Condition-based Low-Degree Approximation of Real Polynomial Systems

Josué TONELLI-CUETO
UT San Antonio
Elias TSIGARIDAS
Inria Paris & IMJ-PRG

Slides at:
https://tonellicueto.xyz/pdf/JMM2023_slides.pdf

WARNING

There will be one result
on Eigenvalue Computations,
the talk will focus
on real polynomial systems

PROBLEM

Given a real polynomial system

$$f_1(x) = \cdots = f_q(x) = 0$$

in n variables,

1) what can we say about the conditioning of solving?
2) what does the condition say about the zero set?

Conditioning

- The condition number depends on the metric
  — how we measure errors —
- The condition number depends on the encoding
  — how we write the problem —

Condition à la Demmel-Renegar — conic framework

$\mathcal{I}$ input space

$\Sigma \subseteq \mathcal{I}$ ill-posed inputs

$$C(i) := \frac{1}{d(i, \Sigma)}$$

WEYL

Setting

— projective

Set Up

$$\mathcal{H}_d := \prod_{i=1}^{n} \mathbb{R}[X_0, \dots, X_n]_d \,;$$

$\mathbb{P}^n$ Real Projective Space

$$\Sigma = \{ f \in \mathcal{H}_d \mid f \text{ has a singular zero} \}$$

Weyl Norm

$$\|g\|_W := \sqrt{\sum_{i=1}^{n} \sum_{|\alpha|=d_i} \binom{d_i}{\alpha}^{-1} |g_{i,\alpha}|^2}$$

where

$$g_i = \sum g_{i,\alpha} X^\alpha \qquad \left(X^\alpha := X_0^{\alpha_0} \cdots X_n^{\alpha_n} \right)$$

$$\binom{d_i}{\alpha} = \frac{d_i!}{\alpha_0! \cdots \alpha_n!}$$

Weyl Condition Number

(Cucker, Krick, Malajovich, Wschebor)

$$\kappa_W(g) := \frac{\|g\|_W}{\operatorname{dist}_W(g, \Sigma)}$$

$$= \sup_{x \in S^n} \frac{\|g\|_W}{\sqrt{\|g(x)\|^2 + \left\| D_x g^{-1} \Delta^{1/2} \right\|^{-2}}}$$

where $\Delta = \operatorname{diag}(d_i)$ & $D_x g : T_x S^n \to \mathbb{R}^n$

Main Theorem (Simple form)

$$\#Z(\mathcal{S}, \mathbb{P}^n) \leq D^{n/2}\, \mathrm{poly}(n, \log D)^n \log^n(2\, \mathcal{K}_w(\mathcal{S}))$$

where $D = \max d_i$

MAIN THEOREM (COMPLEX FORM)

There is a cover

$\{B(x, 1/c\sqrt{D})\}_{x\in\mathcal{G}}$ of size $\mathcal{O}(D^{n/2})$

of $\mathbb{S}^n$ s.t. for all $x\in\mathcal{G}$ & $\mathfrak{f}\in\mathcal{H}_d$,

there is $\Phi_{x,\mathfrak{f}}\in\mathbb{R}[X_0,\dots,X_n]^n$ of degree

$$\leq \mathrm{poly}(n, \log D)\log(2\mathcal{K}_w(\mathfrak{f}))$$

s.t. $\#\mathcal{Z}(\mathfrak{f}, T_x\mathbb{S}^n\cap B(x, 1/c\sqrt{D}))\leq\#\mathcal{Z}(\Phi_{x,\mathfrak{f}}, T_x\mathbb{S}^n)$.

Moreover, for all $\mathfrak{g}\in\mathcal{Z}(\mathfrak{f}, T_x\mathbb{S}^n\cap B(x, 1/c\sqrt{D}))$

there is $z\in\mathcal{Z}(\Phi_{x,\mathfrak{f}}, T_x\mathbb{S}^n)$ converging quadratically

under Newton's method.

Corollary 1 of Main Result

- If $\#Z(\mathcal{S}) \geq D^n$, then

$$\kappa_w(\mathcal{S}) \geq 2^{D^{n/2}/\mathrm{poly}(n,\log D)^n}$$

- Real systems with many zeros are badly-conditioned

COROLALLY 2 OF MAIN THEOREM

Let $f \in \mathcal{H}_d$ be random such that for all $x \in \mathbb{S}^n$, the $f_i(x)$ are independent, subgaussian and with anti-concentration. Then:

$$\left(\mathbb{E}\,\#Z(f,\mathbb{P}^n)^{\ell}\right)^{1/\ell} \leq D^{n/2}\,\mathrm{poly}(n,\log D)^n \ell^n$$

COROLALLY 2 OF MAIN THEOREM

Let $f \in \mathcal{H}_d$. Under very general random hypotheses,

$$\# Z(f, \mathbb{P}^n)^{1/n}$$

is subexponential with constant

$$D^{1/2}\,\mathrm{poly}(n, \log D)$$

I.e.

$$\mathbb{P}\left(\# Z(f, \mathbb{P}^n)^{1/n} \geq t\right) \leq e^{1 - \frac{t}{D^{1/2}\mathrm{poly}(n, \log D)}}$$

COROLALLY 2 OF MAIN THEOREM

A KSS random polynomial system

$$f \in \mathcal{H}_{(D,\dots,D)}$$

has its number of real zeros concentrated around

$$D^{n/2} = \mathbb{E}\,\#Z(f,\mathbb{P}^n)$$

COMPARISSON WITH LERARIO ET. AL.

OUR APPROACH	LERARIO'S APPROACH
Taylor Approx.	S. Harmonic Approx.
Many local approx.	A global approx.
Control the moments	Control only probability
Robust	Only KSS
Exploits analycity!	
(Lerario, Diatta; '22)	(Breiding, Keneshlou, Lerario; '22)

1-Norm Setting

Set Up

$$\mathcal{P}_A := \{ f \in \mathbb{R}[X_0, \dots, X_n] \mid 0 \in \operatorname{supp} f \subseteq A \}^n$$

$$I^n := [-1, 1]^n$$

$$\Sigma = \{ f \in \mathcal{P}_A \mid f \text{ has a singular zero in } I^n \}$$

1-Norm

$$\|g\|_1 := \max_i \sum_{\alpha \in A} |g_{i,\alpha}|$$

where

$$g_i = \sum g_{i,\alpha} X^\alpha \qquad \left(X^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n} \right)$$

Cubic Condition Number

(TC, Tsigaridas)

$$C_1(g) := \sup_{x \in I^n} \frac{\|g\|_1}{\max\{\|g(x)\|_\infty, \|D_x g^{-1} \Delta\|_{\infty,\infty}^{-1}\}}$$

$$\sim \frac{\|g\|_1}{\mathrm{dist}_1(g, \Sigma)}$$

where $\Delta = \mathrm{diag}(d_i)$ & $D_x g : \mathbb{R}^n \to \mathbb{R}^n$

Main Theorem (Simple form)

$$\# Z(g, I^n) \leq \mathrm{poly}(n)^n \log^{2n}(2D) \log^n(2C_1(g))$$

where $D = \max d_i$

MAIN THEOREM (COMPLEX FORM)

There is a partition into boxes

$$\{B\}_{B \in \mathcal{B}} \text{ of size } \mathcal{O}(\log^n(2D))$$

of I^n s.t. for all $B \in \mathcal{B}$ & $f \in \mathcal{H}_d$,

there is $\Phi_{B,f} \in \mathbb{R}[X_0, \dots, X_n]^n$ of degree

$$\leq \mathrm{poly}(n) \log D \log(2\, C_1(f))$$

s.t. $\# Z(f, B) \leq \# Z(\Phi_{B,f})$.

Moreover, for all $\zeta \in Z(f, B)$, there is $z \in Z(\Phi_{B,f})$, converging quadratically under Newton's method.

Corollary 1 of Main Result

- If $\#Z(\mathfrak{g}, I^n) \geq D^K$, then
$$C_1(\mathfrak{g}) \geq 2^{D^K / \mathrm{poly}(n, \log D)^n}$$
- Real systems with many zeros are badly-conditioned

COROLALLY 2 OF MAIN THEOREM

Let $f \in \mathcal{P}_A$ be random such that for all $x \in I^n$, the $f_i(x)$ are independent, subgaussian and with anti-concentration. Then:

$$\left(\mathbb{E} \, \# Z(f, I^n)^{\ell} \right)^{1/\ell} \leq \mathrm{poly}(n)^n \log^{2n}(2D)\, \ell^n$$

COROLALLY 2 OF MAIN THEOREM

Let $f \in \mathcal{P}_A$. Under very general random hypotheses,

$$\# Z(f, I^n)^{1/n}$$

is subexponential with constant

$$\log^2(2D)\,\mathrm{poly}(n)$$

I.e.

$$\mathbb{P}\left(\# Z(f, \mathbb{P}^n)^{1/n} \geq t\right) \leq e^{1 - \frac{t}{\log^2(2D)\,\mathrm{poly}(n)}}$$

The Details

Main Tricks

- Smale's α-theory is stable under analytic truncation
- Well-conditioned polynomials converge fast around zeros — as fast as a geometric series

Smale's α-Theory

$$\alpha(g,x) := \beta(g,x)\,\gamma(g,x)$$

$$\beta(g,x) := \| D_x g^{-1} g(x) \| = \| x - N_g(x) \|$$

$$\gamma(g,x) := \max\left\{ 1, \sup_{k \geq 2} \left\| D_x g^{-1} \frac{1}{k!} D_x^k g \right\|^{\frac{1}{k-1}} \right\}$$

Smales α-Theorem There is absolute $\alpha_* > 0$, s.t. if $\alpha(g,x) < \alpha_*$, then the Newton method at x converges quadratically. More concretely, $\operatorname{dist}(N_g^k(x), Z(g)) = O(2^{-2^k})$

Truncation Theorem (One version)

Let $g \in \mathbb{R}[X_1, \dots, X_n]^n$, $\delta \in \mathbb{N}$, $x \in \mathbb{B}^n$ &

$$T(g, x; \delta) := \sup_{k \geq \delta+1} \left\| D_x g^{-1} \frac{2^k}{k!} D_0^k g \right\|$$

Consider

$$g_{|\delta}(X) := \sum_{k=0}^{\delta} \frac{1}{k!} D_0^k g(X, \dots, X)$$

Then for

$$\delta - \log(\delta+2) \geq \log T(g, x; \delta),$$

$$\alpha(g_{|\delta}, x) \leq \frac{2\,\alpha(g, x) + 2^{1-\delta}\,\gamma(g, x)\,T(g, x; \delta)}{\left(1 - 2^{-\delta}(\delta+2)\,T(g, x; \delta)\right)^2}$$

I.e.

approximate zero of g à la Smale

$$\Downarrow$$

approximate zero of $g|_\delta$ à la Smale

+ reverse & more ineqs.

Moroz's Lemma

W-Lemma: For $g \in \mathbb{R}[X_0, X_1]_d$ & $(x_0, x_1) \in S^1$,

$$\left| \frac{1}{k!} \frac{d^k}{dt^k}\Big|_{t=0} g\big((x_0, x_1) + t(x_1, -x_0)\big) \right| \leq \sqrt{\binom{d}{k}} \, \|g\|_W$$

1-Lemma: For $g \in \mathbb{R}[X]_{\leq d}$, $a \in I$ and $\rho > 0$, if either $2|a| < 1 - \rho$ or $\rho < 1/2d$

then $$\left| \frac{1}{k!} \frac{d^k}{dt^k}\Big|_{t=0} g(a + \rho t) \right| \leq 1/2^k \, \|g\|_1$$

Multivariate Moroz's 1-Lemma

Let $g \in \mathbb{R}[X_1, \ldots, X_n]^n_{\leq D}$, $a \in I^n$ & $\rho \in (0,1]^n$

Consider

$$g_{a,\rho} := \left(g_i(a + PX) / \|g_i\|_1 \right)$$

where $P = \operatorname{diag}(\rho)$.

If for all i,

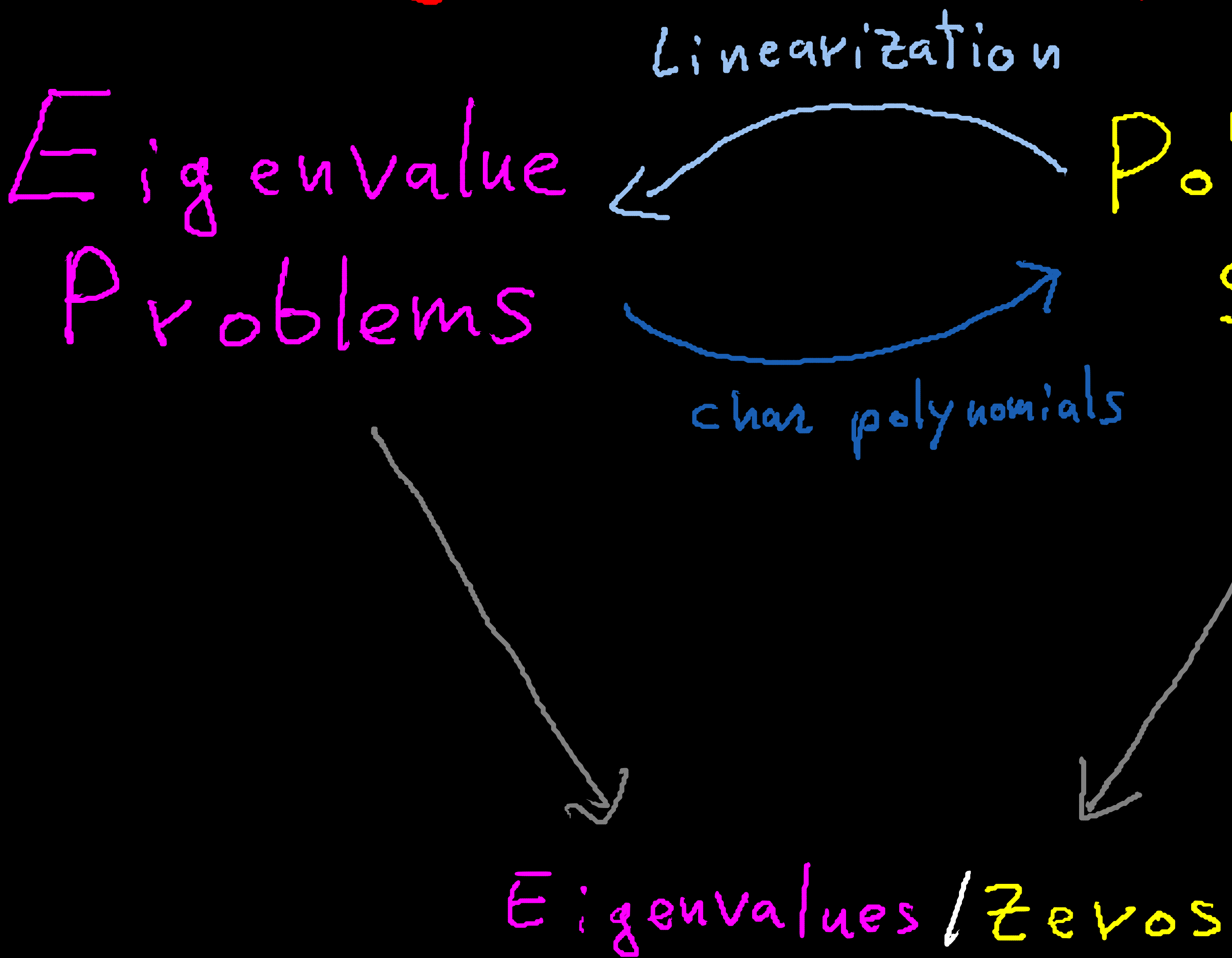
either $2|a_i| \leq 1 - \rho_i$ or $\rho_i \leq 1/2D$

then for all e,

$$\left\| \frac{2^e}{e!} D_0^e \, g_{a,\rho} \right\|_{\infty,\infty} \leq \binom{D+n-1}{n-1} \leq \left(1 + \frac{\ln(n-1)}{n-1} D\right)^{n-1}$$

A Theorem about Eigenvalues

Triangle of competition



Numerical Analist's Rule

NEVER USE CHARACTERISTIC POLYNOMIALS TO COMPUTE EIGENVALUES

A formalization for Hermitian Matrices

THM. Let $A \in \mathrm{Herm}_d$. Then

$$\kappa_w(\chi_A^h) \geq 2^{\sqrt{d}/\mathrm{polylog}(d)}$$

$$\& \quad C_1(\chi_A) \geq 2^{d/\mathrm{polylog}(d)}$$

I.e. characteristic polynomials of Hermitian matrices are badly conditioned.

Future Work

- Can we make all this into fast algorithms?
  — avoid condition estimation —
- Generalize it beyond zero-dim systems
  — volume & Betti numbers —

Thank You for your attention!

Imagen de Freepik