Low-rank Sum of Squares has no spurious local minima

Sum of squares : $p(x) = \sum_{i=1}^{r} u_i(x)^2$

- certifies that $p(x) \geq 0$
- enables poly. optimization $\max \gamma$ s.t. $p(x) - \gamma \geq 0$
- Applications in control, comb. opt., signal processing

How to solve?

Semidefinite programming

$\langle A_i, X \rangle = b_i$, $X \succeq 0$

$\downarrow$ SOS $X \in S^n$

$b(x) = \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ \vdots \end{bmatrix} \in \mathbb{R}^n$ $p(x) = b(x)^T Q\, b(x)$

$Q \succeq 0$

Linear constraints on PSD cone.

Typically solved with interior-point methods

Atleast $O(n^3)$ iteration complexity ($n \approx 1000$)

B-M method

$F(U) = \sum_i (\langle A_i, UU^T \rangle - b_i)^2$ $U \in \mathbb{R}^{n \times r}$

$\downarrow$ SOS

$F_p(\vec{u}) = \| \sum_{i=1}^{r} u_i(x)^2 - p(x) \|^2 \to$ inner product on polynomials

Solved with nonlinear optimization methods

$\nabla F(U)$ fast to compute when r small.

Nonconvex problem so might be stuck in local minima

Smaller r, more non-convex

When does second-order critical point $\Rightarrow$ global optimum?

$r \geq n$: always

$r \gtrsim \sqrt{m}$: almost always (m: no. of constraints)

$\exists$ instance that has spurious SOCP when $r = n-1$.

low rank: matrix completion/sensing, planted statistical problems ...

Fact: If univariate $p(x) \geq 0$, then $p(x) = u_1(x)^2 + u_2(x)^2$.

Thm: $\forall$ univariate $p(x) \geq 0$, $f_p(\vec{u}) = \| u_1^2 + u_2^2 - p \|^2$ has no spurious SOCPs. ($r=2$!)

+ $\nabla f_p(u)$ can be computed in $O(n \log n)$-time if suitable $\|\cdot\|^2$ chosen.

+ first-order method for univariate sos of million-degree polynomials (~30 mins)

Proof:

$$A_{\vec{u}}(\vec{v}) = \sum_{i=1}^{r} u_i(x) v_i(x), \quad \sigma(\vec{u}) = A_{\vec{u}}(\vec{u}) \text{ so } f_p(\vec{u}) = \| \sigma(\vec{u}) - p \|^2$$

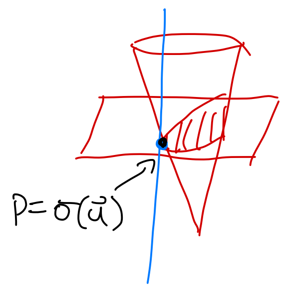
$$\left.\begin{aligned} \nabla f_p(\vec{u})(\vec{v}) &= \langle A_{\vec{u}}(\vec{v}), \sigma(\vec{u}) - p \rangle = 0 \quad \forall \vec{v} \\ \nabla^2 f_p(\vec{u})(\vec{v},\vec{v}) &= \langle \sigma(\vec{v}), \sigma(\vec{u}) - p \rangle + 2 \| A_{\vec{u}}(\vec{v}) \|^2 \geq 0 \quad \forall \vec{v} \end{aligned}\right\} \Rightarrow \vec{u} \text{ is SOCP}$$

Want to show: $\forall p, \forall \vec{u}, \ \nabla f_p(\vec{u}) = 0, \ \nabla^2 f_p(\vec{u}) \succeq 0 \Rightarrow f_p(\vec{u}) = 0 \qquad (*)$

Insight: for fixed $\vec{u}$, finding p counterexample to $(*)$ is convex problem

$\{ p \mid \nabla f_p(\vec{u}) = 0 \}$

$\{ p \mid \nabla^2 f_p(\vec{u}) \succeq 0 \}$



Certificate: $\forall \vec{u}$, find $\vec{\lambda}, \vec{v}_i$ s.t.

$$\nabla f_p(\vec{u})(\vec{\lambda}) + \sum_i \nabla^2 f_p(\vec{u})(\vec{v}_i, \vec{v}_i) = - f_p(\vec{u})$$

$$\Rightarrow -f_p(\vec{u}) \geq 0 \Rightarrow f_p(\vec{u}) = 0.$$

$$\nabla f_p(\vec{u})(\vec{\lambda}) = \langle u_1\lambda_1 + u_2\lambda_2,\ u_1^2+u_2^2-p\rangle$$

if $u_1(x)$ and $u_2(x)$ are coprime, then $\exists\, \lambda_1, \lambda_2$ s.t.

$$u_1\lambda_1 + u_2\lambda_2 = (\text{any polynomial}) = p - u_1^2 - u_2^2$$

$$\nabla f_p(\vec{u})(\vec{\lambda}) = -\|u_1^2+u_2^2-p\|^2 \quad \checkmark$$

On the other hand, suppose $u_1 = u_2$.

Let $p = \sum_i 2t_i^2$, $\vec{\lambda} = (-u_1, -u_2)$, $\vec{v}_i = (t_i, -t_i)$.

$$\nabla f_p(\vec{u})(\vec{\lambda}) = -\langle \sigma(\vec{u}), \sigma(\vec{u}) - p\rangle$$

$$\nabla^2 f_p(\vec{u})(\vec{v}_i, \vec{v}_i) = \langle 2t_i^2, \sigma(\vec{u}) - p\rangle + 2\|\cancel{u_1 t - u_2 t}\|^2 \to 0$$

$$(*) = \langle \underbrace{\sum_i 2t_i^2}_{p} - \sigma(\vec{u}), \sigma(\vec{u}) - p\rangle = -f_p(\vec{u}) \quad \checkmark$$

u_1 and u_2 share common factor: $u_1 = u_1' g$, $u_2 = u_2' g$.

$\forall w$, can always find $\vec{\lambda}$ s.t.

$$\vec{\nabla} f_p(\vec{u})(\vec{\lambda}) = \langle wg, \sigma(\vec{u}) - p\rangle$$

To make second term of Hesian zero, choose $\vec{v} = \begin{pmatrix} -u_2' t \\ u_1' t \end{pmatrix}$

$$\nabla^2 f_p(\vec{u})(\vec{v}, \vec{v}) = \langle t^2(u_1'^2 + u_2'^2), \sigma(\vec{u}) - p\rangle$$

Need to find w, t_i, s.t. $p = wg + \left(\sum_i t_i^2\right)\underbrace{(u_1'^2 + u_2'^2)}_{q}$

$$p \equiv Sq \pmod{g}$$

suppose $\gcd(q, g) = 1$, then $\exists\, a, b$ s.t. $aq + bg = 1$.

Lemma: if $a(x) > 0 \quad \forall x \in \{x \in \mathbb{R} \mid g(x) = 0\}$, then

$\exists \lambda, s$ s.t. $a = s^2 + \lambda g$.

$$s^2 q + \lambda' g = 1 \qquad \underbrace{s^2 p}_{\sum_i t_i^2} q + wg = p$$

Otherwise, $\gcd(q, g) = h$