Symmetry groups, semidefinite programs, and sums of squares

Breki Pálsson and Simon Wegan

20. April 2023

1 Introduction

We will look at a fundamental problem in real algebraic geometry i.e. the existence and computation of a representation of a multivariate polynomial as a sum of squares (SOS). In other words, the question of finding $p_i \in \mathbb{R}[x], i = 1, \cdots, N$ such that

$$f(x) = \sum_{i=1}^{N} (p_i(x))^2.$$

This problem has applications in many fields of applied mathematics, such as continuous and combinatorial optimization as well as being theoretically interesting.

We will show a method that exploits symmetries in polynomials and semidefinite programming (SDP) in order to get a reduction in the problem size. Two reasons are for this, firstly we get a faster solution since the time complexity of SDF has been shown to be polynomial do we have to cite this?, secondly smaller problem size give more accurate solutions may that be do to numerical conditioning or numerical errors.

The paper outlines the theoretical background and gives the reader examples explaining the definitions step by step in order to present an algorithm that is able to use the symmetric properties of a polynomial that is invariant with respect to a certain representation and produces a solution to a semidefinite program given certain constraints.

The problem

Given a polynomial that has symmetries we want to use symmetry reduction techniques to reduce our problem to a smaller semidefinite program.

We will use representation theory to define a good notion of symmetries. In particular we will look at representations $\sigma : G \to \mathrm{Aut}(S^N)$ that preserve $\mathcal{S}_+^N$ and are *induced* by a representation $\rho : G \to \mathrm{GL}(\mathbb{R}^N)$, that is

- $\sigma(g)(\mathcal{S}_+^N) \subseteq \mathcal{S}_+^N$ for all $g \in G$, and
- $\sigma(g)(X) := \rho(g)^T X \rho(g)$ for all $X \in \mathcal{S}, g \in G$.

These representation will be of this kind for most practical instances. The paper shows that with a convenient change of coordinates every invariant matrix will be block diagonal enabling us to go from one big problem to a couple of smaller ones.

Definition 1.1. Given a finite group G, and an associated linear representation $\sigma : G \to \mathrm{Aut}(\mathcal{S}^N)$, a semidefinite optimization problem of the form $F^* := \min_{\mathcal{L} \cap \mathcal{S}_+^N} \langle C, X \rangle$. is called invariant with respect to σ, if the set of feasible matrices $\mathcal{L} \cap \mathcal{S}_+^N$ and the cost function $F(X) = \langle C, X \rangle$ are invariant with respect to σ

Definition 1.2. We define the fixed-point subspace of $\mathcal{S}^N$ as the subspace of all invariant matrices,

$$\mathcal{F} := \{X \in \mathcal{S}^N | X = \sigma(g)(X) \ \forall g \in G\}$$

and the associated semidefinite program

$$F_\sigma := \min_{X \in \mathcal{F} \cup \mathcal{L} \cap \mathcal{S}_+^N} \langle C, X \rangle$$

The paper proves that the solution a SDP which is invariant with respect to a linear representation σ of the kind we have discussed and it's sigma SDP outlined above give the same solution. Thus we can restrict our set of feasible matrices from $\mathcal{L} \cap \mathcal{S}_+^N$ to $X \in \mathcal{F} \cup \mathcal{L} \cap \mathcal{S}_+^N$ making the problem simpler.

Furthermore every linear representation of a finite group G has a canonical decomposition as a direct sum of irreducible representations, i.e.

$$\rho = m_1\vartheta_i \oplus m_2\vartheta_2 \oplus \cdots \oplus m_h\vartheta_h.$$

The paper showed that with these constraints we can actually find a change of coordinates so that all the matrices in the SDP have block diagonal form, the problem therefore collapses into a collection of smaller optimization problems, which are much easier to solve as outlined below

$$F = \min_{X \in \mathcal{L},\ X_i \in \mathcal{S}_+^{m_i}} \sum_{i=1}^{h} n_i \langle C_i, X_i \rangle .$$

Next we look at some invariant theory in order to simplify the problem even further.

Assume we are interested in finding the sum of squares decomposition of a polynomial $f(\mathbf{x})$ of degree $2d$ in n variables which is invariant with respect to a linear representation $\vartheta : G \to \mathrm{GL}(\mathbb{R}^n)$, i.e. $f(\mathbf{x}) = f(\vartheta(g)\mathbf{x})$ for all $g \in G$. The set of all such invariant polynomials is the invariant ring, denoted by $\mathbb{R}[\mathbf{x}]^G$. We could jump to the conclusion that if a $f \in \mathbb{R}[\mathbf{x}]^G$ can be written as a sum of squares $f(\mathbf{x}) = \sum_{i=1}^h (p_i(\mathbf{x}))^2$ then the polynomials p_i must also be in $\mathbb{R}[\mathbf{x}]^G$. Unfortunately this turns out not to be true. We can however use a technique that allows us to couple together squares so they decompose into invariant components. This means we can write

$$f(\mathbf{x}) = \sum_{i=1}^{h} \langle Q_i, P_i(\mathbf{x}) \rangle ,$$

with $P_i \in (\mathbb{R}[\mathbf{x}]^G)^{m_i \times m_i}$ where the Q_i are as always positive semidefinite.

Furthermore using the Hironaka decomposition of the invariant ring, i.e.

$$\mathbb{R}[\mathbf{x}]^G = \bigoplus_{j=1}^{t} \eta_j(\mathbf{x})\mathbb{R}[\theta_1(\mathbf{x}), \cdots, \theta_n(\mathbf{x})]$$

where $\theta_i(\mathbf{x})$, $\eta_i(\mathbf{x})$ are called primary and secondary invariants, respectively, we can actually represent $f(\mathbf{x})$ uniquely as a function with the invariants as inputs. i.e.

$$\tilde{f}(\theta, \eta) = \sum_{i=1}^{h} \left\langle Q_i, \tilde{P}_i(\theta, \eta) \right\rangle ,$$

with $\tilde{P}_i \in (\tilde{T})^{m_i \times m_i}$, where $\tilde{T} = \sum_{i=1}^h \eta_j \mathbb{R}[\theta]$ and the Q_i are as always positive semidefinite.
In practice very often the group representations are so-called reflection groups. In this situation the invariant ring is isomorphic to a polynomial ring and the secondary invariants are not needed.

By a change of coordinates we can assume that $\tilde{P}_i$ has a special form where each $r_i \times r_i$ subblock is a monomial multiple of the first *principal* $r_i \times r_i$ subblock Π_i. Then $\tilde{f}$ can be written as

$$\tilde{f}(\theta, \eta) = \sum_{i=1}^{h} \langle S_i(\theta), \Pi_i(\theta, \eta) \rangle ,$$

where the S_i are SOS matrices. Note that by our construction above, the Π_i do not depend on the choice of the polynomial f.

2 Algorithm

Here we will present an algorithm that is the result of the paper and later explain certain concepts that we need to define in order to understand the algorithm

Algorithm I

Input: Linear representation ϑ of a finite group G on $\mathbb{R}^n$.

1. Determine all real irreducible representations of G.
2. Compute primary and secondary invariants θ_i, η_j.
3. For each non-trivial irreducible representation compute the basis $b_1^i, \cdots, b_{r_i}^i$ of the module of equivariants.
4. For each irreducible representation i compute the corresponding matrix Π_i.

Output: Primary and secondary invariants θ, η and the matrices Π_i.

Algorithm II

Input: Primary and secondary invariants θ, η, matrices Π_i and $f \in \mathbb{R}[\theta]^G$.

1. Rewrite f in fundamental invariants giving $\tilde{f}(\theta, \eta)$.
2. For each irreducible representation determine $w_i(\theta)$ and thus the structure of the matrices $S_i \in \mathbb{R}[\theta]$.
3. Find a feasible solution of the semidefinite program corresponding to the constraints.

Output: SOS matrices S_i providing a generalized sum of squares decomposition of $\tilde{f}$.

Algorithm I does the preprocessing for our problem while the second one, only having a linear representation as an input this means that we can run the first algorithm once and then use the outputs for finding solution for all polynomials that are invariant with respect to this particular representation.

3 Example

We will demonstrate the efficacy of this algorithm with an example.

4 Conclusion

Although it might seem cumbersome to find all the invariants of a representation....

$$\min_{X \in \mathcal{L} \cap \mathcal{S}_+^N} \langle C, X \rangle \to \min_{X \in \mathcal{F} \cap \mathcal{L} \cap \mathcal{S}_+^N} \langle C, X \rangle \to \min_{X \in \mathcal{L}, X_i \in \mathcal{S}_+^{m_i}} \sum_{i=1}^{h} n_i \langle C_i, X_i \rangle$$

$$f(\mathbf{x}) = \sum_{i=1}^{r} (p_i(\mathbf{x}))^2 = \sum_{i=1}^{h} \langle Q_i, P_i(\mathbf{x}) \rangle = \sum_{i=1}^{h} \left\langle Q_i, \tilde{P}_i(\mathbf{x}) \right\rangle = \sum_{i=1}^{h} \langle S_i(\theta), \Pi_i(\theta, \eta) \rangle \to f_j(\theta) = \sum_{i=1}^{h} \left\langle S_i(\theta), \Pi_i^j(\theta) \right\rangle$$

Definition 4.1. Let $S \in \mathbb{R}[\mathbf{x}]^{m \times m}$ be a symmetric matrix, and $\mathbf{y} = [y_1, \cdots, y_m]$ be new indeterminants. The matrix S is a sum of squares (SOS) matrix if the scalar polynomial $\mathbf{y}^T S \mathbf{y}$ is a sum of squares in $\mathbb{R}[\mathbf{x}, \mathbf{y}]$.

Definition 4.2. A polynomial maping $f : \mathbb{R}^n \to \mathbb{R}^{n_i}$ with $f_i \in \mathbb{R}[\mathbf{x}]$ and

$$f(\vartheta(g)\mathbf{x}) = \vartheta_i f(\mathbf{x}) \quad \forall g \in G,$$

is called ϑ-ϑ_i-equivariant.

Theorem 4.1. *For each irreducible representation ϑ_i the module of ϑ-ϑ_i-equivariants is a free module over the ring in the primary invariants:*

$$M_i = \mathbb{R}[\theta]\dot{\{}b_1^i, \cdots, b_{r_i}^i\},$$

where $b_\nu^i \in \mathbb{R}[\mathbf{x}]^{n_i}$ denote the elements in the module basis and r_i is the rank of the module.

[Gatermann_2004]
Hello World