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 semidefinite programming in order to find

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

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

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 $\sigma : G \to Aut(\mathcal{S}^N)$, 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 $\langle C, X\rangle$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 The paper showed that with these constraints we can actually change 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 $\vartheta : G \to GL(\mathbb{R}^n)$,

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, μ_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 $\prod_i$.

Output: Primary and secondary invariants θ, μ and the matrices $\prod_i$.

Algorithm II

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

1. Rewrite f in fundamental invariants giving $\tilde{f}(\theta, \mu)$.
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 lgor, only having a linear representation as an input this means that we can run algorithm I one time and then

une 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....

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

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 $\mathbb{R}[\mathbf{x}, \mathbf{y}]$.

[Gatermann_2004]
Hello World