Polynomial Matrix Optimization, Matrix-Valued Moments, and Sparsity

Jie Wang

Academy of Mathematics and Systems Science, CAS

Joint work with Jared Miller and Feng Guo

ICCOPT, 07/22/2025

Outline

1. Polynomial matrix optimization and the matrix Moment-SOS hierarchy

2. Improving scalability by exploiting sparsity

Outline

The problem

- The polynomial matrix optimization problem:

$$\lambda^\star := \begin{cases} \inf_{\boldsymbol{x}\in\mathbb{R}^n} & \lambda_{\min}(F(\boldsymbol{x})) \\ \text{s.t.} & G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0 \end{cases} \qquad \text{(PMO)}$$

where $F \in \mathbb{S}[\boldsymbol{x}]^p$ and $G_k \in \mathbb{S}^{q_k}[\boldsymbol{x}], k = 1, \ldots, m$ are polynomial matrices

- Example of polynomial matrices:

$$F(x_1, x_2) = \begin{bmatrix} x_1^2 + x_2^2 & 2 + x_1x_2 + x_3^2 \\ 2 + x_1x_2 + x_3^2 & x_2x_3 \end{bmatrix}$$

The problem

- The polynomial matrix optimization problem:

$$\lambda^\star := \begin{cases} \inf\limits_{\boldsymbol{x}\in\mathbb{R}^n} & \lambda_{\min}(F(\boldsymbol{x})) \\ \text{s.t.} & G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0 \end{cases} \tag{PMO}$$

where $F \in \mathbb{S}[\boldsymbol{x}]^p$ and $G_k \in \mathbb{S}^{q_k}[\boldsymbol{x}]$, $k = 1, \ldots, m$ are polynomial matrices

- Example of polynomial matrices:

$$F(x_1, x_2) = \begin{bmatrix} x_1^2 + x_2^2 & 2 + x_1x_2 + x_3^2 \\ 2 + x_1x_2 + x_3^2 & x_2x_3 \end{bmatrix}$$

Polynomial matrix optimization

- Generalization of (scalar) polynomial optimization problems
- Applications in control theory, topology optimization, system verification, quantum information...
- Non-convex, NP-hard to find the global optimum
- Traditional optimization algorithms not applicable

Previous studies

- Moment relaxations for the case where $F(\boldsymbol{x})$ is a scalar polynomial

☞ D. Henrion and J. B. Lasserre. "Convergent relaxations of polynomial matrix inequalities and static output feedback." IEEE Transactions on Automatic Control 51.2 (2006): 192-202.

- Matrix SOS relaxations for the general case where $F(\boldsymbol{x})$ is a also matrix polynomial

☞ C. Scherer and C. Hol. "Matrix sum-of-squares relaxations for robust semi-definite programs." Mathematical programming 107.1 (2006): 189-211.

But...

The dual moment side for the general case is missing!

Previous studies

- Moment relaxations for the case where $F(\boldsymbol{x})$ is a scalar polynomial

☞ D. Henrion and J. B. Lasserre. “Convergent relaxations of polynomial matrix inequalities and static output feedback.” IEEE Transactions on Automatic Control 51.2 (2006): 192-202.

- Matrix SOS relaxations for the general case where $F(\boldsymbol{x})$ is a also matrix polynomial

☞ C. Scherer and C. Hol. “Matrix sum-of-squares relaxations for robust semi-definite programs.” Mathematical programming 107.1 (2006): 189-211.

But...

The dual moment side for the general case is missing!

Previous studies

- Moment relaxations for the case where $F(\boldsymbol{x})$ is a scalar polynomial

☞ D. Henrion and J. B. Lasserre. "Convergent relaxations of polynomial matrix inequalities and static output feedback." IEEE Transactions on Automatic Control 51.2 (2006): 192-202.

- Matrix SOS relaxations for the general case where $F(\boldsymbol{x})$ is a also matrix polynomial

☞ C. Scherer and C. Hol. "Matrix sum-of-squares relaxations for robust semi-definite programs." Mathematical programming 107.1 (2006): 189-211.

But...

The dual moment side for the general case is missing!

A scalarization approach

- By introducing auxiliary variables $\boldsymbol{y}$, one may scalarize the objective:

$$\begin{cases} \inf\limits_{\mathbf{K}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^p} & \boldsymbol{y}^\intercal F(\mathbf{K})\boldsymbol{y} \\ \text{s.t.} & G_1(\mathbf{K}) \succeq 0, \ldots, G_m(\mathbf{K}) \succeq 0 \\ & \|\boldsymbol{y}\|^2 = 1 \end{cases}$$

But...

This increases the number of variables by p and the matrix structure of $F(\boldsymbol{x})$ is destroyed.

A scalarization approach

- By introducing auxiliary variables $\boldsymbol{y}$, one may scalarize the objective:

$$\begin{cases} \inf\limits_{\mathbf{K}\in\mathbb{R}^n, \boldsymbol{y}\in\mathbb{R}^p} & \boldsymbol{y}^\intercal F(\mathbf{K})\boldsymbol{y} \\ \text{s.t.} & G_1(\mathbf{K}) \succeq 0, \ldots, G_m(\mathbf{K}) \succeq 0 \\ & \|\boldsymbol{y}\|^2 = 1 \end{cases}$$

But...

This increases the number of variables by p and the matrix structure of $F(\boldsymbol{x})$ is destroyed.

Fill in the gap

Fill in the gap

Develop a matrix version of the Moment-SOS hierarchy designed for general polynomial matrix optimization.

☞ scalar measure $\Rightarrow$ matrix-valued measure

Fill in the gap

Fill in the gap

Develop a matrix version of the Moment-SOS hierarchy designed for general polynomial matrix optimization.

☞ scalar measure $\Rightarrow$ matrix-valued measure

Reformulation with PMIs

- Let $\mathbf{K} := \{\boldsymbol{x} \in \mathbb{R}^n \mid G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0\}$

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} \lambda_{\min}(F(\boldsymbol{x})) \quad \text{s.t.} \quad G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0$$

$$\Updownarrow$$

$$\sup \lambda \quad \text{s.t.} \quad F(\boldsymbol{x}) - \lambda I_p \succeq 0, \quad \forall \boldsymbol{x} \in \mathbf{K}$$

- Require tractable approximations for $\{P(\boldsymbol{x}) \in \mathbb{S}[\mathbf{x}]^p \mid P(\boldsymbol{x}) \succeq 0, \forall \boldsymbol{x} \in \mathbf{K}\}$

Reformulation with PMIs

- Let $\mathbf{K} := \{\boldsymbol{x} \in \mathbb{R}^n \mid G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0\}$

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} \lambda_{\min}(F(\boldsymbol{x})) \quad \text{s.t.} \quad G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0$$

$$\Updownarrow$$

$$\sup \lambda \quad \text{s.t.} \quad F(\boldsymbol{x}) - \lambda I_p \succeq 0, \quad \forall \boldsymbol{x} \in \mathbf{K}$$

- Require tractable approximations for $\{P(\boldsymbol{x}) \in \mathbb{S}[\mathbf{x}]^p \mid P(\boldsymbol{x}) \succeq 0, \forall \boldsymbol{x} \in \mathbf{K}\}$

Reformulation with PMIs

- Let $\mathbf{K} \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n \mid G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0\}$

$$\inf_{\boldsymbol{x} \in \mathbb{R}^n} \lambda_{\min}(F(\boldsymbol{x})) \quad \text{s.t.} \quad G_1(\boldsymbol{x}) \succeq 0, \ldots, G_m(\boldsymbol{x}) \succeq 0$$

$$\Updownarrow$$

$$\sup \lambda \quad \text{s.t.} \quad F(\boldsymbol{x}) - \lambda I_p \succeq 0, \quad \forall \boldsymbol{x} \in \mathbf{K}$$

- Require tractable approximations for $\{P(\boldsymbol{x}) \in \mathbb{S}[\mathbf{x}]^p \mid P(\boldsymbol{x}) \succeq 0, \forall \boldsymbol{x} \in \mathbf{K}\}$

Sum-of-squares (SOS) matrices

- $P(\mathbf{x})$ is an SOS matrix if $P(\mathbf{x}) = R(\mathbf{x})^\intercal R(\mathbf{x})$
- Define the bilinear mapping $\langle \cdot, \cdot \rangle_p \colon \mathbb{S}[\mathbf{x}]^{pq} \times \mathbb{S}[\mathbf{x}]^q \to \mathbb{S}[\mathbf{x}]^p$ with

$$\langle A, B \rangle_p := \begin{bmatrix} \langle A_{11}, B \rangle & \cdots & \langle A_{1p}, B \rangle \\ \vdots & \ddots & \vdots \\ \langle A_{p1}, B \rangle & \cdots & \langle A_{pp}, B \rangle \end{bmatrix}$$

Sum-of-squares (SOS) matrices

- $P(\mathbf{x})$ is an SOS matrix if $P(\mathbf{x}) = R(\mathbf{x})^\intercal R(\mathbf{x})$

- Define the bilinear mapping $\langle \cdot, \cdot \rangle_p \colon \mathbb{S}[\mathbf{x}]^{pq} \times \mathbb{S}[\mathbf{x}]^q \to \mathbb{S}[\mathbf{x}]^p$ with

$$\langle A, B \rangle_p := \begin{bmatrix} \langle A_{11}, B \rangle & \cdots & \langle A_{1p}, B \rangle \\ \vdots & \ddots & \vdots \\ \langle A_{p1}, B \rangle & \cdots & \langle A_{pp}, B \rangle \end{bmatrix}$$

Matrix quadratic module

- Matrix quadratic module:

$$\mathcal{Q}^p(\mathbf{G}) := \left\{ S_0(\mathbf{x}) + \sum_{k=1}^{m} \langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p \;\middle|\; \begin{array}{l} S_0 \in \mathbb{S}[\mathbf{x}]^p, S_k \in \mathbb{S}[\mathbf{x}]^{pq_k} \\ S_0, \ldots, S_m \text{ are SOS matrices} \end{array} \right\}$$

- Truncated matrix quadratic module $\mathcal{Q}_r^p(\mathbf{G})$:

$$\deg(S_0(\mathbf{x})) \leq 2r, \deg(\langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p) \leq 2r$$

-

$$\mathcal{Q}_1^p(\mathbf{G}) \subseteq \mathcal{Q}_2^p(\mathbf{G}) \subseteq \cdots \subseteq \mathcal{Q}^p(\mathbf{G})$$

Matrix quadratic module

- Matrix quadratic module:

$$\mathcal{Q}^p(\mathbf{G}) := \left\{ S_0(\mathbf{x}) + \sum_{k=1}^{m} \langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p \;\middle|\; \begin{array}{l} S_0 \in \mathbb{S}[\mathbf{x}]^p, S_k \in \mathbb{S}[\mathbf{x}]^{pq_k} \\ S_0, \ldots, S_m \text{ are SOS matrices} \end{array} \right\}$$

- Truncated matrix quadratic module $\mathcal{Q}_r^p(\mathbf{G})$:

$$\deg(S_0(\mathbf{x})) \le 2r, \deg(\langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p) \le 2r$$

-

$$\mathcal{Q}_1^p(\mathbf{G}) \subseteq \mathcal{Q}_2^p(\mathbf{G}) \subseteq \cdots \subseteq \mathcal{Q}^p(\mathbf{G})$$

Matrix quadratic module

- Matrix quadratic module:

$$\mathcal{Q}^p(\mathbf{G}) \coloneqq \left\{ S_0(\mathbf{x}) + \sum_{k=1}^{m} \langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p \;\middle|\; \begin{array}{l} S_0 \in \mathbb{S}[\mathbf{x}]^p, S_k \in \mathbb{S}[\mathbf{x}]^{pq_k} \\ S_0, \ldots, S_m \text{ are SOS matrices} \end{array} \right\}$$

- Truncated matrix quadratic module $\mathcal{Q}_r^p(\mathbf{G})$:

$$\deg(S_0(\mathbf{x})) \le 2r, \deg(\langle S_k(\mathbf{x}), G_k(\mathbf{x}) \rangle_p) \le 2r$$

-

$$\mathcal{Q}_1^p(\mathbf{G}) \subseteq \mathcal{Q}_2^p(\mathbf{G}) \subseteq \cdots \subseteq \mathcal{Q}^p(\mathbf{G})$$

Archimedean Positivstellensatz for polynomial matrices

- Archimedean condition: $\exists N > 0$ and SOS matrices $S_i(\mathbf{x}) \in \mathbb{S}[\mathbf{x}]^{pq_k}$ s.t.

$$N - \|\mathbf{x}\|^2 - \sum_{k=1}^{m} \langle S_k(\mathbf{x}), G(\mathbf{x}) \rangle_p \in \mathcal{Q}^p(\mathbf{G})$$

Theorem (Scherer & Hol, 2006)

Under Archimedean condition, if $F(\mathbf{x}) \in \mathbb{S}[\mathbf{x}]^p$ is positive definite on $\mathbf{K}$, then $F(\mathbf{x}) \in \mathcal{Q}^p(\mathbf{G})$.

The SOS hierarchy

- The hierarchy of SOS relaxations:

$$\lambda_r^* := \begin{cases} \sup\limits_{\lambda} & \lambda \\ \text{s.t.} & F(\mathbf{x}) - \lambda I_p \in \mathcal{Q}_r^p(\mathbf{G}) \end{cases}$$

- $\cdots \leq \lambda_r^* \leq \lambda_{r+1}^* \leq \cdots \leq \lambda^\star$

The SOS hierarchy

- The hierarchy of SOS relaxations:

$$\lambda_r^* := \begin{cases} \sup\limits_{\lambda} & \lambda \\ \text{s.t.} & F(\mathbf{x}) - \lambda I_p \in \mathcal{Q}_r^p(\mathbf{G}) \end{cases}$$

- $\cdots \leq \lambda_r^* \leq \lambda_{r+1}^* \leq \cdots \leq \lambda^\star$

Basics for matrix-valued measures

- A $p \times p$ matrix-valued measure $\Phi \colon \mathscr{B}(\mathbf{K}) \to \mathbb{R}^{p \times p}$

$$\Phi(\mathbf{A}) := [\phi_{ij}(\mathbf{A})] \in \mathbb{R}^{p \times p}, \quad \forall \mathbf{A} \in \mathscr{B}(\mathbf{K}),$$

where each ϕ_{ij} is a Borel measure on $\mathbf{K}$

- Φ is PSD if $\Phi(\mathbf{A}) \in \mathbb{S}_+^p$ for all $\mathbf{A} \in \mathscr{B}(\mathbf{K})$

- The support of Φ is $\operatorname{supp}(\Phi) := \bigcup_{i,j=1}^{p} \operatorname{supp}(\phi_{ij})$

- $\mathfrak{M}_+^p(\mathbf{K})$: $p \times p$ PSD matrix-valued measures supported on $\mathbf{K}$

Basics for matrix-valued measures

- A $p \times p$ matrix-valued measure $\Phi\colon \mathscr{B}(\mathbf{K}) \to \mathbb{R}^{p\times p}$

$$\Phi(\mathbf{A}) := [\phi_{ij}(\mathbf{A})] \in \mathbb{R}^{p\times p}, \quad \forall \mathbf{A} \in \mathscr{B}(\mathbf{K}),$$

where each ϕ_{ij} is a Borel measure on $\mathbf{K}$

- Φ is PSD if $\Phi(\mathbf{A}) \in \mathbb{S}_+^p$ for all $\mathbf{A} \in \mathscr{B}(\mathbf{K})$

- The support of Φ is $\mathrm{supp}(\Phi) := \bigcup_{i,j=1}^{p} \mathrm{supp}(\phi_{ij})$

- $\mathfrak{M}_+^p(\mathbf{K})$: $p \times p$ PSD matrix-valued measures supported on $\mathbf{K}$

Basics for matrix-valued measures

- A $p \times p$ matrix-valued measure $\Phi \colon \mathscr{B}(\mathbf{K}) \to \mathbb{R}^{p\times p}$

$$\Phi(\mathbf{A}) := [\phi_{ij}(\mathbf{A})] \in \mathbb{R}^{p\times p}, \quad \forall \mathbf{A} \in \mathscr{B}(\mathbf{K}),$$

where each ϕ_{ij} is a Borel measure on $\mathbf{K}$

- Φ is PSD if $\Phi(\mathbf{A}) \in \mathbb{S}_+^p$ for all $\mathbf{A} \in \mathscr{B}(\mathbf{K})$
- The support of Φ is $\mathrm{supp}(\Phi) := \bigcup_{i,j=1}^{p} \mathrm{supp}(\phi_{ij})$
- $\mathfrak{M}_+^p(\mathbf{K})$: $p \times p$ PSD matrix-valued measures supported on $\mathbf{K}$

Basics for matrix-valued measures

- A $p \times p$ matrix-valued measure $\Phi\colon \mathscr{B}(\mathbf{K}) \to \mathbb{R}^{p\times p}$

$$\Phi(\mathbf{A}) := [\phi_{ij}(\mathbf{A})] \in \mathbb{R}^{p\times p}, \quad \forall \mathbf{A} \in \mathscr{B}(\mathbf{K}),$$

where each ϕ_{ij} is a Borel measure on $\mathbf{K}$

- Φ is PSD if $\Phi(\mathbf{A}) \in \mathbb{S}_+^p$ for all $\mathbf{A} \in \mathscr{B}(\mathbf{K})$
- The support of Φ is $\mathrm{supp}(\Phi) := \bigcup_{i,j=1}^p \mathrm{supp}(\phi_{ij})$
- $\mathfrak{M}_+^p(\mathbf{K})$: $p \times p$ PSD matrix-valued measures supported on $\mathbf{K}$

Reformulation with matrix-valued measures

- The moment of the matrix-valued measure Φ:

$$\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\Phi(\mathbf{x}) := \left[\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\phi_{ij}(\mathbf{x}) \right]_{i,j=1,\ldots,p} \in \mathbb{R}^{p \times p}$$

- The integral w.r.t. a matrix-valued measure:

$$\int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle = \sum_{\boldsymbol{\alpha} \in \mathrm{supp}(F)} \left\langle F_{\boldsymbol{\alpha}}, \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\Phi(\mathbf{x}) \right\rangle$$

- (PMO) is equivalent to

$$\inf_{\phi \in \mathfrak{M}_{+}^{p}(\mathbf{K})} \left\{ \int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle : \Phi(\mathbf{K}) = I_p \right\}$$

Reformulation with matrix-valued measures

- The moment of the matrix-valued measure Φ:

$$\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\Phi(\mathbf{x}) := \left[\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\phi_{ij}(\mathbf{x}) \right]_{i,j=1,\ldots,p} \in \mathbb{R}^{p \times p}$$

- The integral w.r.t. a matrix-valued measure:

$$\int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle = \sum_{\boldsymbol{\alpha} \in \mathrm{supp}(F)} \left\langle F_{\boldsymbol{\alpha}}, \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\Phi(\mathbf{x}) \right\rangle$$

- (PMO) is equivalent to

$$\inf_{\phi \in \mathfrak{M}_{+}^{p}(\mathbf{K})} \left\{ \int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle : \Phi(\mathbf{K}) = I_p \right\}$$

Reformulation with matrix-valued measures

- The moment of the matrix-valued measure Φ:

$$\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \,\mathrm{d}\Phi(\mathbf{x}) := \left[\int_{\mathbf{K}} \boldsymbol{x}^{\boldsymbol{\alpha}} \,\mathrm{d}\phi_{ij}(\mathbf{x})\right]_{i,j=1,\ldots,p} \in \mathbb{R}^{p\times p}$$

- The integral w.r.t. a matrix-valued measure:

$$\int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x})\rangle = \sum_{\boldsymbol{\alpha}\in\mathrm{supp}(F)} \left\langle F_{\boldsymbol{\alpha}}, \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\Phi(\mathbf{x})\right\rangle$$

- (PMO) is equivalent to

$$\inf_{\phi\in\mathfrak{M}_{+}^{p}(\mathbf{K})} \left\{\int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x})\rangle : \Phi(\mathbf{K}) = I_p\right\}$$

Reformulation with matrix-valued moments

$$\inf_{\phi \in \mathfrak{M}_+^p(\mathbf{K})} \left\{ \int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle : \Phi(\mathbf{K}) = I_p \right\}$$

$$\Updownarrow$$

$$\inf_{\mathbf{S}=\{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n}} \left\{ \mathscr{L}_{\mathbf{S}}(F) = \sum_{\boldsymbol{\alpha}\in\mathrm{supp}(F)} \langle F_{\boldsymbol{\alpha}}, S_{\boldsymbol{\alpha}} \rangle : \exists \Phi \in \mathfrak{M}_+^p(\mathbf{K}) \text{ s.t. } \mathbf{S} \sim \Phi \text{ and } S_{\mathbf{0}} = I_p \right\}$$

The matrix-valued K-moment problem

When does a matrix-valued sequence $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{S}^p$ admit a representing measure $\Phi \in \mathfrak{M}_+^p(\mathbf{K})$?

Reformulation with matrix-valued moments

$$\inf_{\phi \in \mathfrak{M}_+^p(\mathbf{K})} \left\{ \int_{\mathbf{K}} \langle F(\mathbf{x}), \mathrm{d}\Phi(\mathbf{x}) \rangle : \Phi(\mathbf{K}) = I_p \right\}$$

$$\Updownarrow$$

$$\inf_{\mathbf{S}=\{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n}} \left\{ \mathscr{L}_{\mathbf{S}}(F) = \sum_{\boldsymbol{\alpha}\in\mathrm{supp}(F)} \langle F_{\boldsymbol{\alpha}}, S_{\boldsymbol{\alpha}} \rangle : \exists \Phi \in \mathfrak{M}_+^p(\mathbf{K}) \text{ s.t. } \mathbf{S} \sim \Phi \text{ and } S_{\mathbf{0}} = I_p \right\}$$

The matrix-valued K-moment problem

When does a matrix-valued sequence $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{S}^p$ admit a representing measure $\Phi \in \mathfrak{M}_+^p(\mathbf{K})$?

The matrix-valued $\mathbf{K}$-moment problem

Given $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{S}^p$ and $G(\mathbf{x}) = \sum_{\gamma\in\mathrm{supp}(G)} G_\gamma \mathbf{x}^\gamma \in \mathbb{S}[\mathbf{x}]^q$:

- The d-th order moment matrix: $M_d(\mathbf{S}) = [S_{\boldsymbol{\alpha}+\boldsymbol{\beta}}]_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^n_d}$
- The d-th order localizing matrix:

$$M_d(G\mathbf{S}) = \left[\sum_{\gamma\in\mathrm{supp}(G)} S_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}} \otimes G_\gamma\right]_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^n_d}$$

Theorem (Cimprič and Zalar, 2013)

Under Archimedean condition, a matrix-valued sequence $\mathbf{S} = \{S_\alpha\}_{\alpha\in\mathbb{N}^n}$ admits a representing measure $\Phi \in \mathfrak{M}^p_+(\mathbf{K})$ iff $M_d(\mathbf{S}) \succeq 0, M_d(G_k\mathbf{S}) \succeq 0$ for all $d \geq 0$ and $k = 1, \ldots, m$.

The matrix-valued $\mathbf{K}$-moment problem

Given $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{S}^p$ and $G(\mathbf{x}) = \sum_{\gamma\in\mathrm{supp}(G)} G_\gamma \mathbf{x}^\gamma \in \mathbb{S}[\mathbf{x}]^q$:

- The d-th order moment matrix: $M_d(\mathbf{S}) = [S_{\boldsymbol{\alpha}+\boldsymbol{\beta}}]_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^n_d}$
- The d-th order localizing matrix:

$$M_d(G\mathbf{S}) = \left[\sum_{\gamma\in\mathrm{supp}(G)} S_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\gamma} \otimes G_\gamma\right]_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{N}^n_d}$$

Theorem (Cimprič and Zalar, 2013)

Under Archimedean condition, a matrix-valued sequence $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n}$ admits a representing measure $\Phi \in \mathfrak{M}^p_+(\mathbf{K})$ iff $M_d(\mathbf{S}) \succeq 0$, $M_d(G_k\mathbf{S}) \succeq 0$ for all $d \geq 0$ and $k = 1, \ldots, m$.

The truncated matrix-valued $\mathbf{K}$-moment problem

Theorem (Guo and Wang, 2024)

Given a truncated matrix-valued sequence $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n_{2d}} \subseteq \mathbb{S}^p$, *the following statements are equivalent:*

- $\mathbf{S}$ *admits an atomic representing measure* $\Phi = \sum_{i=1}^t W_i\delta_{\mathbf{x}^{(i)}}$ *with* $W_i \in \mathbb{S}^p_+$, $\mathbf{x}^{(i)} \in \mathbf{K}$, *and* $\sum_{i=1}^t \mathrm{rank}(W_i) = \mathrm{rank}(M_d(\mathbf{S}))$
- $M_d(\mathbf{S}) \succeq 0$ *and* $\mathbf{S}$ *admits an extension* $\tilde{\mathbf{S}} = \{\tilde{S}_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n_{2(d+d_{\mathbf{G}})}}$ *such that* $M_{d+d_{\mathbf{G}}}(\tilde{\mathbf{S}}) \succeq 0$, $M_d(G\tilde{\mathbf{S}}) \succeq 0$ *and* $\mathrm{rank}(M_d(\mathbf{S})) = \mathrm{rank}(M_{d+d_{\mathbf{G}}}(\tilde{\mathbf{S}}))$

☞ There is a linear algebra procedure for extracting $\mathbf{x}^{(i)} \in \mathbf{K}$ and $W_i \in \mathbb{S}^p_+$

The truncated matrix-valued $\mathbf{K}$-moment problem

Theorem (Guo and Wang, 2024)

Given a truncated matrix-valued sequence $\mathbf{S} = \{S_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n_{2d}} \subseteq \mathbb{S}^p$, *the following statements are equivalent:*

- $\mathbf{S}$ *admits an atomic representing measure* $\Phi = \sum_{i=1}^t W_i \delta_{\mathbf{x}^{(i)}}$ *with* $W_i \in \mathbb{S}^p_+$, $\mathbf{x}^{(i)} \in \mathbf{K}$, *and* $\sum_{i=1}^t \mathrm{rank}(W_i) = \mathrm{rank}(M_d(\mathbf{S}))$
- $M_d(\mathbf{S}) \succeq 0$ *and* $\mathbf{S}$ *admits an extension* $\tilde{\mathbf{S}} = \{\tilde{S}_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n_{2(d+d_{\mathbf{G}})}}$ *such that* $M_{d+d_{\mathbf{G}}}(\tilde{\mathbf{S}}) \succeq 0$, $M_d(G\tilde{\mathbf{S}}) \succeq 0$ *and* $\mathrm{rank}(M_d(\mathbf{S})) = \mathrm{rank}(M_{d+d_{\mathbf{G}}}(\tilde{\mathbf{S}}))$

☞ There is a linear algebra procedure for extracting $\mathbf{x}^{(i)} \in \mathbf{K}$ and $W_i \in \mathbb{S}^p_+$

The moment hierarchy

- The hierarchy of moment relaxations:

$$\lambda_r := \begin{cases} \inf_{\mathbf{S}} & \mathscr{L}_{\mathbf{S}}(F) \\ \text{s.t.} & M_r(\mathbf{S}) \succeq 0 \\ & M_{r-d_k}(G_k\mathbf{S}) \succeq 0, \quad k = 1, \ldots, m \\ & S_{\mathbf{0}} = I_p \end{cases}$$

- $\cdots \leq \lambda_r \leq \lambda_{r+1} \leq \cdots \leq \lambda^\star$

The moment hierarchy

- The hierarchy of moment relaxations:

$$\lambda_r := \begin{cases} \inf\limits_{\mathbf{S}} & \mathscr{L}_{\mathbf{S}}(F) \\ \text{s.t.} & M_r(\mathbf{S}) \succeq 0 \\ & M_{r-d_k}(G_k\mathbf{S}) \succeq 0, \quad k = 1, \ldots, m \\ & S_{\mathbf{0}} = I_p \end{cases}$$

- $\cdots \le \lambda_r \le \lambda_{r+1} \le \cdots \le \lambda^\star$

Asymptotical convergence and global optimality

- Under Archimedean condition, asymptotical convergence holds:

$$\lambda_r \nearrow \lambda^\star \text{ and } \lambda_r^* \nearrow \lambda^\star \text{ as } r \to \infty$$

- Global optimality is certified (i.e., $\lambda_r = \lambda^\star$) whenever

$$\operatorname{rank}(M_{r-d_{\mathbf{G}}}(\mathbf{S})) = \operatorname{rank}(M_r(\mathbf{S}))$$

Asymptotical convergence and global optimality

- Under Archimedean condition, asymptotical convergence holds:

$$\lambda_r \nearrow \lambda^\star \text{ and } \lambda_r^* \nearrow \lambda^\star \text{ as } r \to \infty$$

- Global optimality is certified (i.e., $\lambda_r = \lambda^\star$) whenever

$$\mathrm{rank}(M_{r-d_{\mathbf{G}}}(\mathbf{S})) = \mathrm{rank}(M_r(\mathbf{S}))$$

Improving scalability by exploiting sparsity

- The matrix Moment-SOS hierarchy scales badly!

☞ Computational burden rapidly grows with the number of polynomial variables and the relaxation order

☞ Can rescue by exploiting various sparsities

Improving scalability by exploiting sparsity

- The matrix Moment-SOS hierarchy scales badly!

☞ Computational burden rapidly grows with the number of polynomial variables and the relaxation order

☞ Can rescue by exploiting various sparsities

Improving scalability by exploiting sparsity

- The matrix Moment-SOS hierarchy scales badly!

☞ Computational burden rapidly grows with the number of polynomial variables and the relaxation order

☞ Can rescue by exploiting various sparsities

Term sparsity

- There is an iterative procedure for exploiting term sparsity of polynomial matrices
- Leading to a bilevel hierarchy of lower bounds $\left\{\lambda_r^{(s)}\right\}_{r,s}$ on $\lambda^\star$

☞ Fixing a relaxation order r, $\left\{\lambda_r^{(s)}\right\}_s$ is monotonically non-decreasing

☞ Fixing a sparse order s, $\left\{\lambda_r^{(s)}\right\}_r$ is monotonically non-decreasing

- When the maximal chordal extension is chosen, $\left\{\lambda_r^{(s)}\right\}_s$ converges to λ_r in finitely many steps

Term sparsity

- There is an iterative procedure for exploiting term sparsity of polynomial matrices

- Leading to a bilevel hierarchy of lower bounds $\left\{\lambda_r^{(s)}\right\}_{r,s}$ on $\lambda^\star$

☞ Fixing a relaxation order r, $\left\{\lambda_r^{(s)}\right\}_s$ is monotonically non-decreasing

☞ Fixing a sparse order s, $\left\{\lambda_r^{(s)}\right\}_r$ is monotonically non-decreasing

- When the maximal chordal extension is chosen, $\left\{\lambda_r^{(s)}\right\}_s$ converges to λ_r in finitely many steps

Term sparsity

- There is an iterative procedure for exploiting term sparsity of polynomial matrices

- Leading to a bilevel hierarchy of lower bounds $\left\{\lambda_r^{(s)}\right\}_{r,s}$ on $\lambda^\star$

☞ Fixing a relaxation order r, $\left\{\lambda_r^{(s)}\right\}_s$ is monotonically non-decreasing

☞ Fixing a sparse order s, $\left\{\lambda_r^{(s)}\right\}_r$ is monotonically non-decreasing

- When the maximal chordal extension is chosen, $\left\{\lambda_r^{(s)}\right\}_s$ converges to λ_r in finitely many steps

PMI sign symmetries

- A PMI sign symmetry of $P(\boldsymbol{x}) \in \mathbb{S}^p[\boldsymbol{x}]$ is a binary vector $\theta \in \{-1, 1\}^n$ such that either $P(\theta \circ \boldsymbol{x}) = P(\boldsymbol{x})$ or there exists a complete bipartite graph $\mathcal{G}^\theta(\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = [p]$ and satisfying
  - $[P(\theta \circ \boldsymbol{x})]_{ij} = [P(\boldsymbol{x})]_{ij}$ if $i = j$ or $\{i, j\} \notin \mathcal{E}$
  - $[P(\theta \circ \boldsymbol{x})]_{ij} = -[P(\boldsymbol{x})]_{ij}$ if $\{i, j\} \in \mathcal{E}$

☞ For a PMI sign symmetry θ of $P(\boldsymbol{x})$, $P(\boldsymbol{x}) \succeq 0 \Rightarrow P(\theta \circ \boldsymbol{x}) \succeq 0$

PMI sign symmetries

- A PMI sign symmetry of $P(\boldsymbol{x}) \in \mathbb{S}^p[\boldsymbol{x}]$ is a binary vector $\theta \in \{-1, 1\}^n$ such that either $P(\theta \circ \boldsymbol{x}) = P(\boldsymbol{x})$ or there exists a complete bipartite graph $\mathcal{G}^\theta(\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = [p]$ and satisfying
  - $[P(\theta \circ \boldsymbol{x})]_{ij} = [P(\boldsymbol{x})]_{ij}$ if $i = j$ or $\{i, j\} \notin \mathcal{E}$
  - $[P(\theta \circ \boldsymbol{x})]_{ij} = -[P(\boldsymbol{x})]_{ij}$ if $\{i, j\} \in \mathcal{E}$

☞ For a PMI sign symmetry θ of $P(\boldsymbol{x})$, $P(\boldsymbol{x}) \succeq 0 \Rightarrow P(\theta \circ \boldsymbol{x}) \succeq 0$

Term sparsity and PMI sign symmetries

Theorem (Miller, Wang, and Guo, 2024)

The block structures produced by the term sparsity iterations with maximal chordal extensions converge to the one determined by the common PMI sign symmetries of $F(\boldsymbol{x}), G_1(\boldsymbol{x}), \ldots, G_m(\boldsymbol{x})$.

- Consider the example $\inf_{\boldsymbol{x}\in\mathbb{R}^2} \lambda_{\min}(F(\boldsymbol{x}))$ s.t. $1 - x_1^2 - x_2^2 \geq 0$ with

$$F(\boldsymbol{x}) = \begin{bmatrix} x_1^2 & x_1 + x_2 \\ x_1 + x_2 & x_2^2 \end{bmatrix}$$

☞ $F(\boldsymbol{x})$ has a PMI sign symmetry $\theta = (-1, -1)$, giving rise to two blocks

Term sparsity and PMI sign symmetries

Theorem (Miller, Wang, and Guo, 2024)

The block structures produced by the term sparsity iterations with maximal chordal extensions converge to the one determined by the common PMI sign symmetries of $F(\boldsymbol{x}), G_1(\boldsymbol{x}), \ldots, G_m(\boldsymbol{x})$.

- Consider the example $\inf_{\boldsymbol{x}\in\mathbb{R}^2} \lambda_{\min}(F(\boldsymbol{x}))$ s.t. $1 - x_1^2 - x_2^2 \geq 0$ with

$$F(\boldsymbol{x}) = \begin{bmatrix} x_1^2 & x_1 + x_2 \\ x_1 + x_2 & x_2^2 \end{bmatrix}$$

☞ $F(\boldsymbol{x})$ has a PMI sign symmetry $\theta = (-1, -1)$, giving rise to two blocks

Term sparsity and PMI sign symmetries

Theorem (Miller, Wang, and Guo, 2024)

The block structures produced by the term sparsity iterations with maximal chordal extensions converge to the one determined by the common PMI sign symmetries of $F(\boldsymbol{x}), G_1(\boldsymbol{x}), \ldots, G_m(\boldsymbol{x})$.

- Consider the example $\inf_{\boldsymbol{x}\in\mathbb{R}^2} \lambda_{\min}(F(\boldsymbol{x}))$ s.t. $1 - x_1^2 - x_2^2 \geq 0$ with

$$F(\boldsymbol{x}) = \begin{bmatrix} x_1^2 & x_1 + x_2 \\ x_1 + x_2 & x_2^2 \end{bmatrix}$$

☞ $F(\boldsymbol{x})$ has a PMI sign symmetry $\theta = (-1, -1)$, giving rise to two blocks

The correlative sparsity hierarchy not necessarily converge!

- A counterexample:

$$F(\mathbf{x}) = \begin{bmatrix} 2(x_1-1)^2+(x_2-1)^2+(x_2-2)^2 & 3-2x_2 \\ 3-2x_2 & 2(x_1-2)^2+(x_2-1)^2+(x_2-2)^2 \end{bmatrix}$$

and

$$\mathbf{K} = \left\{(x_1,x_2)\in\mathbb{R}^2 : 4-x_1^2\ge 0,\ 4-x_2^2\ge 0\right\}$$

- The correlative sparsity hierarchy terminates at a lower bound 0
- But $\inf_{\mathbf{x}\in\mathbf{K}} \lambda_{\min}(F(\mathbf{x})) = \frac{1}{2}$!

The correlative sparsity hierarchy not necessarily converge!

- A counterexample:

$$F(\boldsymbol{x}) = \begin{bmatrix} 2(x_1-1)^2+(x_2-1)^2+(x_2-2)^2 & 3-2x_2 \\ 3-2x_2 & 2(x_1-2)^2+(x_2-1)^2+(x_2-2)^2 \end{bmatrix}$$

and

$$\mathbf{K} = \left\{(x_1, x_2) \in \mathbb{R}^2 : 4 - x_1^2 \geq 0,\, 4 - x_2^2 \geq 0\right\}$$

- The correlative sparsity hierarchy terminates at a lower bound 0
- But $\inf_{\boldsymbol{x}\in\mathbf{K}} \lambda_{\min}(F(\boldsymbol{x})) = \frac{1}{2}$!

The correlative sparsity hierarchy not necessarily converge!

- A counterexample:

$$F(\mathbf{x}) = \begin{bmatrix} 2(x_1-1)^2+(x_2-1)^2+(x_2-2)^2 & 3-2x_2 \\ 3-2x_2 & 2(x_1-2)^2+(x_2-1)^2+(x_2-2)^2 \end{bmatrix}$$

and

$$\mathbf{K} = \left\{(x_1, x_2) \in \mathbb{R}^2 : 4 - x_1^2 \geq 0,\ 4 - x_2^2 \geq 0\right\}$$

- The correlative sparsity hierarchy terminates at a lower bound 0

- But $\inf_{\mathbf{x}\in\mathbf{K}} \lambda_{\min}(F(\mathbf{x})) = \frac{1}{2}$!

Objective matrix sparsity

Theorem (Zheng-Fantuzzi 23', Jared-Wang-Guo 24')

Let $F(\boldsymbol{x})$ be a polynomial matrix whose sparsity graph is chordal and has maximal cliques $\mathcal{C}_1, \ldots, \mathcal{C}_t$. If $F(\boldsymbol{x})$ is strictly positive definite on $\mathbf{K}$, then there exist SOS matrices $S_{k,i}(\boldsymbol{x})$ of size $q_k|\mathcal{C}_i| \times q_k|\mathcal{C}_i|$ such that

$$F(\boldsymbol{x}) = \sum_{i=1}^{t} E_{\mathcal{C}_i}^{\intercal} \left(S_{0,i}(\boldsymbol{x}) + \sum_{k=1}^{m} \langle S_{k,i}(\boldsymbol{x}), G_k(\boldsymbol{x}) \rangle_p \right) E_{\mathcal{C}_i}.$$

Constraint matrix sparsity

- Consider the PMI constraint

$$G(\boldsymbol{x}) = \begin{bmatrix} 1 - x_1^2 - x_2^2 - x_3^2 & x_1x_2x_3 & 0 \\ x_1x_2x_3 & x_3 & x_3x_4x_5 \\ 0 & x_3x_4x_5 & 1 - x_3^2 - x_4^2 - x_5^2 \end{bmatrix} \succeq 0$$

- By introducing a new variable y, $G(\boldsymbol{x}) \succeq 0$ splits as

$$G_1(x_1, x_2, x_3, y) = \begin{bmatrix} 1 - x_1^2 - x_2^2 - x_3^2 & x_1x_2x_3 \\ x_1x_2x_3 & y^2 \end{bmatrix} \succeq 0$$

$$G_2(x_3, x_4, x_5, y) = \begin{bmatrix} x_3 - y^2 & x_3x_4x_5 \\ x_3x_4x_5 & 1 - x_3^2 - x_4^2 - x_5^2 \end{bmatrix} \succeq 0$$

Constraint matrix sparsity

- Consider the PMI constraint

$$G(\boldsymbol{x}) = \begin{bmatrix} 1 - x_1^2 - x_2^2 - x_3^2 & x_1x_2x_3 & 0 \\ x_1x_2x_3 & x_3 & x_3x_4x_5 \\ 0 & x_3x_4x_5 & 1 - x_3^2 - x_4^2 - x_5^2 \end{bmatrix} \succeq 0$$

- By introducing a new variable y, $G(\boldsymbol{x}) \succeq 0$ splits as

$$G_1(x_1, x_2, x_3, y) = \begin{bmatrix} 1 - x_1^2 - x_2^2 - x_3^2 & x_1x_2x_3 \\ x_1x_2x_3 & y^2 \end{bmatrix} \succeq 0$$

$$G_2(x_3, x_4, x_5, y) = \begin{bmatrix} x_3 - y^2 & x_3x_4x_5 \\ x_3x_4x_5 & 1 - x_3^2 - x_4^2 - x_5^2 \end{bmatrix} \succeq 0$$

Software and papers

☞ Fully implemented in TSSOS:

https://github.com/wangjie212/TSSOS

☞ F. Guo and J. Wang, **A Moment-SOS Hierarchy for Robust Polynomial Matrix Inequality Optimization with SOS-Convexity**, Mathematics of Operations Research, 2024.

☞ J. Miller, J. Wang, and F. Guo, **Sparse Polynomial Matrix Optimization**, arXiv:2411.15479, 2024.

Thank You!

https://wangjie212.github.io/jiewang