14 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

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

To find a Lyapunov function we can thus search for a function V that satisfies:

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

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:

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

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

A hierarchy of relaxations for polynomial optimization Let $p(\mathbf{x})$ be a polynomial and consider the problem of minimizing p on $\mathbb{R}^n$. We saw last time that we can formulate the following semidefinite program:

$$\max \ \gamma \ : \ p(\mathbf{x}) - \gamma \text{ is a sum-of-squares.} \tag{3}$$

If we call v_0 the value of (3) then it is clear that $v_0 \leq \min_{\mathbf{x}\in\mathbb{R}^n} p(\mathbf{x})$. In the case of univariate polynomials we have equality because nonnegative polynomials are sums of squares. However in general we can have a strict inequality.

Recall the Motzkin polynomial from last lecture:

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

We saw that M is nonnegative but $M - \gamma$ is not a sum-of-squares for any γ. This means that $v_0 = -\infty$! The lower bound obtained from the semidefinite relaxation is not useful.

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

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

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*:

$$v_r := \max \quad \gamma \quad : \quad (1 + x_1^2 + \cdots + x_n^2)^r (p(\mathbf{x}) - \gamma) \text{ is a sum-of-squares.} \tag{5}$$

It is not hard to show that the sequence (v_r) is monotonic nondecreasing and satisfies

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

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 $\min p(x) \geq \gamma$.

Note that one can define another hierarchy of semidefinite relaxations where the multiplier $(1 + x_1^2 + \cdots + x_n^2)^r$ is replaced by another nonnegative polynomial. This will yield in general a *different* hierarchy.

A natural question is to ask whether the sequence v_r defined in (5) 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 14.1 (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.*

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.

Constrained polynomial optimisation Let $p(\mathbf{x})$ be a polynomial and consider the problem of deciding nonnegativity of p on a set

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_k(\mathbf{x}) \geq 0\} \tag{6}$$

where $g_1, \ldots, g_k$ are polynomials. One way to certify that $p \geq 0$ on S is to write

$$p = s_0 + s_1 g_1 + \cdots + s_k g_k \tag{7}$$

where $s_0, s_1, \ldots, s_k$ are sums of squares. More generally one can consider certificates of the form:

$$p = s_0 + \sum_{\substack{I \subseteq \{1, \ldots, k\} \\ I \neq \emptyset}} s_I \prod_{i \in I} g_i \tag{8}$$

where the s_I (for $I \subseteq [n]$, $I \neq \emptyset$) are sums of squares. Do such representations exist, for any nonnegative polynomial p on S? The answer in general is no. However under some mild conditions on p and the description of S, one can guarantee the existence of such a representation. This is the content of so-called *Positivstellensatz* results:

Theorem 14.2 (Schmüdgen Positivstellensatz)**.** *Assume S defined in (6) is compact and p a polynomial such that $p(\mathbf{x}) > 0$ for all $\mathbf{x} \in S$. Then there exist s_0 and s_I (for $I \subseteq [k]$, $I \neq \emptyset$) sums of squares, such that (8) holds.*

Theorem 14.3 (Putinar's Positivstellensatz)**.** *Let S as defined in (6) and assume there exists $i \in \{1, \ldots, k\}$ such that $\{\mathbf{x} : g_i(\mathbf{x}) \geq 0\}$ is compact. Assume p is a polynomial such that $p(\mathbf{x}) > 0$ for all $\mathbf{x} \in S$. Then there exist $s_0, \ldots, s_k$ sums of squares such that (7) holds.*

For any integer r consider the following optimisation problem:

$$v_r := \max \quad \gamma \quad : \quad p - \gamma = s_0 + s_1 g_1 + \cdots + s_k g_k, \quad \text{where } \begin{cases} s_0, s_1, \ldots, s_k \text{ sums of squares} \\ \deg(s_0) \leq 2r \\ \deg(s_i g_i) \leq 2r \ \forall i = 1, \ldots, k. \end{cases}$$

For each r, the optimisation problem defining v_r can be formulated as a semidefinite program. Furthermore the sequence (v_r) is monotone nondecreasing and we have $v_r \leq \min_{x \in S} p(x)$ for all r. If we assume there is i such that $\{\mathbf{x} : g_i(x) \geq 0\}$ is compact, then Putinar's Positivstellensatz tells us that $v_r \to \min_{x \in S} p(x)$ as $r \to \infty$.

Remark 1. *One can also define another hierarchy using the Schmüdgen type representation (8). The resulting semidefinite programs however will be much bigger since we need to search over 2^k sum-of-squares polynomials (rather than just $k+1$ in the Putinar representation (7)).*

There are also other Positivstellensatz that do not rely on sums of squares. One such result is Polya's theorem.

Theorem 14.4 (Polya's theorem)**.** *Let $p(\mathbf{x})$ be a homogeneous polynomial in n variables and assume that $p(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbb{R}^n_+ \setminus \{0\}$. Then there exists an integer N such that the coefficients of the polynomial $(x_1 + x_2 + \cdots + x_n)^N p(\mathbf{x})$ are all nonnegative.*

Application to optimization: consider the problem of minimizing a homogeneous polynomial $p(\mathbf{x})$ of degree d on the unit simplex $\{\mathbf{x} \in \mathbb{R}^n_+ : x_1 + \cdots + x_n = 1\}$. One can define the following linear programming hierarchy:

$$v_N := \max \quad \gamma \quad : \quad (x_1 + \cdots + x_n)^N (p(\mathbf{x}) - \gamma \cdot (x_1 + \cdots + x_n)^d) \text{ has nonnegative coefficients}$$

Note that for any fixed N the coefficients of the polynomial

$$(x_1 + \cdots + x_n)^N (p(\mathbf{x}) - \gamma \cdot (x_1 + \cdots + x_n)^d)$$

are all linear in γ. This means that computing v_N amounts to solving a linear program. Polya's theorem guarantees that v_N converges to the minimum of p on the unit simplex as $N \to \infty$.

References

[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. 2