Lower bounds for matrix factorization ranks via noncommutative polynomial optimization

Monique Laurent

CWI
Centrum Wiskunde & Informatica

Tilburg University

Joint with Sander Gribling (CWI) and David de Laat (MIT)

Hierarchies, Extended Formulations and Matrix-Analytic Techniques Workshop

Simons Institute, 6–9 November 2017

Four matrix factorization ranks

- For a nonnegative $m \times n$ matrix A
  - nonnegative rank $\mathrm{rank}_+(A)$: smallest d for which $A = (\langle x_i, y_j \rangle)$ with $x_1, \ldots, x_m, y_1, \ldots, y_n \in \mathbb{R}^d_+$
  - positive semidefinite rank $\mathrm{psd\text{-}rank}(A)$: smallest d for which $A = (\langle X_i, Y_j \rangle)$ with $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ $d \times d$ Hermitian PSD
- **Symmetric ranks** for a symmetric $n \times n$ matrix A
  - completely positive rank $\mathrm{cp\text{-}rank}(A)$: smallest d for which $A = (\langle x_i, x_j \rangle)$ with $x_1, \ldots, x_n \in \mathbb{R}^d_+$ when A is **completely positive** (CP)
  - completely positive semidefinite rank $\mathrm{cpsd\text{-}rank}(A)$: smallest d for which $A = (\langle X_i, X_j \rangle)$ with $X_1, \ldots, X_n$ $d \times d$ Hermitian PSD when A is **completely positive semidefinite** (CPSD)

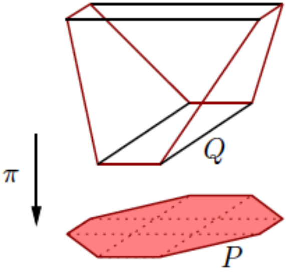
$$\mathrm{CP}^n \subseteq \mathrm{CPSD}^n \subseteq \mathrm{PSD}^n$$

Common approach to lower bound these four matrix factorization ranks using (noncommutative tracial) polynomial optimization

Motivation for rank_+ and psd-rank

rank_+ and psd-rank are used in (quantum) communication complexity, linear/semidefinite extension complexity

[Yannakakis 1991, Gouveia-Parrilo-Thomas 2013]



Motivation for CP and CPSD

- CP is used to model discrete optimization problems [de Klerk-Pasechnik'02, Burer'09]
- CPSD is used to model quantum graph parameters [L-Piovesan'15]
- CPSD used to model bipartite quantum correlations in $C_q(m,k)$

$p = (p(a,b|s,t) := \langle \Psi, A_s^a \otimes B_t^b \Psi \rangle)$, with $d \in \mathbb{N}$, $\Psi \in \mathbb{C}^d \otimes \mathbb{C}^d$ unit vector, A_s^a, B_t^b $d \times d$ Hermitian PSD, $\sum_{a=1}^k A_s^a = \sum_{b=1}^k B_t^b = I$ for $s,t \in [m]$

Smallest such d = **entanglement dimension** of p

- $C_q(m,k)$ is an affine slice of CPSD^{2mk} [Mancinska-Roberson'14] [Sikora-Varvitsiotis'15]
- If p is **synchronous**: $p(a,b|s,s) = 0$ whenever $a \neq b$, then its **entanglement dimension** is equal to $\text{cpsd-rank}(A_p)$, where $(A_p)_{(a,s),(b,t)} = p(a,b|s,t)$ [G-dL-L'17]
- $C_q(m,k)$ is not closed [Slofstra'17] [Dykema-Paulsen-Prakash'17]
  $\rightsquigarrow$ CPSD^n is not closed for $n \geq 1942$, for $n \geq 10$

Basic bounds

Upper bounds:

- For $A \in \mathbb{R}_+^{m\times n}$: $\text{psd-rank}(A) \leq \text{rank}_+(A) \leq \min\{m, n\}$
- For $A \in \text{CP}^n$: $\text{cp-rank}(A) \leq \binom{n+1}{2}$
- For $A \in \text{CPSD}^n$: No upper bound exists on cpsd-rank in terms of n

Lower bounds:

- $\text{rank}(A) \leq \text{rank}_+(A), \text{cp-rank}(A)$
- $\sqrt{\text{rank}(A)} \leq \text{psd-rank}(A), \text{cpsd-rank}(A)$

More lower bounds on rank_+ and cp-rank

- Fawzi-Parrilo (2016) define lower bounds $\tau_+(\cdot)$ and $\tau_{cp}(\cdot)$ based on the **atomic definition** of rank_+ and cp-rank:

$$\text{rank}_+(A) = \min\ d \ \text{ s.t. } A = u_1 v_1^T + \ldots + u_d v_d^T \ \text{ with } u_i, v_i \in \mathbb{R}^n_+$$

$$\tau_+(A) = \min\ \alpha \ \text{ s.t. } \frac{1}{\alpha} A \in \text{conv}(R : 0 \le R \le A,\ \text{rank}(R) \le 1)$$

$$\text{cp-rank}(A) = \min\ d \ \text{ s.t. } A = u_1 u_1^T + \ldots + u_d u_d^T \ \text{ with } u_i \in \mathbb{R}^n_+$$

$$\tau_{cp}(A) = \min\ \alpha \ \text{ s.t. } \frac{1}{\alpha} A \in \text{conv}(R : 0 \le R \le A, \text{rank}(R) \le 1, R \preceq A)$$

- Fawzi-Parrilo (2016) define SDP lower bounds $\tau_+^{sos}(\cdot)$ and $\tau_{cp}^{sos}(\cdot)$:

$$\tau_+^{sos}(A) \le \tau_+(A) \le \text{rank}_+(A),\ \ \text{rank}(A) \le \tau_{cp}^{sos}(A) \le \tau_{cp}(A) \le \text{cp-rank}(A)$$

- Link to the combinatorial 'rectangle covering' bound on rank_+:

$\text{rank}_+(A) \ge \chi(RG(A)) =$ coloring number of 'rectangle graph' $RG(A)$

$$\text{rank}_+(A) \ge \tau_+(A) \ge \chi_f(RG(A)), \qquad \text{rank}_+(A) \ge \tau_+^{sos}(A) \ge \vartheta(\overline{RG(A)})$$

New approach to bound all four factorization ranks

since no atomic definition exists for psd-rank and cpsd-rank

Commutative polynomial optimization [Lasserre, Parrilo,...]
Noncommutative eigenvalue optimization [Pironio, Navascués, Acín,...]
Noncommutative tracial optimization
[Burgdorf, Cafuta, Klep, Povh, Schweighofer,...]

$$f_*^c = \inf f(x) \text{ s.t. } x \in \mathbb{R}^n,\ g(x) \geq 0\ (g \in S) \qquad [d=1]$$

$$f_*^{nc} = \inf \mathrm{Tr}(f(\mathbf{X}))/d \text{ s.t. } d \in \mathbb{N},\ \mathbf{X} \in (H^d)^n,\ g(\mathbf{X}) \succeq 0\ (g \in S)$$

$$f_\infty^{nc} = \inf \tau(f(\mathbf{X})) \text{ s.t. } \mathcal{A}\ C^*\text{-algebra}, \tau \text{ trace}, \mathbf{X} \in \mathcal{A}^n,\ g(\mathbf{X}) \succeq 0\ (g \in S)$$

$$f_\infty^{nc} \leq f_*^{nc} \leq f_*^c$$

- SDP lower bounds: $\inf L(f)$ over $L \in \mathbb{R}[\mathbf{x}]_{2t}^*$, $L(1) = 1$..., $L \in \mathbb{R}\langle \mathbf{x} \rangle_{2t}^*$
  Under Archimedean condition: $f_t^c \longrightarrow f_*^c$, $f_t^{nc} \longrightarrow f_\infty^{nc}$ as $t \to \infty$
- Equality: $f_t^{nc} = f_*^{nc}$, $f_t^c = f_*^c$ if order t bound has flat optimal sol.

For lower bounding matrix factorization ranks: use the same framework, but now minimize $L(1)$ with L **not normalized** s.t. ...

Polynomial optimization approach for cpsd-rank

Assume $A = (\mathrm{Tr}(X_i X_j))$ has a factorization by $d \times d$ Hermitian PSD matrices $\mathbf{X} = (X_1, \ldots, X_n)$ and $d = \text{cpsd-rank}(A)$.
Let $L \in \mathbb{R}\langle x_1, \ldots, x_n \rangle^*$ be the real part of the trace evaluation $L_{\mathbf{X}}$ at $\mathbf{X}$:

$$L_{\mathbf{X}}(p) = \mathrm{Tr}(p(\mathbf{X})), \quad L(p) = \mathrm{Re}(\mathrm{Tr}(p(\mathbf{X}))) \quad \text{for } p \in \mathbb{R}\langle x_1, \ldots, x_n \rangle$$

(0) $L(1) = d$

(1) $L(x_i x_j) = A_{ij}$ for all $i, j \in [n]$

(2) L is symmetric $(L(p^*) = L(p))$, tracial $(L(pq) = L(qp))$

(3) L is positive $(L(p^* p) \geq 0)$

(4) L positive on localizing polynomials: $L(p^*(\sqrt{A_{ii}} x_i - x_i^2) p) \geq 0 \; \forall i$

$$L \geq 0 \text{ on } \underbrace{\mathrm{cone}\{p^* g p : g \in \{1\} \cup \overbrace{\{\sqrt{A_{ii}} x_i - x_i^2 : i \in [n]\}}^{S_A^{\mathrm{cpsd}}}, p \in \mathbb{R}\langle \mathbf{x} \rangle\}}_{\mathcal{M}(S_A^{\mathrm{cpsd}})}$$

Get lower bounds by minimizing $L(1)$ over $L \in \mathbb{R}\langle \mathbf{x} \rangle_{2t}^*$ satisfying (1)-(4).

Lower bounds for cpsd-rank

For an integer $t \in \mathbb{N} \cup \{\infty\}$

$$\xi_t^{cpsd}(A) = \min L(1) \;\text{ s.t. }\; L \in \mathbb{R}\langle \mathbf{x} \rangle_{2t}^* \text{ symmetric, tracial, } A = (L(x_i x_j))$$
$$L \geq 0 \text{ on } \mathcal{M}_{2t}(S_A^{\text{cpsd}})$$

$\xi_*^{cpsd}(A)$ is $\xi_\infty^{cpsd}(A)$ with extra constraint $\text{rank}(M(L) = (L(u^* v))) < \infty$

$$\xi_1^{cpsd}(A) \leq \ldots \leq \xi_t^{cpsd}(A) \leq \ldots \leq \xi_\infty^{cpsd}(A) \leq \xi_*^{cpsd}(A) \leq \text{cpsd-rank}(A)$$

- Asymptotic convergence: $\xi_t^{cpsd}(A) \to \xi_\infty^{cpsd}(A)$ as $t \to \infty$

  $\xi_\infty^{cpsd}(A) = \min \;\alpha$ s.t. $\frac{1}{\alpha} A = (\tau(X_i X_j))$, where $\mathcal{A}$ C^*-algebra with trace τ, $\mathbf{X} \in \mathcal{A}^n$ s.t. $\sqrt{A_{ii}} X_i - X_i^2 \succeq 0$ for $i \in [n]$

- $\xi_*^{cpsd}(A) = \min \;\alpha$ s.t. ... $\mathcal{A}$ finite dimensional ...

  $= \min \; L(1)$ s.t. L conic combination of trace evaluations ...

- $\xi_t^{cpsd}(A) = \xi_*^{cpsd}(A)$ if $\xi_t^{cpsd}(A)$ has a flat optimal solution

Strengthening and extending the bounds

One can strengthen the basic bounds by adding constraints on L:

1. $L(p^*(v^\mathsf{T} A v - (\sum_i v_i x_i)^2)p) \geq 0$ for all $v \in \mathbb{R}^n$ [v-constraints]
2. $L(p^* g p g') \geq 0$ for g, g' localizing for A [Berta et al.'16]
3. $L(p x_i x_j) = 0$ if $A_{ij} = 0$ [zeros propagate]
4. $L(p(\sum_i v_i x_i)) = 0$ for all $v \in \ker A$ [kernel vectors propagate]

One can extend the bounds:

- Asymmetric setting (for rank_+ and psd-rank): use two sets of variables $x_1, \ldots, x_m, y_1, \ldots, y_n$
- Commutative setting (for rank_+ and cp-rank): use polynomials in commutative variables, after viewing nonnegative vectors as diagonal PSD matrices

Small example

Consider $A = \begin{pmatrix} 1 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 1 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1 & 1/2 \\ 1/2 & 0 & 0 & 1/2 & 1 \end{pmatrix}$

- $\text{cpsd-rank}(A) \leq 5$

  because if $\mathbf{X} = \text{Diag}(1,1,0,0,0)$ and its cyclic shifts
  then $\mathbf{X}/\sqrt{2}$ is a factorization of A

- $L = \frac{1}{2} L_{\mathbf{X}}$ is feasible for $\xi_*^{cpsd}(A)$, with value $L(1) = 5/2$

  Hence $\xi_*^{cpsd}(A) \leq 5/2$, in fact $\xi_1^{cpsd}(A) = \xi_*^{cpsd}(A) = 5/2$

- $\xi_{2,V}^{cpsd}(A) = 5 = \text{cpsd-rank}(A)$
  with the v-constraints for $v = (1,-1,1,-1,1)$ and its cyclic shifts

Lower bounds for cp-rank

$$\xi_t^{cp}(A) = \min L(1) \;\text{ s.t. }\; L \in \mathbb{R}[\mathbf{x}]_{2t}^*,\ A = (L(x_i x_j)),\ L \geq 0 \text{ on } \mathcal{M}_{2t}(S_A^{\mathrm{cp}})$$

where $S_A^{\mathrm{cp}} = \{\sqrt{A_{ii}} x_i - x_i^2 : i \in [n]\} \cup \{A_{ij} - x_i x_j : i, j \in [n]\}$

$\xi_{t,\dagger}^{cp}(A)$ has the additional constraints:

(P) $L(ug) \geq 0$ for $g \in \{1\} \cup S_A^{\mathrm{cp}}$ and monomials u with $\deg(ug) \leq 2t$

(T) $A^{\otimes l} - (L(u^* v))_{u,v \in \langle \mathbf{x} \rangle_{=l}} \succeq 0 \quad$ for $2 \leq l \leq t$

Comparison to the bounds τ_{cp}^{sos} and τ_{cp} of Fawzi-Parrilo (2016):

- $\xi_t^{cp}(A) \leq \xi_\infty^{cp}(A) = \xi_*^{cp}(A) \leq \tau_{cp}(A)$
- $\tau_{cp}^{sos}(A) \leq \xi_{2,\dagger}^{cp}(A) \leq \xi_{\infty,\dagger}^{cp}(A) \leq \xi_{*,\dagger}^{cp}(A) = \tau_{cp}(A)$
- $\tau_{cp}(A)$ is also reached as asymptotic limit when using the v-constraints for a dense subset of $\mathbb{S}^{n-1}$ instead of (P)-(T)

Example: $A_{a,b} = \begin{pmatrix} (q+a)I_p & J \\ J & (p+b)I_q \end{pmatrix} \in \mathcal{S}^{p+q} \quad$ for $a, b \in [0,1]^2$

- $\xi_{2,\dagger}^{cp}(A_{a,b}) \geq pq$
- $\xi_{2,\dagger}^{cp}(A_{a,b}) = 6 = \mathrm{cp\text{-}rank}(A_{a,b})$ is tight for $(p,q) = (2,3)$

  $5 \leq \tau_{cp}^{sos} < 6$ for all nonzero $(a,b) \in [0,1]^2$, equal to 5 on subregion

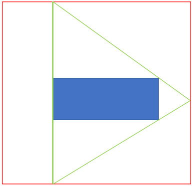
Lower bounds for rank_+ and psd-rank

Same approach: as **no a priori bound** on the eigenvalues of the factors
... **rescale** the factors to get such bounds and thus localizing constraints
Get now $\tau_+(A) = \xi_\infty^+(A)$ directly as asymptotic limit of the SDP bounds

Example for rank_+: [Fawzi-Parrilo'16]

$$S_{a,b} = \begin{pmatrix} 1-a & 1+a & 1+a & 1-a \\ 1+a & 1-a & 1-a & 1+a \\ 1-b & 1-b & 1+b & 1+b \\ 1+b & 1+b & 1-b & 1-b \end{pmatrix} \quad \text{for } a, b \in [0,1]$$

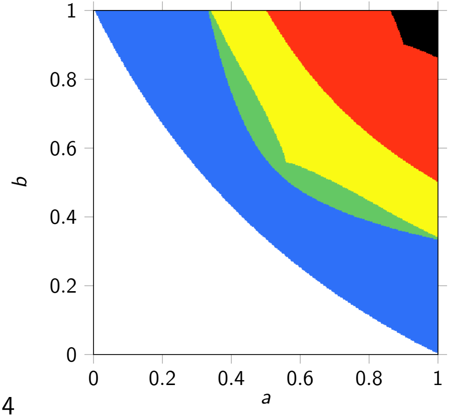
slack matrix of nested rectangles: $R = [-a,a] \times [-b,b] \subseteq P = [-1,1]^2$



$\exists$ triangle T s.t. $R \subseteq T \subseteq P \iff \mathrm{rank}_+(S_{a,b}) = 3$

Extension complexity: Nested rectangle problem

White region: $\operatorname{rank}_+(S_{a,b}) = 3 \iff (1+a)(1+b) \leq 2$



Colored: $\operatorname{rank}_+ = 4$

Top right: $\xi^+_{1,\dagger} > 3$

Two top right regions: $\tau^{\mathrm{sos}}_+ > 3$

Three top right regions: $\xi^+_{2,\dagger} > 3$

Four top right regions: $\xi^+_{3,\dagger} > 3$

$$\begin{pmatrix} 1-a & 1+a & 1+a & 1-a \\ 1+a & 1-a & 1-a & 1+a \\ 1-b & 1-b & 1+b & 1+b \\ 1+b & 1+b & 1-b & 1-b \end{pmatrix}$$

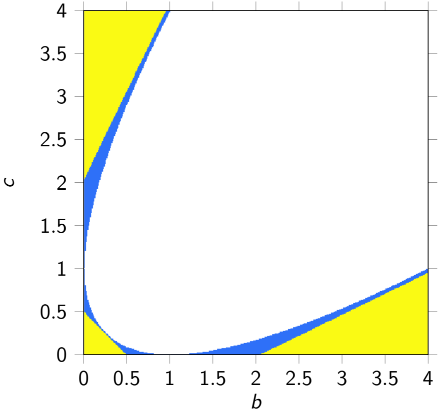
Small example for psd-rank

[Fawzi et al.'15] For $M_{b,c} = \begin{pmatrix} 1 & b & c \\ c & 1 & b \\ b & c & 1 \end{pmatrix}$

White region: $\text{psd-rank}_{\mathbb{R}} \leq 2 \iff b^2 + c^2 + 1 \leq 2(b + c + bc)$

Colored region: $\text{psd-rank}_{\mathbb{R}} = 3$

Yellow region: $\xi_2^{psd} > 2$



Concluding remarks

- Bounds via (tracial nc) polynomial optimization: arXiv:1708.01573

commutative	tracial noncommutative
completely positive cone CP^n	completely positive semidefinite cone CPSD^n
cp-rank, rank_+	cpsd-rank, psd-rank

- 'Minimizing $L(1)$' was used by [Tang-Sha'15, Nie'16] to get bounds converging to the tensor nuclear norm (commutative setting)

- The approach extends to the nonnegative tensor rank, also considered by Fawzi-Parrilo (2016) (commutative setting)

- The bounds apply to the **complex** ranks (using Hermitian factors). How to tailor the bounds for **real** ranks?

- Extension to lower bound the entanglement dimension of a (non-synchronous) quantum correlation in arXiv:1708.09696