MIT 6.7230 - Algebraic techniques and semidefinite optimization	March 15, 2024

Lecture 9

Lecturer: Pablo A. Parrilo *Scribe: ???*

In this lecture we continue our study of SOS polynomials. After presenting a couple of applications, we discuss the dual side, and provide a natural probabilistic interpretation of the corresponding problem. We further introduce a natural geometric description, in terms of approximations to the convex hull of a certain algebraic variety.

1 Applications of sum of squares

1.1 Lyapunov functions

Expressing conditions for a polynomial to be a sum-of-squares as an SDP is very useful, since we can use the SOS property as a convenient "replacement" for polynomial nonnegativity. In the dynamical systems context, recent work has applied the sum-of-squares approach to the problem of finding a Lyapunov function for nonlinear systems [Par00, PP02].

This approach enables the search over affinely parametrized polynomial or rational Lyapunov functions for systems with dynamics of the form

$$\dot{x}_i(t) = f_i(x(t)) \qquad \text{for all } i = 1, \ldots, n \tag{1}$$

where the functions f_i are polynomials or rational functions. Recall that for a system to be globally asymptotically stable, it is sufficient to prove the existence of a Lyapunov function that satisfies

$$V(x) > 0, \qquad \dot{V}(x) = \left(\frac{\partial V}{\partial x}\right)^T f(x) < 0$$

for all $x \in \mathbb{R}^n \setminus \{0\}$, where without loss of generality we have assumed that the system (1) has an equilibrium at the origin (see, e.g., [Kha92]). Then the condition that the Lyapunov function be positive, and that its Lie derivative be negative, are both directly imposed as sum-of-squares constraints in terms of the coefficients of the Lyapunov function.

As an example, consider the following system:

$$\begin{aligned} \dot{x} &= -x + (1+x)y \\ \dot{y} &= -(1+x)x. \end{aligned} \tag{2}$$

It is known that this system has no quadratic Lyapunov function. However, using SOSTOOLS [PPP05] we easily find a quartic polynomial Lyapunov function, which after rounding (for purely cosmetic reasons) is given by

$$V(x,y) = 6x^2 - 2xy + 8y^2 - 2y^3 + 3x^4 + 6x^2y^2 + 3y^4.$$

The corresponding phase diagram, showing both the trajectories and the level sets of the Lyapunov function V, is given in Figure 1. It can be readily verified that both $V(x,y)$ and $(-\dot{V}(x,y))$ are SOS, since

$$V = \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 6 & -1 & 0 & 0 & 0 \\ -1 & 8 & 0 & 0 & -1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & -1 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}, \qquad -\dot{V} = \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix}^T \begin{bmatrix} 10 & 1 & -1 & 1 \\ 1 & 2 & 1 & -2 \\ -1 & 1 & 12 & 0 \\ 1 & -2 & 0 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix},$$

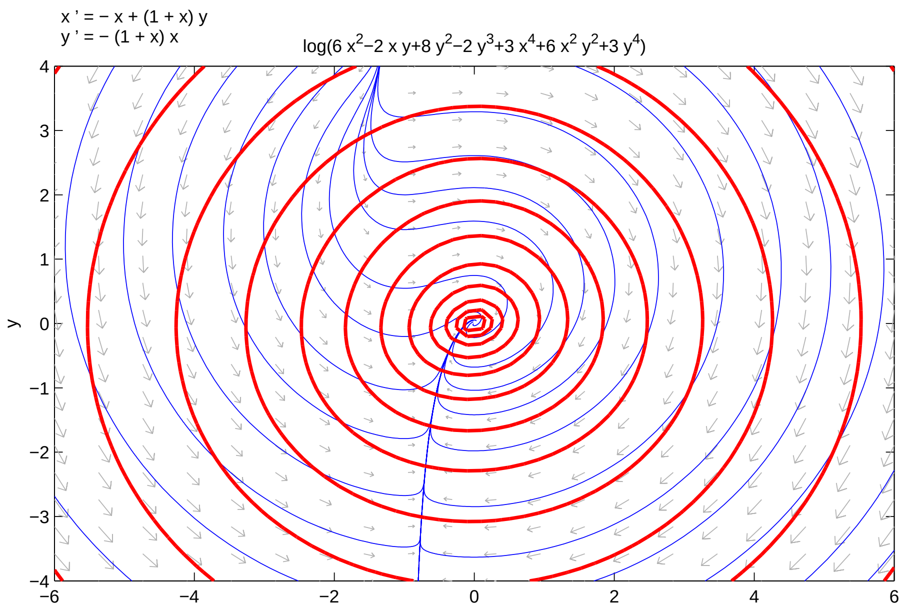

Figure 1: Trajectories of the dynamical system (2) and level sets of the SOS Lyapunov function $V(x, y)$.

and the matrices in the expression above are positive definite. Similar approaches may also be used for finding Lyapunov functionals for certain classes of hybrid systems.

Remark 1. *In stark constrast to the linear case (where quadratic Lyapunov always exist), polynomial dynamical systems that are globally asymptotically stable may not admit polynomial Lyapunov functions; see for instance [AKP11] for a simple counterexample. Nevertheless, the basic technique described above can be easily modified to deal with stability over compact regions – in this case (essentially, by the Stone-Weierstrass approximation theorem) polynomial Lyapunov functions that prove stability always exist.*

1.2 Entangled states in quantum mechanics

The state of a finite-dimensional quantum system can be described in terms of a positive semidefinite Hermitian matrix, called the *density matrix*. An important property of a bipartite quantum state ρ is whether or not it is *separable*, which means that it can be written as a convex combination of tensor products of rank one matrices, i.e.,

$$\rho = \sum_i p_i \, (x_i x_i^T) \otimes (y_i y_i^T), \qquad p_i \geq 0, \quad \sum_i p_i = 1,$$

where for simplicity we have restricted ρ, x_i, y_i to be real. Here $x_i \in \mathbb{R}^{n_1}$, $y_i \in \mathbb{R}^{n_2}$, and $\rho \in \mathcal{S}_+^{n_1 n_2}$. If the state is not separable, then it is said to be *entangled*.

A question of interest is the following: Given the density matrix ρ of a quantum state, how to recognize whether the state is entangled or not? How can we certify that the state is entangled? It has been shown by Gurvits that in general this is an NP-hard question [Gur03].

A natural mathematical object to study in this context is the set of *positive maps*, i.e., the linear operators $\Lambda : \mathcal{S}^{n_1} \to \mathcal{S}^{n_2}$ that map positive semidefinite matrices into positive semidefinite matrices. Notice that to any such Λ, we can associate a unique "observable" $L \in \mathcal{S}^{n_1 n_2}$, that satisfies $y^T \Lambda(xx^T) y = (x \otimes y)^T L (x \otimes y)$. Furthermore, if Λ is a positive map, then the pairing between the observable L and any separable state will always give a nonnegative number, since

$$\begin{aligned}\langle L, \rho \rangle &= \operatorname{Tr} L \cdot (\sum_i p_i \, (x_i x_i^T) \otimes (y_i y_i^T)) = \sum_i p_i \operatorname{Tr} L \cdot (x_i \otimes y_i) \cdot (x_i \otimes y_i)^T \\ &= \sum_i p_i \, (x_i \otimes y_i)^T L (x_i \otimes y_i) = \sum_i p_i \, y_i^T \Lambda(x_i x_i^T) y_i \geq 0.\end{aligned}$$

In other words, every positive map yields a *separating hyperplane* for the convex set of separable states. It can further be shown that this is in fact a complete characterization (and thus, these sets are dual to each other).

The set of positive maps can be exactly characterized in terms of a multivariate polynomial nonnegativity condition, since the map $\Lambda : \mathcal{S}^{n_1} \to \mathcal{S}^{n_2}$ is positive if and only if the polynomial $p(x, y) = y^T \Lambda(xx^T) y$ is nonnegative for all x, y (why?). Replacing nonnegativity with sum of squares based conditions, we can obtain a family of efficiently computable criteria that certify entanglement.

For more background and details about this problem, see [DPS02, DPS04] and the references therein.

2 Dual side: moments

Consider a nonnegative measure μ on $\mathbb{R}$ (or if you prefer, a real-valued random variable X). We can then define the *moments*, which are the expectation of powers of X.

$$\mu_k := \mathbf{E}[X^k] = \int x^k d\mu \tag{3}$$

What constraints, if any, should the μ_k satisfy? Is is true that for any set of numbers $\mu_0, \mu_1, \ldots, \mu_k$, there always exists a nonnegative measure having exactly these moments? This is the classical (truncated) moment problem [Akh65].

It should be apparent that some conditions are required. For instance, consider (3) for an even value of k. Since the measure μ is nonnegative, it is clear that in this case we have $\mu_k \geq 0$.

However, that's clearly not enough, and more restrictions should hold. A simple one can be derived by recalling the relationship between the first and second moments and the variance of a random variable, i.e., $\operatorname{var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \mu_2 - \mu_1^2$. Since the variance is always nonnegative, we should have $\mu_2 - \mu_1^2 \geq 0$.

How to systematically derive conditions of this kind? Notice that the previous inequality can be obtained by noticing that for all a, b,

$$0 \leq \mathbf{E}[(a + bX)^2] = a^2 \mathbf{E}[1] + 2ab \mathbf{E}[X] + b^2 \mathbf{E}[X^2] = \begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix},$$

which implies that the 2×2 matrix above must be positive semidefinite. Interestingly, the variance inequality obtained earlier corresponds to the determinant of this matrix.

Exactly the same procedure can be done for higher-order moments. Proceeding this way, by considering expectations of squares of higher-order polynomials, we have

$$0 \leq \mathbf{E}[(c_0 + c_1 X + \cdots + c_d X^d)^2] = \sum_{j=0}^{d} \sum_{k=0}^{d} c_j c_k \mathbf{E}[X^{j+k}]$$

This is a quadratic form in the variables $(c_0, c_1, \ldots, c_d)$, and thus the higher order moments must always satisfy the semidefinite condition on a Hankel matrix given by

$$\begin{bmatrix} 1 & \mu_1 & \mu_2 & \cdots & \mu_d \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{d+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_d & \mu_{d+1} & \mu_{d+2} & \cdots & \mu_{2d} \end{bmatrix} \succeq 0. \tag{4}$$

Notice that the diagonal elements correspond to even-order moments, which should obviously be nonnegative.

As we will see below, this condition is "almost" necessary and sufficient in the univariate case. In the multivariate case, however, there will be more serious problems (just like for polynomial nonnegativity vs. sums of squares).

Remark 2. *For unbounded intervals, the SDP conditions characterize the* closure *of the set of moments, but not necessarily the whole set. As an example, consider the set of moments given by* $\mu = (1, 0, 0, 0, 1)$*, corresponding to the Hankel matrix*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Although the matrix above is PSD, it is not hard to see that there is no nonnegative measure corresponding to those moments. However, the parametrized atomic measure given by

$$\mu_\varepsilon = \frac{\varepsilon^4}{2} \cdot \delta(x + \frac{1}{\varepsilon}) + (1 - \varepsilon^4) \cdot \delta(x) + \frac{\varepsilon^4}{2} \cdot \delta(x - \frac{1}{\varepsilon})$$

has as first five moments $(1, 0, \varepsilon^2, 0, 1)$*, and thus as* $\varepsilon \to 0$ *the corresponding Hankel matrix is the one given above.*

2.1 Nonnegative measures on intervals

Just like we did for the case of polynomials nonnegative on intervals, we can similarly obtain necessary and sufficient characterizations for moments of measures supported on intervals. For simplicity, we present below only one particular case, corresponding to the interval $[-1, 1]$.

Lemma 3. *There exists a nonnegative measure supported on* $[-1, 1]$ *with moments* $(\mu_0, \mu_1, \ldots, \mu_{2d+1})$ *if and only if*

$$\begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_d \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{d+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_d & \mu_{d+1} & \mu_{d+2} & \cdots & \mu_{2d} \end{bmatrix} \pm \begin{bmatrix} \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{d+2} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{d+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{d+1} & \mu_{d+2} & \mu_{d+3} & \cdots & \mu_{2d+1} \end{bmatrix} \succeq 0. \tag{5}$$

The necessity of this condition is clear, since it follows from consideration of the quadratic form (in the c_i):

$$0 \le \mathbf{E}\left[(1 \pm X)(c_0 + c_1 X + \cdots + c_d X^d)^2\right] = \sum_{j=0}^{d} \sum_{k=0}^{d} (\mu_{j+k} \pm \mu_{j+k+1}) c_j c_k,$$

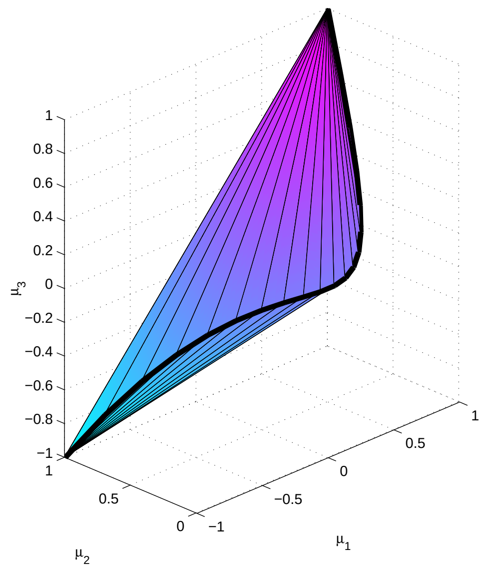

Figure 2: Set of valid moments (μ_1, μ_2, μ_3) of a probability measure on $[-1, 1]$. This is the convex hull of the moment curve (t, t^2, t^3), for $-1 \leq t \leq 1$. An explicit SDP representation is given in (5).

where the first inequality follows since $1 \pm X$ is always nonnegative, since X is supported on $[-1, 1]$. Notice the similarities (in fact, the duality) with the conditions for polynomial nonnegativity discussed in a previous lecture.

2.2 Convex hull of the moment curve

An appealing geometric interpretation of the set of valid moments is in terms of the so-called *moment curve*, which is the parametric curve in $\mathbb{R}^{d+1}$ given by $t \mapsto (1, t, t^2, \ldots, t^d)$. Indeed, it is easy to see that every point on the curve can be associated to a Dirac measure where all the probability is concentrated on a given point, since if the pdf of X is $\delta(x-a)$, then

$$\mu_k = \mathbf{E}[X^k] = a^k.$$

Thus, every finite (or infinite) measure on the interval corresponds to a point in the convex hull. In Figure 2 we present an illustration of the set of valid moments, for the case $d = 3$.

Remark 4. *Consider again the situation described in Remark 2. By restricting to even measures (i.e., $\mu(x) = \mu(-x)$), we can see that the moment curve corresponds to the half-parabola $C = \{(a, b) : a \geq 0, a^2 = b\}$. The point $(0, 1)$ is in the closure of* $\operatorname{conv}(C)$*, but it is not in* $\operatorname{conv}(\mathrm{C})$ *(since the convex hull is not closed). Notice that this situation cannot happen if the measure is supported on a compact set, since in this case the set of valid moments is always compact.*

3 Bridging the gap

What to do in the cases where the set of nonnegative polynomials is no longer equal to the SOS ones? As we will see in much more detail later, it turns out that we can approximate *any* semialgebraic problem (including the simple case of a single polynomial being nonnegative) by sum of squares techniques.

As a preview, and a hint at some of the possibilities, let's consider how to prove nonnegativity of a particular polynomial which is not a sum of squares. Recall that the Motzkin polynomial was defined as:

$$M(x, y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2.$$

and is a nonnegative polynomial that is not SOS. We can try multiplying it by another polynomial which is known to be positive, and check whether the resulting product is SOS. For instance, for the Motzkin example, multiplying by the factor $(x^2 + y^2)$ we can find the decomposition

$$(x^2 + y^2) \cdot M(x, y) = y^2(1 - x^2)^2 + x^2(1 - y^2)^2 + x^2y^2(x^2 + y^2 - 2)^2,$$

which clearly certifies that $M(x, y) \geq 0$.

More details will follow...

References

[Akh65] N. I. Akhiezer. *The classical moment problem.* Hafner Publishing Company, New York, 1965.

[AKP11] A. A. Ahmadi, M. Krstic, and P. A. Parrilo. A globally asymptotically stable polynomial vector field with no polynomial Lyapunov function. In *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 7579–7580. IEEE, 2011.

[DPS02] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing separable and entangled states. *Physical Review Letters*, 88(18), 2002.

[DPS04] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Complete family of separability criteria. *Physical Review A*, 69:022308, 2004.

[Gur03] L. Gurvits. Classical deterministic complexity of Edmonds' problem and quantum entanglement. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 10–19, New York, NY, USA, 2003. ACM.

[Kha92] H. Khalil. *Nonlinear Systems.* Macmillan Publishing Company, 1992.

[Par00] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, May 2000.

[PP02] A. Papachristodoulou and S. Prajna. On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings of the 41^{th} IEEE Conference on Decision and Control*, 2002.

[PPP05] S. Prajna, A. Papachristodoulou, and P. A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, 2002-05. Available from `http://www.cds.caltech.edu/sostools` and `http://www.mit.edu/~parrilo/sostools`.