card}\left\{(a,b) \in (\{0,1\}^n)^2 \;:\; a^T b = 0 \text{ and } M_{a,b} > 0\right\}.$$

Lower bound on $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$ follows from two facts:
→ Fact 1: $\text{val}(UDISJ) = 3^n$ (simple calculation)
→ Fact 2 (main part): Any $M \in \mathcal{A}_{\mathcal{S}_+^d}$ has $\text{val}(M) \leq t(d)^n$ where $t(d) < 3$

Atoms

$$\mathcal{A}_{\mathcal{S}_+^d}(n) = \{M \in \mathbb{R}_+^{2^n \times 2^n} \ : \ \text{rank}_{psd}(M) \leq d \text{ and } M_{a,b} = 0 \ \text{ for } a^T b = 1\}.$$

- In any $(\mathcal{S}_+^d)^r$-factorization of *UDISJ*, all terms must be in $\mathcal{A}_{\mathcal{S}_+^d}$
- To prove lower bound on $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$, we show that matrices in $\mathcal{A}_{\mathcal{S}_+^d}(n)$ must be very sparse, and thus $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$ has to be large to be able to fill up the > 0 entries of *UDISJ*

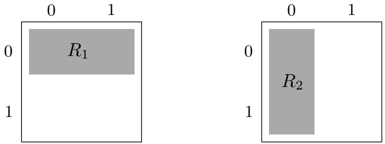
Precisely, let

$$\text{val}(M) = \text{card}\left\{(a,b) \in (\{0,1\}^n)^2 \ : \ a^T b = 0 \text{ and } M_{a,b} > 0\right\}.$$

Lower bound on $\text{rank}_{\mathcal{S}_+^d}(UDISJ)$ follows from two facts:
→ Fact 1: $\text{val}(UDISJ) = 3^n$ (simple calculation)
→ Fact 2 (main part): Any $M \in \mathcal{A}_{\mathcal{S}_+^d}$ has $\text{val}(M) \leq t(d)^n$ where $t(d) < 3$

Induction argument of Kaibel and Weltge for $d = 1$ (I)

- Quantity $\text{val}(M)$ introduced by Kaibel and Weltge [2013] for LP-lifts ($d = 1$). They give elementary induction proof that $\text{val}(M) \leq 2^n$ for all $M \in \mathcal{A}(n)$
- Proof relies on a simple fact about 2×2 matrices in $\mathcal{A}(n = 1)$: Consider rectangles $R_1 = \{0\} \times \{0, 1\}$ and $R_2 = \{0, 1\} \times \{0\}$.



Let M be a 2×2 rank-one matrix with $M_{1,1} = 0$. Then M has at most 2 nonzero entries and these can be associated in a one-to-one way to the rectangles R_1 and R_2:

Two possible sparsity patterns: $\begin{bmatrix} 1 & \\ 2 & \end{bmatrix}$ or $\begin{bmatrix} 2 & 1 \\ & \end{bmatrix}$

Induction argument of Kaibel and Weltge for $d = 1$ (II)

- Bootstrap the previous observation, to show that for any $n \geq 1$ and any $M \in \mathcal{A}(n)$

$$\mathrm{val}(M) \leq \mathrm{val}(\underbrace{M^{0,0} + M^{0,1}}_{\in \mathcal{A}(n-1)}) + \mathrm{val}(\underbrace{M^{0,0} + M^{1,0}}_{\in \mathcal{A}(n-1)})$$

  where $M^{x,y}$ blocks of M:

$$M = \begin{bmatrix} M^{0,0} & M^{0,1} \\ M^{1,0} & M^{1,1} \end{bmatrix}$$

- By induction this gives $\mathrm{val}(M) \leq 2^n$

Generalize the induction argument to arbitrary d

Generalize this induction argument to arbitrary d using the notion of a *uniform covering* with rectangles.

Theorem

Let $K = \mathcal{S}^d_+$. If $\mathcal{A}_K(d')$ has a uniform-covering with k rectangles then any $M \in \mathcal{A}_K(n)$ for $n \geq d'$ satisfies:

$$val(M) \leq k^{\lfloor (n-1)/d' \rfloor + 1}$$

- Kaibel and Weltge [2013] correspond to $d = d' = 1$. The two rectangles R_1 and R_2 from before give a uniform covering of $\mathcal{A}(1)$ with 2 rectangles.

Constructing uniform coverings (I)

- To finish the proof we will show that $\mathcal{A}_{\mathcal{S}_+^d}(d)$ has a uniform covering with $k < 3^d$ rectangles.
- We need to understand the sparsity pattern of elements in $\mathcal{A}_{\mathcal{S}_+^d}(d)$.
- For $d = 2$ we can do things by hand:

Lemma

Any 4×4 matrix $M \in \mathcal{A}_{\mathcal{S}_+^2}(2)$ has one of the following six sparsity patterns below:

$$(1) \begin{bmatrix} \times & \times & \times & \times \\ \times & 0 & 0 & 0 \\ \times & 0 & 0 & 0 \\ \times & 0 & 0 & ? \end{bmatrix} \text{ or } (2) \begin{bmatrix} \times & \times & \times & 0 \\ \times & 0 & \times & 0 \\ \times & \times & 0 & 0 \\ 0 & 0 & 0 & ? \end{bmatrix} \text{ or } (3) \begin{bmatrix} \times & \times & \times & \times \\ 0 & 0 & 0 & 0 \\ \times & \times & 0 & 0 \\ \times & 0 & 0 & ? \end{bmatrix}$$

$$\text{or } (4) \begin{bmatrix} \times & \times & \times & \times \\ \times & 0 & \times & 0 \\ 0 & 0 & 0 & 0 \\ \times & 0 & 0 & ? \end{bmatrix} \text{ or } (5) \begin{bmatrix} \times & 0 & \times & \times \\ \times & 0 & \times & 0 \\ \times & 0 & 0 & 0 \\ \times & 0 & 0 & ? \end{bmatrix} \text{ or } (6) \begin{bmatrix} \times & \times & 0 & \times \\ \times & 0 & 0 & 0 \\ \times & \times & 0 & 0 \\ \times & 0 & 0 & ? \end{bmatrix}$$

In particular $val(M) \leq 7$.

Constructing uniform coverings (II)

- For $d = 2$ one can construct a uniform covering with $k = 7$ rectangles by hand.
- What about general d? A detailed analysis of sparsity pattern seems difficult, but one can show the following fact, which will turn out to be enough:

Lemma

If M is a $2^d \times 2^d$ matrix with $M_{a,b} = 0$ when $a^T b = 1$ and $\mathrm{rank}_{psd}(M) \leq d$, then M has at least one zero on the antidiagonal, i.e., $M_{\alpha,\bar{\alpha}} = 0$ for some $\alpha \in \{0,1\}^d$.

- Using this lemma, one can use a recursive construction to show that $\mathcal{A}_{\mathcal{S}^d_+}(d)$ has a uniform covering with $k = 3^d - 1 < 3^d$ rectangles.

Recap

Using uniform coverings we showed that for any $M \in \mathcal{A}_{\mathcal{S}_+^d}(n)$

$$\text{val}(M) \leq t(d)^n$$

where $t(d) < 3$. Thus we get the following exponential lower bound:

$$\text{rank}_{\mathcal{S}_+^d}(UDISJ) \geq \frac{\text{val}(UDISJ)}{t(d)^n} = (3/t(d))^n.$$

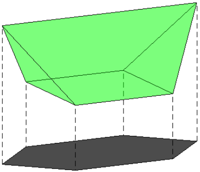
The End

Thank You!

Want to know more?

- J. Gouveia, P.A. Parrilo, R. Thomas, Lifts of convex sets and cone factorizations, *Mathematics of Operations Research*, 38:2, 2013. `arXiv:1111.3164`.
- H. Fawzi, P.A. Parrilo, New lower bounds on nonnegative rank using conic programming, `arXiv:1210.6970`.
- H. Fawzi, P.A. Parrilo, Exponential lower bounds on fixed-size psd rank and semidefinite extension complexity, `arXiv:1311.2571`.

Example: hexagon (III)



A nonnegative factorization:

$$S_H = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$