Lower bounds for matrix factorization ranks via noncommutative polynomial optimization

Monique Laurent

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 $\text{rank}_+(A)$: smallest $d$ for which $A = (\langle x_i, y_j \rangle)$ with $x_1, \ldots, x_m, y_1, \ldots, y_n \in \mathbb{R}^d_+$
  - positive semidefinite rank $\text{psd-rank}(A)$: 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 $\text{cp-rank}(A)$: 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 $\text{cpsd-rank}(A)$: 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)

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

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 $C_q(m, k)$

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

Smallest such $d = \textbf{entanglement dimension}$ 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]

$\Rightarrow CPSD^n$ 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 $A \in \mathbb{CP}^n$: $\text{cp-rank}(A) \leq (n+1)$
- For $A \in \mathbb{CPSD}^n$: No upper bound exists on $\text{cpsd-rank}$ 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:

\[
\text{rank}_+(A) = \min d \quad \text{s.t.} \quad A = u_1v_1^T + \ldots + u_dv_d^T \quad \text{with} \quad u_i, v_i \in \mathbb{R}_+^n \\
\tau_+(A) = \min \alpha \quad \text{s.t.} \quad \frac{1}{\alpha} A \in \text{conv}(R : 0 \leq R \leq A, \text{rank}(R) \leq 1)
\]

\[
\text{cp-rank}(A) = \min d \quad \text{s.t.} \quad A = u_1u_1^T + \ldots + u_du_d^T \quad \text{with} \quad u_i \in \mathbb{R}_+^n \\
\tau_{cp}(A) = \min \alpha \quad \text{s.t.} \quad \frac{1}{\alpha} A \in \text{conv}(R : 0 \leq R \leq A, \text{rank}(R) \leq 1, R \preceq A)
\]

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

\[
\tau_{sos}^+(A) \leq \tau_+(A) \leq \text{rank}_+(A), \quad \text{rank}(A) \leq \tau_{sos}^{cp}(A) \leq \tau_{cp}(A) \leq \text{cp-rank}(A)
\]

- Link to the combinatorial ‘rectangle covering’ bound on rank\(_+\):

\[
\text{rank}_+(A) \geq \chi(RG(A)) = \text{coloring number of ‘rectangle graph’ } RG(A)
\]

\[
\text{rank}_+(A) \geq \tau_+(A) \geq \chi_f(RG(A)), \quad \text{rank}_+(A) \geq \tau_{sos}^+(A) \geq \vartheta(RG(A))
\]
New approach to bound all four factorization ranks since no atomic definition exists for \textit{psd-rank} and \textit{cpsd-rank}

\textbf{Commutative polynomial optimization} \ \ [\text{Lasserre, Parrilo,}...]

\textbf{Noncommutative eigenvalue optimization} \ \ [\text{Pironio, Navascués, Acín,}...]

\textbf{Noncommutative tracial optimization} \ \ [\text{Burgdorf, Cafuta, Klep, Povh, Schweighofer,}...]

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

\[
f_{nc}^* = \inf \frac{\text{Tr}(f(X))}{d} \quad \text{s.t.} \quad d \in \mathbb{N}, \quad X \in (H^d)^n, \quad g(X) \succeq 0 \ (g \in S)
\]

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

\[
f_{nc}^\infty \leq f_{nc}^* \leq f_c^*
\]

- SDP lower bounds: \( \inf L(f) \) over \( L \in \mathbb{R}[x]_t^*, \quad L(1) = 1... \). \( L \in \mathbb{R}\langle x \rangle_2^t \)
  
  Under Archimedean condition: \( f_t^c \rightarrow f_c^*, \quad f_t^{nc} \rightarrow f_{nc}^\infty \) as \( t \rightarrow \infty \)

- Equality: \( f_t^{nc} = f_{nc}^*, \quad f_t^c = f_c^* \) 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 $A = (\text{Tr}(X_i X_j))$ has a factorization by $d \times d$ Hermitian PSD matrices $X = (X_1, \ldots, X_n)$ 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 $L_X$ at $X$:

\[
L_X(p) = \text{Tr}(p(X)), \quad L(p) = \text{Re}(\text{Tr}(p(X))) \quad \text{for } p \in \mathbb{R}\langle x_1, \ldots, x_n \rangle
\]

(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: $L(p^* (\sqrt{A_{ii}} x_i - x_i^2) p) \geq 0 \ \forall i$

\[
L \geq 0 \text{ on } \text{cone}\{p^* gp : g \in \{1\} \cup \{\sqrt{A_{ii}} x_i - x_i^2 : i \in [n]\}, p \in \mathbb{R}\langle x \rangle\}
\]

Get lower bounds by minimizing $L(1)$ over $L \in \mathbb{R}\langle x \rangle^*_{2t}$ satisfying (1)-(4).
Lower bounds for \( \text{cpsd-rank} \)

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

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

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

\[
\xi^{\text{cpsd}}_1 (A) \leq \ldots \leq \xi^{\text{cpsd}}_t (A) \leq \ldots \leq \xi^{\text{cpsd}}_\infty (A) \leq \xi^*_\text{cpsd}(A) \leq \text{cpsd-rank}(A) 
\]

- Asymptotic convergence: \( \xi^{\text{cpsd}}_t (A) \to \xi^{\text{cpsd}}_\infty (A) \) as \( t \to \infty \)

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

- \( \xi^*_\text{cpsd}(A) = \min \alpha \text{ s.t. } ... \mathcal{A} \text{ finite dimensional } ... 
\]

\[
= \min L(1) \text{ s.t. } L \text{ conic combination of trace evaluations } ...
\]

- \( \xi^{\text{cpsd}}_t (A) = \xi^*_\text{cpsd}(A) \) if \( \xi^{\text{cpsd}}_t (A) \) has a flat optimal solution
Strengthening and extending the bounds

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

1. $L(p^*(v^T A v - (\sum_i v_i x_i)^2)p) \geq 0$ for all $v \in \mathbb{R}^n$ [v-constraints]
2. $L(p^* gpg') \geq 0$ for $g, g'$ localizing for $A$ [Berta et al.’16]
3. $L(px_i x_j) = 0$ 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 $\text{rank}_+$ and psd-rank): use two sets of variables $x_1, \ldots, x_m, y_1, \ldots, y_n$
- Commutative setting (for $\text{rank}_+$ 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 $X = \text{Diag}(1, 1, 0, 0, 0)$ and its cyclic shifts then $X/\sqrt{2}$ is a factorization of $A$

- $L = \frac{1}{2}L_X$ is feasible for $\xi^{\text{cpsd}}_*(A)$, with value $L(1) = 5/2$

Hence $\xi^{\text{cpsd}}_*(A) \leq 5/2$, in fact $\xi^{\text{cpsd}}_1(A) = \xi^{\text{cpsd}}_*(A) = 5/2$

- $\xi^{\text{cpsd}}_{2, V}(A) = 5 = \text{cpsd-rank}(A)$

with the $v$-constraints for $v = (1, -1, 1, -1, 1)$ and its cyclic shifts
Lower bounds for \( \xi^{cp}_t(A) \) = \( \min L(1) \) s.t. \( L \in \mathbb{R}[x]_{2t}^* \), \( A = (L(x_i x_j)) \), \( L \geq 0 \) on \( M_{2t}(S^A) \)

where \( S^A = \{ \sqrt{A_{ii}} x_i - x_i^2 : i \in [n] \} \cup \{ A_{ij} - x_i x_j : i, j \in [n] \} \)

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

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

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

Comparison to the bounds \( \tau^{sos}_{cp} \) and \( \tau_{cp} \) of Fawzi-Parrilo (2016):

- \( \xi^{cp}_t(A) \leq \xi^{cp}_{\infty}(A) = \xi^{*}_{cp}(A) \leq \tau_{cp}(A) \)
- \( \tau^{sos}_{cp}(A) \leq \xi^{cp}_{2,\dagger}(A) \leq \xi^{cp}_{\infty,\dagger}(A) \leq \xi^{cp*}_{s,\dagger}(A) = \tau_{cp}(A) \)
- \( \tau_{cp}(A) \) is also reached as asymptotic limit when using the \( \nu \)-constraints for a dense subset of \( S^{n-1} \) instead of (P)-(T)

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

- \( \xi^{cp}_{2,\dagger}(A_{a,b}) \geq pq \)
- \( \xi^{cp}_{2,\dagger}(A_{a,b}) = 6 = \text{cp-rank}(A_{a,b}) \) is tight for \( (p, q) = (2, 3) \)
- \( 5 \leq \tau^{sos}_{cp} < 6 \) for all nonzero \( (a, b) \in [0, 1]^2 \), equal to 5 on subregion
Lower bounds for \( \text{rank}_+ \) and \( \text{psd-rank} \)

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

\textbf{Example for rank}_+: [Fawzi-Parrilo’16]

\[
S_{a,b} = \begin{pmatrix}
1 - a & 1 + a & 1 + a & 1 - a \\
1 + a & 1 - a & 1 - a & 1 + a \\
1 - b & 1 - b & 1 + b & 1 + b \\
1 + b & 1 + b & 1 - b & 1 - b \\
\end{pmatrix}
\]

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

\( \exists \) triangle \( T \) s.t. \( R \subseteq T \subseteq P \iff \text{rank}_+(S_{a,b}) = 3 \)
Extension complexity: Nested rectangle problem

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

Colored: \( \text{rank}_+ = 4 \)
Top right: \( \xi_{1,\dagger}^+ > 3 \)
Two top right regions: \( \tau_{\dagger}^{\text{sos}} > 3 \)
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: \( \xi_{psd}^2 > 2 \)
Concluding remarks

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

<table>
<thead>
<tr>
<th></th>
<th>commutative</th>
<th>tracial noncommutative</th>
</tr>
</thead>
<tbody>
<tr>
<td>completely positive cone</td>
<td>( \text{CP}^n )</td>
<td>completely positive semidefinite cone ( \text{CPSD}^n )</td>
</tr>
<tr>
<td>( \text{cp-rank}, \text{rank}_+ )</td>
<td>( \text{cpsd-rank}, \text{psd-rank} )</td>
<td></td>
</tr>
</tbody>
</table>

- ‘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