A Symmetric TSSOS Hierarchy for Polynomial Optimization

Igor Klep, Victor Magron,
Tobias Metzlaff, Jie Wang

European Conference on Computational Optimization
Session: Combinatorial Optimization

September/October 2025, Klagenfurt

Let A be a graded real *-algebra.
A **sum of squares** (SOS) is an element of the form

$$q = \sum_{t \in T} q_t \, q_t^*$$

with T a finite index set and $q_t \in A$.

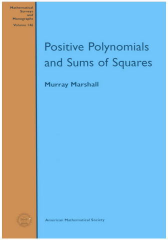
In this talk: Exploit algebraic structures for verification of existence and computation of such representations.

Content

1. Historical Motivation and Applications
2. Symmetry Reduction
3. Sparsity Exploitation

Historical Motivation and Applications

1. Hilbert, 1888:
   "Given $n, r \in \mathbb{N}$, *every* nonnegative polynomial $f \in A := \mathbb{R}[X_1, \dots, X_n]$ of degree $2r$ is a sum of squares."
   $\Leftrightarrow (n, 2r) \in \{(1, 2r), (n, 2), (2, 4)\}$
2. Motzkin, 1967:
   $f = X_1^4 X_2^2 + X_1^2 X_2^4 + 1 - 3 X_1^2 X_2^2$
   is nonnegative but *not* a sum of squares.
3. *Positivstellensatz*:
   A theorem that states the existence of an SOS-representation.
   Schmüdgen, 1991; Putinar, 1996.



Marshall, 2008:
POSITIVE POLYNOMIALS AND SUMS OF SQUARES.
https://bookstore.ams.org/surv-146/

Let $f, g_1, \dots, g_\ell \in \mathbb{R}[X]$, $K := \{X \in \mathbb{R}^n \mid g_1(X), \dots, g_\ell(X) \geq 0\}$.

$$
\begin{aligned}
f^* = \quad \min \quad & f(X) & = \quad \max \quad & \lambda \\
\text{s.t.} \quad & X \in K & \text{s.t.} \quad & \lambda \in \mathbb{R}, \\
& & & f - \lambda \geq 0 \text{ on } K
\end{aligned}
\qquad \text{(POP)}
$$

Truncated Quadratic Module

$$
\mathrm{QM}_r(g) := \{q_0 + \sum_{k=0}^{\ell} q_k\, g_k \mid q_k \text{ is SOS of degree } \leq 2r\}
$$

Lasserre Hierarchy, 2001

$$
\begin{aligned}
f^* \geq f_{\mathrm{sos}}^r := \quad \max \quad & \lambda \\
\text{s.t.} \quad & \lambda \in \mathbb{R}, \\
& f - \lambda \in \mathrm{QM}_r(g)
\end{aligned}
$$

with $f_{\mathrm{sos}}^r \to f^*$ for $r \to \infty$ under certain assumptions (Putinar).

Applications

- Korda, Henrion, Jones, 2013:
  Computing a maximal positive invariant (MPI) set of a dynamical system $\dot{X}(t) = F(X(t))$.
- Ozawa, 2016:
  "A finitely generated group $\mathfrak{G}$ has *Kazhdan's property* (T)"
  $\Leftrightarrow \exists \lambda > 0 : \Delta^2 - \lambda \Delta$ is SOS in $\mathbb{R}[\mathfrak{G}]$ with Laplacian Δ.

To summarize...

- Explicit SOS-certificates give not only an optimal solution, but also an optimizer, in which the solution is attained.
- In practice: converging hierarchy of semidefinite (numerical) lower bounds by restriction of degree.
- Goal: Handle size of computation through exploitation of algebraic structures.

Symmetry Reduction

Symmetry in Nature and Science

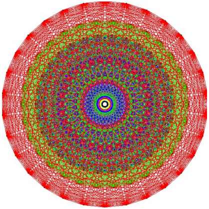
Photo credit: Matteo Fieni

Maryna Viazovska

For the proof that the E_8 lattice provides the densest packing of identical spheres in 8 dimensions, and further contributions to related extremal problems and interpolation problems in Fourier analysis.

citation | video | popular scientific exposition | CV/publications

interview | laudatio | proceedings | Plus magazine! article (intro)



Source: Wikipedia/AMS

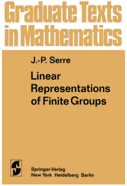
Gosset polytope drawn BY HAND (!) by Peter McMullen, 1960s

Let $\mathcal{G}$ be a finite group.

1. Two elements $\sigma, \tilde{\sigma} \in \mathcal{G}$ are called **conjugate**, if $\sigma\tau = \tau\tilde{\sigma}$ for some $\tau \in \mathcal{G}$.
2. A **$\mathcal{G}$-module** W is a vector space together with a group homomorphism $\rho_W : \mathcal{G} \to \mathrm{GL}(W)$, called **representation**.
3. A $\mathcal{G}$-module W is called **irreducible**, if its only $\mathcal{G}$-submodules are 0 and W itself.

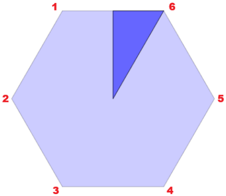
Fact

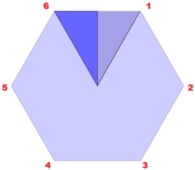
$\#$ nonisom. irred. $\mathcal{G}$-modules $= \#$ conjugacy classes $=: h$

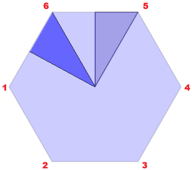


Serre, 1977:
LINEAR REPRESENTATIONS OF FINITE GROUPS.
https://link.springer.com/book/10.1007/978-1-4684-9458-7

Decomposition into irreducibles





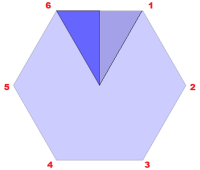


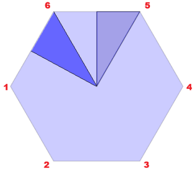
$s = (1,6)(2,5)(3,4)$ $\qquad$ $r = (1,2,3,4,5,6)$

$\mathfrak{D}_{2\cdot 6} = \langle s, r \mid s^2 = r^6 = (sr)^2 = e \rangle$ "dihedral group of order 12"

```
               _
   _       _ _(_)_     |  Documentation: https://docs.julialang.org
  (_)     | (_) (_)    |
   _ _   _| |_  __ _   |  Type "?" for help, "]?" for Pkg help.
  | | | | | | |/ _` |  |
  | | |_| | | | (_| |  |  Version 1.11.5 (2025-04-14)
 _/ |\__'_|_|_|\__'_|  |  Official https://julialang.org/ release
|__/                   |

julia> using Oscar
  ___   ____   ____    _    ____
 / _ \ / ___| / ___|  / \  |  _ \   |  Combining ANTIC, GAP, Polymake, Singular
| | | |\___ \| |     / _ \ | |_) |  |  Type "?Oscar" for more information
| |_| | ___) | |___ / ___ \|  _ <   |  Manual: https://docs.oscar-system.org
 \___/ |____/ \____/_/   \_\_| \_\  |  Version 1.3.0

julia> n=6
6

julia> G=dihedral_group(2*n)
Pc group of order 12

julia> character_table(G)
Character table of pc group of order 12

  2  2  2  1  1  2  2
  3  1  .  1  1  .  1

    1a 2a 6a 3a 2b 2c
 2P 1a 1a 3a 3a 1a 1a
 3P 1a 2a 2c 1a 2b 2c
 5P 1a 2a 6a 3a 2b 2c

X_1  1  1  1  1  1  1
X_2  1 -1 -1  1  1 -1
X_3  1 -1  1  1 -1  1
X_4  1  1 -1  1 -1 -1
X_5  2  .  1 -1  . -2
X_6  2  . -1 -1  .  2
```

Decomposition into irreducibles





$s = (1,6)(2,5)(3,4) \qquad r = (1,2,3,4,5,6)$

Consider the 6-dimensional representation

$$\rho : \mathfrak{D}_{2\cdot 6} \to \mathrm{GL}(\mathbb{R}^6),\ s \mapsto \begin{pmatrix} & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \\ 1 & & & & & \end{pmatrix},\ r \mapsto \begin{pmatrix} & & & & & 1 \\ & & & & 1 & \\ & & & 1 & & \\ & & 1 & & & \\ & 1 & & & & \\ 1 & & & & & \end{pmatrix}.$$

The $\mathfrak{D}_{2\cdot 6}$-module $\mathbb{R}^6$ can be decomposed into irreducibles

$$\mathbb{R}^6 = \left\langle \begin{pmatrix}1\\1\\1\\1\\1\\1\end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix}1\\-1\\1\\-1\\1\\-1\end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix}2\\1\\-1\\-2\\-1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\\-1\\-1\end{pmatrix} \right\rangle \oplus \left\langle \begin{pmatrix}2\\-1\\-1\\2\\-1\\-1\end{pmatrix}, \begin{pmatrix}0\\1\\-1\\0\\1\\-1\end{pmatrix} \right\rangle.$$

The induced action on $\mathbb{R}[X_1, \ldots, X_6]$ is $f^\sigma(X) := f(\rho(\sigma^{-1}) \cdot X)$.

Setup

Let $A = A_0 \oplus A_1 \oplus A_2 \oplus \dots$ with A_r *finite dimensional* and $\mathcal{G} \subseteq \mathrm{GL}(A)$ be a *finite* group with a linear action

$$\mathcal{G} \times A_r \to A_r, (\sigma, f) \mapsto f^\sigma.$$

Isotypic Decomposition

$$A_r \otimes_{\mathbb{R}} \mathbb{C} = \bigoplus_{i=1}^{h} \bigoplus_{j=1}^{m_r^{(i)}} V_j^{(i)}$$

h number of irreducible characters of $\mathcal{G}$ with multiplicities $m_r^{(i)}$ and $V_1^{(i)}, \dots, V_{m_r^{(i)}}^{(i)}$ pairwise isomorphic irreducible $\mathcal{G}$-modules.

Reynolds Operator

$$\mathcal{R}^{\mathcal{G}}(f) := \frac{1}{|\mathcal{G}|} \sum_{\sigma \in \mathcal{G}} f^\sigma$$

Symmetric SOS

Observation (for fixed degree r)

If S is a basis for A_r and $\mathbf{Q} = (\mathbf{Q}^*)^t \succeq 0$ is a Hermitian psd matrix of size $|S|$ with entries in A_0, then

$$f = (\mathbf{S}^*)^t \cdot \mathbf{Q} \cdot \mathbf{S} \in A_{2r}$$

is a sum of squares, where $\mathbf{S}$ is the vector of basis elements.

Proposition

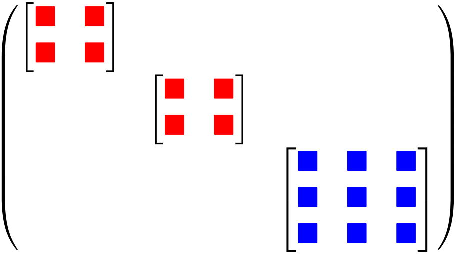
Let $f \in A^{\mathcal{G}} \cap A_{2r}$ and $S^{(i)} \subset A_r$ contain exactly one nonzero element of each $V_j^{(i)}$, with $1 \leq j \leq m_r^{(i)}$. If f is a sum of squares in A, then there exist $\mathbf{Q}^{(i)} = ((\mathbf{Q}^{(i)})^*)^t \succeq 0$ of size $m_r^{(i)}$, such that

$$f = \sum_{i=1}^{h} \mathcal{R}^{\mathcal{G}} \left(((\mathbf{S}^{(i)})^*)^t \cdot \mathbf{Q}_r^{(i)} \cdot \mathbf{S}^{(i)} \right).$$

Informal Consequence

"The matrix of a symmetric SOS-certificate over A_r has h blocks, each consisting of d_i many identical blocks of size $m_r^{(i)}$."

$$\dim(A_r) = \sum_{i=1}^{h} d_i \, m_r^{(i)} \qquad d_i := \dim(V_1^{(i)}) = \ldots = \dim(V_{m_r^{(i)}}^{(i)})$$



Remark

Change of basis does not effect the trace. Hence, dense and symmetric relaxations (SDP) have the same theoretical value.

Sparsity Exploitation

What is sparsity and where does it appear?

- Correlative Sparsity:
  $f = X_1 X_2 + X_2 X_3 + \ldots + X_{99} X_{100}$
- Term Sparsity:
  $f = X_1 X_2^{99} + X_1^{99} X_2$

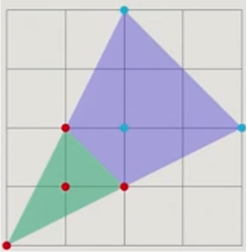
- Deep learning (robustness, computer vision)
- Power systems (optimal power flow, stability)
- Quantum systems (condensed matter)

Motzkin is not SOS

Recall: $f_{\mathrm{Motzkin}} = 1 \underline{-3\, X_1^2\, X_2^2} + X_1^4\, X_2^2 + X_1^2\, X_2^4$
is nonnegative but *not* a sum of squares (Motzkin, 1967).

Reznick, 1978

If $f = \sum_t q_t^2$, then $\mathrm{NewtonPoly}(q_t) \subseteq \frac{1}{2}\mathrm{NewtonPoly}(f)$.



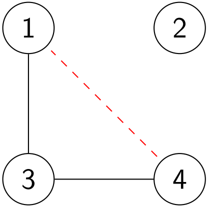
Hence, if f_{Motzkin} was SOS, then

$$f_{\mathrm{Motzkin}} = \sum (a\, 1 + \underline{b\, X_1\, X_2} + c\, X_1^2\, X_2 + d\, X_1\, X_2^2)^2$$

and thus $-3 = \sum b^2$:(

Encoding and Exploiting Sparsity

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \text{ with } \overline{\mathbf{B}} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$



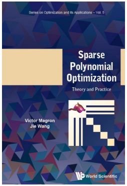
Idea: Instead of sums of squares of the form

$$f = (\mathbf{S})^t \cdot \mathbf{Q} \cdot (\mathbf{S})^*,$$

consider

$$\tilde{f} = (\mathbf{S})^t \cdot (\mathbf{B} \circ \mathbf{Q}) \cdot (\mathbf{S})^*,$$

where **S** is a vector of basis elements and **B** a binary matrix.



Magron & Wang, 2023:
SPARSE POLYNOMIAL OPTIMIZATION.
https://www.worldscientific.com/worldscibooks/10.1142/q0382

Symmetry-adapted TSSOS Hierarchy

$$f^* = \max \lambda \quad \text{s.t.} \quad \lambda \in \mathbb{R}, \quad f - \lambda \geq 0 \text{ on } K \qquad \geq f_{\mathrm{sos}}^{r,s} := \max \; f_1 - \sum_{k,i} \mathrm{tr}(\mathbf{A}_{r,s,k,1}^{(i)} \cdot \mathbf{Q}_k^{(i)}) \quad \text{s.t.} \quad \mathbf{Q}_k^{(i)} \in \mathrm{Sym}_{r-d_k}^{(i)}(\mathbf{B}_{r,s,k}^{(i)}), \quad \mathbf{Q}_k^{(i)} \succeq 0, \forall j \geq 2 : \quad f_j = \sum_{k,i} \mathrm{tr}(\mathbf{A}_{r,s,k,j}^{(i)} \cdot \mathbf{Q}_k^{(i)}),$$

r: degree of approximation

s: level of sparsity

$\mathbf{A}_{r,s,k,j}^{(i)}$: sparse coefficient matrices in the symmetry basis

$\mathbf{B}_{r,s,k}^{(i)}$: binary matrices encoding sparsity

Theorem

For fixed degree $r \geq r_{\min}$, the sequence $(f_{\mathrm{sos}}^{r,s})_{s\geq 1}$ is monotonously nondecreasing with $f_{\mathrm{sos}}^{r,*} = f_{\mathrm{sos}}^{r}$.
For fixed sparsity order $s \geq 1$, the sequence $(f_{\mathrm{sos}}^{r,s})_{r\geq r_{\min}}$ is monotonously nondecreasing.

Outlook: Correlative sparsity, Complex Noncommutative Variables

Thank You.

https://github.com/wangjie212/TSSOS