Condition-based Bounds on the Number of $\mathbb{R}$eal Zeros

UTSA The University of Texas at San Antonio™

Josué Tonelli-Cueto
UT San Antonio
josue.tonelli.cueto@bizkaia.eu

Elias Tsigaridas
Inria Paris & IMJ-PRG
elias.tsigaridas@inria.fr

Inria

Condition Number

Let $f = (f_1, \dots, f_n)$ be a real polynomial system in n variables with f_k of degree at most d_k, its *condition number* is

$$\mathsf{C}(f) := \sup_{x \in [-1,1]^n} \frac{\|f\|}{\max\{\|f(x)\|_\infty, \|\mathrm{D}_x f^{-1}\Delta\|_{\infty,\infty}^{-1}\}}$$

where $\|f\| := \max_k \sum_\alpha |f_{k,\alpha}|$ is the 1-norm, $\|\ \|_\infty$ the ∞-norm, $\|\ \|_{\infty,\infty}$ the matrix norm induced by the ∞-norm and $\Delta := \operatorname{diag}(d_1, \dots, d_n)$.

Meaning? Measure of the numerical sensitivity of the real zeros of f with respect perturbations of f. It becomes ∞ when f has a singular zero in $[-1,1]^n$.

Geometric Interpretation

Discriminant Variety:

$$\Sigma := \{g \mid \text{there is } x \in [-1,1]^n \text{ s.t. } g(x) = 0,\ \det \mathrm{D}_x g = 0\}.$$

Condition Number Theorem

Let $f = (f_1, \dots, f_n)$ be a real polynomial system in n variables with f_i of degree at most d_i, then

$$\frac{\|f\|}{\operatorname{dist}(f,\Sigma)} \leq \mathsf{C}(f) \leq \left(1 + \max_k d_k\right) \frac{\|f\|}{\operatorname{dist}(f,\Sigma)}$$

where dist is the distance induced by $\|\ \|$.

A New $\mathbb{R}$eal Phenomenon!

Main Theorem (T.-C., Ts.; '23 +)

Let $f = (f_1, \dots, f_n)$ be a real polynomial system in n variables. Then

$$\#\mathcal{Z}(f, [-1,1]^n) \leq \mathcal{O}\left(\log \mathbf{D} \max\{n \log \mathbf{D}, \log \mathsf{C}(f)\}\right)^n$$

where $\mathbf{D}$ is the maximum degree.

Corollary:
Well-posed Real Polynomial Systems have few Real Zeros

Observation!
If $\#\mathcal{Z}(f, [-1,1]^n) \geq \Omega(\mathcal{D})$, then $\mathsf{C}(f) \geq 2^{\Omega\left(\frac{\mathcal{D}}{\log \mathbf{D}}\right)}$

Example: Hermitian Matrices

The Question: Given an Hermitian matrix $A \in \mathbb{C}^{d\times d}$, its characteristic polynomial

$$\chi_A := \det(X\mathbb{I} - A)$$

is real and has all its roots real. Is it recommendable to compute the characteristic polynomial of A and then its real roots to compute the eigenvalues of A?

Numerical Analyst's Answer:

NO!

Our Answer: Effectively no, because, by the theorem below, the characteristic polynomial is ill-posed with respect the perturbation of its coefficients.

Theorem (Moroz, 22) (T.-C., Ts.; '23 +)

Let $A \in \mathbb{C}^{d\times d}$ be an Hermitian matrix, then, for some absolute constant c,

$$\mathsf{C}(\chi_A) \geq 2^{cd/\log d}.$$

Probabilistic Consequences

Prob. Theorem (Ver. A) (T.-C., Ts.; '23 +)

Let $\mathfrak{f} = (\mathfrak{f}_1, \dots, \mathfrak{f}_n)$ be a random real polynomial system in n variables whose coefficients are independent and uniformly distributed in $[-1,1]$. Then for $\ell \geq 1$,

$$\mathbb{E}_{\mathfrak{f}} \#\mathcal{Z}_r(\mathfrak{f}, \mathbb{R}^n)^\ell \leq \mathcal{O}\left(n\ell \log^2 \mathbf{D}\right)^{n\ell}$$

where $\mathcal{Z}_r(\mathfrak{f}, \mathbb{R}^n)$ is the set of real zeros of $\mathfrak{f}_1 = \dots = \mathfrak{f}_n = 0$, and $\mathbf{D}$ is the maximum degree.

Prob. Theorem (Ver. B) (T.-C., Ts.; '23 +)

Let $\mathfrak{f} = (\mathfrak{f}_1, \dots, \mathfrak{f}_n)$ be a random real polynomial system in n variables whose coefficients are independent and uniformly distributed in $[-1,1]$. Then there is an absolute constant C such that, for $t \geq 1$,

$$\mathbb{P}_{\mathfrak{f}}\left(\sqrt[n]{\#\mathcal{Z}(\mathfrak{f}, \mathbb{R}^n)} \geq t\right) \leq \exp\left(\frac{-t}{Cn \log^2 \mathbf{D}}\right)$$

where $\mathcal{Z}_r(\mathfrak{f}, \mathbb{R}^n)$ is the set of real zeros of $\mathfrak{f}_1 = \dots = \mathfrak{f}_n = 0$, and $\mathbf{D}$ is the maximum degree.

Corollary:
Fewnomial systems with many zeros are very improbable

More general...
We can cover a wide range of probabilistic assumptions

Why these bounds?

Condition numbers have nice probabilistic properties!

Prob. Theorem (T.-C., Ts.; '23 +)

Let $\mathfrak{f} = (\mathfrak{f}_1, \dots, \mathfrak{f}_n)$ be a random real polynomial system in n variables whose coefficients are independent and uniformly distributed in $[-1,1]$. Then for $\ell \geq 1$,

$$\mathbb{E}_{\mathfrak{f}} \log^\ell \mathsf{C}(\mathfrak{f}) \leq \mathcal{O}\left(n\ell \log \mathbf{D}\right)^{n\ell}$$

where $\mathcal{Z}(\mathfrak{f}, \mathbb{R}^n)$ is the set of real zeros of $\mathfrak{f}_1 = \dots = \mathfrak{f}_n = 0$, and $\mathbf{D}$ is the maximum degree.

Other Valid Distributions:

- Exponential.
- Gaussian.
- Integer variables uniformly distributed on an interval.

Evaluation Reduction
If $\mathsf{C}(f)$ is large, then $\mathbb{P}(|f(\mathfrak{x})|$ is small) is large, where $\mathfrak{x} \in [-1,1]^n$ random.

Algorithmic Consequences

Proof is fully constructible!

Alg. Theorem (T.-C., Ts.; '23 +)

There is a explicit partition

$$\mathcal{B}$$

of $[-1,1]^n$ into

$$\mathcal{O}(\log \mathbf{D})^n$$

boxes such that for all real polynomial system

$$f = (f_1, \dots, f_n)$$

in n variables of degree at most $\mathbf{D}$ and all $\mathrm{B} \in \mathcal{B}$, there is a polynomial

$$\phi_{f,\mathrm{B}}$$

of degree $\mathcal{O}(\max\{n \log \mathbf{D}, \log \mathsf{C}(f)\})$ such that

$$\#\mathcal{Z}(f, \mathrm{B}) \leq \#\mathcal{Z}(\phi_{f,\mathrm{B}}, \mathbb{R}^n).$$

Moreover, every real zero of f in B has a zero of $\phi_{f,\mathrm{B}}$ that converges quadratically to it under Newton's method.

Proof idea: Well-conditioned polynomials are fast converging Taylor series

What's the issue?
We need an estimate of $\mathsf{C}(f)$ to make the scheme effective, can we get the estimate fast? Or can we go around it?

A New Goal

Is there a Montecarlo numerical algorithm that, given a real polynomial system f, outputs an approximation of

$$\mathcal{Z}(f, [-1,1]^n)$$

with run-time at most

$$\mathcal{O}_{n,\mathbf{D}}\left(\log \mathsf{C}(f) + \log\log \frac{1}{\varepsilon}\right)^{\mathcal{O}(n)} \mathrm{L}(f)$$

with ε being the failure probability and $\mathrm{L}(f)$ the evaluation cost of f?

On the sphere...
We can cover also random polynomial system with the Weyl scaling. However, in this case, we are unable to cover fewnomials, and we get probabilistic bounds of the form $\mathcal{O}\left(\sqrt{\mathbf{D}} \log \mathbf{D}\right)^n$ appears in the bound.

Poster supported by NSF grant DMS 2232812