15 Sum-of-squares hierarchies

Application: Dynamical systems and Lyapunov functions Consider a dynamical system

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x})$$

where f is a polynomial. Assume that the origin $\mathbf{x} = 0 \in \mathbb{R}^n$ is an equilibrium of the system, i.e., $f(0) = 0$. We would like to understand whether all the trajectories $\mathbf{x}(t)$ converge to 0 as $t \to \infty$. One way to check this is to find a Lyapunov function, which is a positive energy function that decreases along trajectories, i.e., $V(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$ and $\frac{d}{dt}V(\mathbf{x}(t)) < 0$. Note that

$$\frac{d}{dt}V(\mathbf{x}(t)) = \left\langle \frac{d}{dt}\mathbf{x}(t), \nabla V(\mathbf{x}(t)) \right\rangle = \langle f(\mathbf{x}(t)), \nabla V(\mathbf{x}(t))\rangle.$$

We can thus impose the following conditions on V:

$$\begin{cases} V(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \{0\} \\ \langle \nabla V(\mathbf{x}), f(\mathbf{x})\rangle < 0 \quad \forall \mathbf{x} \in \mathbb{R}^n \setminus \{0\}. \end{cases} \tag{1}$$

The second condition ensures that the value $V(\mathbf{x}(t))$ decreases along trajectories. If we assume V to be a polynomial then the conditions (1) are polynomial positivity conditions. Consider the following sum-of-squares relaxation:

$$\text{Find polynomial } V(x_1, \ldots, x_n) \text{ such that } \quad \begin{cases} V(\mathbf{x}) \text{ is a sum-of-squares} \\ -\langle \nabla V(\mathbf{x}), f(\mathbf{x})\rangle \text{ is a sum-of-squares.} \end{cases} \tag{2}$$

If we impose a bound on the degree of V, then solving (2) amounts to a semidefinite feasibility problem.

Motzkin polynomial We saw in the previous lecture that not all nonnegative polynomials are sums of squares. In particular we saw that the "minimal" cases where this happens is $(n, 2d) = (2, 6)$ and $(n, 2d) = (3, 4)$ where n is the number of variables and $2d$ is the degree. We are now going to look at a concrete example of polynomial in the case $(n, 2d) = (2, 6)$ that is nonnegative but not a sum-of-squares. Consider the *Motzkin polynomial* defined by:

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

One can show that $M(x, y)$ is nonnegative via the arithmetic-geometric mean inequality. Indeed we have, for any $x, y \in \mathbb{R}$

$$\frac{1}{3}(x^4y^2 + x^2y^4 + 1) \geq (x^6y^6)^{1/3} = x^2y^2.$$

On the other hand one can show that $M(x, y)$ is not a sum of squares. In fact one can prove even more generally that $M(x, y) - \gamma$ is not a sum-of-squares for any $\gamma \in \mathbb{R}$.

Proposition 15.1. *$M(x, y) - \gamma$ is not a sum of squares for any $\gamma \in \mathbb{R}$.*

Proof. This proof is based on [Lau09, Example 3.7]. Assume $M(x,y)-\gamma = \sum_k q_k^2$ where $q_k(x,y) = a_k x^3 + b_k y^3 + c_k x^2 y + d_k x y^2 + e_k x^2 + f_k y^2 + g_k xy + h_k x + i_k y + j_k$. Since the coefficient of x^6 in $M-\gamma$ is zero we get $\sum_k a_k^2 = 0$ i.e., $a_k = 0$ for all k. Similarly we get $b_k = 0$ for all k. The coefficient of x^4 in $M-\gamma$ is also zero and so we now get $\sum_k a_k h_k + e_k^2 = 0$ which yields $e_k = 0$ for all k since we have $a_k = 0$. Similarly by looking at the coefficient of y^4 we get $f_k = 0$. Now looking at the coefficient of x^2 we get $\sum_k e_k j_k + h_k^2 = 0$ which again yields $h_k = 0$ for all k. Similarly by looking at the coefficient of y^2 we get $i_k = 0$ for all k. Finally our polynomials q_k must look like $q_k = c_k x^2 y + d_k x y^2 + g_k xy + j_k$. Now looking at the coefficient of x^2y^2 we get that $-3 = \sum_k g_k^2$ which is impossible. □

Sum of squares hierarchy Even though $M(x,y)$ is not a sum-of-squares it turns out that the polynomial $(1+x^2+y^2)M(x,y)$ is a sum-of-squares. Indeed one can verify that

$$\begin{aligned}(1+x^2+y^2)M(x,y) \;&=\; y^2(1-x^2)^2 + x^2(1-y^2)^2 + (x^2y^2-1)^2 \\ &\quad + x^2y^2\Big(\frac{3}{4}(x^2+y^2-2)^2 + \frac{1}{4}(x^2-y^2)^2\Big).\end{aligned} \tag{3}$$

The previous equation clearly shows that $M(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^2$.

If we are interested in minimising a polynomial $p(\mathbf{x})$ we can thus define the following *sum-of-squares hierarchy*:

$$v_r \;:=\; \max \quad \gamma \quad : \quad (1+x_1^2+\cdots+x_n^2)^r(p(\mathbf{x})-\gamma) \text{ is a sum-of-squares.} \tag{4}$$

For the Motzkin polynomial we know that $v_0 = -\infty$ and $v_1 = 0 = \min M(x,y)$. In general the sequence (v_r) is monotonic nondecreasing and satisfies

$$v_0 \leq v_1 \leq v_2 \leq \cdots \leq \min_{\mathbf{x}\in\mathbb{R}^n} p(x).$$

Indeed $v_r \leq v_{r+1}$ because if for some $\gamma \in \mathbb{R}$, $(1+x_1^2+\cdots+x_n^2)^r(p(\mathbf{x})-\gamma)$ is a sum-of-squares then $(1+x_1^2+\cdots+x_n^2)^{r+1}(p(\mathbf{x})-\gamma) = (1+x_1^2+\cdots+x_n^2)\cdot(1+x_1^2+\cdots+x_n^2)^r(p(\mathbf{x})-\gamma)$ is a sum-of-squares as a product of two sums of squares. Also $v_r \leq \min p(\mathbf{x})$ for any r because if $(1+x_1^2+\cdots+x_n^2)^r(p(\mathbf{x})-\gamma)$ is a sum-of-squares then this means that $p(\mathbf{x})-\gamma \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and so in particular $\min p(x) \geq \gamma$.

A natural question is to ask whether the sequence v_r converges to the minimum of p. Some results can be used to prove this under some conditions on p, like for example the following theorem of Reznick stated for homogeneous polynomials (a homogeneous polynomial of degree $2d$ is a polynomial only involving monomials of degree exactly $2d$):

Theorem 15.1 (Reznick, [Rez95])**.** *Assume $p \in \mathbb{R}[x_0,\ldots,x_n]$ is a homogeneous polynomial of degree $2d$ such that $p(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^{n+1} \setminus \{0\}$. Then there exists $r \in \mathbb{N}$ such that $(x_0^2+x_1^2+\cdots+x_n^2)^r p(\mathbf{x})$ is a sum of squares.*

Exercise 15.1. *Let $p \in \mathbb{R}[\mathbf{x}]$ be a polynomial of degree $2d$ (not necessarily homogeneous) with $\min_{\mathbf{x}\in\mathbb{R}^n} p(\mathbf{x}) = 0$. Using Theorem 15.1, show that if there is a constant $\epsilon > 0$ such that $p(\mathbf{x}) - \epsilon(\sum_{i=1}^n x_i^2)^d \geq 0$ for all $\mathbf{x}$, then the sequence (v_r) defined in (4) converges to $0 = \min p(\mathbf{x})$ as $r \to \infty$.*

Exercise 15.2. *We have been using the notation* $\min p(\mathbf{x})$ *throughout however strictly speaking it is possible for a polynomial not to attain its infimum. Show that* $p(x, y) = x^2 + (1 - xy)^2$ *is one such polynomial.*

Note that if p is a nonnegative polynomial, then expressing $(1 + x_1^2 + \cdots + x_n^2)p(\mathbf{x})$ as a sum of squares amounts to writing p as a sum of squares of *rational functions*. Hilbert's 17th problem asks whether any nonnegative polynomial can be written as a sum of squares of rational functions. This question was answered positively first by Artin in 1927. See [Rez00] for more on this question.

References

[Lau09] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, pages 157–270. Springer, 2009. 2

[Rez95] Bruce Reznick. Uniform denominators in Hilbert's seventeenth problem. *Mathematische Zeitschrift*, 220(1):75–97, 1995. 2

[Rez00] Bruce Reznick. Some concrete aspects of hilbert's 17th problem. *Contemporary Mathematics*, 253:251–272, 2000. 3