A hierarchy of eigencomputations for polynomial optimization on the sphere

Nathaniel Johnston[1] Benjamin Lovitz[2] Aravindan Vijayaraghavan[3]

1. Mount Allison University
2. NSF Postdoc, Northeastern University
3. Northwestern University

MEGA 2024

August 2, 2024

Notation

- $V_{\mathbb{R}} = \mathbb{R}^n$ real inner product space
- $V_{\mathbb{C}} = \mathbb{C}^n$ complex inner product space
- $S^d(V) := $ homogeneous degree-d polynomials
- $\mathrm{End}\, S^d(V) \coloneqq S^d(V) \otimes S^d(V^*) =$ bi-hom deg-(d, d) polynomials
- $M \in \mathrm{End}\, S^d(V)$ positive semidefinite (psd) if $\langle q, Mq \rangle \geq 0$ for all $q \in S^d(V)$
- Inner product determines (anti-)linear isomorphism $V \simeq V^*$ $\quad v^* = \langle -, v \rangle$
- $\mathrm{Herm}\, S^d(V_{\mathbb{C}}) \subseteq \mathrm{End}\, S^d(V_{\mathbb{C}})$ Hermitian endos: $H(z, w^*) = \overline{H(w, z^*)}$
- $H \in \mathrm{Herm}\, S^d(V_{\mathbb{C}}) \Rightarrow H(z, z^*) = \overline{H(z, z^*)}$ real

Real vs. Hermitian optimization (on the sphere)

Real

Given $p \in S^{2d}(V_{\mathbb{R}})$

compute $p_* := \min_{\|x\|=1} p(x)$

- Hierarchies of lower bounds on p_*:
  - SOS hierarchy [Reznick 95, Lasserre 01]
  - DSOS hierarchy [Ahmadi-Majumdar 19]
  - Harmonic hrchy [Cristancho-Velasco 22]
  - …

Hermitian

Given $H \in \mathrm{Herm}\, S^d(V_{\mathbb{C}})$

compute $H_* := \min_{\|z\|=1} H(z, z^*)$

- Hierarchies of lower bounds on H_*:
  - HSOS hierarchy [Quillen 68, Hudson-Moody 76, Catlin-D'Angelo 96]
  - …?

[JLV]

Main result [JLV]: Using Hermitian optimization for real optimization.

Outline

1. (Real) sum of squares optimization (SOS)
2. Hermitian sum of squares optimization (HSOS)
3. [JLV]: HSOS for SOS (RHSOS)
4. Comparing to other optimizers
5. Constrained optimization:
    I. Optimization over Segre-Veronese
    II. Optimization over an algebraic set

Gram matrix of a real form

$\mathrm{End}\, S^d(V) \coloneqq S^d(V) \otimes S^d(V^*) =$ bi-hom deg-(d, d) polynomials

Def: $M \in \mathrm{End}\, S^d(V)$ positive semidefinite (psd) if $\langle q, Mq \rangle \geq 0$ for all $q \in S^d(V)$

Def: A Gram matrix for $p \in S^{2d}(V_{\mathbb{R}})$ is an $M \in \mathrm{End}\, S^d(V_{\mathbb{R}})$ s.t. $p(x) = M(x, x^*)$

Example: Canonical gram matrix of p / central catalecticant $P \in \mathrm{End}\, S^d(V_{\mathbb{R}})$ is the image of p under the inclusion $S^{2d}(V_{\mathbb{R}}) \hookrightarrow S^d(V_{\mathbb{R}}) \otimes S^d(V_{\mathbb{R}}) \cong \mathrm{End}\, S^d(V_{\mathbb{R}})$.

Example: If $p(x, y) = \sum_i c_i \binom{2d}{i} x^i y^{2d-i}$ then

$$P = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_d \\ c_1 & c_2 & & & \\ c_2 & & & & c_{2d-2} \\ \vdots & & & c_{2d-2} & c_{2d-1} \\ c_d & & c_{2d-2} & c_{2d-1} & c_{2d} \end{bmatrix}$$

Real sums of squares (SOS)

$\mathrm{End}\, S^d(V) \coloneqq S^d(V) \otimes S^d(V^*)$ = bi-hom deg-(d,d) polynomials

Def: $M \in \mathrm{End}\, S^d(V)$ positive semidefinite (psd) if $\langle q, Mq \rangle \geq 0$ for all $q \in S^d(V)$

Def: $M \in \mathrm{End}\, S^d(V_{\mathbb{R}})$ is a Gram matrix for $p \in S^{2d}(V_{\mathbb{R}})$ if $p(x) = M(x, x^*)$

Def: $\Sigma_d^{\mathbb{R}} = \{\sum_i q_i^2\} \subseteq S^{2d}(V_{\mathbb{R}})$ sums-of-squares polynomials

Fact: $\Sigma_d^{\mathbb{R}} = \{ p \in S^{2d}(V_{\mathbb{R}})$ s.t. $\exists$ a psd Gram matrix for M $\}$

So: Checking if $p \in \Sigma_d^{\mathbb{R}} \Leftrightarrow$ Checking if affine sub. of Gram matrices for p intersects the psd cone

(semidefinite program)

Hermitian sums of squares (HSOS)

$\mathrm{End}\, S^d(V) \coloneqq S^d(V) \otimes S^d(V^*)$ = bi-hom deg-(d,d) polynomials

Def: $M \in \mathrm{End}\, S^d(V)$ positive semidefinite (psd) if $\langle q, Mq \rangle > 0$ for all $q \in S^d(V)$.

Fact (polarization): $H \in \mathrm{Herm}\, S^d(V_{\mathbb{C}})$ is uniquely determined by $H(z, z^*)$

Def: $\Sigma_d^{\mathbb{C}} = \{H \;:\; H(z, z^*) = \sum_i |q_i(z)|^2\} \subseteq \mathrm{Herm}\, S^d(V_{\mathbb{C}})$ Hermitian SOS

Fact: $H \in \Sigma_d^{\mathbb{C}} \iff H$ is psd

Takeaway: Checking containment in $\Sigma_d^{\mathbb{C}}$ is easier than checking containment in $\Sigma_d^{\mathbb{R}}$

(checking psd-ness of single matrix) (searching an affine subspace for a psd matrix)

Real vs. Hermitian optimization (on the sphere)

Real

Given $p \in S^{2d}(V_{\mathbb{R}})$

compute $p_* := \min_{\|x\|=1} p(x)$

Theorem (SOS hierarchy) [Reznick, Lasserre]:

$p_* > 0$

$\Rightarrow p_k := p(x)\|x\|^{2(k-d)} \in \Sigma_k^{\mathbb{R}}$ for $k \gg 0$

Convergence [Fang-Fawzi 19]:

$p_k \notin \Sigma_k^{\mathbb{R}} \Rightarrow p_* \lesssim 1/k^2$

Hermitian

Given $H \in \mathrm{Herm}\, S^d(V_{\mathbb{C}})$

compute $H_* := \min_{\|z\|=1} H(z, z^*)$

Theorem (HSOS hierarchy) [Quillen]:

$H_* > 0$

$\Rightarrow H_k := h(z, \overline{z})\|z\|^{2(k-d)} \in \Sigma_k^{\mathbb{C}}$ for $k \gg 0$

Convergence [To-Yeung 06, Christandl-König-Mitchison-Renner 08]:

$H_k \notin \Sigma_k^{\mathbb{C}} \Rightarrow H_* \lesssim 1/k$

"Trade-off:"

SOS hierarchy has fast convergence $1/k^2$, but each level is slow to compute

HSOS hierarchy has slow convergence $1/k$, but each level is fast to compute

Real vs. Hermitian optimization (on the sphere)

Real	Hermitian
Given $p \in S^{2d}(V_{\mathbb{R}})$	Given $H \in \mathrm{Herm}\, S^{d}(V_{\mathbb{C}})$
compute $p_* := \min_{\|x\|=1} p(x)$	compute $H_* := \min_{\|z\|=1} H(z, z^*)$

Def: Canonical gram matrix $P \in \mathrm{End}\, S^d(V_{\mathbb{R}})$ is the image of p under $S^{2d}(V_{\mathbb{R}}) \hookrightarrow \mathrm{End}\, S^d(V_{\mathbb{R}})$

Key idea: Regard $P \in \mathrm{Herm}\, S^d(V_{\mathbb{C}})$ and run HSOS.

Theorem [JLV]: For any $p \in S^{2d}(V_{\mathbb{R}})$, it holds that $P_* \leq p_* \leq \frac{P_*}{c_d}$ where $c_d = \sqrt{\frac{1}{2^d}\sum_{j=0}^{d} \frac{\binom{d}{j}^2}{\binom{2d}{2j}^2}}$

In particular, $p_* > 0 \iff P_* > 0$

Proof: $P_* \leq p_*$: $P_* = \min_{\|z\|=1} P(z, z^*) \leq \min_{\|x\|=1} P(x, x^*) = \min_{\|x\|=1} p(x) = p_*$

$p_* \leq \frac{P_*}{c_d}$: Let $z = x + iy$ be s.t. $P_* = P(z, z^*) = \langle z^d, P z^d \rangle \underset{1}{=} \langle w, p \rangle \underset{2}{\geq} c_d p_*$,

where $w^* = (x^2 + y^2)^{\otimes d}$

2: • $w|_{S^{2d}(V_{\mathbb{R}})} \propto (x^2 + y^2)^d$ has norm at least c_d

• $(x^2 + y^2)^d \in \{\sum_i l_i^{2d}\}$ [Reznick]

1: Tensor gymnastics

Real vs. Hermitian optimization (on the sphere)

Real	Hermitian
Given $p \in S^{2d}(V_{\mathbb{R}})$	Given $H \in \mathrm{Herm}\, S^{d}(V_{\mathbb{C}})$
compute $p_* := \min_{\lVert x\rVert=1} p\,(x)$	compute $H_* := \min_{\lVert z\rVert=1} H\,(z, z^*)$

Def: Canonical gram matrix $P \in \mathrm{End}\, S^{d}(V_{\mathbb{R}})$ is the image of p under $S^{2d}(V_{\mathbb{R}}) \hookrightarrow \mathrm{End}\, S^{d}(V_{\mathbb{R}})$

Key idea: Regard $P \in \mathrm{Herm}\, S^{d}(V_{\mathbb{C}})$ and run HSOS.

Theorem [JLV]: For any $p \in S^{2d}(V_{\mathbb{R}})$, it holds that $P_* \leq p_* \leq \frac{P_*}{c_d}$ where $c_d = \sqrt{\frac{1}{2^d}\sum_{j=0}^{d} \frac{\binom{d}{j}^2}{\binom{2d}{2j}^2}}$

In particular, $p_* > 0 \iff P_* > 0$

Cor (RHSOS hierarchy) [JLV]: $p_* > 0 \Rightarrow P_k := P(z, \bar{z})\lVert z\rVert^{2(k-d)} \in \Sigma_k^{\mathbb{C}}$ for $k \gg 0$

Different view: $P_k \in \mathrm{End}\, S^{d}(V_{\mathbb{R}})$ is a Gram matrix for p_k: $P_k(x, x^*) = p(x)\lVert x\rVert^{2(k-d)} = p_k(x)$

- RHSOS says one can check positivity of the *single* Gram matrix P_k, rather than searching over all Gram matrices as in SOS. (P_k is called $M_k(p)$ in the preprint)

Real vs. Hermitian optimization (on the sphere)

Real	Hermitian
Given $p \in S^{2d}(V_{\mathbb{R}})$	Given $H \in \mathrm{Herm}\ S^d(V_{\mathbb{C}})$
compute $p_* := \min_{\|x\|=1} p(x)$	compute $H_* := \min_{\|z\|=1} H(z, z^*)$

Def: Canonical gram matrix $P \in \mathrm{End}\ S^d(V_{\mathbb{R}})$ is the image of p under $S^{2d}(V_{\mathbb{R}}) \hookrightarrow \mathrm{End}\ S^d(V_{\mathbb{R}})$

Key idea: Regard $P \in \mathrm{Herm}\ S^d(V_{\mathbb{C}})$ and run HSOS.

Theorem [JLV]: For any $p \in S^{2d}(V_{\mathbb{R}})$, it holds that $P_* \leq p_* \leq \frac{P_*}{c_d}$ where $c_d = \sqrt{\frac{1}{2^d}\sum_{j=0}^{d} \frac{\binom{d}{j}^2}{\binom{2d}{2j}^2}}$

In particular, $p_* > 0 \iff P_* > 0$

Cor (RHSOS hierarchy) [JLV]: $p_* > 0 \Rightarrow P_k := P(z, \bar{z})\|z\|^{2(k-d)} \in \Sigma_k^{\mathbb{C}}$ for $k \gg 0$

Convergence [JLV]: $P_k \notin \Sigma_k^{\mathbb{C}} \Rightarrow p_* \lesssim 1/k$

- Proof uses the Theorem and $1/k$ convergence of HSOS [

Comparing to other hierarchies

Hierarchies of lower bounds on p_*:	Convergence	Computation at each level
• SOS hierarchy [Reznick 95, Lasserre 01]	$1/k^2$	Semidefinite program
• DSOS hierarchy [Ahmadi-Majumdar 19]	None	Linear program
• Harmonic hrchy [Cristancho-Velasco 22]	$1/k^2$	Min over $\sim k^n$-vertex polytope
• RHSOS hierarchy [JLV 24]	$1/k$	Min eigenvalue computation

Table: Time for SOS, DSOS and RHSOS hierarchies to compute at level k for a random n-variate quartic polynomial.

Run using MATLAB on a desktop computer with 16GB of RAM.

	SOS	DSOS		Ours		
n	$k-d=0$	$k-d=0$	$k-d=1$	$k-d=0$	$k-d=1$	$k-d=2$
10	0.24 s	0.30 s	0.92 s	0.16 s	0.86 s	6.61 s
15	5.60 s	0.38 s	6.26 s	0.99 s	8.80 s	4.24 min
20	1.37 min	0.74 s	38.0 s	4.42 s	1.47 min	1.90 h
25	17.8 min	15.51 s	6.15 min	12.7 s	10.5 min	27.7 h
30	∞	7.88 s	∞	31.6 s	1.07 h	∞
40	∞	10.7 s	∞	2.56 min	21.1 h	∞
50	∞	26.0 s	∞	9.39 min	∞	∞
60	∞	58.1 s	∞	31.2 min	∞	∞
70	∞	5.71 min	∞	1.55 h	∞	∞
80	∞	∞	∞	4.56 h	∞	∞
90	∞	∞	∞	11.4 h	∞	∞

Constraints: Segre-Veronese optimization

Given $p \in S^{D_1}(V_{1,\mathbb{R}}) \otimes \cdots \otimes S^{D_m}(V_{m,\mathbb{R}})$

compute $p_* := \min_{\|x_i\|=1} p\,(x_1, \ldots, x_m)$

Theorem (m-RHSOS Hierarchy) [JLV]: Relate p_* to a multi-Hermitian optimization problem. Solve by hierarchy of eigencomputations with convergence $1/k$.

Corollary (tensor spectral norm) [JLV]: Given $p \in V_{1,\mathbb{R}} \otimes \cdots \otimes V_{m,\mathbb{R}}$, we compute

$$\|p\|_{\sigma,\mathbb{R}} := \max_{\|x_i\|=1} p\,(x_1, \ldots, x_m)$$

by hierarchy of eigencomputations converging as $1/k$.

General constraints

Real

Given $p \in S^{2d}(V_{\mathbb{R}})$ and $q_1, \dots, q_\ell \in S(V_{\mathbb{R}})$ homogns, compute $p_*^{(q)} := \min_{\substack{\|x\|=1 \\ q_1(x)=\cdots=q_\ell(x)=0}} p(x)$

Hermitian

Given $H \in \mathrm{End}\, S^d(V_{\mathbb{C}})$ Herm and $q_1 \dots q_\ell \in S(V_{\mathbb{C}})$ hom, cmpt $H_*^{(q)} := \min_{\substack{\|z\|=1 \\ q_1(z)=\cdots=q_\ell(z)=0}} H(z, z^*)$

Theorem [Catlin-D'Angelo 99, DJLV 24+]: Let $I = (q_1, \dots, q_\ell) \subseteq S(V_{\mathbb{C}})$ and $I_k^\perp \in S^k(V_{\mathbb{C}}^*)$

Then $H_*^{(q)} > 0 \Rightarrow H_k := h(z, \overline{z})\|z\|^{2(k-d)}$ has a p.d. restriction to $I_k^\perp$ for $k \gg 0$.

We might hope $p_*^{(q)}$ and $P_*^{(q)}$ are related up to $c > 0$, so we can get an RHSOS hierarchy for $p_*^{(q)}$

Unfortunately, this is not always the case:

For $p(x, y, z) = xz$ and $q(x, y, z) = x^2 + y^2 - \frac{1}{\sqrt{3}}xz$ we have $p_*^{(q)} = 0 > \frac{-1}{2\sqrt{3}} \geq P_*^{(q)}$

Conclusion

- [JLV]: Using Hermitian optimization for real optimization on the sphere
  - Comparable performance to other optimizers
  - Generalized to optimization over spherical Segre-Veronese
- More general constraints?
- Motivates development of other Hermitian optimization techniques (beyond HSOS) for use in real optimization.

A hierarchy of eigencomputations for polynomial optimization on the sphere

Nathaniel Johnston[1] Benjamin Lovitz[2] Aravindan Vijayaraghavan[3]

1. Mount Allison University
2. NSF Postdoc, Northeastern University
3. Northwestern University

MEGA 2024

August 2, 2024

Northeastern University