

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

$$g_1(x) = \dots = g_r(x) = 0$$

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

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

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

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

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

Weyl Condition Number

(Cucker, Krick, Malajovich, Wschebor)

$$\kappa_w(\gamma) := \frac{\|\gamma\|_w}{\text{dist}_w(\gamma, \Sigma)}$$

$$= \sup_{x \in S^n} \frac{\|\gamma\|_w}{\sqrt{\|\gamma(x)\|^2 + \|D_x \gamma^{-1} \Delta^{1/2}\|^{-2}}}$$

where $\Delta = \text{diag}(d_i)$ & $D_x \gamma: T_x S^n \rightarrow \mathbb{R}^n$

Main Theorem (Simple form)

$$\#Z(g, \mathbb{P}^n)$$

$$\leq D^{n/2} \text{poly}(n, \log D)^n \log^n(2K_w(g))$$

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

there is $\Phi_{x,g} \in \mathbb{R}[x_0, \dots, x_n]^n$ of degree

$$\leq \text{poly}(n, \log D) \log(2\kappa_w(g))$$

s.t. $\#z(g, T_x S^n \cap B(x, 1/c\sqrt{D})) \leq \#z(\Phi_{x,g}, T_x S^n)$.

Moreover, for all $g \in z(g, T_x S^n \cap B(x, 1/c\sqrt{D}))$ there is $z \in z(\Phi_{x,g}, T_x S^n)$ converging quadratically under Newton's method.

Corollary 1 of Main Result

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

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

• Real systems with many zeros
are badly-conditioned

COROLLARY 2 OF MAIN THEOREM

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

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

COROLLARY 2 OF MAIN THEOREM

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

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

is subexponential with constant $D^{1/2} \text{poly}(n, \log D)$

I.e.

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

COROLLARY 2 OF MAIN THEOREM

A KSS random polynomial system

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

has its number of real zeros
concentrated around

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

COMPARISON WITH LERARIO ET.AL.

OUR APPROACH

Taylor Approx.

Many local approx.

Control the moments

Robust

Exploits analyticity!

LERARIO'S APPROACH

S. Harmonic Approx.

A global approx.

Control only probability

only KSS

(Lerario, Diatta; '22) (Breiding, Keneshlou, Lerario; '22)

1-Norm

Setting

Set Up

$$\mathcal{P}_A := \{f \in \mathbb{R}[x_0, \dots, x_n] \mid 0 \in \text{supp } f \subseteq A\}^n$$

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

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

Weyl Norm

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

where

$$g_i := \sum_{\alpha} g_{i,\alpha} X^{\alpha} \quad (X^{\alpha} := X_1^{\alpha_1} \cdots X_n^{\alpha_n})$$

Cubic Condition Number

(TC, Tsigeridas)

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

$$\sim \frac{\|\mathcal{F}\|_1}{\text{dist}_1(\mathcal{F}, \Sigma)}$$

where $\Delta = \text{diag}(d_i)$ & $\mathcal{D}_x \mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^n$

Main Theorem (Simple form)

$$\#Z(\mathcal{F}, I^n) \leq \text{poly}(n)^n \log^{2n}(2D) \log(2C_1(\mathcal{F}))$$

where $D = \max d_i$

MAIN THEOREM (COMPLEX FORM)

There is a partition into boxes

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

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

there is $\Phi_{B,\mathbf{g}} \in \mathbb{R}[x_0, \dots, x_n]^n$ of degree

$$\leq \text{poly}(n) \log D \log(2C_1(\mathbf{g}))$$

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

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

Corollary 1 of Main Result

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

$$C_1(\mathcal{S}) \geq 2^{D^K / \text{poly}(n, \log D)^n}$$

• Real systems with many zeros
are badly-conditioned

COROLLARY 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:

$$\left(\mathbb{E} \#Z(f, I^n) e \right)^{1/e} \leq \text{poly}(n)^n \log^2 n (20) e^n$$

COROLLARY 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 $\log^2(2D) \text{poly}(n)$

I.e.

$$\mathbb{P}(\# z(f, I^n)^{1/n} \geq t) \leq e^{1 - \frac{t}{\log^2(2D) \text{poly}(n)}}$$

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

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

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

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

Smale's α -Theorem There is absolute $\alpha_* > 0$, s.t. if $\alpha(f, x) < \alpha_*$, then the Newton method at x converges quadratically. More concretely, $\text{dist}(N_f^k(x), z(f)) = \mathcal{O}(2^{-2^k})$

Truncation Theorem (One version)

Let $g \in \mathbb{R}[x_1, \dots, x_n]^n$, $\delta \in \mathbb{N}$, $x \in B^n$ &

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

Consider

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

Then for

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

$$\alpha(g|_{\delta}, x) \leq \frac{2 \alpha(g, x) + 2^{1-\delta} \alpha(g, x) \tau(g, x; \delta)}{(1 - 2^{-\delta} (\delta+2) \tau(g, x; \delta))^2}$$

I.e.

approximate zero of $\&$ à la Smale



approximate zero of $\& \mid \delta$ à la Smale

+ reverse & more ineqs.

Moroz's Lemma

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

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

1-Lemma: For $g \in \mathbb{R}[x]_{\leq d}$, $a \in \mathbb{I}$

and $\rho > 0$, if

either $2|a| < 1 - \rho$ or $\rho < \frac{1}{2}d$

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

Multivariate Moroz's 1-Lemma

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

Consider

$$g_{a, \rho} := (g; (a + \rho X) / \|g\|_1)$$

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

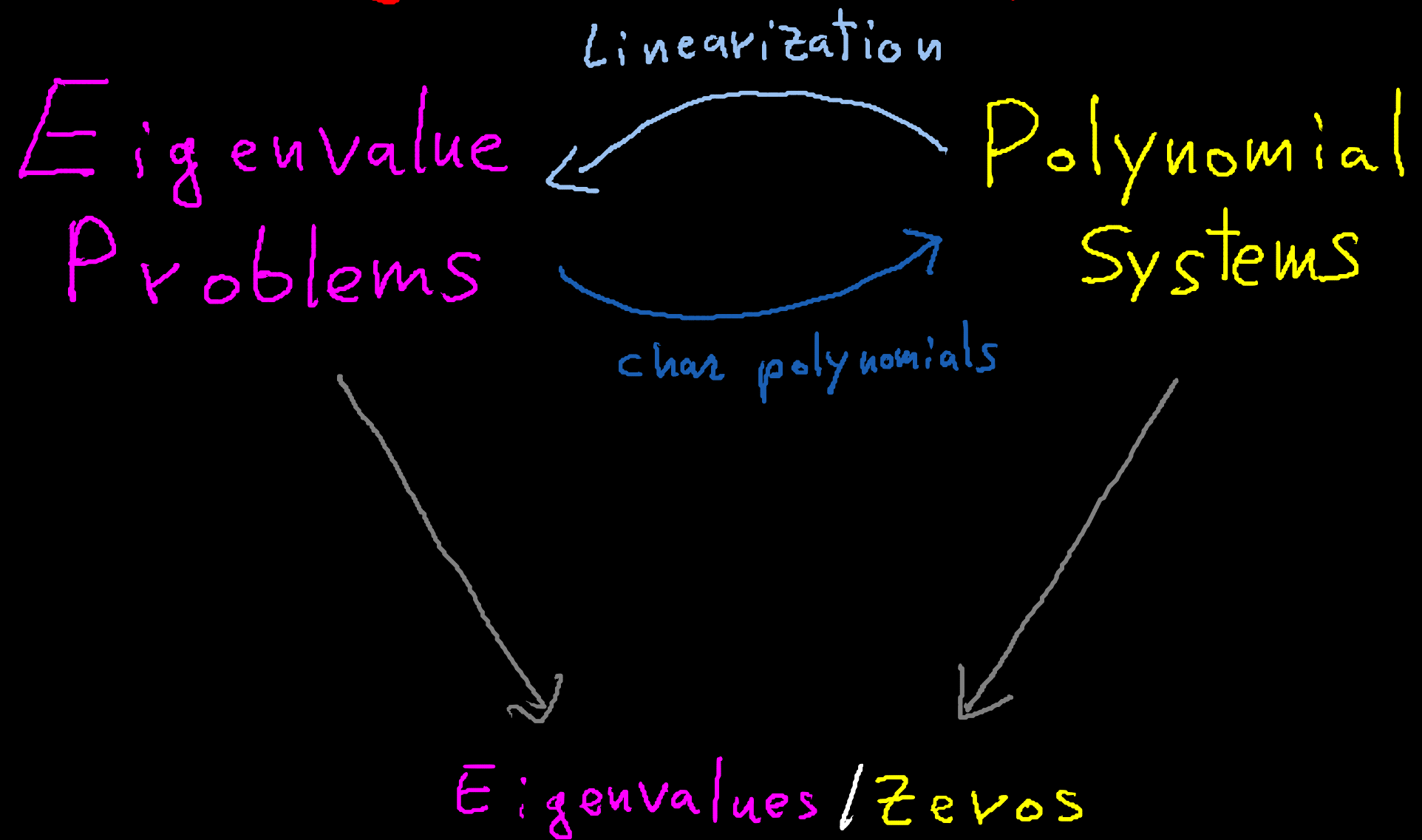
If for all i ,

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

then for all e ,

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

Triangle of competition



Numerical Analyst's Rule

NEVER USE
CHARACTERISTIC
POLYNOMIALS
TO COMPUTE
EIGENVALUES

A Formalization for Hermitian Matrices

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

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

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

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!

