Polynomial optimization bounds for the (sparse) completely positive rank

Andries Steenkamp

Joint work with S. Gribling, M. Korda, M. Laurent, and V. Magron

30^{th} August 2022

Workshop on Semidefinite and Polynomial Optimization

Centrum Wiskunde & Informatica

Overview

For $A \in \mathbb{R}_+^{n \times n}$

$$\mathsf{rank}_{\mathsf{cp}}(A) \geq \tau_{\mathsf{cp}}(A) = \mathsf{VAL}^{\mathsf{id}}(A) \xleftarrow[\infty]{t} \xi_t^{\mathsf{id}}(A)$$

Overview

For $A \in \mathbb{R}_+^{n \times n}$

$$\mathrm{rank_{cp}}(A) \geq \tau_{\mathrm{cp}}(A) = \mathrm{VAL^{id}}(A) \xleftarrow[\infty]{t} \xi_t^{\mathrm{id}}(A)$$

Main Message

- A new type of constraint in $\xi_t^{\mathrm{id}}(A)$
  - $L\left((A - xx^T) \otimes [x]_{t-1}[x]_{t-1}^T\right) \succeq 0$
- Glimpses of "Ideal Sparsity" when A has zero entries.

Completely Positive (CP) rank

For $A \in \mathbb{R}_+^{\mathrm{n}\times\mathrm{n}}$

$$\mathrm{rank_{cp}}(A) := \min\left\{ r \in \mathbb{N} : \exists \;\; a_1, \cdots, a_r \in \mathbb{R}_+^n \text{ s.t } A = \sum_{\ell=1}^{r} a_\ell a_\ell^T \right\}$$

Completely Positive (CP) rank

For $A \in \mathbb{R}_{+}^{\mathrm{n} \times \mathrm{n}}$

$$\mathrm{rank}_{\mathrm{cp}}(A) := \min \left\{ r \in \mathbb{N} : \exists \;\; a_1, \cdots, a_r \in \mathbb{R}_{+}^{n} \text{ s.t } A = \sum_{\ell=1}^{r} a_\ell a_\ell^T \right\}$$

Applications in:

- Quantum information theory
- Copositive optimization
- Experimental block design

- Hard to compute
- One among many M.F.R

Convexification of $\text{rank}_{\text{cp}}(A)$

$$\text{rank}_{\text{cp}}(A) := \min\left\{ r \in \mathbb{N} : \exists \;\; a_1, \cdots, a_r \in \mathbb{R}^n_+ \text{ s.t } A = \sum_{\ell=1}^{r} a_\ell a_\ell^T \right\}$$

$$\tau_{\text{cp}}(A) := \inf\left\{ \lambda : \; \frac{1}{\lambda} A \in \text{conv}\left\{ xx^T : \begin{array}{c} x \geq 0 \\ A - xx^T \geq 0 \\ A - xx^T \succeq 0 \end{array} \right\} \right\}$$

Hamza Fawzi and Pablo A Parrilo. "Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank". Mathematical Programming 158(1) (2016), pp. 417–465.

Convexification of $\text{rank}_{\text{cp}}(A)$

$$\text{rank}_{\text{cp}}(A) := \min \left\{ r \in \mathbb{N} : \exists \;\; a_1, \cdots, a_r \in \mathbb{R}^n_+ \text{ s.t } A = \sum_{\ell=1}^{r} a_\ell a_\ell^T \right\}$$

$$\tau_{\text{cp}}(A) := \inf \left\{ \lambda : \; \frac{1}{\lambda} A \in \text{conv}\left\{ xx^T : \begin{array}{c} x \geq 0 \\ A - xx^T \geq 0 \\ A - xx^T \succeq 0 \end{array} \right\} \right\}$$

- No clear way to compute.
- Example of $\tau_{\text{cp}}(A) < \text{rank}_{\text{cp}}(A)$
- Can be formulated as "GMP".

Hamza Fawzi and Pablo A Parrilo. "Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank". Mathematical Programming 158(1) (2016), pp. 417–465.

$\tau_{\text{cp}}(A)$ as a GMP

$$\tau_{\text{cp}}(A) := \inf\left\{ \lambda : \frac{1}{\lambda} A \in \text{conv}\left\{ xx^T : x \geq 0, A - xx^T \geq 0,\ A - xx^T \succeq 0 \right\} \right\}$$

$$= \inf_{\mu \in \mathcal{M}(\mathbb{R}^n)} \left\{ \int 1 \; d\mu \ : \ \int x_i x_j \; d\mu = A_{ij}, \ \ \text{supp}(\mu) \subseteq K \right\}$$

$$K := \left\{ x \in \mathbb{R}^n \ : \begin{array}{lll} \sqrt{A_{ii}}\, x_i - x_i^2 & \geq 0 & (i \in [n]) \\ A_{ij} - x_i x_j & \geq 0 & (i \neq j \in [n]) \\ A - xx^T & \succcurlyeq 0 & \end{array} \right\}$$

Generalized Moment Problem (GMP)

$$\mathrm{VAL}^{id} := \inf_{\mu \in \mathcal{M}(\mathbb{R}^n)} \left\{ \int p_0 d\mu \ : \ \begin{array}{l} \int p_i d\mu \ = \ b_i \quad (i \in [N]) \\ \mathrm{supp}(\mu) \subseteq K \end{array} \right\}$$

where $p_0, p_1, ..., p_N \in \mathbb{R}[x]$, $b_1, ..., b_N \in \mathbb{R}$, and

$$K := \left\{ x \in \mathbb{R}^n \ : \ \begin{array}{ll} g_j(x) \geq 0 & (j \in [m]) \\ \prod_{i \in s} x_i = 0 & (s \in S) \end{array} \right\}$$

The moment method in a nutshell

$$L \in \mathbb{R}[x]_{2t}^{*}, \quad L : \mathbb{R}[x]_{2t} \ni p \mapsto \int p d\mu \in \mathbb{R}$$

The moment method in a nutshell

$$L \in \mathbb{R}[x]_{2t}^*, \quad L : \mathbb{R}[x]_{2t} \ni p \mapsto \int p d\mu \in \mathbb{R}$$

$$\text{supp}(\mu) \subseteq \left\{ x \in \mathbb{R}^n \;:\; \begin{array}{ll} g_j(x) \geq 0 & (j \in [m]) \\ \prod_{i \in s} x_i = 0 & (s \in S) \end{array} \right\}$$

$$\Longleftrightarrow$$

$$L \geq 0 \text{ on } \mathcal{M}(\boldsymbol{g})_{2t}, \; L = 0 \;\text{ on } \mathcal{I}_{S,2t}$$

$$\mathcal{M}(\boldsymbol{g})_{2t} := \left\{ \sum_{j=0}^{m} \sigma_i g_i : \sigma_j \in \Sigma, \; \deg(\sigma_j g_j) \leq 2t \right\}$$

$$\mathcal{I}_{S,2t} := \left\{ \sum_{s \in S} u_s \prod_{i \in S} x_i : u_s \in \mathbb{R}[x]_{2t-|s|} \right\}$$

Hierarchy

$$\xi_t^{\mathrm{id}} := \inf \Big\{ L(p_0) \ : \ L \in \mathbb{R}[x]_{2t}^* \\ L(p_i) = b_i \\ L \geq 0 \ \text{ on } \mathcal{M}(\boldsymbol{g})_{2t} \\ L = 0 \ \text{ on } \mathcal{I}_{S,2t} \Big\}$$

$$\mathcal{M}(\boldsymbol{g})_{2t} := \left\{ \sum_{j=0}^{m} \sigma_i g_i : \sigma_j \in \Sigma, \ \deg(\sigma_j g_j) \leq 2t \right\}$$

$$\mathcal{I}_{S,2t} := \left\{ \sum_{s \in S} u_s \prod_{i \in S} x_i : u_s \in \mathbb{R}[x]_{2t-|s|} \right\}$$

Convergence results

Theorem

If $\mathcal{M}(\boldsymbol{g})$ is Archimedean then $\xi_t^{\mathsf{id}} \xrightarrow[\infty]{t} \xi_\infty^{\mathsf{id}}$ and $\xi_\infty^{\mathsf{id}} = \mathsf{VAL}^{\mathsf{id}}$

$$\xi_1^{\mathsf{id}} \leq \xi_2^{\mathsf{id}} \leq \ldots \leq \xi_\infty^{\mathsf{id}} = \mathsf{VAL}^{\mathsf{id}} = \tau_{\mathrm{cp}}(A)$$

Convergence results

Theorem

If $\mathcal{M}(\boldsymbol{g})$ is Archimedean then $\xi_t^{\text{id}} \xrightarrow[\infty]{t} \xi_\infty^{\text{id}}$ and $\xi_\infty^{\text{id}} = \text{VAL}^{\text{id}}$

$$\xi_1^{\text{id}} \leq \xi_2^{\text{id}} \leq \ldots \leq \xi_\infty^{\text{id}} = \text{VAL}^{\text{id}} = \tau_{\text{cp}}(A)$$

$$\text{VAL}^{id} := \inf_{\mu \in \mathcal{M}(\mathbb{R}^n)} \left\{ \int p_0 d\mu \ : \ \begin{array}{l} \int p_i d\mu = b_i \quad (i \in [N]) \\ \text{supp}(\mu) \subseteq K \end{array} \right\}$$

has a finite atomic optimal solution if $\exists\ z_0, z_1, \ldots, z_m \in \mathbb{R}$ such that $\sum_{j=0}^{m} z_j p_j(x) > 0$ for all $x \in K$

Encoding $A - xx^T \succcurlyeq 0$

Idea 1: $L\left(p(x)[x][x]^T\right) \succcurlyeq 0$ for all principal minors $p(x)$ of $A - xx^T$

$$\Updownarrow$$

Idea 2: $L\left((v^T(A - xx^T)v)[x][x]^T\right) \succcurlyeq 0$ for $v \in \mathbb{S}^n$

$$\Uparrow$$

Our Contribution: $L\left((A - xx^T) \otimes [x][x]^T\right) \succcurlyeq 0$

Sander Gribling, David de Laat, and Monique Laurent. "Lower Bounds on Matrix Factorization Ranks via Noncommutative Polynomial Optimization".Foundations of Computational Mathematics (2019), pp. 1–58.

Sander Gribling, Monique Laurent, and Andries Steenkamp. "Bounding the separable rank via polynomial optimization". Linear Algebra and its Applications 648 (2022), pp. 1–55

The dense CP-Hierarchy

$$\begin{aligned}\xi_t^{\mathrm{id}}(A) = \min\{L(1) :& L \in \mathbb{R}[x]_{2t}^*,\\ & L\left(x_i x_j\right) = A_{ij} \quad (i,j \in V),\\ & L\left([x]_t [x]_t^T\right) \succeq 0\\ & L\left(\left(\sqrt{A_{ii}} x_i - x_i^2\right)[x]_{t-1}[x]_{t-1}^T\right) \succeq 0 \text{ for } i \in V,\\ & L\left(\left(A_{ij} - x_i x_j\right)[x]_{t-1}[x]_{t-1}^T\right) \succeq 0 \text{ for } i \neq j,\\ & L\left(\left(A - xx^T\right) \otimes [x]_{t-1}[x]_{t-1}^T\right) \succeq 0\}\end{aligned}$$

Why does it work?

Give $a_1, \ldots, a_r \in \mathbb{R}^n_+$ s.t $A = \sum_{\ell=1}^r a_\ell a_\ell^T$

Define evaluation functionals $L_{a_\ell} : \mathbb{R}[x] \ni p(x) \mapsto p(a_\ell) \in \mathbb{R}$.

Then $L := \sum_{\ell=1}^r L_{a_\ell}$ satisfies :

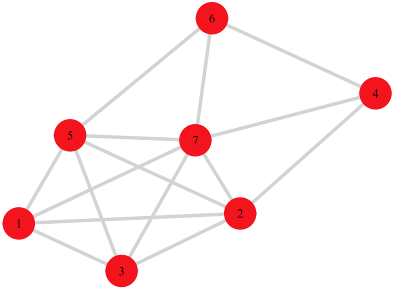
- $L(1) = \sum_{\ell=1}^r L_{a_\ell}(1) = \sum_{\ell=1}^r 1 = r$
- $L(x_i x_j) = \sum_{\ell=1}^r (a_\ell)_i (a_\ell)_j = A_{ij}$
- $L\left((A_{ij} - x_i x_j)[x]_{t-1}[x]_{t-1}^T\right)$ $= \sum_{\ell=1}^r \left(A_{ij} - (a_\ell)_i (a_\ell)_j\right)[a_\ell]_{t-1}[a_\ell]_{t-1}^T \succeq 0$
- $L\left(\left(A - xx^T\right) \otimes [x]_{t-1}[x]_{t-1}^T\right)$ $= \sum_{\ell=1}^r \left(A - a_\ell a_\ell^T\right) \otimes [a_\ell]_{t-1}[a_\ell]_{t-1}^T \succeq 0$

A touch of sparsity (a small example)

A:=

$$\begin{pmatrix} 10 & 7 & 4 & 0 & 7 & 0 & 3 \\ 7 & 10 & 3 & 4 & 4 & 0 & 4 \\ 4 & 3 & 10 & 0 & 7 & 0 & 1 \\ 0 & 4 & 0 & 10 & 0 & 1 & 7 \\ 7 & 4 & 7 & 0 & 10 & 4 & 3 \\ 0 & 0 & 0 & 1 & 4 & 10 & 5 \\ 3 & 4 & 1 & 7 & 3 & 5 & 10 \end{pmatrix}$$

Support Graph $G_A := (V, E_A)$



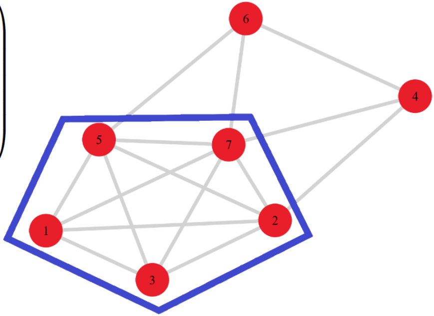
$S = \overline{E}_A := \{\{1,4\},\{1,6\},\{2,6\},\{3,4\},\{3,6\},\{4,5\}\}$
$V_1 := \{1,2,3,5,7\},\ V_2 := \{2,4,7\},\ V_3 := \{5,6,7\},\ V_4 := \{4,6,7\}$

A touch of sparsity (a small example)

A:=

$$\begin{pmatrix} 10 & 7 & 4 & 0 & 7 & 0 & 3 \\ 7 & 10 & 3 & 4 & 4 & 0 & 4 \\ 4 & 3 & 10 & 0 & 7 & 0 & 1 \\ 0 & 4 & 0 & 10 & 0 & 1 & 7 \\ 7 & 4 & 7 & 0 & 10 & 4 & 3 \\ 0 & 0 & 0 & 1 & 4 & 10 & 5 \\ 3 & 4 & 1 & 7 & 3 & 5 & 10 \end{pmatrix}$$

Support Graph G_A



$S = \overline{E}_A := \{\{1,4\},\{1,6\},\{2,6\},\{3,4\},\{3,6\},\{4,5\}\}$

$V_1 := \{1,2,3,5,7\}$, $V_2 := \{2,4,7\}$, $V_3 := \{5,6,7\}$, $V_4 := \{4,6,7\}$

A touch of sparsity (general)

Let

$$E_A := \{\{i,j\} \mid i \neq j,\ A_{ij} \neq 0\}$$

Then $V_1, \cdots, V_q$ are the maximal cliques of the support graph $G_A := (V := [n]\ ,\ E_A)$ of the matrix A.

$$K := \left\{ x \in \mathbb{R}^n \ : \begin{array}{lll} \sqrt{A_{ii}}\, x_i - x_i^2 & \geq 0 & (i \in [n]) \\ A_{ij} - x_i x_j & \geq 0 & (\{i,j\} \in E_A) \\ x_i x_j & = 0 & (\{i,j\} \in \overline{E}_A) \\ A - xx^T & \succcurlyeq 0 & \end{array} \right\}$$

Observe: $A_{i,j} = 0 \implies x_i x_j = 0$ for $x \in K$

The sparse CP-Hierarchy

$$\xi_t^{\mathrm{sp}}(A) = \min \Bigg\{ \sum_{k=1}^{q} L_k(1) :$$

$$L_k \in \mathbb{R}\left[x\left(V_k\right)\right]_{2t}^* \quad (k \in [q]),$$

$$\sum_{k \in [q]: i,j \in V_k} L_k\left(x_i x_j\right) = A_{ij} \quad \left(i, j \in V_k\right),$$

$$L_k\left(\left[x\left(V_k\right)\right]_t \left[x\left(V_k\right)\right]_t^T\right) \succeq 0 \quad (k \in [q]),$$

$$L_k\left(\left(\sqrt{A_{ii}} x_i - x_i^2\right)\left[x\left(V_k\right)\right]_{t-1}\left[x\left(V_k\right)\right]_{t-1}^T\right) \succeq 0,$$

$$L_k\left(\left(A_{ij} - x_i x_j\right)\left[x\left(V_k\right)\right]_{t-1}\left[x\left(V_k\right)\right]_{t-1}^T\right) \succeq 0,$$

$$L_k\left(\left(A - xx^T|_{V_k}\right) \otimes \left[x\left(V_k\right)\right]_{t-1}\left[x\left(V_k\right)\right]_{t-1}^T\right) \succeq 0 \Bigg\}$$

Conclusion

$$\begin{array}{ccc}
\mathrm{rank}_{cp} & & \\
\rotatebox{90}{$\geq$} & & \\
\tau_{\mathrm{cp}} & & \\
\| & & \\
\mathsf{VAL}^{\mathsf{id}} & = & \mathsf{VAL}^{\mathsf{sp}} \\
\mathsf{t}\uparrow\infty & & \mathsf{t}\uparrow\infty \\
\xi_t^{\mathsf{id}} & \leq & \xi_t^{\mathsf{sp}}
\end{array}$$

References

- **Milan Korda, Monique Laurent, Victor Magron, and Andries Steenkamp.** "Exploiting ideal-sparsity in the generalized moment problem with application to matrix factorization ranks" (2022+).
- **Hamza Fawzi and Pablo A Parrilo.** "Self-scaled bounds for atomic cone ranks: applications to nonnegative rank and cp-rank". Mathematical Programming 158(1) (2016), pp. 417–465.
- **Sander Gribling, David de Laat, and Monique Laurent.** "Lower Bounds on Matrix Factorization Ranks via Noncommutative Polynomial Optimization". Foundations of Computational Mathematics (2019), pp. 1–58.
- **Sander Gribling, Monique Laurent, and Andries Steenkamp.** "Bounding the separable rank via polynomial optimization". Linear Algebra and its Applications 648 (2022), pp. 1–55

Thank You!

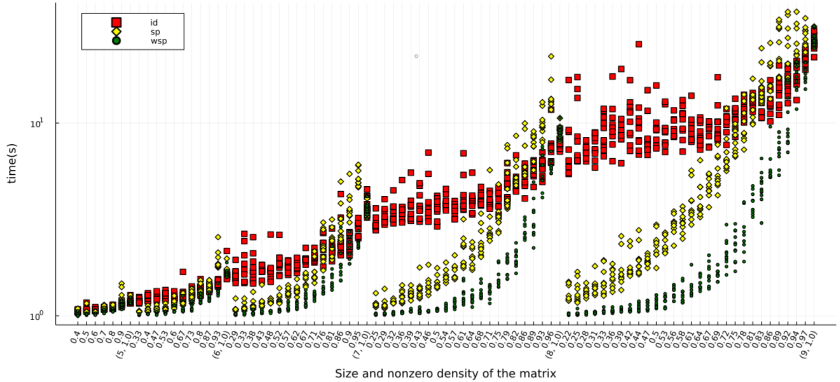
Bound on some hich cp rnak matrices

Table: Bounds for completely positive rank at level t=3.

A	rank(A)	n	$\lfloor \frac{n^2}{4} \rfloor$	$\xi^{\mathrm{id}}_{3,(2019)}(A)$	$\xi^{\mathrm{id}}_{3}(A)$	$\mathrm{rank}_{\mathrm{cp}}(A)$
M_7	7	7	12	10.5	**11.4**	14
$\tilde{M}_7$	7	7	12	10.5	10.5	14
$\tilde{M}_8$	8	8	16	13.82	**14.5**	18
$\tilde{M}_9$	9	9	20	17.74	**18.4**	26

I. M. Bomze, W. Schachinger, and R. Ullrich. " From seven to eleven: Completely positive matrices with high cp-rank.". Linear Algebra and its Applications (2014), 459:208–221.

Random cp-mats with increasing sparsity (order $t = 2$)



Bonus Slide

$$A := \begin{pmatrix} 91 & 0 & 0 & 0 & 19 & 24 & 24 & 24 & 19 & 24 & 24 & 24 \\ 0 & 42 & 0 & 0 & 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 \\ 0 & 0 & 42 & 0 & 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 \\ 0 & 0 & 0 & 42 & 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 \\ 19 & 24 & 24 & 24 & 91 & 0 & 0 & 0 & 19 & 24 & 24 & 24 \\ 24 & 6 & 6 & 6 & 0 & 42 & 0 & 0 & 24 & 6 & 6 & 6 \\ 24 & 6 & 6 & 6 & 0 & 0 & 42 & 0 & 24 & 6 & 6 & 6 \\ 24 & 6 & 6 & 6 & 0 & 0 & 0 & 42 & 24 & 6 & 6 & 6 \\ 19 & 24 & 24 & 24 & 19 & 24 & 24 & 24 & 91 & 0 & 0 & 0 \\ 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 & 0 & 42 & 0 & 0 \\ 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 & 0 & 0 & 42 & 0 \\ 24 & 6 & 6 & 6 & 24 & 6 & 6 & 6 & 0 & 0 & 0 & 42 \end{pmatrix}$$

$p = 64$, $\mathrm{rank}(A) = 10$, $\mathrm{c}(A) = 16$, $\mathrm{rank_{cp}}(A) = 37$
$\xi_1^{\mathrm{id}}(A) = 4.85$ $\xi_1^{\mathrm{sp}}(A) = 29.66$ $\xi_1^{\mathrm{wsp}}(A) = 29.63$,
$\xi_2^{\mathrm{id}}(A) = 29.57$, $\xi_2^{\mathrm{sp}}(A) = 29.66$, $\xi_2^{\mathrm{wsp}}(A) = 29.66$,
$\xi_3^{\mathrm{wsp}}(A) = 29.66$,

I. M. Bomze, W. Schachinger, and R. Ullrich. "New lower bounds and asymptotics for the cp-rank.". SIAM J. Matrix Anal. Appl.(2015), 36:20–37.