Farewell to Weyl: Condition-based analysis with a Banach norm in numerical algebraic geometry

Josué Tonelli-Cueto
(Berlin Mathematical School and Technische Universität Berlin)

July 11, 2019

SIAM Conference on Applied Algebraic Geometry

This is joint work with

- Felipe Cucker (CityU Hong Kong) and
- Alperen Ergür (Technische Universität Berlin)

to be preprinted this year as

Whatever title we agree on at the end

and funded by the

Einstein Stiftung Berlin
Einstein Foundation Berlin,

within the Einstein Visiting Fellowship ”Complexity and accuracy of numerical algorithms in algebra and geometry” of Felipe Cucker.

Table of contents

Motivation

$A \in \mathbb{C}^{m\times n}$ a matrix and $m \leq n$.
Two norms:

1. **Spectral norm**.

$$\|A\| := \max_{x\in\mathbb{S}(\mathbb{C}^n)} \|Ax\|$$

2. **Fröbenius norm**.

$$\|A\|_F := \sqrt{\sum_{i,j} \left|A^i_j\right|^2}$$

$A \in \mathbb{C}^{m\times n}$ a matrix and $m \leq n$

$$\Sigma := \{B \in \mathbb{C}^{m\times n} \mid \operatorname{rank} B < m\}$$

...and two conic condition numbers:

1. $\kappa(A) := \frac{\|A\|}{\operatorname{dist}(A,\Sigma)} = \|A\|\|A^\dagger\|$
2. $\kappa_F(A) := \frac{\|A\|_F}{\operatorname{dist}_F(A,\Sigma)}$

Curiously,

$$\frac{\|A\|}{\kappa(A)} = \operatorname{dist}(A,\Sigma) = \operatorname{dist}_F(A,\Sigma) = \frac{\|A\|_F}{\kappa(A)_F}$$

In general,

$$\frac{1}{m}\|A\|_F \leq \|A\| \leq \|A\|_F$$

but for random A,

$$\mathbb{E}_A \frac{\|A\|}{\|A\|_F} = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$$

Also,

$$\frac{\kappa(A)}{\kappa_F(A)} = \frac{\|A\|}{\|A\|_F}$$

So...

changing the norm improves the condition of large matrices!

Norms on polynomials

Notation

- $X_0, X_1, \ldots, X_n$ variables
- $n+1 :=$ number of variables
- $q :=$ number of distinct polynomials
- $\boldsymbol{d} = (d_1, \ldots, d_q)$ tuple of degrees
- $D := \max\{d_1, \ldots, d_q\}$
- $\mathcal{H}_{\boldsymbol{d}}[q]$ space of q-tuples f, where f_i is homogeneous polynomial of degree d_i in the $n+1$ variables $X_0, X_1, \ldots, X_n$
- $N := \sum_{i=1}^{q} \binom{n+d_i}{n} = q \min\left\{\mathcal{O}(D^n), \mathcal{O}(n^D)\right\} = \dim \mathcal{H}_{\boldsymbol{d}}[q]$
- $\Delta := \mathrm{diag}(\sqrt{\boldsymbol{d}})$
- $\mathrm{D}_x f$ tangent map $\mathrm{T}_x\mathbb{S}^n \to \mathbb{R}^q$ or $\mathrm{T}_{[x]}\mathbb{P}^n \to \mathbb{C}^q$

Weyl norm

$$\|f\|_W := \sqrt{\sum_{i=1}^{q} \|f_i\|_W^2}$$

where

$$\|f_i\|_W = \sqrt{\sum_{\alpha} \binom{d_i}{\alpha}^{-1} |f_{i,\alpha}|^2} \quad \text{and} \quad f_i = \sum_{\alpha} f_{i,\alpha} X^{\alpha}$$

Some properties:

1. Invariant under orthogonal/unitary transformations
2. It controls evaluation: $\|f(x)\| \leq \|f\|_W$
3. It controls the norm of the derivative: $\|\partial f\|_W \leq D\|f\|_W$
4. It comes from an inner product

Max norm

$$\|f\|_\infty := \max_{x\in\mathbb{S}^n} \|f(x)\|$$

and

$$\|f\|_{\mathfrak{m}} := \max_{x\in\mathbb{S}^n} \sqrt{\|f(x)\|^2 + \|\Delta^{-1}\mathrm{D}_x f\|^2}$$

Some properties:

1. Invariant under orthogonal/unitary transformations
2. It controls evaluation: $\|f(x)\| \leq \|f\|_\infty \leq \|f\|_{\mathfrak{m}}$
3. It controls the norm of the derivative: $\|\partial f\|_\infty \leq \sqrt{2}D\|f\|_\infty$ (Kellogs' Theorem)
4. $\|f\|_\infty$ better for computation and polynomial inequalities and $\|f\|_{\mathfrak{m}}$ better for condition inequalities, but they are computationally equivalent

$$\|f\|_\infty \leq \|f\|_{\mathfrak{m}} \leq \sqrt{2}\min\{D, \sqrt{qD}\}\|f\|_\infty$$

Considerations I

Example

$f \in \mathcal{H}_{\mathbb{1}}[q]$, i.e., f linear map given by $A \in \mathbb{C}^n$

$$\|f\|_\infty = \|A\|.$$

$$\|f\|_{\mathfrak{m}} = \sqrt{\|A\|^2 + \sigma_2(A)^2}$$

Proposition

Let $f \in \mathcal{H}_{\boldsymbol{d}}[q]$. *Then*

$$\|f\|_\infty \leq \|f\|_{\mathfrak{m}} \leq \|f\|_W \leq \sqrt{qN}\|f\|_\infty^{\mathbb{C}}.$$

Considerations II

Theorem

Let $f \in \mathcal{H}_{\mathbf{d}}[q]$ be a KSS random polynomial tuple and c_0 an absolute constant. Then

$$\mathbb{P}\left(\|f\|_W \geq c_0 N t\right) \leq \exp(1 - N t^2),$$

and

$$\mathbb{P}\left(\|f\|_\infty \geq c_0 \sqrt{n} \log(D) t\right) \leq \exp(1 - n \log(D) t^2)$$

Remark

We can also make this for dobro random polynomials...

Condition numbers

Old condition number

$$\mu(f,x) := \frac{\|f\|_W}{\sigma_q(\Delta^{-1}\mathrm{D}_x f)}$$

$$\kappa(f,x) := \frac{\|f\|_W}{\sqrt{\|f(x)\|^2 + \sigma_q(\Delta^{-1}\mathrm{D}_x f)^2}}$$

Theorem (Condition Number Theorem)

$$\kappa(f,x) = \|f\|_W / \mathrm{dist}_W(f, \Sigma_x)$$

where

$$\Sigma_x := \{g \mid g \textit{ singular at } x\}.$$

Why does it work?

1. Higher Derivative Estimate: It controls Smale's Gamma,

$$\gamma(f,x) := \sup_{k\geq 2}\left\|\frac{1}{k!}\mathrm{D}_x f^{\dagger} D_x^k f\right\|^{\frac{1}{k-1}} \leq \frac{1}{2}D^{3/2}\mu(f,x)$$

2. It's inverse is Lipschitz with respect to f,

$$\left|\frac{\|f\|_W}{\mu(f,x)} - \frac{\|g\|_W}{\mu(g,x)}\right| \leq \|f-g\|_W \text{ and } \left|\frac{\|f\|_W}{\kappa(f,x)} - \frac{\|g\|_W}{\kappa(g,x)}\right| \leq \|f-g\|_W;$$

3. and with respect to x,

$$\left|\frac{\|f\|_W}{\mu(f,x)} - \frac{\|f\|_W}{\mu(f,y)}\right| \leq D\|x-y\| \text{ and } \left|\frac{\|f\|_W}{\kappa(f,x)} - \frac{\|g\|_W}{\kappa(g,x)}\right| \leq D\|x-y\|.$$

These are what makes everything work!

New condition numbers?

$$\mathsf{M}(f,x) := \frac{\|f\|_{\mathfrak{m}}}{\sigma_q(\Delta^{-1}\mathrm{D}_x f)}$$

$$\mathsf{K}(f,x) := \frac{\|f\|_{\mathfrak{m}}}{\sqrt{\|f(x)\|^2 + \sigma_q(\Delta^{-1}\mathrm{D}_x f)^2}}$$

Theorem (Condition Number Theorem)

$$\mathsf{K}(f,x) = \|f\|_{\mathfrak{m}} / \mathrm{dist}_{\mathfrak{m}}(f, \Sigma_x)$$

where

$$\Sigma_x := \{g \mid g \textit{ singular at } x\}.$$

Do they still work?

1. Higher Derivative Estimate: It controls Smale's Gamma,

$$\gamma(f,x) := \sup_{k\geq 2}\left\|\frac{1}{k!}\mathrm{D}_x f^\dagger D_x^k f\right\|^{\frac{1}{k-1}} \leq \min\{\sqrt{q},\sqrt{D}\}D^{3/2}\mathrm{M}(f,x)$$

2. It's inverse is Lipschitz with respect to f,

$$\left|\frac{\|f\|_{\mathfrak{m}}}{\mathrm{M}(f,x)}-\frac{\|g\|_{\mathfrak{m}}}{\mathrm{M}(g,x)}\right| \leq \|f-g\|_{\mathfrak{m}} \text{ and } \left|\frac{\|f\|_{\mathfrak{m}}}{\mathrm{K}(f,x)}-\frac{\|g\|_{\mathfrak{m}}}{\mathrm{K}(g,x)}\right| \leq \|f-g\|_{\mathfrak{m}};$$

3. and with respect to x,

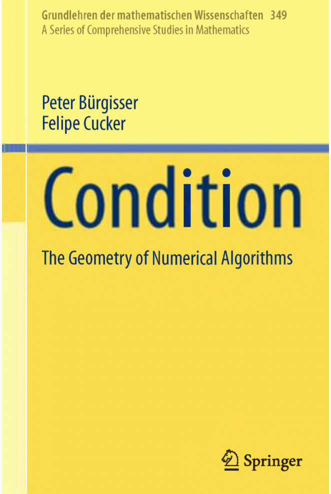
$$\left|\frac{\|f\|_{\mathfrak{m}}}{\mathrm{M}(f,x)}-\frac{\|f\|_{\mathfrak{m}}}{\mathrm{M}(f,y)}\right| \leq D\|x-y\| \text{ and } \left|\frac{\|f\|_{\mathfrak{m}}}{\mathrm{K}(f,x)}-\frac{\|g\|_{\mathfrak{m}}}{\mathrm{K}(g,x)}\right| \leq \sqrt{2}D\|x-y\|.$$

This means that...

We can carry,
up to parameters and constants,
the same condition-based
complexity analysis!

How?

Just follow the book!



...and some other papers!

(Proof-analysis of all it)

Case of linear homotopy

State-of-the-art

	Expected number of iterations	
Beltrán, Pardo; 2011	$\mathcal{O}\left(D^{3/2} n N\right)$	
Armentano, Beltrán, Bürgisser, Cucker, Shub; 2016	$\mathcal{O}\left(D^{3/2} n N^{1/2}\right)$	
Lairez; 2017	$\mathcal{O}\left(D^2 n^5\right)$	Not for linear homotopy!
Cucker, Ergür, T-C; $\leq$ 2020	$\mathcal{O}\left(D^{5/2} \log(D)^2 n^{5/2}\right)$	Some work to do...

Work to do

1. Can we compute $\|f\|_\infty$ up to a $\mathrm{poly}(D, n)$-factor in $\mathcal{O}(N)$-time?
    - To make the complexity bound effective, we need to be able to approximate the max norm fast
    - It can be with $\mathcal{O}(D)^n$ parallel evaluations and $\mathcal{O}(n \log(D))$ comparisons (Non-adaptive grid)
2. More general distributions
3. More general functions?

Case of grid and subdivision methods

Grid and subdivision methods

Based on a simple idea:

1. Subdivide region (or refine grid),
2. evaluate, and
3. compare.

Two types of subdivisions:

- Uniform subdivisions $\rightarrow$ effective (weak complexity)
  - Zero location (Cucker, Krick, Malajovich, Wschebor; 2008-12)
  - Homology computation of semialgebraic sets (Cucker, Krick, Shub; 2017), (Bürgisser, Cucker, Lairez; 2018) and (Bürgisser, Cucker, T.-C.; 2018&19)
- Adaptive subdivisions $\rightarrow$ efficient (average complexity) – recent!
  - Plantinga-Vegter algorithm (Next slide. . .)
  - Real condition estimation (Jiadong, Lairez; 2018)

Moreover, we can compute max norms on the way!

Plantinga-Vegter algorithm I

1. (Plantinga, Vegter; 2004)
    - Determination of isotopy type of smooth implicit curves inside a square and smooth implicit surfaces inside a box
    - Certification via interval arithmetic
    - No complexity analysis
2. (Burr, Gao, Tsigaridas; 2017)
    - Generalization of subdivision to arbitrary dimensions
    - Local size bound and continuous amortization
    - Worst-case bound for integer polynomials of degree D
3. (Cucker, Ergür, T.-C.; 2019)
    - Condition-based analysis (using Weyl norm) of the local size bound
    - Average and smoothed analysis for dobro polynomials, obtaining
      $$\tilde{\mathcal{O}}\left(D^{\frac{n^2+3n}{2}}\right)$$
      subdivisions on average
    - More at ISSAC19 next week in Beijing!

Plantinga-Vegter algorithm II

With the new norm…

$$\tilde{\mathcal{O}}\left(D^{\frac{n^2+3n}{2}})\right) \to \tilde{\mathcal{O}}\left(D^{\frac{3n}{2}} \log^{n+1} D\right)$$

So for curves…

$$\mathcal{O}\left(D^3 \log^3 D\right),$$

i.e., a lot better on average that many symbolic algorithms ($\mathcal{O}(D^{16} \log^5 D)$ c.f. (González-Vega, El Kahou; 1996))

Bere arretagatik eskerrik asko!

Galderak?