

Sum of Squares

If a real polynomial can be written as a sum of squares, then this sum provides a certificate that the polynomial has no real roots. In this lecture, mainly following [10] we will see how with semidefinite programming we may decide whether a polynomial can be written as a sum of squares. A recent survey is in [4].

1 Positive Semidefinite Matrices and Nonnegative Polynomials

Consider a real polynomial $p \in \mathbb{R}[x]$. If p can be written as a sum of squares, i.e.: if there exist r polynomials $q_i \in \mathbb{R}[x]$, $i = 1, 2, \dots, r$, such that $p = q_1^2 + q_2^2 + \dots + q_r^2$, then $p(x) \geq 0$, for all $\mathbf{x} \in \mathbb{R}$. Moreover, for any $\epsilon > 0$, $p + \epsilon > 0$ is positive and has no real roots. Not every nonnegative polynomial is a sum of squares.

We will make the connection between positive semidefinite matrices and sum of squares. A matrix A is positive semidefinite if for all vectors \mathbf{x} : $\mathbf{x}^T A \mathbf{x} \geq 0$. A symmetric positive semidefinite matrix A admits a decomposition as $A = R^T R$, where R is an upper triangular matrix. This is known as the Cholesky decomposition (the command `chol` in Octave) of A .

Matrices arise in the representation of polynomials as follows. Take as p for example a fourth degree polynomial, which we may represent as

$$p(x) = [x^2 \ x \ 1] \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ x \\ 1 \end{bmatrix} = \mathbf{x}^T A \mathbf{x}. \quad (1)$$

If A is positive semidefinite, then its Cholesky decomposition $A = R^T R$ gives

$$p(x) = \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T R^T R \mathbf{x} = (R \mathbf{x})^T (R \mathbf{x}) = y_1^2 + y_2^2 + y_3^2, \quad \mathbf{y} = R \mathbf{x}. \quad (2)$$

We are interested in the positive definite cone (or PSD cone) of all real symmetric positive semidefinite n -by- n matrices

$$PSD(\mathbb{R}^n) = \{ A \in \mathbb{R}^{n \times n} \mid A^T = A \text{ and } A \text{ is positive semidefinite} \}. \quad (3)$$

That $PSD(\mathbb{R}^n)$ is a cone follows from the inequalities $\mathbf{x}^T A \mathbf{x} \geq 0$.

Semidefinite programming (or SDP for short) offers efficient algorithms for the following problems:

SDP decision: given r linear functions $\ell_1, \ell_2, \dots, \ell_r$, does there exist an $A \in PSD(\mathbb{R}^n)$ such that $\ell_i(A) = 0$, for $i = 1, 2, \dots, r$?

SDP optimization: for $r + 1$ linear functions $\ell_0, \ell_1, \dots, \ell_r$, minimize $\ell_0(A)$ subject to $A \in PSD(\mathbb{R}^n)$ and $\ell_i(A) = 0$, for $i = 1, 2, \dots, r$.

While this clearly generalizes the classical linear programming models, interior-point methods provide efficient algorithms to solve these problems [11].

For the polynomial $p(x) = x^4 - x^2 - 2x + 2$ (example taken from [10]), we consider

$$\mathbf{x}^T A \mathbf{x} = a_{11}x^4 + 2a_{12}x^3 + (2a_{13} + a_{22})x^2 + a_{23}x + a_{33}. \quad (4)$$

Now we want to find those $A \in PSD(\mathbb{R}^n)$ satisfying the linear equations $a_{11} = 1$, $a_{12} = 0$, $2a_{13} + a_{22} = -1$, $a_{23} = -1$, and $a_{33} = 2$. It turns out that this SDP decision problem has a unique solution so that p can be written as a sum of squares.

Consider for example the set defined by $1 \leq x^2 + y^2 \leq 2$. This is clearly not a convex set. Can we use convex optimization to compute with this set? The idea is to view the set defined by $1 \leq x^2 + y^2 \leq 2$ as the projection of the convex set $\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq z, 1 \leq z \leq 2 \}$. This lifting is a useful relaxation technique.

2 Control of Nonlinear Systems via Lyapunov Functions

Following [6] (see also [8]), consider a nonlinear system

$$\frac{\partial \mathbf{x}(t)}{\partial t} = f(\mathbf{x}(t)), \quad (5)$$

where $f(\mathbf{x})$ is a real polynomial system in \mathbf{x} . To prove asymptotic stability of $\mathbf{x} = \mathbf{0}$ as a fixed point of this nonlinear system, we must find a Lyapunov function $V(\mathbf{x})$ satisfying

$$\text{for all } \mathbf{x} \neq \mathbf{0}: \quad V(\mathbf{x}) > \mathbf{0} \quad \text{and} \quad \frac{\partial V(\mathbf{x}(t))}{\partial t} = \left(\frac{\partial V}{\partial \mathbf{x}} \right)^T f(\mathbf{x}) < 0, \quad (6)$$

for all nonzero \mathbf{x} in a neighborhood of the origin.

Expressing $V(\mathbf{x})$ as a sum of squares gives an explicit way of showing its positivity. To automate the search for a Lyapunov function, parameters λ are introduced, so $V = V(\mathbf{x}, \lambda)$. For \mathbf{z} a monomial vector, we then represent V as a quadratic form in \mathbf{z} : $V(\mathbf{x}, \lambda) = \mathbf{z}^T Q(\lambda) \mathbf{z}$. The parameters λ are then determined so that $Q(\lambda)$ is positive semidefinite.

3 Global Optimization

Via semidefinite programming we may find the global minimum of a polynomial function f on \mathbb{R}^n . Without semidefinite programming, we could compute all critical points of the system defined by all partial derivatives $\frac{\partial f}{\partial x_i}(\mathbf{x}) = 0$, for $i = 1, 2, \dots, n$. The number of solutions of this system however grows exponentially and solving this system of partial derivatives is feasible only for a very modest number of variables and for low degree polynomials.

Instead, we consider a relaxation of the problem:

SOS Relaxation: Find the largest $\lambda \in \mathbb{R}$ such that $f(x_1, x_2, \dots, x_n) - \lambda$ is a sum of squares.

The dimension of the vector space N for which we consider $PSD(\mathbb{R}^N)$ is proportional to the number of monomials in f , i.e. it grows like $\binom{n+d}{d}$, where $d = \deg(f)$. Although also this dimension grows exponentially, the growth is more moderate compared to the system of partial derivatives.

4 Sum of Squares and Radical Ideals

The definition of radical of an ideal I is

$$\sqrt{I} = \{ f \in \mathbb{R}[\mathbf{x}] \mid f^k \in I \text{ for some integer } k \geq 1 \}. \quad (7)$$

Hilbert's Nullstellensatz states $\sqrt{I} = I(V)$ for a variety V . In words, the radical of an ideal is the vanishing ideal of the solutions. Lemma 4.1 gives a criterion for an ideal to be radical, limiting the powers to squares.

Lemma 4.1 *Let I be an ideal in $\mathbb{R}[\mathbf{x}]$. I is radical if and only if*

$$\text{for all } f \in \mathbb{R}[\mathbf{x}]: f^2 \in I \Rightarrow f \in I. \quad (8)$$

Proof. First, consider the \Rightarrow in the *if and only if*, assuming I is radical. If I is radical, then for all integer $k \geq 1$: $f^k \in I$ implies $f \in I$. Take $k = 2$ and (8) follows.

Second, to show the \Leftarrow in the *if and only if*, assuming (8) holds, we show that $f \in I$, if $f^k \in I$, for some integer $k \geq 1$. For example, let $k = 4$: if $f^4 \in I$, then $(f^2)^2 \in I$ implies $f^2 \in I$, using (8). Applying (8) once more results in $f \in I$. For $k = 2^\ell$, applying (8) ℓ times yields $f \in I$. If k is not a power of 2, we multiply f^k sufficiently many times with f till we have $f^{2^\ell} \in I$. \square

The real radical of an ideal I is

$$\sqrt[\mathbb{R}]{I} = \left\{ f \in \mathbb{R}[\mathbf{x}] \mid f^{2k} + \sum_{j=1}^m p_j^2 \in I \text{ for some } k \geq 1, p_j \in \mathbb{R}[\mathbf{x}] \right\}. \quad (9)$$

Similar to Lemma 4.1, we have a criterion for an ideal to be real radical in Lemma 4.2.

Lemma 4.2 *Let I be an ideal in $\mathbb{R}[\mathbf{x}]$. I is real radical if and only if*

$$\text{for all } p_j \in \mathbb{R}[\mathbf{x}] : \sum_{j=1}^m p_j^2 \in I \Rightarrow p_j \in I. \quad (10)$$

Proof. First, consider the \Rightarrow in the *if and only if*, assuming I is real radical. If I is real radical, i.e.: $I = \sqrt[\mathbb{R}]{I}$, then for all integer $k \geq 1$ and $p_j \in \mathbb{R}[\mathbf{x}]$: $f^{2k} + \sum_{j=1}^m p_j^2 \in I$ implies $f \in I$. Take $k = 2$ and (10) follows.

Second, to show the \Leftarrow in the *if and only if*, assuming (10) holds. Let $f, p_j \in \mathbb{R}[\mathbf{x}]$ such that $f^{2k} + \sum_{j=1}^m p_j^2 \in I$ holds. By (10), we have $f^k, p_j \in I$. As (10) implies (8), we apply Lemma 4.2 to deduce $f \in I$. \square

For a polynomial system $f(\mathbf{x}) = \mathbf{0}$, with $f = (f_1, f_2, \dots, f_N)$, $f_i \in \mathbb{C}[\mathbf{x}]$, for $i = 1, 2, \dots, N$. Hilbert's Nullstellensatz implies

$$f^{-1}(\mathbf{0}) = \emptyset \Leftrightarrow \langle f \rangle = \langle 1 \rangle. \quad (11)$$

So if the system $f(\mathbf{x}) = \mathbf{0}$ has no solutions, then there are polynomials $g_i \in \mathbb{C}[\mathbf{x}]$, $i = 1, 2, \dots, N$, such that $1 = g_1 f_1 + g_2 f_2 + \dots + g_N f_N$. The polynomials g_i provide a certificate that $f(\mathbf{x}) = \mathbf{0}$ has no solutions. See [2] for more on Hilbert's Nullstellensatz.

Either a system of real polynomial equations and inequalities has a real solution, or there is a certificate that no solution exists. Such certificate consists in a polynomial identity listed in the theorem [10] below.

Theorem 4.1 (Real Nullstellensatz) *The system of real polynomial equations and inequalities*

$$\begin{cases} f_1(\mathbf{x}) = 0, f_2(\mathbf{x}) = 0, \dots, f_r(\mathbf{x}) = 0 \\ g_1(\mathbf{x}) \geq 0, g_2(\mathbf{x}) \geq 0, \dots, g_s(\mathbf{x}) \geq 0 \\ h_1(\mathbf{x}) > 0, h_2(\mathbf{x}) > 0, \dots, h_t(\mathbf{x}) > 0 \end{cases} \quad (12)$$

has a solution for $\mathbf{x} \in \mathbb{R}^n$, or there exists a polynomial identity

$$\sum_{i=1}^r \alpha_i f_i + \sum_{\nu \in \{0,1\}^s} \left(\sum_j b_{j\nu}^2 \right) \cdot g_1^{\nu_1} g_2^{\nu_2} \cdots g_s^{\nu_s} + \sum_{\nu \in \{0,1\}^t} \left(\sum_j c_{j\nu}^2 \right) \cdot h_1^{\nu_1} h_2^{\nu_2} \cdots h_t^{\nu_t} + \sum_k d_k^2 + \prod_{l=1}^t h_l^{u_l} = 0, \quad (13)$$

where $u_j \in \mathbb{N}$ and $\alpha_i, b_{j\nu}, c_{j\nu}$, and d_k are polynomials.

To get a better understanding of this theorem it is instructive to consider certain special cases, for example $s = 0 = t$. In real algebraic geometry the focus is on real roots and one studies semi-algebraic sets [1]. In [6], it is stated that a Real Nullstellensatz certificate of bounded degree can be computed efficiently by semidefinite programming.

5 Moment Matrices

In this section, we follow [5] and [9].

We can represent a polynomial with support A as $f(\mathbf{x}) = \sum_{\mathbf{a} \in A} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}} \in \mathbb{R}[\mathbf{x}]$ by its coefficient vector \mathbf{c}_A with respect to some fixed monomial basis $(\mathbf{x}^{\mathbf{a}})_{\mathbf{a} \in A}$. The corresponding dual basis uses differentials $(\partial^{\mathbf{a}})_{\mathbf{a} \in A}$. Given a sequence of numbers $\mathbf{y} = (y_{\mathbf{a}})_{\mathbf{a} \in A} \in \mathbb{R}^m$, for $m = \#A$, we consider a linear functional $\Lambda \in (\mathbb{R}[\mathbf{x}])^*$ as $\Lambda = \sum_{\mathbf{a} \in A} y_{\mathbf{a}} \partial^{\mathbf{a}}$. Then \mathbf{y} is *the coordinate sequence* $\mathbf{y} = (\Lambda(\mathbf{x}^{\mathbf{a}}))_{\mathbf{a} \in A}$.

Given the linear form Λ , we define the *moment matrix* as

$$M(\mathbf{y}) = (\Lambda(\mathbf{x}^{\mathbf{a}+\mathbf{b}}))_{\mathbf{a}, \mathbf{b}} = (y_{\mathbf{a}+\mathbf{b}})_{\mathbf{a}, \mathbf{b}}, \quad (14)$$

with rows and columns indexed by $\mathbf{a}, \mathbf{b} \in A$. For a polynomial f as denoted above, we have the inner product

$$\langle f, f \rangle_{\Lambda} = \mathbf{c}_A^T M(\mathbf{y}) \mathbf{c}_A = \Lambda(f^2). \quad (15)$$

By construction, the positive semidefiniteness of $M(\mathbf{y})$ is equivalent to $\Lambda(f^2) \geq 0$.

An interpretation of the moment matrix is the matrix representing the quasi-Hankel operator:

$$H_{\Lambda} : \mathbb{R}[\mathbf{x}] \rightarrow (\mathbb{R}[\mathbf{x}])^* : h \mapsto h \cdot \Lambda. \quad (16)$$

We have $h \cdot \Lambda(f) = \Lambda(hf)$ for all $f \in \mathbb{R}[\mathbf{x}]$.

The kernel of the moment matrix is

$$\ker M(\mathbf{y}) = \{ f \in \mathbb{R}[\mathbf{x}] \mid M(\mathbf{y}) \mathbf{c}_A = \mathbf{0} \}. \quad (17)$$

The kernel $\ker M(\mathbf{y})$ is an ideal in $\mathbb{R}[\mathbf{x}]$. If $M(\mathbf{y})$ is positive semidefinite, then $\ker M(\mathbf{y})$ is real radical.

In computations, we consider polynomials of degree at most d , denoted by $\mathbb{R}[\mathbf{x}]_d$. Given a linear form $\Lambda = \sum_{\mathbf{a} \in A} y_{\mathbf{a}} \partial^{\mathbf{a}}$, the matrix

$$M_{\lfloor d/2 \rfloor}(\mathbf{y}) = (\Lambda(\mathbf{x}^{\mathbf{a}+\mathbf{b}}))_{\mathbf{a}, \mathbf{b}} = (y_{\mathbf{a}+\mathbf{b}})_{\mathbf{a}, \mathbf{b}} \quad (18)$$

is *the truncated moment matrix*.

The method is to find \mathbf{y} for which the rank of the moment matrix is maximum. This rank equals the number of real solutions.

Denote $V_{\mathbb{C}}(I)$ the solutions in \mathbb{C}^n and by $V_{\mathbb{R}}(I) = V_{\mathbb{C}}(I) \cap \mathbb{R}^n$ the real solutions. We end with an algorithm of [9].

Algorithm 5.1 (Moment-Matrix algorithm)

Input: $I = \langle f_1, f_2, \dots, f_N \rangle$, with $\#V_{\mathbb{R}}(I) < \infty$.

Output: a border or Gröbner basis for $J := \ker M_d(\mathbf{y})$, with $V_{\mathbb{C}}(J) = V_{\mathbb{C}}(I)$.

1. $d := D := \max_{i=1}^N \deg(f_i)$;
2. Find a generic \mathbf{y} ;
3. if $(\text{rank} M_s(\mathbf{y}) = \text{rank} M_{s-1}(\mathbf{y}), \text{ for some } D \leq s \leq \lfloor d/2 \rfloor$
or $\text{rank} M_s(\mathbf{y}) = \text{rank} M_{s-t}(\mathbf{y}) \text{ for some } t \leq s \leq \lfloor d/2 \rfloor)$
4. then return a basis for the column space of $M_{d-1}(\mathbf{y})$;
5. else $d := d + 1$; go to step 2;
6. end if.

6 Exercises

1. Verify that we may write every fourth degree polynomial as (1), although not necessarily in a unique way. Can you generalize this representation for polynomials of any degree?
2. Suppose we want to compute the global minimum of a cubic polynomial f in 10 variables. Use Bézout's theorem to bound the number of solutions of the system of all partial derivatives. Compare this number with the number of monomials in f .
3. Install SOSTOOLS on your computer and use it for example through its Macaulay 2 interface [7]. Report running times for the examples in [7].
4. Consider the family of homogeneous polynomials M_{jk} of [3]:

$$M_{jk}(x_1, x_2, x_3) = jx_3^6 + x_1^2x_2^2(jx_1^2 + jx_2^2 - kx_3^2), \quad (19)$$

where j and k are positive integers.

Give examples of numerical evaluations of M_{jk} for some instances of j and k at well chosen values for x_1 , x_2 , and x_3 to illustrate the importance of positivity on the accuracy of the evaluation.

References

- [1] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer-Verlag, 2003.
- [2] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, second edition, 1997.
- [3] J. Demmel, I. Dumitriu, and O. Holtz. Toward accurate polynomial evaluation in rounded arithmetic. In L.M. Pardo, A. Pinkus, E. Süli, and M.J. Todd, editors, *Foundations of Computational Mathematics, Santander 2005*, volume 331 of *London Mathematical Society Lecture Note Series*, pages 36–105. Cambridge University Press, 2006. [arXiv:math/0508350v2 \[math.NA\] 18 Jan 2006](#).
- [4] J.-B. Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM Review*, 49(4):651–669, 2007.
- [5] J.-B. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Foundations of Computational Mathematics*, 8(5):607–647, 2008.
- [6] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.
- [7] H. Peyrl and P.A. Parrilo. A Macaulay 2 package for computing sum of squares decompositions of polynomials with rational coefficients. In J. Verschelde and S.M. Watt, editors, *SNC'07. Proceedings of the 2007 International Workshop on Symbolic-Numeric Computation*, pages 207–208. ACM, 2007.
- [8] S. Prajna, P.A. Parrilo, and A. Rantzer. Nonlinear control synthesis by convex optimization. *IEEE Transactions on Automatic Control*, 42(2):310–314, 2004.
- [9] P. Rostalski. *Algebraic Moments. Real Root Finding and Related Topics*. PhD thesis, ETH Zürich, 2009.
- [10] B. Sturmfels. *Solving Systems of Polynomial Equations*. Number 97 in CBMS Regional Conference Series in Mathematics. AMS, 2002.
- [11] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.