Exactness of SOS relaxations in copositive programming

Research Project Report - "Modelling seminar project" course

Mohamed Amine Abdellaziz
mohamed-amine.abdellaziz@grenoble-inp.org
Suraj Ghimire
suraj.ghimire@grenoble-inp.org
Anna Laskina
laskina.as@phystech.edu
Sofiane Tanji
sofiane.tanji@grenoble-inp.org

Abstract

For our research project, we studied the Parrilo relaxations for certifying copositivy of a given matrix of size 6×6. In this small project report, we first present quickly the framework of Parrilo relaxations with its hypotheses and the usual notations, along with a sum-up of the sums of squares literature. Our main contribution is to certify the exactness of Parrilo relaxations on unit diagonal copositive matrices for $n = 6$. We first present the framework of Parrilo relaxations quickly while recapitulating notations. In section 2, we present the computational approach taken to check the exactness of Parrilo relaxations and the results obtained. Then, we present the structure of the sums-of-squares decomposition obtained numerically and explain it in section 3. In section 4, we prove the 1-Parrilo cone to be a tractable approximation of the unit-diagonal copositive cone for $n = 6$.

If needed, see on-line at `https://github.com/sofianetanji/copositive-matrices` for additional resources (slides, code, figures, complete bibliography etc), open-sourced under the GNU GPL v3 License.

Outline

Project Advisor: Roland Hildebrand (DAO team, Laboratoire Jean Kuntzmann)
Course: *"Modeling seminar and projects"*, by Emmanuel Maître, in 2020,
Master 2 **program:** Master of Science in Industrial and Applied Mathematics (MSIAM) at Grenoble INP - Ensimag.

1 Presentation and motivation

Here are detailed the main results on copositive matrices and its applications along with common approximations of the copositive cone.

1.1 The copositive cone

1.2 Parrilo cones and sums of squares

1.3 Quick overview of the goals of this project

The next sections aim to prove that every matrix $A \in \mathcal{COP}^6$ with unit diagonal is in $\mathcal{K}_6^1$.

2 Numerical certificate of exactness

This section presents the computational approach taken to have a numerical certificate. We detail how we obtained the special copositive matrices and how we checked existence of a sums of squares decomposition.

2.1 Generation of random instances

We first focus on generating random instances of the extreme rays of $\mathcal{COP}^6$. This is first motivated by [1] which provides a complete classification of the extreme rays and by the strong result that an inner approximation of a convex cone is exact if and only if it contains all extreme rays. Namely, checking the exactness of Parrilo relaxations on the extreme rays is equivalent to checking it on the whole $\mathcal{COP}^6$ cone.

1. Pick a family of special matrices.
2. Generate random angles $(\phi_i)_i$.
3. Generate the special matrix A parametrized by the angles.

Figure 1: RANDOM GENERATION OF A SPECIAL MATRIX (randomgen.py)

The family of special matrices are all defined in [1] and corresponds to stratums 13.1, 13.2, 16, 17, 19.
The number of operations (*"flops"*) necessary for this algorithm is about :

$$T_{\text{RandomSpecialInstance}}(n_\phi) = n_\phi^2 + 7n_\phi + 48 = \mathcal{O}(n_\phi^2). \tag{1}$$

where n_ϕ is the number of angles needed to parametrize one special matrix, namely 5, 6 or 7 in our case. We provide in appendix the analytical expression of the parametrization of each special matrix. The corresponding code is in `src/randomgen.py`.

2.2 Certify a matrix is in the Parrilo cone

If each special matrix A, built using the previous algorithm in Figure 1, belongs to a Parrilo cone, we will have the numerical certificate we are searching for. Therefore, if for a given r, we are able to check if a given matrix belongs to the Parrilo cone $\mathcal{K}_6^r$, we will

have our numerical proof. We explain below how this was implemented.
The r-th Parrilo condition says that the product polynomial below has a sums of squares representation.

$$p_{r,A}(x) = \left(\sum_{k=1}^{n} x_k^2\right)^r \cdot \sum_{k,l=1}^{n} A_{kl} x_k^2 x_l^2 \tag{2}$$

Our approach is then as such, for given r and A :

1. Construct the Parrilo polynomial.
2. Implement a generic SOS representation, depending on a variable C.
3. Try to build C through a semi-definite program.

Figure 2: PARRILO POLYNOMIAL AND SOS REPRESENTABILITY (parrilo.py)

We now discuss the process used for stage 1 and 2. For stage 1, we build separately the coefficients for expressions:

$$\left(\sum_{k=1}^{n} x_k^2\right)^r \tag{3}$$

using function `additional_polynomial_creation(r)` and:

$$\sum_{k,l=1}^{n} A_{kl} x_k^2 x_l^2 \tag{4}$$

using function `matrix_polynomial_creation(A)`. The final coefficients are obtained with the method `multiplication_of_polynomials(poly1, poly2)` which computes the coefficients of the product of two polynomials.
We now detail the implementation of stage 2 as it was trickier and made use of notions we did not discuss yet.
We want to build a SOS representation of the form $\mathbf{x}^T C \mathbf{x}$ where C is a positive semi-definite matrix. As stated in section 1, comparing the coefficients of the polynomial with this representation gives the following linear equality constraints on the entries of C :

$$a_\gamma = \sum_{\alpha+\beta=\gamma} C_{\alpha\beta} \tag{5}$$

for each multi-index γ of modulus the degree of the polynomial. To compute it numerically, we first need to re-write it as such :

$$\begin{aligned} a_\gamma = \sum_{\alpha+\beta=\gamma} C_{\alpha\beta} &= \sum_{\alpha,\beta} \mathbf{1}_{\alpha+\beta=\gamma} \cdot C_{\alpha\beta} \\ &= \langle v_\gamma, \widetilde{C} \rangle \\ &= \langle M, \widetilde{C} \rangle. \end{aligned}$$

where $v_\gamma = \{\mathbf{1}_{\alpha+\beta=\gamma}\}_{\alpha\beta}$, the matrix M has rows v_γ for each γ and $\widetilde{C}$ corresponds to matrix C reshaped as a vector. Therefore, these constraints can be re-written as as dot product between M and C.

Finally, computing the SOS representation is equivalent to enforcing the constraint above so we only need to build the coefficients of matrix M.
This is done using function `creation_matrix_m(max_power)`. Both stages are implemented in file `src/parrilo.py` while stage 3 is implemented in `src/solver.py` using `cvxpy` library.

2.3 Implementation of the semi-definite program and results

We now detail the implementation of the semi-definite program (i.e. stage 3) and interpret the results obtained.
In order to verify whether the polynomial $p_{r,A}$ is a Sums Of Squares, we have to check the existence of a positive semi-definite matrix C such that:

$$p_{r,A}(x) = \mathbf{x}^T C \mathbf{x}. \tag{6}$$

A way to solve the feasibility problem above is therefore to solve the associated semi-definite program (SDP):

$$\begin{aligned} \underset{t,C}{\text{minimize}} \quad & \text{t} \\ \text{subject to} \quad & p_{r,A}(x) = \mathbf{x}^T C \mathbf{x}, \\ & C + tI \succeq 0. \end{aligned} \tag{7}$$

If the solver returns $t = 0$, this would mean that $C \succeq 0$ is on the frontier of the positive semi-definite cone. If it returns a non-positive value then this would also mean that the polynomial p is representable as a Sum Of Squares (precisely, we would be inside the cone) and finally, if it returns a positive value, then t is the distance from the polynomial to the cone of SOS polynomials in a certain direction.

2.4 Solving the SDP : issues tackled

We quickly present the issues faced to solve the SDP.

3 Structure of the SOS decomposition

In this section, we detail the structure of the SOS decompositions obtained numerically for each family, discuss about the monomials participating to the decomposition for each family and what this tells us on the structure of the problem.

3.1 Family 1

3.2 Family 2

3.3 Family 3

3.4 Family 4

3.5 Family 5

4 Analytical certificate of exactness

We now present the main contribution of our project, an analytical certificate of exactness of Parrilo relaxations in the case $n = 6$.

4.1 SOS decomposition as a function of angles

4.2 SOS decomposition and the kernels of the special matrices

4.3 Certificate of exactness for each family

4.4 Questions still not answered

Hopefully, we'll be able to delete this subsection ☺

5 Conclusion

A Special matrices

B Introduction

B.1 Copositive matrices

Definition B.1. *A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is a **copositive** matrix if $\forall x \in \mathbb{R}^n_+, \mathbf{x}^T A\mathbf{x} \geq 0$. The set of copositive matrices*

$$\mathcal{COP}^n = \{A \in \mathcal{M}_{n\times n}\mathbb{R} \mid A \text{ is copositive}\}$$

*is called the **copositive cone**.*

Let us define

$$\begin{aligned} \mathcal{S}^n_+ &: \text{ the set of positive semi-definite matrices} \\ \mathcal{N}_n &: \text{ the set of element-wise non-negative symmetric matrices} \end{aligned}$$

Theorem B.2 (Dianonda)**.** *For $n \leq 4$*

$$\mathcal{S}^n_+ + \mathcal{N}_n = \mathcal{COP}^n$$

Theorem B.3 (A. Horn)**.** *For $n \geq 5$*

$$\mathcal{S}^n_+ + \mathcal{N}_n \subsetneq \mathcal{COP}^n$$

Theorem B.4. *If $A \in \mathcal{COP}^5$ and A has a unit diagonal then $A \in \mathcal{K}^1_5$.*

The copositive cone $\mathcal{COP}^n$ is invariant under the action of the multiplicative group $\mathbb{R}^n_{++}, A \mapsto DAD$ where $D = diag(d), d \in \mathbb{R}^n_{++}$. Thus $A \in \mathcal{COP}^n$ with a non unit positive diagonal can be scaled to a copositive matrix with unit diagonal by setting $d = diagA$ and applying

$$\tilde{A} = diag\left(d^{-\frac{1}{2}}\right) Adiag\left(d^{-\frac{1}{2}}\right).$$

Consequently for $n = 5$ a work-around strategy to check if a matrix A is copositive would be to scale it to have its diagonal a unit vector and then check if it belongs to $\mathcal{K}^1_5$.
A matrix A belongs to the r-th Parrilo cone $\mathcal{K}^r_n$ if the polynomial of degree $2r + 4$

$$\left(\sum_{k=1}^{n} x_k^r\right)^r \sum_{K,l=1}^{n} A_{k,l} x_k^2 x_l^2$$

can be represented as a SOS of polynomials of degree $2 + r$.
The goal of the project is to check the exactness of the Parrilo relaxations on unit diagonal copositive matrices in the $n = 6$ case.

Write a proper introduction, usefulness of copositive matrices, what do they represent, what are we going to do, the layout of the report

fill the blank spaces with comments to create a smooth flow

Put the references

B.2 Sum of Squares

Let $p : \mathbb{R}^n \to \mathbb{R}$ be any polynomial of degree $2d$. If $p(x) \geq 0, \forall x \in \mathbb{R}^n$, the question is whether can the polynomial be represented as a sum of squares (SOS), i.e. can be decomposed as

$$p(x) = \sum_{n=1}^{m} q_j^2(x)$$

where $q_j(x)$ is a homogeneous polynomial of degree d, $i = 1, \ldots, m$.
Let $\mathbf{x} = \{x^\alpha\}_{|\alpha|=d}$ the sequence of all monomials of degree d,

$$x^\alpha = \prod_{i=1}^{n} x_i^{\alpha_i}$$

with $\alpha = (\alpha_i)_{i=1,\ldots,d} \in \mathbb{N}^n$ and $|\alpha| = \sum_{i=1}^{n} \alpha_i$ is the degree .

Proposition B.5. *A polynomial $p(x)$ is a SOS if and only if there exists a PSD matrix C such that $p(x) = \mathbf{x}^T C\mathbf{x}$.*

Proof. We have

$$q_j(x) = c_j^T \mathbf{x} = \sum_{|\alpha|=d} c_{j,\alpha}^T x^\alpha \tag{8}$$

where $c_j = \{c_{j,\alpha}\}_{|\alpha|=d}$ is the vector of coefficients of q_j. The component $c_j^T \mathbf{x}$ is a scalar so

$$\left(c_j^T \mathbf{x}\right)^T = \mathbf{x}^T c_j = c_j^T \mathbf{x} \tag{9}$$

and as a result

$$\begin{aligned}
\sum_{j=1}^{m} q_j(x)^2 &= \sum_{j=1}^{m} \left(c_j^T \mathbf{x}\right)^2 = \sum_{j=1}^{m} c_j^T \mathbf{x} c_j^T \mathbf{x} \\
&= \sum_{j=1}^{m} \mathbf{x}^T c_j c_j^T \mathbf{x} = \mathbf{x}^T \left(\sum_{j=1}^{m} c_j c_j^T \right) \mathbf{x} \\
&= \mathbf{x}^T C \mathbf{x}
\end{aligned}$$

where $C = \sum_{j=1}^{m} c_j c_j^T$ is a PSD matrix.
On the other hand any semi-definite matrix $C \succeq 0$ can be written as the sum $C = \sum_{j=1}^{m} c_j c_j^T$ and equivalently $\mathbf{x}^T C\mathbf{x}$ is a SOS. □

Verifying that $p(x)$ is a SOS is equivalent to solving the SDP

$$\min_{t,C} t : p(x) = \mathbf{x}^T C\mathbf{x} \text{ and } C + tI \succeq 0 \tag{10}$$

If we write

$$p(x) = \sum_{|\gamma|=2d} a_\gamma x^\gamma$$

then the equality constraint $p(x) = \mathbf{x}^T C\mathbf{x}$ can be divided into simpler constraints on the entries of C by comparing the coeficients at the powers of x on both sides, i.e.

$$a_\gamma = \sum_{\alpha+\beta=\gamma} = C_{\alpha,\beta}$$

Example B.6. We take a polynomial p of degree 2

$$p(x) = 9x_1^2 + 4x_2^2 + 12x_1x_2$$

that can be written as the square of a homogeneous matrix $q(x) = 3x_1 + 2x_2$

$$p(x) = (q(x))^2 .$$

We shall find the matrix $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ such that

$$p(x) = \mathbf{x}^T C\mathbf{x}$$

We distribute

$$\begin{aligned}
\mathbf{x}^T C\mathbf{x} &= (x_1, x_2) \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= (c_{11}x_1 + c_{21}x_2, c_{12}x_1 + c_{22}x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= c_{11}x_1^2 + (c_{21} + c_{12})\, x_1x_2 + c_{22}x_2^2
\end{aligned}$$

The component c_{11} corresponds to x_1 the first component of both $\mathbf{x}$ and $\mathbf{x}^T$ so we can note $c_{11} = c_{(1,0),(1,0)}$. Also c_{12} corresponds to the second component of $\mathbf{x}$ and to the first component of $\mathbf{x}^T$ so we can write it as $c_{12} = c_{(1,0),(0,1)}$. Equivalently we get $c_{21} = c_{(0,1),(1,0)}$ and $c_{22} = c_{(0,1),(0,1)}$. So

$$\mathbf{x}^T C\mathbf{x} = c_{(1,0),(1,0)}x_1^2 + \left(c_{(0,1),(1,0)} + c_{(1,0),(0,1)}\right) x_1x_2 + c_{22} = c_{(0,1),(0,1)}x_2^2$$

We get the system

$$\begin{cases}
a_{(2,0)} = \sum_{\alpha+\beta=(2,0)} C_{\alpha\beta} = c_{(1,0),(1,0)} = 9 \\
a_{(1,1)} = \sum_{\alpha+\beta=(1,1)} C_{\alpha\beta} = c_{(1,0),(0,1)} + c_{(0,1),(1,0)} = 4 \\
a_{(0,2)} = \sum_{\alpha+\beta=(0,2)} C_{\alpha\beta} = c_{(0,1),(0,1)} = 12
\end{cases}$$

And a solution would be

$$C = \begin{pmatrix} 9 & 6 \\ 6 & 4 \end{pmatrix}$$

B.3 Sum of Squares

Let p be temporary quadratic function $p(x, y, z)$

$$p(x, y, z) = x^4 + y^4 + z^4 - 2x^2yz - -2xy^2z - -2xyz^2$$

$$X = (x^2, y^2, z^2, yz, xz, xy)^T$$

please complete this part

$$\max_{t,c} t : C - tI \geq 0 \tag{11}$$

$$a_r = \sum_{\alpha+\beta=r} c_{\alpha\beta} \forall r \tag{12}$$

If $t \geq 0$ then p is SOS representable if $t < 0$ then p is not SOS
$A \in S^n$. A is copositive if $X^T AX \geq 0 \forall X \in \mathbb{R}^n_+$
Question Is A Co-positive?

$$\text{A copositive } \Leftrightarrow P_A = \sum_{i,j=1}^{n} A_{ij} X_i^2 X_j^2 \geq 0$$

References

[1] Andrey Afonin, Roland Hildebrand, and Peter Dickinson. The extreme rays of the 6×6 copositive cone. 2020.