Sums of Squares

Lecture 1

Let $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ denote the polynomial ring in the indeterminates $x_1, \dots, x_n$ with real coefficients.

Let $\Sigma^2 \subset \mathbb{R}[x]$ the convex cone of Sums of Squares polynomials, that is:

$$\Sigma^2 := \left\{ \sigma \in \mathbb{R}[x] \;\middle|\; \exists r \in \mathbb{N} \;\; \exists h_1, \dots, h_r \in \mathbb{R}[x] \text{ s.t. } \sigma = h_1^2 + \dots + h_r^2 \right\}$$

Example $(n = 1)$

Let $f(x) = x^4 - 2x^3 + 2x^2 - 2x + 1 = (x^2 - x)^2 + (x - 1)^2$

- $f(x) \geq 0 \;\; \forall x \in \mathbb{R}$, since $f \in \Sigma^2$
- $f(x) = (1 \;\; x \;\; x^2) \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} (0, -1, 1) \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} +$

$$+ (1 \;\; x \;\; x^2) \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} (-1 \;\; 1 \;\; 0) \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} =$$

$$= (1 \;\; x \;\; x^2) \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}}_{\succeq 0} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} + (1 \;\; x \;\; x^2) \underbrace{\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\succeq 0} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} =$$

$$= (1 \;\; x \;\; x^2) \overbrace{\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}}^{\succeq 0} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}$$

Lemma

0) If $\sigma \in \Sigma^2$, $\sigma(x) \geq 0 \ \forall x \in \mathbb{R}^n$. ← This is why they are interesting!

1) If $\sigma = h_1^2 + \dots + h_r^2 \in \Sigma^2$ with $\max_i \deg h_i = d$, then $\deg \sigma = 2d$

2) Σ^2 is a pointed convex cone: $\Sigma^2 \cap -\Sigma^2 = \{0\}$ (in other words, Σ^2 contains no lines)

3) $\Sigma^2 - \Sigma^2 := \{\sigma_1 - \sigma_2 \in \mathbb{R}[x] \mid \sigma_1, \sigma_2 \in \Sigma^2\} = \mathbb{R}[x]$

Proof Exercise.

Proposition Let $f \in \mathbb{R}[x]$, $\deg f = 2d$, $N := \dim \mathbb{R}[x]_{\leq d}$ then $f \in \Sigma^2$ if and only if $\exists$ a positive semidefinite $G \in \mathbb{R}^{N \times N}$ (Gram matrix) such that, if $[x]_d \in \mathbb{R}^N$ denotes the vector of the monomial basis of $\mathbb{R}[x]_{\leq d}$,

$$f(x) = [x]_d^t \ G \ [x]_d .$$

Proof ($\Rightarrow$) As in the example.

($\Leftarrow$)

$$G = U^t U = \begin{pmatrix} | & & | \\ u_1 & \dots & u_N \\ | & & | \end{pmatrix} \begin{pmatrix} - & u_1 & - \\ & \vdots & \\ - & u_N & - \end{pmatrix}$$

$G \succeq 0$, Cholesky

$$\Rightarrow f(x) = [x]_d^t \begin{pmatrix} | & & | \\ u_1 & \dots & u_N \\ | & & | \end{pmatrix} \begin{pmatrix} - & u_1 & - \\ & \vdots & \\ - & u_N & - \end{pmatrix} [x]_d$$

$$= \left([x]_d^t \begin{pmatrix} | \\ u_1 \\ | \end{pmatrix}\right)^2 + \dots + \left([x]_d^t \begin{pmatrix} | \\ u_N \\ | \end{pmatrix}\right)^2$$

$\square$

Proposition The convex cone Σ^2 is closed in $\mathbb{R}[x]$, in the finest locally convex topology, i.e. $\Sigma^2_{2d,n}$ * is closed in $\mathbb{R}[x]_{\leq 2d}$ $\forall d$.

* See next page for def

Proof Let $N = \dim \mathbb{R}[x]_{\leq d}$. If $f \in \Sigma^2_{n,2d}$, then $f = p_1^2 + \dots + p_N^2$ (we can bound the number of addends by $N = N(d,n)$, see previous proposition) for some $p_i \in \mathbb{R}[x]_{\leq d}$.

Chose good norm $\|\cdot\|$ on $\mathbb{R}[x]$. Say for $p \in \mathbb{R}[x]$
$\|p\| := \int p \, d\mu$, where μ has normal (Gaussian) distribution. Up to scaling, assume WLOG that $\|f\| = 1$.

We have:

$$\|p_i\|^2 \leq \sum_{i=1}^N \|p_i\|^2 = \int \sum_{i=1}^N p_i^2 \, d\mu = \int f d\mu = \|f\| = 1.$$

thus each p_i lives in the (compact) unit ball of $(\mathbb{R}[x]_{\leq d}, \|\cdot\|)$.

If we pick a converging sequence $\{f_k\}_{k \in \mathbb{N}} \subset \Sigma^2_{2d,n}$, then writing $\forall k$ a SoS representation for f_k, each SoS addenda lives in a compact set, and thus we can extract a subsequence s.t. every SoS addenda converges. Thus $\lim\limits_{k \to \infty} f_k \in \Sigma^2_{2d,n}$, and $\Sigma^2_{2d,n} \subset \mathbb{R}[x]_{\leq 2d}$ is closed.

Question: If $f(x) \geq 0 \ \forall x \in \mathbb{R}^n$, then $f \in \Sigma^2$?

Theorem (Hilbert)

Denote: • $\Sigma^2_{2d,n}$ the SOS n-variate polynomials of deg $\leq 2d$

• $P_{2d,n}$ the n-variate polynomials of deg $\leq 2d$ that are non-negative on $\mathbb{R}^n$.

then $\Sigma^2_{2d,n} = P_{2d,n} \iff \big(n=1 \text{ or } d=1 \text{ or } (n,d)=(2,2)\big)$

Examples

- Every univariate non-negative polynomial can be written as a sum of 2 squares. (Ex.)
- Let $M(x,y) = 1 - 3x^2y^2 + x^2y^4 + x^4y^2$.
  - $M(x,y) \notin \Sigma^2$ (Ex.)
  - $(x^2+y^2)^2 M(x,y) = x^2y^2(1+x^2+y^2)(-2+x^2+y^2)^2 + (x^2-y^2)^2$

    $\Rightarrow M(x,y) \geq 0$

    (Alternative proof: AM-GM inequality, Ex.)
- **SoS optimization**: Consider $f \in \mathbb{R}[X]$ and

  $f^* := \inf\{f(x) : x \in \mathbb{R}^n\} = \sup\{\lambda \in \mathbb{R} : f(x) - \lambda \geq 0 \ \forall x \in \mathbb{R}^n\}$.

  then $f^*_{SOS} := \sup\{\lambda \in \mathbb{R} : f - \lambda \in \Sigma^2\} \leq f^*$.

We write SoS optimization as an SDP more explicitly. Write $f = [x]_d^t \, G \, [x]_d$, $G \succeq 0$.

Consider f as a 1×1 matrix with entries in $\mathbb{R}[x]$.

then :

$$f = tr(f) = tr([x]_d^t \, G [x]_d) = tr(G [x]_d [x]_d^t).$$

Write $A = [x]_d [x]_d^t = \sum_{|\alpha| \leq 2d} A_\alpha x^\alpha$

for some $A_\alpha \in \mathbb{R}^{N \times N}$.

$$\left(\text{e.g.:} \quad \begin{pmatrix} 1 \\ x \end{pmatrix} (1, x) = \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A_0} \cdot 1 + \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{A_1} x + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{A_2} x^2 \right)$$

Therefore:

$$f = \sum_{|\alpha| \leq d} tr(G A_\alpha) x^\alpha \qquad \left(f_0 = tr(G A_0) \right)$$

and :

$$f^*_{SoS} - f_0 = \sup_\lambda \lambda - f_0 \quad \text{s.t.} \quad f - \lambda \in \Sigma^2_{n,2d} =$$

$$= \sup -tr(G A_0) \quad \text{s.t.}$$
$$G \succeq 0, \; tr(G A_\alpha) = f_\alpha \quad \forall \, 0 \neq |\alpha| \leq 2d.$$

$$= \sup \langle G, -A_0 \rangle \quad \text{s.t.}$$
$$G \succeq 0, \; \langle G, A_\alpha \rangle = f_\alpha \quad \forall \, 0 \neq |\alpha| \leq 2d.$$

(A_α, f_α for $|\alpha| \leq 2d$: data of the SDP. $G \succeq 0$: variable)

Representation theorems

Since $\Sigma^2_{2d,n} \neq P_{2d,n}$, we cannot certify always non-negativity using an SoS representation.

Question (Hilbert's 17th problem):
Can we always represent non-negative polynomials using ratios of SoS? Answer: yes!

Theorem (Artin) If $f \in \mathbb{R}[x]$ is ≥ 0 on $\mathbb{R}^n$, $\exists\, p_1, \dots, p_m, q_1, \dots, q_m \in \mathbb{R}[x]$ s.t.

$$f = \left(\frac{p_1}{q_1}\right)^2 + \dots + \left(\frac{p_m}{q_m}\right)^2.$$

Remarks

- Unknown denominator(s) $\rightarrow$ difficult to write as an SDP.
- Unknown degrees: the degrees of p_i, q_j can be higher than $\deg f$.

Theorem (Reznick) If $f \in \mathbb{R}[x]$ is homogeneous and $f(x) > 0$ for $x \neq 0$, $d = \deg f$, and denoting

$$\varepsilon(f) = \frac{\inf\{f(u) : u \in S^{n-1}\}}{\sup\{f(u) : u \in S^{n-1}\}} > 0,$$

then $(x_1^2 + \dots + x_n^2)^r f \in \Sigma^2$ if $r \geq \dfrac{n d(d-1)}{(4\log 2)\,\varepsilon(f)} - \dfrac{n+d}{2}$

Other kind of effective representation theorems:

Theorem (Polyà, Powers, Reznick) If:

- $f \in \mathbb{R}[X]$ is homogeneous, $d = \deg f$;
- $f(x) > 0$ for $x \in \Delta_n = \{x \in \mathbb{R}^n : x_i \geq 0, \sum_i x_i = 1\}$
- $f = \sum_{|\alpha| = d} c(\alpha) \frac{d!}{\alpha_1! \dots \alpha_n!} X^\alpha$, $\|f\| := \max_\alpha |c(\alpha)|$;
- $\varepsilon(f) := \frac{\inf \{f(u) : u \in \Delta_n\}}{\|f\|}$;

Then $(x_1 + \dots + x_n)^r f$ has positive coefficients if $r > \frac{d(d-1)}{2\varepsilon(f)} - d$

For representations of polynomials over $\mathbb{C}$, the parameter $\varepsilon(f)$ does not appear in the complexity estimates! $\longrightarrow$ More about $\varepsilon(f)$ in the upcoming lectures.