Lecture 11

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

In this lecture we continue our study of SOS polynomials. After presenting a couple of applications, we focus here on the dual side, and provide a natural probabilistic interpretation of the corresponding problem. We further present some recent results on the density of the cone of SOS polynomials relative to that of the nonnegative polynomials.

1   SOS applications

1.1   Lyapunov functions

The possibility of reformulating conditions for a polynomial to be a sum-of-squares as an SDP is very useful, since we can use the SOS property in a control context as a convenient sufficient condition for polynomial nonnegativity. Recent work has applied the sum-of-squares approach to the problem of finding a Lyapunov function for nonlinear systems [Par00, PP02]. This approach allows one to search over affinely parametrized polynomial or rational Lyapunov functions for systems with dynamics of the form

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

where the functions f_i are polynomials or rational functions. 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:

$$\dot{x} = -x + (1 + x)y$$
$$\dot{y} = -(1 + x)x.$$

Using SOSTOOLS [PPP05] we easily find a quartic polynomial Lyapunov function, which after rounding (for purely cosmetic reasons) is given by

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

It can be readily verified that both $V(x, y)$ and $(-\dot{V}(x, y))$ are SOS, since

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

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.

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 $\rho \in \mathcal{S}_+^{n_1 n_2}$.

How to recognize if a state is entangled or not?

<div style="border:1px solid">Complete</div>   ToDo

2   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.

$$\mu_k := E[X^k] = \int x^k d\mu \tag{1}$$

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?

It should be apparent that some conditions are required. For instance, consider (1) 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., $\text{var}(X) = E[X^2] - E[X]^2 = \mu_2 - \mu_1^2$. 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,

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

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

Exactly the same procedure can be done for higher-order moments. Proceeding this way, we have that the higher order moments must always satisfy:

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

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

<div style="border:1px solid">Necessary and sufficient, multivariate case</div>   ToDo

Remark 1. *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 a necessary and sufficient characterization for moments. For simplicity, we present below only one particular case, corresponding to the interval $[-1, 1]$.

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

$$\begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_d \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{d+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_d & \mu_{d+1} & \mu_{d+2} & \cdots & \mu_{2d} \end{bmatrix} \pm \begin{bmatrix} \mu_1 & \mu_2 & \mu_4 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_5 & \cdots & \mu_{d+2} \\ \mu_3 & \mu_4 & \mu_6 & \cdots & \mu_{d+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{d+1} & \mu_{d+2} & \mu_{d+3} & \cdots & \mu_{2d+1} \end{bmatrix} \succeq 0. \tag{3}$$

Notice that the necessity is clear, since it follows from consideration of the quadratic form (in the a_i):

$$0 \leq E\left[(1 \pm X)(\sum_{i=0}^{d} a_i X^i)^2\right] = \sum_{j=0}^{d}\sum_{k=0}^{d}(\mu_{j+k} \pm \mu_{j+k+1})a_j a_k,$$

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.

2.2 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 corresponds to a Dirac measure, where all the probability is concentrated on a given point. Thus, every finite (or infinite) measure on the interval corresponds to a point in the convex hull. In Figure 1 we present an illustration of the set of valid moments, for the case $d = 3$.

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:

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

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. In this case, multiplying by the factor $(x^2 + y^2)$, we can find the decomposition

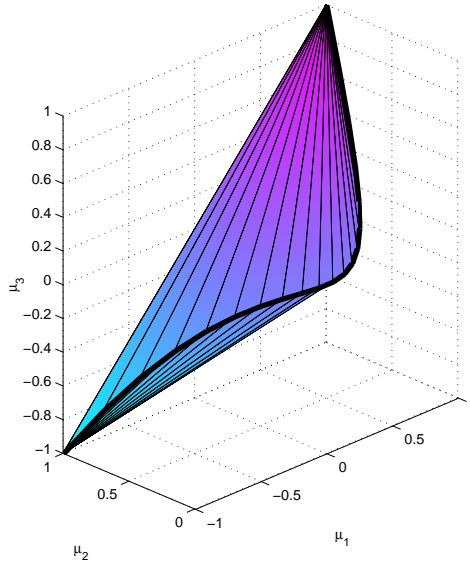$$(x^2 + y^2) \cdot M(x, y) = y^2(1 - x^2)^2 + x^2(1 - y^2)^2 + x^2 y^2(x^2 + y^2 - 2)^2,$$

Figure 1: 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 (3).

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

More details will follow...

References

[Par00] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, May 2000. Available at http://resolver.caltech.edu/CaltechETD:etd-05062004-055516.

[PP02] A. Papachristodoulou and S. Prajna. On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings of the 41th 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.