Math 591 – Real Algebraic Geometry and Convex Optimization

Lecture 8: Sums of squares and SDPs
Cynthia Vinzant, Spring 2019

Today we move into the 21st century with a nice connection between sums of squares and semidefinite programming. These connections were pioneered by Pablo Parrilo and Jean Bernard Lasserre in the early 2000's.

Let $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ and let $m_d \in (\mathbb{R}[x_1, \ldots, x_n]_{\leq d})^N$ denote the vector of all monomials in $x_1, \ldots, x_n$ of degree at most d.

Proposition. *f is a sum of squares if and only if there exists $A \in \mathrm{PSD}_N$ such that*

$$f = m_d^T A m_d.$$

Proof. Any polynomial $h \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d}$ has a unique representation as $h = v^T m_d$ where $v \in \mathbb{R}^N$. Using this and the fact that a matrix is positive semidefinite if and only if it can be written as a sum of rank one matrices we find that:

$$\begin{aligned} f = \sum_{i=1}^k h_i^2 \text{ for some } h_i \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d} \quad \Leftrightarrow \quad f &= \sum_{i=1}^k (v_i^T m_d)^2 \text{ for some } v_i \in \mathbb{R}^N \\ &= \sum_{i=1}^k m_d^T v_i v_i^T m_d \\ &= m_d^T \left(\sum_{i=1}^k v_i v_i^T \right) m_d \\ \Leftrightarrow \quad f &= m_d^T A m_d \text{ for some } A \in \mathrm{PSD}_N \end{aligned}$$

□

Note that for a given polynomial f, the condition $f = m_d^T A m_d$ affine linear constraints on the matrix A. Therefore testing whether or not f is a sum of squares is equivalent to testing whether or not the intersection of the PSD cone with an affine linear space is empty (i.e. whether a certain semidefinite program is feasible).

Example. ($n = 1$, $d = 2$) We have $m_d^T = \begin{pmatrix} 1 & x & x^2 \end{pmatrix}$ and $N = 3$. Consider $f = 2 + 2x^3 + x^4$. To test if f is a sum of squares we write the condition $f = m_d^T A m_d$:

$$\begin{aligned} f = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} &= \left\langle \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}, \begin{pmatrix} 1 & x & x^2 \\ x & x^2 & x^3 \\ x^2 & x^3 & x^4 \end{pmatrix} \right\rangle \\ &= a_{11} + 2a_{12}x + (a_{22} + 2a_{13})x^2 + 2a_{23}x^3 + a_{33}x^4 \end{aligned}$$

Matching up the coefficients of f with $m_d^T A m_d$ we find the affine-linear conditions:

$$a_{11} = 2, \quad 2a_{12} = 0, \quad a_{22} + 2a_{13} = 0, \quad 2a_{23} = 2, \quad \text{and } a_{33} = 1.$$

This gives as affine line in $\mathbb{R}^{3\times 3}_{\text{sym}}$. By parametrizing it, we find that f is a sum of squares if and only if there exists $a \in \mathbb{R}$ for which the matrix

$$A = \begin{pmatrix} 2 & 0 & -a \\ 0 & 2a & 1 \\ -a & 1 & 1 \end{pmatrix}$$

is positive semidefinite. Indeed, one can check that A is positive semidefinite if and only if $(-1+\sqrt{5})/2 \leq a \leq 1$. For example, choosing $a = 1$, we can decompose the matrix A as

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \end{pmatrix}.$$

This writes f as $(1+x)^2 + (-1+x+x^2)^2$.

Note that in addition to the different possibilities for a, there are different ways of decomposing the matrix A! For any 2×2 orthogonal matrix U, we have

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} U^T U \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$$

Taking the rows of $U \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$ as coefficient vectors would also give a representation of f as a sum of squares.

Corollary. $\mathrm{SOS}_{n,\leq 2d}$ *is the image of* PSD_N *under a linear projection (namely* $A \mapsto m_d^T A m_d$*).* $\mathcal{P}_{n,\leq 2d}$ *is the image of* PSD_N *under a linear projection for* $n = 1, 2d = 2$ *and* $(n, 2d) = (2, 4)$*.*

We can use this for polynomial optimization as well. Note that for $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$,

$$f^* = \min_{x \in \mathbb{R}^n} f(x) = \max_{c \in \mathbb{R}} c \quad \text{such that} \quad f - c \in \mathcal{P}_{n,\leq 2d}.$$

From this, we get the following immediate corollary:

Corollary. *For* $n = 1, 2d = 2$ *and* $(n, 2d) = (2, 4)$*, minimizing a polynomial* $f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq 2d}$ *is a semidefinite program. Namely*

$$f^* = \max_{c \in \mathbb{R}} c \quad \textit{such that} \quad f - c = m_d^T A m_d, \quad A \in \mathrm{PSD}_N.$$

Example. $(n = 1, 2d = 4)$ For $f = 2 + 2x^3 + x^4$, as above, we find that

$$\min_{x \in \mathbb{R}} f(x) = \max_{a,c \in \mathbb{R}} c \quad \text{such that} \quad \begin{pmatrix} 2-c & 0 & -a \\ 0 & 2a & 1 \\ -a & 1 & 1 \end{pmatrix}$$

We can check either by the usual tricks for optimizing univariate polynomials or with a semidefinite program solver that the maximum is attained for $c = 5/16$.

Of course, for any n, d, we could run this semidefinite program! Let's define

$$f^*_{sos} = \max_{c \in \mathbb{R}} c \quad \text{such that} \quad f - c \in \mathrm{SOS}_{n,\leq 2d}.$$

Note that if $f - c$ is a sum of squares, then it is nonnegative, meaning that $c \leq f^*$. Although the sum-of squares program may not give the tru minimum, this does give a bound:

$$f^*_{sos} \leq f^*.$$

Constrained polynomial optimization. One might also try to use sums of squares to find or approximate the minimum of a polynomial over a semialgebraic set. Consider

$$S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_s(x) \geq 0\}$$

where $g_1, \ldots, g_s \in \mathbb{R}[x_1, \ldots, x_n]$ and take $f \in \mathbb{R}[x_1, \ldots, x_n]$. Then

$$\begin{aligned} f^* = \min_{x \in S} \ f(x) \ &\geq \ \max \ c \quad \text{s.t. } f - c \in \mathrm{PO}(g_1, \ldots, g_s) \\ &\geq \ \max \ c \quad \text{s.t. } f - c = \sum_{\alpha \in \{0,1\}^s} \sigma_\alpha g_1^{\alpha_1} \cdots g_s^{\alpha_s} \text{ where } \sigma_\alpha \in \mathrm{SOS}_{n, \leq 2d} \quad := f^*_{sos,d}. \end{aligned}$$

For any fixed d, the last maximization problem can be written as a semidefinite program, assigning one positive semidefinite matrix for each sum of squares σ_α.

Example. Consider $S = [-1, 1] \subset \mathbb{R}$ and $f = 1 - 2x + 3x^2 - 2x^3$. We might try to find f^* by bounding is using this semidefinite program:

$$\begin{aligned} f^* = \min_{x \in \mathbb{R}} f(x) \ &\geq \ \max_{c \in \mathbb{R}} c \ \text{ s.t. } f - c = \sigma_0 + \sigma_1(1 - x^2) \ \text{ where } \ \sigma_0 \in \mathrm{SOS}_{1, \leq 4}, \ \sigma_1 \in \mathrm{SOS}_{1, \leq 2} \\ &= \max_{c \in \mathbb{R}} c \ \text{ s.t. } f - c = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} A \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} + \begin{pmatrix} 1 & x \end{pmatrix} B \begin{pmatrix} 1 \\ x \end{pmatrix} \cdot (1 - x^2) \\ &\qquad\qquad\qquad\qquad \text{where } A \in \mathrm{PSD}_3 \ \text{ and } B \in \mathrm{PSD}_2. \end{aligned}$$

Note that just like in the unconstrained case, the polynomial equality on $f - c$ gives affine-linear constraints on the entries of A and B. For example, if $A = (a_{ij})_{ij}$ and $B = (b_{ij})_{ij}$, then comparing the coefficient of x^4 on each side gives that

$$0 \ = \ a_{33} - b_{22}.$$

In total comparing the coefficients of $1, x, x^2, x^3, x^4$ gives five affine conditions on the entries of A, B, and $c \in \mathbb{R}$. In this case, it turns out that there is equality: $f^* = f^*_{sos,4} = 0$.

On the one hand, $f(1) = 0$, so $f^* \leq 0$. On the other hand,

$$f \ = \ \frac{1}{2}(1 - x)^4 + \frac{1}{2}(1 + x^2)(1 - x^2),$$

which shows that $f^* \geq f^*_{sos,4} \geq 0$.

In fact, general speaking, things behave nicely for univariate polynomials.

Theorem. *For $a, b \in \mathbb{R}$, a polynomial $f \in \mathbb{R}[x]$ is nonnegative on the interval $[a, b]$ if and only if*

$$f(x) = \begin{cases} \sigma_0 + (x - a)(b - x)\sigma_1 \ \textit{where } \sigma_0 \in \mathrm{SOS}_{1, \leq 2d}, \sigma_1 \in \mathrm{SOS}_{1, \leq 2d-2} & \textit{if } \deg(f) \leq 2d \\ (x - a)\sigma_0 + (b - x)\sigma_1 \ \textit{where } \sigma_0, \sigma_1 \in \mathrm{SOS}_{1, \leq 2d}, & \textit{if } \deg(f) \leq 2d + 1. \end{cases}$$

This says that not only does every univariate polynomial nonnegative on $[a, b]$ have a representation using sums of squares, but that the certificate has the lowest degree possible.

For example, there is an even lower degree certificate that the univariate polynomial $f = 1 - 2x + 3x^2 - 2x^3$ is nonnegative on $[-1, 1]$, namely

$$f = (1 - x) \cdot \left(\left(\frac{1 - 4x}{\sqrt{8}} \right)^2 + \left(\sqrt{\frac{7}{8}} \right)^2 \right),$$

which shows that $f \geq 0$ on the larger set $\{x \in \mathbb{R} : x \leq 1\}$.

Another example of sums-of-squares representations doing well is when S is compact.

Theorem (Schmüdgen)**.** *If $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_s(x) \geq 0\}$ is compact and $f > 0$ on S, then $f \in \mathrm{PO}(g_1, \ldots, g_s)$.*

The difference between $f \geq 0$ and $f > 0$ here is important. We know that $f - f^*$ is nonnegative on S, so for any $\epsilon > 0$, $f - f^* + \epsilon$ is positive on S, meaning that there is a representation

$$f - f^* + \epsilon \;=\; \sum_{\alpha \in \{0,1\}^s} \sigma_\alpha g_1^{\alpha_1} \cdots g_s^{\alpha_s} \quad \text{where} \quad \sigma_\alpha \in \mathrm{SOS_n}$$

Therefore for high enough d (namely $d \geq \deg(\sigma_\alpha)/2$), we have $f^* \geq f^*_{sos,d} \geq f^* - \epsilon$.

Corollary. *If S is compact, then $f^*_{sos,d} \to f^*$ as $d \to \infty$.*

The caveat here is that the degrees of σ_α might need to grow arbitrarily large, and indeed this can happen when S is high dimensional.

In addition to univariate polynomials, one case where the degrees do not grow arbitrarily high, is when S is finite.

Theorem. *Let $S = \{p_1, \ldots, p_N\} \subset \mathbb{R}^n$. Then $f \in \mathbb{R}[x_1, \ldots, x_n]$ is nonnegative on S if and only if*

$$f \;=\; \sigma + g \quad \textit{where } \sigma \in \mathrm{SOS}_{n,\leq 2N} \quad \textit{and} \quad g \in \mathcal{I}(S).$$

Proof. Find polynomials $h_1, \ldots, h_N$ of degree $\leq N$ so that $h_i(p_j)$ is 1 if $i = j$ and 0 otherwise. Note that h_i^2 has the same evaluations at these points. Consider the polynomial

$$g \;=\; f - \sum_{i=1}^{N} f(p_i) h_i^2.$$

One can check that $g(p_j) = 0$ for all $j = 1, \ldots, N$, meaning that $g \in \mathcal{I}(S)$. Then

$$f \;=\; \sum_{i=1}^{N} f(p_i) h_i^2 + g$$

is the promised representation. □