

Symmetry Reduction for Sum-of-Squares Optimization

Benoît Legat Marek Kaluba Tillmann Weisser

January, 2022

INFORMS 2022

Nonnegative polynomial into sum of squares (SOS)

$$(x^3, x^2, x) \quad \text{not unique}$$
$$p(x) = X^T Q X$$

$$x^6 - 2x^4 + 2x^2 \quad (1) = X^T \begin{pmatrix} 1 & 0 & -1 \\ . & 0 & 0 \\ . & . & 2 \end{pmatrix} X$$

$$p(x) \geq 0 \quad \forall x \iff Q \geq 0 \quad \downarrow \text{cholesky}$$

$$x^2 + (-x^3 + x)^2 \longleftarrow X^T \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}^T \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} X$$

¹ Gatermann, Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, 2004: Example 5.1

Semidefinite program

$$\begin{aligned} \min_{X \in \mathcal{S}^N} \quad & \langle C, Q \rangle \\ \text{subject to:} \quad & \langle A_i, Q \rangle = b_i, \quad i = 1, 2, \dots, m \\ & Q \geq 0. \end{aligned}$$

Is $p(x) = x^6 - 2x^4 + 2x^2$ SOS² ? $p(x) = X^\top Q X = \langle X X^\top, Q \rangle$:

$$X = \begin{bmatrix} x^3 \\ x^2 \\ x \end{bmatrix}, X X^\top = \begin{bmatrix} x^6 & x^5 & x^4 \\ \cdot & x^4 & x^3 \\ \cdot & \cdot & x^2 \end{bmatrix}, A_{x^6} = \begin{bmatrix} 1 & 0 & 0 \\ \cdot & 0 & 0 \\ \cdot & \cdot & 0 \end{bmatrix}, A_{x^5} = \begin{bmatrix} 0 & 1 & 0 \\ \cdot & 0 & 0 \\ \cdot & \cdot & 0 \end{bmatrix},$$

$$A_{x^4} = \begin{bmatrix} 0 & 0 & 1 \\ \cdot & 1 & 0 \\ \cdot & \cdot & 0 \end{bmatrix}, A_{x^3} = \begin{bmatrix} 0 & 0 & 0 \\ \cdot & 0 & 1 \\ \cdot & \cdot & 0 \end{bmatrix}, A_{x^2} = \begin{bmatrix} 0 & 0 & 0 \\ \cdot & 0 & 0 \\ \cdot & \cdot & 1 \end{bmatrix}.$$

²Gatermann, Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, 2004: Example 5.1

For n -variate polynomial of degree $2d$: $N = \binom{n+d}{n}$, $m = \binom{n+2d}{n}$.

Complexity of second order interior point solvers³

Time: $\mathcal{O}((mN + m^2)N^{5/2} \log(\varepsilon))$. Space: $\mathcal{O}(m^2)$.

E.g. CSDP, DSDP, Hypatia, MOSEK, SDPA, SDPT3, SeDuMi.

Alternatives:

- Newton conjugate gradient (Newton-CG), e.g., SDPNAL.
- ADMM, e.g., CDCS, COSMO, SCS.
- Barvinok–Pataki bound: $\exists Q$ of rank $r \leq \sqrt{2m}$.
SOSP: $\exists Q$ of rank $r \leq \text{pythagoras number}$.
 - Burer–Monteiro method: $Q = UU^T$, $U \in \mathbb{R}^{N \times r}$, e.g. SDPLR.
 - Primal-dual hybrid gradient (PDHG), e.g., ProxSDP.

³Monteiro, Todd. "Path-following methods." *Handbook of semidefinite programming*. 2000. 267–306.

Schur lemma

$(\rho_i)_i \subseteq \mathbb{R}^{d \times d}$ *reducible* if $\exists \{0\} \subset \mathcal{V} \subseteq \mathbb{R}^d$ s.t. $\forall i, \rho_i \mathcal{V} \subseteq \mathcal{V}$.

$(\rho_i^{(1)})_i, (\rho_i^{(2)})_i \subseteq \mathbb{R}^{d \times d}$ *equivalent* if $\exists T$ s.t. $\forall i, \rho_i^{(1)} = T \rho_i^{(2)} T^{-1}$.

If $(\rho_i^{(1)})_i, (\rho_i^{(2)})_i$ are irreducible and inequivalent, then

$$\forall i, Q \begin{bmatrix} \rho_i^{(1)} & 0 \\ 0 & \rho_i^{(2)} \end{bmatrix} = \begin{bmatrix} \rho_i^{(1)} & 0 \\ 0 & \rho_i^{(2)} \end{bmatrix} Q \Rightarrow \exists c_1, c_2 \text{ s.t. } Q = \begin{bmatrix} c_1 I & 0 \\ 0 & c_2 I \end{bmatrix}$$

If $(\rho_i)_i$ is irreducible, then (permuting rows/columns give $I_d \otimes C$)

$$Q \begin{bmatrix} \rho_i & 0 \\ 0 & \rho_i \end{bmatrix} = \begin{bmatrix} \rho_i & 0 \\ 0 & \rho_i \end{bmatrix} Q \Rightarrow \exists C \in \mathcal{S}^2 \text{ s.t. } Q = C \otimes I_d = \begin{bmatrix} c_{1,1} I & c_{1,2} I \\ c_{1,2} I & c_{2,2} I \end{bmatrix}$$

$$\forall i, Q(\rho_i \otimes I_m) = (\rho_i \otimes I_m)Q \Rightarrow Q = C \otimes I_d, C \in \mathcal{S}^m \quad \text{Note } m \leftrightarrow d$$

Sagan, *The symmetric group*, Springer Science & Business Media, 2001: Definition 1.4.5, Definition 1.6.2,

Example 1.7.2, Example 1.7.3 (adapted from $\mathbb{C}^{d \times d}$ to symmetric matrices $\mathcal{S}^n \subseteq \mathbb{R}^{d \times d}$)

Invariance and convexity

Theorem⁴

Objective and feasible sets are **invariant** and **convex** $\rightarrow$ can restruct to invariant solutions.

Invariant SDP: Q feasible $\Rightarrow \rho_i^\top Q \rho_i$ feasible w/ same cost:

$$\forall i, j, \rho_i A_j \rho_i^\top = A_j \text{ and } \rho_i C \rho_i^\top = C.$$

Invariant SOSP: $p(g_i(x)) = p(x)$, let ρ_i s.t. $X(g_i(x)) = \rho_i X(x)$.

Invariant solution: $Q = \rho_i^\top Q \rho_i$. If $\rho_i^\top = \rho_i^{-1}$, $\rho_i Q = Q \rho_i$.

How to **simultaneously** block-diagonalize $(\rho_i)_i$?

⁴Gatermann, Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, 2004: Theorem 3.3

Group Theory to the rescue!

Definition

A **group** $(G, *)$ is a set closed under the multiplication $*$, together with the existence of unique inverse elements.

Example: The set of all bijections of the set $\{1, \dots, n\}$ with composition is a group (known as the full symmetric group on n letters S_n).

$S_2 = \{(1)(2), (1, 2)\}$ **acts** on polynomial ring $\mathbb{R}[x]$ by

$$((1)(2), x) \mapsto x$$

$$((1, 2), x) \mapsto -x.$$

Observe: $f(x) = x^6 - 2x^4 + 2x^2 = x^2 + (x - x^3)^2$ is invariant, but $x - x^3$ isn't!

Isotypical projections

“Types” of actions $\leftrightarrow$ blocks in the block-diagonalization:

$$V \cong V_1 \oplus \cdots \oplus V_k.$$

Each block V_i decomposes further as $V = \bigoplus_i \oplus_{m_i} V'_i$. ⁽⁵⁾

Universal (baseless) **projections**⁶ in group ring $\mathbb{R}[G] \leftrightarrow$ projections onto blocks (for every representation V !).

Example

Universal **projections** for S_2 :

$$\pi_1 = \frac{1}{2}((1)(2) + (1, 2)) \rightarrow \text{invariants}$$

$$\pi_2 = \frac{1}{2}((1)(2) - (1, 2)) \rightarrow \text{flipped}$$

⁵ m_i is the multiplicity of the action type in V

⁶related to irreducible characters

Permutation symmetry example

```
using PermutationGroups
G = PermGroup([perm"(1,2,3,4)"])
# permutation group generated by one element (cyclic of order 4)
using DynamicPolynomials
@polyvar x[1:4]
poly = sum(x) + sum(x.^2)
basis = monomials(x, 0:(maxdegree(f)÷2))
using SymbolicWedderburn
symmetry_adapted_basis(G, basis, Symmetry.VariablePermutation())
# G acts by permuting variables
```

Isotypical blocks when acting on $[1, x_1, x_2, x_3, x_4, x_5]$:

$$B_1 = \begin{bmatrix} 1 \\ x_1 + x_2 + x_3 + x_4 \end{bmatrix} \quad B_2 = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_4 \end{bmatrix} \quad B_3 = [x_1 - x_2 + x_3 - x_4]$$

$$V \cong (V'_1 \oplus V'_1) \oplus V'_2 \oplus V'_3, \quad d_2 = m_1 = 2, \quad d_1 = d_3 = m_2 = m_3 = 1.$$

Permutation symmetry example

```
using SumOfSquares
import ECOS # Note: ECOS doesn't support PSD cone,
# but 2x2 psd constraints translate to SOC/quadratic constraints
model = Model(ECOS.Optimizer)
@variable(model, t)
@objective(model, Max, t)
pattern = Symmetry.Pattern(G, Symmetry.VariablePermutation())
@constraint(model, poly - t in SOScone(), symmetry = pattern)
optimize!(model)
```

Naive: 15 vars 15 constrs, Q is 5×5 ; Symmetry: 5 vars 5 constrs,

$$Q = Q_1 \oplus (Q_2 \oplus Q_2) \oplus Q_3 \quad Q_1 \text{ is } 2 \times 2, Q_2 \text{ is } 1 \times 1, Q_3 \text{ is } 1 \times 1$$

Note inversion: multiplicities $\leftrightarrow$ dimensions

$$V \cong (V'_1 \oplus V'_1) \oplus V'_2 \oplus V'_3, \quad d_2 = m_1 = 2, \quad d_1 = d_3 = m_2 = m_3 = 1.$$

Even symmetry example⁷

```
struct OnSign <: Symmetry.OnMonomials end
# <: SymbolicWedderburn.Action, corresponds to  $x \mapsto -x$ 
SymbolicWedderburn.action(::OnSign, p::Permutation,
    mono::AbstractMonomial) = sign(p)^degree(mono)*mono
G = PermGroup([perm"(1,2)"])
@polyvar x
f = x^6 - 2x^4 + 2x^2
basis = monomials(x, 0:(maxdegree(f)÷2))
symmetry_adapted_basis(G, basis, OnSign())
```

Isotypical blocks when acting on $[1, x, x^2, x^3]$:

$$B_1 = \begin{bmatrix} 1 \\ x^2 \end{bmatrix} \quad B_2 = \begin{bmatrix} x \\ x^3 \end{bmatrix} \quad V = (V'_1 \oplus V'_1) \oplus (V'_2 \oplus V'_2)$$

Naive formulation 4×4 -psd; with symmetry: $2 \times 2, 2 \times 2$ -psd.

⁷ Gatermann, Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, 2004: Example 5.1

Dihedral symmetry example⁸

```
@polyvar x y
f = x^6 + y^6 - x^4 * y^2 - y^4 * x^2 - x^4 - y^4 -
    x^2 - y^2 + 3x^2 * y^2 + 1
struct DihedralAction <:
    SymbolicWedderburn.ByLinearTransformation end
```

The action is a mixture of changing sign and permuting variables.

$$\begin{aligned} B_1 &= [x^2 + y^2, 1], & m_1 &= 2, d_1 = 1, \\ B_2 &= [x^3, x^2y, xy^2, y^3, x, y], & m_2 &= 3, d_2 = 2, \\ B_3 &= [xy], & m_3 &= d_3 = 1, \\ B_4 &= [x^2 - y^2], & m_4 &= d_4 = 1, \end{aligned}$$

10 × 10-psd into 2 × 2, 3 × 3, 1 × 1, 1 × 1-psd.

Non-commutative groups allow $d_i > 1 \rightarrow$ better for the symmetrization!

⁸Gatermann, Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, 2004: Example 5.4

Implementing group action

- implement methods from GroupsCore.jl API for DihedralGroup

```
import GroupsCore
struct DihedralGroup <: GroupsCore.Group ... end
struct DihedralElement <: GroupsCore.GroupElement
    ...
end
```

- implement action methods from SymbolicWedderburn for DihedralAction

```
import SymbolicWedderburn
SymbolicWedderburn.action(
    ::DihedralAction,
    ::DihedralElement,
    ::AbstractMonomial) = ...
```

Behind the scenes: SymbolicWedderburn.jl

- compute **character table** of the group $\longrightarrow$ universal projections
- given action e.g. on monomials, **induce** the action to the whole basis
- evaluate the projections in the action $\longrightarrow$ **block diagonalizing** change of basis
- further **decompose** each block into **irreducible** blocks⁽⁹⁾.

⁹for now numerically only

Large scale example

- Optimization problem from geometric group theory¹⁰
- sdp-constraint of size 4641×4641 , $1.1 \cdot 10^7$ constraints
- symmetry group: $S_2 \wr S_5$ (3840 elements)
- After symmetrization:
 - 29-blocks (largest: 58×58), 13232 variables in total
 - 7230 constraints
- Solvable!

¹⁰Kaluba, M., Nowak, P.W. & Ozawa, N. $\text{Aut}(F_5)$ has property (T). *Math. Ann.* **375**, 1169–1191 (2019). <https://doi.org/10.1007/s00208-019-01874-9>



