Optimization with Polynomials in Diverse Forms

Jie Wang

Academy of Mathematics and Systems Science, CAS

May 20, 2025

Outline

1. Optimization with polynomials in diverse forms
2. Positivstellensatz and the moment-SOS hierarchy
3. Improve scalability by exploiting structures

Outline

1. Optimization with polynomials in diverse forms
2. Positivstellensatz and the moment-SOS hierarchy
3. Improve scalability by exploiting structures

Outline

1. Optimization with polynomials in diverse forms
2. Positivstellensatz and the moment-SOS hierarchy
3. Improve scalability by exploiting structures

Optimization with polynomials in diverse forms

Commutative polynomial optimization

- commutative polynomial optimization:

$$\begin{cases} \inf_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \ge 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

- cover a board class of continuous and discrete nonconvex optimization problems

➤ optimal power flow, computer vision, combinatorial optimization, neutral network verification, sensor network localization

Commutative polynomial optimization

- commutative polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \ge 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

- cover a board class of continuous and discrete nonconvex optimization problems

➤ optimal power flow, computer vision, combinatorial optimization, neutral network verification, sensor network localization

Commutative polynomial optimization

- commutative polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

- cover a board class of continuous and discrete nonconvex optimization problems

➤ optimal power flow, computer vision, combinatorial optimization, neutral network verification, sensor network localization

Sum-of-rational-functions optimization

- sum-of-rational-functions optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & \sum_{i=1}^N \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \\ \text{s.t.} & g_i(\mathbf{x}) \ge 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ computer vision, signal processing

Sum-of-rational-functions optimization

- sum-of-rational-functions optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & \sum_{i=1}^N \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \\ \mathrm{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ computer vision, signal processing

Polynomial matrix optimization

- polynomial matrix optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & \lambda_{\min}(F(\mathbf{x})) \\ \text{s.t.} & G_i(\mathbf{x}) \succeq 0, \quad i = 1, \ldots, m \end{cases}$$

➤ control theory, program verification

Polynomial matrix optimization

- polynomial matrix optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & \lambda_{\min}(F(\mathbf{x})) \\ \text{s.t.} & G_i(\mathbf{x}) \succeq 0, \quad i = 1, \ldots, m \end{cases}$$

➤ control theory, program verification

Complex polynomial optimization

- complex polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{z}\in\mathbb{C}^n} & f(\mathbf{z},\overline{\mathbf{z}}) \\ \text{s.t.} & g_i(\mathbf{z},\overline{\mathbf{z}}) \ge 0, \quad i = 1,\ldots,m \\ & h_j(\mathbf{z},\overline{\mathbf{z}}) = 0, \quad j = 1,\ldots,\ell \end{cases}$$

➤ power system, signal processing

Complex polynomial optimization

- complex polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{z}\in\mathbb{C}^n} & f(\mathbf{z},\overline{\mathbf{z}}) \\ \text{s.t.} & g_i(\mathbf{z},\overline{\mathbf{z}}) \ge 0, \quad i=1,\ldots,m \\ & h_j(\mathbf{z},\overline{\mathbf{z}}) = 0, \quad j=1,\ldots,\ell \end{cases}$$

➤ power system, signal processing

Trigonometric polynomial optimization

- trigonometric polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in[0,2\pi)^n} & f(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n) \\ \text{s.t.} & g_i(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n)\ge 0, \quad i=1,\ldots,m \\ & h_j(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n)=0, \quad j=1,\ldots,\ell \end{cases}$$

➤ signal processing

Trigonometric polynomial optimization

- trigonometric polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x}\in[0,2\pi)^n} & f(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n) \\ \text{s.t.} & g_i(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n)\ge 0, \quad i=1,\ldots,m \\ & h_j(\sin x_1,\ldots,\sin x_n,\cos x_1,\ldots,\cos x_n)=0, \quad j=1,\ldots,\ell \end{cases}$$

➤ signal processing

Eigenvalue optimization with noncommutative polynomials

- eigenvalue optimization with noncommutative polynomials:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\cup_{k\geq 1}(\mathbb{S}_k)^n} & \lambda_{\min}(f(\mathbf{x})) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1,\ldots,m \\ & h_j(\mathbf{x}) = 0, \quad j = 1,\ldots,\ell \end{cases}$$

➤ quantum information, ground state energy

Eigenvalue optimization with noncommutative polynomials

- eigenvalue optimization with noncommutative polynomials:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\cup_{k\geq 1}(\mathbb{S}_k)^n} & \lambda_{\min}(f(\mathbf{x})) \\ \quad\ \ \mathrm{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1,\ldots,m \\ & h_j(\mathbf{x}) = 0, \quad j = 1,\ldots,\ell \end{cases}$$

➤ quantum information, ground state energy

Trace optimization with noncommutative polynomials

- (normalized) trace optimization with noncommutative polynomials:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\cup_{k\geq 1}(\mathbb{S}_k)^n} & \mathrm{tr}(f(\mathbf{x})) \quad (\mathrm{tr}(A) := \frac{1}{k}\sum_{i=1}^k A_{ii}) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ Connes' embedding conjecture

Trace optimization with noncommutative polynomials

- (normalized) trace optimization with noncommutative polynomials:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\cup_{k\geq 1}(\mathbb{S}_k)^n} & \mathrm{tr}(f(\mathbf{x})) \quad (\mathrm{tr}(A) := \frac{1}{k}\sum_{i=1}^k A_{ii}) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ Connes' embedding conjecture

Trace polynomial optimization

- trace polynomial: $\mathrm{tr}(x_1^2)x_2x_1 + \mathrm{tr}(x_1)\mathrm{tr}(x_2x_1x_2)$, $x_1, \ldots, x_n \in \mathcal{B}(\mathcal{H})$
- trace polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x} \in \cup_{k \geq 1} (\mathbb{S}_k)^n} & \mathrm{tr}(f(\mathbf{x})) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ quantum information, maximal entanglement states

Trace polynomial optimization

- trace polynomial: $\mathrm{tr}(x_1^2)x_2x_1 + \mathrm{tr}(x_1)\mathrm{tr}(x_2x_1x_2)$, $x_1, \ldots, x_n \in \mathcal{B}(\mathcal{H})$
- trace polynomial optimization:

$$\begin{cases} \inf\limits_{\mathbf{x} \in \cup_{k \geq 1} (\mathbb{S}_k)^n} & \mathrm{tr}(f(\mathbf{x})) \\ \text{s.t.} & g_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ quantum information, maximal entanglement states

State polynomial optimization

- state polynomial: $\varsigma(x_1^2)x_2x_1 + \varsigma(x_1)\varsigma(x_2x_1x_2)$, $x_1, \ldots, x_n \in \mathcal{B}(\mathcal{H})$, ς is a formal state (i.e., a positive unital linear functional) on $\mathcal{B}(\mathcal{H})$
- state polynomial optimization:

$$\begin{cases} \inf\limits_{(\mathcal{H},\varsigma,\mathbf{x})} & \varsigma(f(\varsigma;\mathbf{x})) \\ \text{s.t.} & g_i(\varsigma;\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\varsigma;\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ quantum information, quantum states

State polynomial optimization

- state polynomial: $\varsigma(x_1^2)x_2x_1 + \varsigma(x_1)\varsigma(x_2x_1x_2)$, $x_1, \ldots, x_n \in \mathcal{B}(\mathcal{H})$, ς is a formal state (i.e., a positive unital linear functional) on $\mathcal{B}(\mathcal{H})$
- state polynomial optimization:

$$\begin{cases} \inf_{(\mathcal{H},\varsigma,\mathbf{x})} & \varsigma(f(\varsigma;\mathbf{x})) \\ \text{s.t.} & g_i(\varsigma;\mathbf{x}) \geq 0, \quad i = 1, \ldots, m \\ & h_j(\varsigma;\mathbf{x}) = 0, \quad j = 1, \ldots, \ell \end{cases}$$

➤ quantum information, quantum states

Difficult to solve

- Non-convex
- Various natures of variables
- No classical derivatives
- NP-hard

Positivstellensatz and the moment-SOS hierarchy

Various nonnegativity

- real polynomial $f \geq 0$ if $f(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$
- polynomial matrix $F \geq 0$ if $F(\mathbf{x}) \succeq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$
- complex polynomial $f \geq 0$ if $f(\mathbf{z}, \bar{\mathbf{z}}) \geq 0$, $\forall \mathbf{z} \in \mathbb{C}^n$
- noncommutative polynomial $f \geq 0$ if $f(\mathbf{x}) \succeq 0$, $\forall \mathbf{x} \in \cup_{k\geq 1}(\mathbb{S}_k)^n$
- noncommutative polynomial $\mathrm{tr}(f) \geq 0$ if $\mathrm{tr}(f(\mathbf{x})) \geq 0$, $\forall \mathbf{x} \in \cup_{k\geq 1}(\mathbb{S}_k)^n$
- trace polynomial $\mathrm{tr}(f) \geq 0$ if $\mathrm{tr}(f(\mathbf{x})) \geq 0$, $\forall \mathbf{x} \in \cup_{k\geq 1}(\mathbb{S}_k)^n$
- state polynomial $f \geq 0$ if $\varsigma(f(\varsigma; \mathbf{x})) \geq 0$, $\forall (\mathcal{H}, \varsigma, \mathbf{x})$

Sum of squares

- Sum of squares (SOS):

$$f = f_1^2 + \cdots + f_t^2 \quad \rightsquigarrow \quad f \geq 0$$

- Hilbert, 1888: "nonnegativity = SOS" $\iff$ $n = 1 || d = 2 || n = 2, d = 4$
- Artin, 1927: "nonnegative polynomials = rational SOS"
- Matrix sum of squares: $F(\mathbf{x}) = P(\mathbf{x})^\intercal P(\mathbf{x})$
- Hermitian sum of squares: $f = |f_1|^2 + \cdots + |f_t|^2$
- Sum of Hermitian squares: $f = f_1^* f_1 + \cdots + f_t^* f_t$

Sum of squares

- Sum of squares (SOS):

$$f = f_1^2 + \cdots + f_t^2 \quad \rightsquigarrow \quad f \geq 0$$

- Hilbert, 1888: "nonnegativity = SOS" $\iff n = 1 || d = 2 || n = 2, d = 4$
- Artin, 1927: "nonnegative polynomials = rational SOS"
- Matrix sum of squares: $F(\mathbf{x}) = P(\mathbf{x})^{\intercal} P(\mathbf{x})$
- Hermitian sum of squares: $f = |f_1|^2 + \cdots + |f_t|^2$
- Sum of Hermitian squares: $f = f_1^* f_1 + \cdots + f_t^* f_t$

Sum of squares

- Sum of squares (SOS):

$$f = f_1^2 + \cdots + f_t^2 \quad \rightsquigarrow \quad f \geq 0$$

- Hilbert, 1888: "nonnegativity = SOS" $\iff$ $n = 1 || d = 2 || n = 2, d = 4$
- Artin, 1927: "nonnegative polynomials = rational SOS"
- Matrix sum of squares: $F(\mathbf{x}) = P(\mathbf{x})^{\intercal} P(\mathbf{x})$
- Hermitian sum of squares: $f = |f_1|^2 + \cdots + |f_t|^2$
- Sum of Hermitian squares: $f = f_1^* f_1 + \cdots + f_t^* f_t$

Sum of squares

- Sum of squares (SOS):

$$f = f_1^2 + \cdots + f_t^2 \quad \rightsquigarrow \quad f \geq 0$$

- Hilbert, 1888: "nonnegativity = SOS" $\iff$ $n = 1 || d = 2 || n = 2, d = 4$
- Artin, 1927: "nonnegative polynomials = rational SOS"
- Matrix sum of squares: $F(\mathbf{x}) = P(\mathbf{x})^\intercal P(\mathbf{x})$
- Hermitian sum of squares: $f = |f_1|^2 + \cdots + |f_t|^2$
- Sum of Hermitian squares: $f = f_1^* f_1 + \cdots + f_t^* f_t$

Sum of squares

- Sum of squares (SOS):

$$f = f_1^2 + \cdots + f_t^2 \quad \rightsquigarrow \quad f \geq 0$$

- Hilbert, 1888: "nonnegativity = SOS" $\iff$ $n = 1 || d = 2 || n = 2, d = 4$
- Artin, 1927: "nonnegative polynomials = rational SOS"
- Matrix sum of squares: $F(\mathbf{x}) = P(\mathbf{x})^{\intercal} P(\mathbf{x})$
- Hermitian sum of squares: $f = |f_1|^2 + \cdots + |f_t|^2$
- Sum of Hermitian squares: $f = f_1^* f_1 + \cdots + f_t^* f_t$

Gram matrix and semidefinite programming

- $\deg(f) = 2d$, $[\mathbf{x}]_d := [1, x_1, \ldots, x_n, x_1^d, \ldots, x_n^d]$
- f is an SOS $\iff$ there exists a PSD matrix G s.t.

$$f = [\mathbf{x}]_d \cdot G \cdot [\mathbf{x}]_d^\intercal \quad \rightsquigarrow \quad \text{SDP}$$

- G is called a Gram matrix of f, which is of size $\binom{n+d}{n} = \binom{n+d}{d}$

Gram matrix and semidefinite programming

- $\deg(f) = 2d$, $[\mathbf{x}]_d \coloneqq [1, x_1, \ldots, x_n, x_1^d, \ldots, x_n^d]$
- f is an SOS $\iff$ there exists a PSD matrix G s.t.

$$f = [\mathbf{x}]_d \cdot G \cdot [\mathbf{x}]_d^\intercal \quad \rightsquigarrow \quad \text{SDP}$$

- G is called a Gram matrix of f, which is of size $\binom{n+d}{n} = \binom{n+d}{d}$

Gram matrix and semidefinite programming

- $\deg(f) = 2d$, $[\mathbf{x}]_d := [1, x_1, \ldots, x_n, x_1^d, \ldots, x_n^d]$
- f is an SOS $\iff$ there exists a PSD matrix G s.t.

$$f = [\mathbf{x}]_d \cdot G \cdot [\mathbf{x}]_d^\intercal \quad \rightsquigarrow \quad \text{SDP}$$

- G is called a Gram matrix of f, which is of size $\binom{n+d}{n} = \binom{n+d}{d}$

Quadratic module

- $\Sigma(\mathbf{x}) := \{f \in \mathbb{R}[\mathbf{x}] \mid f = \sum_i f_i^2, f_i \in \mathbb{R}[\mathbf{x}]\}$
- Quadratic module: Given $\mathbf{g} = \{g_i\}_{i=1}^m \subseteq \mathbb{R}[\mathbf{x}]$,

$$\mathcal{Q}(\mathbf{g}) := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), i = 0, 1, \ldots, m\right\}$$

- Truncated quadratic module:

$$\mathcal{Q}(\mathbf{g})_{2r} := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), \deg(\sigma_i g_i) \le 2r, i = 0, 1, \ldots, m\right\}$$

- Archimedean's condition ($\approx$ compactness): there exists $N > 0$ s.t.

$N - \|\mathbf{x}\|^2 \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$

Quadratic module

- $\Sigma(\mathbf{x}) := \{f \in \mathbb{R}[\mathbf{x}] \mid f = \sum_i f_i^2, f_i \in \mathbb{R}[\mathbf{x}]\}$
- Quadratic module: Given $\mathbf{g} = \{g_i\}_{i=1}^m \subseteq \mathbb{R}[\mathbf{x}]$,

$$\mathcal{Q}(\mathbf{g}) := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), i = 0, 1, \ldots, m\right\}$$

- Truncated quadratic module:

$$\mathcal{Q}(\mathbf{g})_{2r} := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), \deg(\sigma_i g_i) \le 2r, i = 0, 1, \ldots, m\right\}$$

- Archimedean's condition ($\approx$ compactness): there exists $N > 0$ s.t.

$N - \|\mathbf{x}\|^2 \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$

Quadratic module

- $\Sigma(\mathbf{x}) := \{f \in \mathbb{R}[\mathbf{x}] \mid f = \sum_i f_i^2, f_i \in \mathbb{R}[\mathbf{x}]\}$
- Quadratic module: Given $\mathbf{g} = \{g_i\}_{i=1}^m \subseteq \mathbb{R}[\mathbf{x}]$,

$$\mathcal{Q}(\mathbf{g}) := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), i = 0, 1, \ldots, m\right\}$$

- Truncated quadratic module:

$$\mathcal{Q}(\mathbf{g})_{2r} := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), \deg(\sigma_i g_i) \le 2r, i = 0, 1, \ldots, m\right\}$$

- Archimedean's condition ($\approx$ compactness): there exists $N > 0$ s.t.

$N - \|\mathbf{x}\|^2 \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$

Quadratic module

- $\Sigma(\mathbf{x}) := \{f \in \mathbb{R}[\mathbf{x}] \mid f = \sum_i f_i^2, f_i \in \mathbb{R}[\mathbf{x}]\}$
- Quadratic module: Given $\mathbf{g} = \{g_i\}_{i=1}^m \subseteq \mathbb{R}[\mathbf{x}]$,

$$\mathcal{Q}(\mathbf{g}) := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), i = 0, 1, \ldots, m\right\}$$

- Truncated quadratic module:

$$\mathcal{Q}(\mathbf{g})_{2r} := \left\{\sigma_0 + \sum_{i=1}^m \sigma_i g_i \,\middle|\, \sigma_i \in \Sigma(\mathbf{x}), \deg(\sigma_i g_i) \leq 2r, i = 0, 1, \ldots, m\right\}$$

- Archimedean's condition ($\approx$ compactness): there exists $N > 0$ s.t.

$N - \|\mathbf{x}\|^2 \in \mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$

Commutative Positivstellensatz

- $S := \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \ldots, h_\ell(\mathbf{x}) = 0\}$

Theorem (Putinar's Positivstellensatz, 1993)

Assume $\mathcal{Q}(\mathbf{g}) + \mathcal{I}(\mathbf{h})$ satisfies Archimedean's condition. If f is positive on S, then

$$f = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m + \tau_1 h_1 + \cdots + \tau_\ell h_\ell,$$

where $\sigma_0, \ldots, \sigma_m$ are SOS, and $\tau_1, \ldots, \tau_\ell$ are polynomials.

Matrix Positivstellensatz

- $S := \{\mathbf{x} \in \mathbb{R}^n \mid G(\mathbf{x}) \succeq 0\}$

Theorem (Scherer and Hol, 2006)

Assume that $\mathcal{Q}(G)$ satisfies Archimedean's condition and $F(\mathbf{x}) \in \mathbb{S}[\mathbf{x}]^p$ is positive definite on S. Then there exist SOS polynomial matrices $\Sigma_0(\mathbf{x}) \in \mathbb{S}[\mathbf{x}]^p$ and $\Sigma_1(\mathbf{x}) \in \mathbb{S}[\mathbf{x}]^{pq}$ such that

$$F(\mathbf{x}) = \Sigma_0(\mathbf{x}) + (\Sigma_1(\mathbf{x}), G(\mathbf{x}))_p.$$

Complex Positivstellensatz

- $S := \{\mathbf{z} \in \mathbb{C}^n \mid g_1(\mathbf{z}, \overline{\mathbf{z}}) \geq 0, \ldots, g_m(\mathbf{z}, \overline{\mathbf{z}}) \geq 0, |z_1|^2 + \cdots + |z_n|^2 = R\}$

Theorem (D'Angelo and Putinar, 2009)

If f is positive on S, then

$$f = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m + \tau(|z_1|^2 + \cdots + |z_n|^2 - R),$$

where $\sigma_0, \ldots, \sigma_m$ are Hermitian SOS, and τ is a self-conjugate complex polynomial.

Noncommutative Positivstellensatz

- $S := \{\mathbf{x} \in \mathcal{B}(\mathcal{H})^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0\}$

Theorem (Helton and McCullough, 2004)

Assume that $\{g_i\}_{i=1}^m$ satisfies Archimedean's condition. If f is positive on S, then

$$f = \sum_{g \in \{1\} \cup \{g_i\}_{i=1}^m} s_g^* g s_g, \quad \textit{for some } \{s_g\}_g \subseteq \mathbb{R}\langle \mathbf{x} \rangle.$$

Reformulate POPs with nonnegative polynomials

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \ldots, h_\ell(\mathbf{x}) = 0\}$

$$f_{\min} \coloneqq \sup_{\lambda} \{\lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S\}$$

- $P_S(\mathbf{x}) \coloneqq \{g \in \mathbb{R}[\mathbf{x}] \mid g \geq 0 \text{ on } S\} \rightsquigarrow$ intractable

- Replace $P_S(\mathbf{x})$ by tractable subsets $\rightsquigarrow \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r}$

Reformulate POPs with nonnegative polynomials

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \ldots, h_\ell(\mathbf{x}) = 0\}$

$$f_{\min} := \sup_{\lambda} \{\lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S\}$$

- $P_S(\mathbf{x}) := \{g \in \mathbb{R}[\mathbf{x}] \mid g \geq 0 \text{ on } S\} \rightsquigarrow$ intractable

- Replace $P_S(\mathbf{x})$ by tractable subsets $\rightsquigarrow \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r}$

Reformulate POPs with nonnegative polynomials

- $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1(\mathbf{x}) \geq 0, \ldots, g_m(\mathbf{x}) \geq 0, h_1(\mathbf{x}) = 0, \ldots, h_\ell(\mathbf{x}) = 0\}$

$$f_{\min} \coloneqq \sup_{\lambda} \{\lambda : f(\mathbf{x}) - \lambda \geq 0, \forall \mathbf{x} \in S\}$$

- $P_S(\mathbf{x}) \coloneqq \{g \in \mathbb{R}[\mathbf{x}] \mid g \geq 0 \text{ on } S\} \rightsquigarrow$ intractable

- Replace $P_S(\mathbf{x})$ by tractable subsets $\rightsquigarrow \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r}$

The hierarchy of dual SOS relaxations

- The hierarchy of dual SOS relaxations (Parrilo 2000 & Lasserre 2001)

$$\theta_r^* := \begin{cases} \sup\limits_{\lambda} & \lambda \\ \text{s.t.} & f - \lambda \in \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r} \end{cases}$$

- $\cdots \le \theta_r^* \le \theta_{r+1}^* \le \cdots \le f_{\min}$

The hierarchy of dual SOS relaxations

- The hierarchy of dual SOS relaxations (Parrilo 2000 & Lasserre 2001)

$$\theta_r^* := \begin{cases} \sup\limits_{\lambda} & \lambda \\ \text{s.t.} & f - \lambda \in \mathcal{Q}(\mathbf{g})_{2r} + \mathcal{I}(\mathbf{h})_{2r} \end{cases}$$

- $\cdots \le \theta_r^* \le \theta_{r+1}^* \le \cdots \le f_{\min}$

Measures and moments

- For a finite Borel measure μ, $y_{\boldsymbol{\alpha}} = \int_S \mathbf{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\mu \rightsquigarrow$ moment
- Reformulation using measures and moments:

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) \, \mathrm{d}\mu : \mu(S) = 1 \right\}$$

$$\Updownarrow$$

$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\alpha \in \mathrm{supp}(f)} f_\alpha y_\alpha : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ and } y_{\mathbf{0}} = 1 \right\}$$

Question: When does a sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ admits a finite Borel measure representation with support contained in S?

Measures and moments

- For a finite Borel measure μ, $y_{\boldsymbol{\alpha}} = \int_S \mathbf{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\mu \rightsquigarrow$ moment
- Reformulation using measures and moments:

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) \, \mathrm{d}\mu : \mu(S) = 1 \right\}$$

$$\Updownarrow$$

$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\boldsymbol{\alpha} \in \mathrm{supp}(f)} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ and } y_{\mathbf{0}} = 1 \right\}$$

Question: When does a sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ admits a finite Borel measure representation with support contained in S?

Measures and moments

- For a finite Borel measure μ, $y_{\boldsymbol{\alpha}} = \int_S \mathbf{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\mu \rightsquigarrow$ moment
- Reformulation using measures and moments:

$$\inf_{\mu \in \mathcal{M}(S)_+} \left\{ \int_S f(\mathbf{x}) \, \mathrm{d}\mu : \mu(S) = 1 \right\}$$

$$\Updownarrow$$

$$\inf_{\mathbf{y}} \left\{ L_{\mathbf{y}}(f) = \sum_{\boldsymbol{\alpha} \in \mathrm{supp}(f)} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} : \exists \mu \in \mathcal{M}(S)_+ \text{ s.t. } \mathbf{y} \sim \mu \text{ and } y_{\mathbf{0}} = 1 \right\}$$

Question: When does a sequence $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ admits a finite Borel measure representation with support contained in S?

Moment matrix and localizing matrix

- $\mathbb{N}^n_r := \{\boldsymbol{\beta} = (\beta_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \beta_i \le r\}$
- r-th order moment matrix $\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\beta\gamma} := y_{\beta+\gamma}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- Given $g = \sum_\alpha g_\alpha \mathbf{x}^\alpha$, r-th order localizing matrix $\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\beta\gamma} := \sum_\alpha g_\alpha y_{\alpha+\beta+\gamma}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- $\mathbf{x} = x$, $g = 1 - x^2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{array}{c|c} & \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{array}{c} 1 \\ x \\ x^2 \end{array} & \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array}, \quad \mathbf{M}_1(g\mathbf{y}) = \begin{array}{c|c} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array}$$

Moment matrix and localizing matrix

- $\mathbb{N}^n_r := \{\boldsymbol{\beta} = (\beta_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \beta_i \le r\}$
- r-th order moment matrix $\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- Given $g = \sum_{\alpha} g_{\alpha} \mathbf{x}^{\alpha}$, r-th order localizing matrix $\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := \sum_{\alpha} g_{\alpha} y_{\alpha+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- $\mathbf{x} = x$, $g = 1 - x^2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{array}{c c} & \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{array}{c} 1 \\ x \\ x^2 \end{array} & \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array}, \quad \mathbf{M}_1(g\mathbf{y}) = \begin{array}{c c} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array}$$

Moment matrix and localizing matrix

- $\mathbb{N}^n_r := \{\boldsymbol{\beta} = (\beta_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \beta_i \le r\}$
- r-th order moment matrix $\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- Given $g = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$, r-th order localizing matrix $\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- $\mathbf{x} = x$, $g = 1 - x^2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{array}{c c} & \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{array}{c} 1 \\ x \\ x^2 \end{array} & \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array}, \quad \mathbf{M}_1(g\mathbf{y}) = \begin{array}{c c} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array}$$

Moment matrix and localizing matrix

- $\mathbb{N}^n_r := \{\boldsymbol{\beta} = (\beta_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \beta_i \le r\}$
- r-th order moment matrix $\mathbf{M}_r(\mathbf{y})$:

$$[\mathbf{M}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- Given $g = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$, r-th order localizing matrix $\mathbf{M}_r(g\mathbf{y})$:

$$[\mathbf{M}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_r$$

- $\mathbf{x} = x$, $g = 1 - x^2$:

$$\mathbf{M}_2(\mathbf{y}) = \begin{array}{c c} & \begin{array}{ccc} 1 & x & x^2 \end{array} \\ \begin{array}{c} 1 \\ x \\ x^2 \end{array} & \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \end{array}, \quad \mathbf{M}_1(g\mathbf{y}) = \begin{array}{c c} & \begin{array}{cc} 1 & x \end{array} \\ \begin{array}{c} 1 \\ x \end{array} & \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} \end{array}$$

Sequences admitting a Borel representing measure

Theorem (The dual to Putinar's Positivstellensatz)

Assume that Archimedean's condition holds. The sequence $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ *admits a finite Borel representing measure with support contained in* S *if and only if* $\mathbf{M}_r(\mathbf{y}) \succeq 0, \mathbf{M}_{r-d_i}(g_i \mathbf{y}) \succeq 0, L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\alpha}} h_j) = 0$ *for all* i, j, r *and* $\boldsymbol{\alpha} \in \mathbb{N}^n$.

- $d_i := \lceil \deg(g_i)/2 \rceil, i = 1, \ldots, m$
- Riesz linear functional $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$, $f = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \mapsto \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}$

The hierarchy of moment relaxations

- The hierarchy of moment relaxations (Lasserre, 2001)

$$\theta_r := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0 \\ & \mathbf{M}_{r-d_i}(g_i\mathbf{y}) \succeq 0, \quad i = 1, \ldots, m \\ & L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\alpha}} h_j) = 0, \quad |\boldsymbol{\alpha}| \le 2r - \deg(h_j), j = 1, \ldots, \ell \\ & y_{\mathbf{0}} = 1 \end{cases}$$

- $\cdots \le \theta_r \le \theta_{r+1} \le \cdots \le f_{\min}$

The hierarchy of moment relaxations

- The hierarchy of moment relaxations (Lasserre, 2001)

$$\theta_r := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0 \\ & \mathbf{M}_{r-d_i}(g_i\mathbf{y}) \succeq 0, \quad i = 1, \ldots, m \\ & L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\alpha}} h_j) = 0, \quad |\boldsymbol{\alpha}| \le 2r - \deg(h_j), j = 1, \ldots, \ell \\ & y_{\mathbf{0}} = 1 \end{cases}$$

- $\cdots \le \theta_r \le \theta_{r+1} \le \cdots \le f_{\min}$

The Moment-SOS/Lasserre's hierarchy

$$
\begin{array}{ccccccc}
 & & & f_{\min} & & & \\
 & & \rotatebox{45}{$\leq$} & & \rotatebox{-45}{$\geq$} & & \\
 & & \vdots & & \vdots & & \\
 & & \geq & & \geq & & \\
\textcolor{red}{\text{(Moment relaxation)}} & & \theta_r & \text{“}=\text{”} & \theta_r^* & & \textcolor{red}{\text{(dual SOS relaxation)}} \\
 & & \geq & & \geq & & \\
 & & \vdots & & \vdots & & \\
 & & \geq & & \geq & & \\
 & & \theta_{r_{\min}} & \text{“}=\text{”} & \theta_{r_{\min}}^* & &
\end{array}
$$

Asymptotical convergence and finite convergence

- Under Archimedean's condition ($\approx$ compactness)

  ➤ $\theta_r \nearrow f_{\min}$ and $\theta_r^* \nearrow f_{\min}$ as $r \to \infty$ (**Putinar's Positivstellensatz**)

  ➤ Finite convergence happens generically (**Nie, 2014**)

Asymptotical convergence and finite convergence

- Under Archimedean's condition ($\approx$ compactness)

  ➤ $\theta_r \nearrow f_{\min}$ and $\theta_r^* \nearrow f_{\min}$ as $r \to \infty$ (**Putinar's Positivstellensatz**)

  ➤ Finite convergence happens generically (**Nie, 2014**)

Asymptotical convergence and finite convergence

- Under Archimedean's condition ($\approx$ compactness)
  - $\theta_r \nearrow f_{\min}$ and $\theta_r^* \nearrow f_{\min}$ as $r \to \infty$ (**Putinar's Positivstellensatz**)
  - Finite convergence happens generically (**Nie, 2014**)

Detecting global optimality

- Global optimality certified ($\theta_r = f_{\min}$) when one of the following conditions holds:
  - There exists r' with $r_0 \leq r' \leq r$, $\operatorname{rank} \mathbf{M}_{r'-r_0}(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$

    $\rightsquigarrow$ A simple algorithm for extracting $\operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$ optimal solutions
  - $\operatorname{rank} \mathbf{M}_{r_{\min}}(\mathbf{y}) = 1$

    $\rightsquigarrow$ Extract one optimal solution

Detecting global optimality

- Global optimality certified $(\theta_r = f_{\min})$ when one of the following conditions holds:
  - There exists r' with $r_0 \le r' \le r$, $\operatorname{rank} \mathbf{M}_{r'-r_0}(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$
    - $\rightsquigarrow$ A simple algorithm for extracting $\operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$ optimal solutions
  - $\operatorname{rank} \mathbf{M}_{r_{\min}}(\mathbf{y}) = 1$
    - $\rightsquigarrow$ Extract one optimal solution

Detecting global optimality

- Global optimality certified $(\theta_r = f_{\min})$ when one of the following conditions holds:
  - ➤ There exists r' with $r_0 \leq r' \leq r$, $\operatorname{rank} \mathbf{M}_{r'-r_0}(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$
    - $\rightsquigarrow$ A simple algorithm for extracting $\operatorname{rank} \mathbf{M}_{r'}(\mathbf{y})$ optimal solutions
  - ➤ $\operatorname{rank} \mathbf{M}_{r_{\min}}(\mathbf{y}) = 1$
    - $\rightsquigarrow$ Extract one optimal solution

Improve scalability by exploiting structures

Rapidly growing size of the moment-SOS hierarchy

- The size of SDP corresponding to the r-th SOS relaxation:
  1. Size of PSD constraints: $\binom{n+r}{r}$
  2. Number of equality constraints: $\binom{n+2r}{2r}$
- $r = 2$, $n < 30$ (Mosek)
- Exploiting structures:
  - POP
  - SDP

Rapidly growing size of the moment-SOS hierarchy

- The size of SDP corresponding to the r-th SOS relaxation:
  1. Size of PSD constraints: $\binom{n+r}{r}$
  2. Number of equality constraints: $\binom{n+2r}{2r}$
- $r = 2$, $n < 30$ (Mosek)
- Exploiting structures:
  - POP
  - SDP

Rapidly growing size of the moment-SOS hierarchy

- The size of SDP corresponding to the r-th SOS relaxation:
  1. Size of PSD constraints: $\binom{n+r}{r}$
  2. Number of equality constraints: $\binom{n+2r}{2r}$
- $r = 2$, $n < 30$ (Mosek)
- Exploiting structures:
  - POP
  - SDP

Quotient ring

- Equality constraints: $h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell$

- Build the moment-SOS hierarchy in the quotient ring

$$\mathbb{R}[\mathbf{x}]/(h_1(\mathbf{x}), \ldots, h_\ell(\mathbf{x}))$$

$\rightsquigarrow$ Reduce SDP sizes by using Gröbner basis

Quotient ring

- Equality constraints: $h_j(\mathbf{x}) = 0, \quad j = 1, \ldots, \ell$
- Build the moment-SOS hierarchy in the quotient ring

$$\mathbb{R}[\mathbf{x}]/(h_1(\mathbf{x}), \ldots, h_\ell(\mathbf{x}))$$

$\rightsquigarrow$ Reduce SDP sizes by using Gröbner basis

Symmetries

- permutation symmetry: $(x_1, \ldots, x_n) \to (x_{\tau(1)}, \ldots, x_{\tau(n)})$
- translation symmetry: $(x_1, \ldots, x_n) \to (x_{1+i}, \ldots, x_{n+i})$, $x_{n+i} = x_i$
- sign symmetry: $(x_{i_1}, \ldots, x_{i_k}) \to (-x_{i_1}, \ldots, -x_{i_k})$
- conjugate symmetry: $(z_{i_1}, \ldots, z_{i_k}) \to (\overline{z}_{i_1}, \ldots, \overline{z}_{i_k})$
- $\mathbb{T}$-symmetry: $(z_{i_1}, \ldots, z_{i_k}) \to (e^{\mathbf{i}\theta} z_{i_1}, \ldots, e^{\mathbf{i}\theta} z_{i_k})$

➤ Use group theory to derive block-diagonal moment-SOS relaxations

Symmetries

- permutation symmetry: $(x_1, \ldots, x_n) \to (x_{\tau(1)}, \ldots, x_{\tau(n)})$
- translation symmetry: $(x_1, \ldots, x_n) \to (x_{1+i}, \ldots, x_{n+i})$, $x_{n+i} = x_i$
- sign symmetry: $(x_{i_1}, \ldots, x_{i_k}) \to (-x_{i_1}, \ldots, -x_{i_k})$
- conjugate symmetry: $(z_{i_1}, \ldots, z_{i_k}) \to (\overline{z}_{i_1}, \ldots, \overline{z}_{i_k})$
- $\mathbb{T}$-symmetry: $(z_{i_1}, \ldots, z_{i_k}) \to (e^{\mathbf{i}\theta} z_{i_1}, \ldots, e^{\mathbf{i}\theta} z_{i_k})$

➤ Use group theory to derive block-diagonal moment-SOS relaxations

Correlative sparsity (Waki et al. 2006)

- Correlative sparsity pattern graph $G^{\mathrm{csp}}(V, E)$:
  - $V := \{x_1, \ldots, x_n\}$
  - $\{x_i, x_j\} \in E \iff x_i, x_j$ appear in the same term of f or in the same

constraint polynomial g_k

- For each maximal clique of $G^{\mathrm{csp}}(V, E)$, apply

$$I_k \longmapsto \mathbf{M}_r(\mathbf{y}, I_k), \mathbf{M}_{r-d_i}(g_i\mathbf{y}, I_k)$$

- Sparse Positivstellensatz: asymptotical convergence

Correlative sparsity (Waki et al. 2006)

- Correlative sparsity pattern graph $G^{\mathrm{csp}}(V, E)$:
  - ➤ $V := \{x_1, \ldots, x_n\}$
  - ➤ $\{x_i, x_j\} \in E \iff x_i, x_j$ appear in the same term of f or in the same

constraint polynomial g_k

- For each maximal clique of $G^{\mathrm{csp}}(V, E)$, apply

$$I_k \longmapsto \mathbf{M}_r(\mathbf{y}, I_k), \mathbf{M}_{r-d_i}(g_i\mathbf{y}, I_k)$$

- Sparse Positivstellensatz: asymptotical convergence

Correlative sparsity (Waki et al. 2006)

- Correlative sparsity pattern graph $G^{\mathrm{csp}}(V, E)$:

  ➤ $V := \{x_1, \ldots, x_n\}$

  ➤ $\{x_i, x_j\} \in E \Longleftrightarrow x_i, x_j$ appear in the same term of f or in the same

constraint polynomial g_k

- For each maximal clique of $G^{\mathrm{csp}}(V, E)$, apply

$$I_k \longmapsto \mathbf{M}_r(\mathbf{y}, I_k), \mathbf{M}_{r-d_i}(g_i\mathbf{y}, I_k)$$

- Sparse Positivstellensatz: asymptotical convergence

Term sparsity (Wang & Magron & Lasserre, 2021)

- Term sparsity pattern graph $G^{\mathrm{tsp}}(V,E)$:
  - $V := [\mathbf{x}]_r = \{1, x_1, \ldots, x_n, x_1^r, \ldots, x_n^r\}$
  - $\{\mathbf{x}^{\boldsymbol{\alpha}}, \mathbf{x}^{\boldsymbol{\beta}}\} \in E \iff \mathbf{x}^{\boldsymbol{\alpha}} \cdot \mathbf{x}^{\boldsymbol{\beta}} = \mathbf{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} \in \mathrm{supp}(f) \cup \bigcup_{i=1}^m \mathrm{supp}(g_i) \cup [\mathbf{x}]_r^2$
- Impose a sparsity pattern on $\mathbf{M}_r(\mathbf{y})$ using $G^{\mathrm{tsp}}(V,E)$

$$\begin{array}{cc} & \begin{array}{ccc} \cdots & \boldsymbol{\alpha} & \cdots \end{array} \\ \begin{array}{c} \vdots \\ \boldsymbol{\beta} \\ \vdots \end{array} & \begin{bmatrix} & \vdots & \\ \cdots & y_{\boldsymbol{\alpha}+\boldsymbol{\beta}} & \cdots \\ & \vdots & \end{bmatrix} \end{array} = \mathbf{M}_r(\mathbf{y}), \quad y_{\boldsymbol{\alpha}+\boldsymbol{\beta}} = 0 \iff \{\mathbf{x}^{\boldsymbol{\alpha}}, \mathbf{x}^{\boldsymbol{\beta}}\} \notin E$$

Structures of the moment problem

- Low-rank optimal solutions: $\mathrm{rank}(\mathbf{M}^{\mathrm{opt}}) \ll n$
- Unit diagonal: $\mathrm{diag}(\mathbf{M}) = \mathbf{1}$
- Unit trace: $\mathrm{tr}(\mathbf{M}) = 1$

Solving the moment problem via manifold optimization

- Low-rank: $\operatorname{rank} \mathbf{M}^{\mathrm{opt}} \ll n \rightsquigarrow \mathbf{M} = YY^{\intercal}$, $Y \in \mathbb{R}^{n\times p}$ **Burer-Monteiro**
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p}\}$
- Unital diagonal: $\operatorname{diag}(\mathbf{M}) = \mathbf{1}$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y(k,:)\| = 1, k = 1,\ldots,n\}$
- Unital trace: $\operatorname{tr}(\mathbf{M}) = 1$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y\|_F = 1\}$

$\rightsquigarrow$ Efficient manifold optimization algorithms

Solving the moment problem via manifold optimization

- Low-rank: $\mathrm{rank}\,\mathbf{M}^{\mathrm{opt}} \ll n \rightsquigarrow \mathbf{M} = YY^\intercal,\ Y \in \mathbb{R}^{n\times p}$ **Burer-Monteiro**
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p}\}$
- Unital diagonal: $\mathrm{diag}(\mathbf{M}) = \mathbf{1}$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y(k,:)\| = 1, k = 1, \ldots, n\}$
- Unital trace: $\mathrm{tr}(\mathbf{M}) = 1$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y\|_F = 1\}$

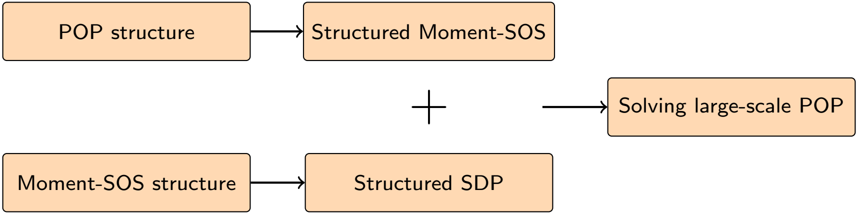
$\rightsquigarrow$ Efficient manifold optimization algorithms

Solving the moment problem via manifold optimization

- Low-rank: $\mathrm{rank}\,\mathbf{M}^{\mathrm{opt}} \ll n \rightsquigarrow \mathbf{M} = YY^\intercal,\ Y \in \mathbb{R}^{n\times p}$ **Burer-Monteiro**
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p}\}$
- Unital diagonal: $\mathrm{diag}(\mathbf{M}) = \mathbf{1}$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y(k,:)\| = 1, k = 1,\ldots,n\}$
- Unital trace: $\mathrm{tr}(\mathbf{M}) = 1$
  - $\mathcal{N} := \{Y \in \mathbb{R}^{n\times p} \mid \|Y\|_F = 1\}$

$\rightsquigarrow$ Efficient manifold optimization algorithms

Solving large-scale polynomial optimization



Software

- TSSOS: based on JuMP, support commutative/complex/matrix polynomial optimization

  https://github.com/wangjie212/TSSOS

- NCTSSOS: based on JuMP, support noncommutative/trace/state polynomial optimization

  https://github.com/wangjie212/NCTSSOS

- ManiSDP: in MATLAB, efficiently solve low-rank SDPs via manifold optimization

  https://github.com/wangjie212/ManiSDP

➤ See our papers for numerical examples and applications.

Software

- TSSOS: based on JuMP, support commutative/complex/matrix polynomial optimization

https://github.com/wangjie212/TSSOS

- NCTSSOS: based on JuMP, support noncommutative/trace/state polynomial optimization

https://github.com/wangjie212/NCTSSOS

- ManiSDP: in MATLAB, efficiently solve low-rank SDPs via manifold optimization

https://github.com/wangjie212/ManiSDP

➤ See our papers for numerical examples and applications.

Summary

- The moment-SOS hierarchy provides a power tool for **global optimization** of many difficult non-convex problems involving polynomials.
- The scalability of the moment-SOS hierarchy can be **significantly improved** by exploiting various structures.
- There are **tons of** applications in many different fields!

Thank You!

https://wangjie212.github.io/jiewang