Numerical Condition in Polynomial Equation Solving

Peter Bürgisser

Technische Universität Berlin

Meeting on Algebraic Vision 2015

Berlin, October 8, 2015

Part I: Definition of condition and general phenomena

Explained for linear equation solving

Condition number as norm of derivative

- Numerical computation problem ("solution map")

$$f : \mathbb{R}^p \to \mathbb{R}^q, \; x \mapsto y = f(x)$$

Fix norms $\| \; \|$ on $\mathbb{R}^p, \mathbb{R}^q$.

Condition number as norm of derivative

- Numerical computation problem ("solution map")

$$f : \mathbb{R}^p \to \mathbb{R}^q, \; x \mapsto y = f(x)$$

  Fix norms $\| \;\|$ on $\mathbb{R}^p, \mathbb{R}^q$.
- Relative error $\|\Delta x\|/\|x\|$ of input x causes relative error $\|\Delta y\|/\|y\|$ of output y

Condition number as norm of derivative

- Numerical computation problem ("solution map")

$$f : \mathbb{R}^p \to \mathbb{R}^q,\ x \mapsto y = f(x)$$

  Fix norms $\| \ \|$ on $\mathbb{R}^p, \mathbb{R}^q$.
- Relative error $\|\Delta x\|/\|x\|$ of input x causes relative error $\|\Delta y\|/\|y\|$ of output y
- Condition number $\kappa(f,x)$ of f at x:

$$\|\Delta y\|/\|y\| \ \lesssim \ \kappa(f,x)\ \|\Delta x\|/\|x\|.$$

Condition number as norm of derivative

- Numerical computation problem ("solution map")

$$f : \mathbb{R}^p \to \mathbb{R}^q,\ x \mapsto y = f(x)$$

Fix norms $\| \ \|$ on $\mathbb{R}^p, \mathbb{R}^q$.

- Relative error $\|\Delta x\|/\|x\|$ of input x causes relative error $\|\Delta y\|/\|y\|$ of output y
- Condition number $\kappa(f,x)$ of f at x:

$$\|\Delta y\|/\|y\| \ \lesssim \ \kappa(f,x)\ \|\Delta x\|/\|x\|.$$

- Formal definition if f is differentiable:

$$\kappa(f,x) \ := \ \|Df(x)\| \frac{\|x\|}{\|f(x)\|}$$

where $\|Df(x)\|$ denotes the operator norm of the derivative of f at x.

Condition number of matrix inversion

- Consider matrix inversion

$$f\colon \mathrm{GL}(m,\mathbb{R}) \to \mathbb{R}^{m\times m}, A \mapsto A^{-1}.$$

We measure errors with the spectral norm.

Condition number of matrix inversion

- Consider matrix inversion

$$f\colon \mathrm{GL}(m,\mathbb{R}) \to \mathbb{R}^{m\times m}, A \mapsto A^{-1}.$$

We measure errors with the spectral norm.

- **Theorem.** The condition number of f at A equals

$$\kappa(A) := \kappa(f, A) = \|A\|\,\|A^{-1}\|.$$

This is well known as "the condition number of the matrix A".

Condition number of matrix inversion

- Consider matrix inversion

$$f : \mathrm{GL}(m, \mathbb{R}) \to \mathbb{R}^{m\times m}, A \mapsto A^{-1}.$$

We measure errors with the spectral norm.

- **Theorem.** The condition number of f at A equals

$$\kappa(A) := \kappa(f, A) = \|A\|\,\|A^{-1}\|.$$

This is well known as "the condition number of the matrix A".
- $\kappa(A)$ was first introduced by A. Turing in 1948.

Condition number of matrix inversion

- Consider matrix inversion

$$f : \mathrm{GL}(m, \mathbb{R}) \to \mathbb{R}^{m \times m}, A \mapsto A^{-1}.$$

We measure errors with the spectral norm.

- **Theorem.** The condition number of f at A equals

$$\kappa(A) := \kappa(f, A) = \|A\| \, \|A^{-1}\|.$$

This is well known as "the condition number of the matrix A".

- $\kappa(A)$ was first introduced by A. Turing in 1948.
- Warnung: a different computational problem related to A has a different condition number.

Condition number of matrix inversion

- Consider matrix inversion

$$f\colon \mathrm{GL}(m,\mathbb{R}) \to \mathbb{R}^{m\times m}, A \mapsto A^{-1}.$$

We measure errors with the spectral norm.

- **Theorem.** The condition number of f at A equals

$$\kappa(A) := \kappa(f, A) = \|A\|\,\|A^{-1}\|.$$

This is well known as "the condition number of the matrix A".

- $\kappa(A)$ was first introduced by A. Turing in 1948.
- Warnung: a different computational problem related to A has a different condition number.
- For computing the eigenvalues $\lambda_1, \dots, \lambda_n$ of A have n condition numbers $\kappa(A,\lambda_1), \dots, \kappa(A,\lambda_n)$ (Wilkinson).

Distance to ill-posedness

- We call the set of singular matrices $\Sigma \subseteq \mathbb{R}^{m \times m}$ the set of ill-posed instances for matrix inversion. Clearly, $A \in \Sigma \Leftrightarrow \det A = 0$.

Distance to ill-posedness

- We call the set of singular matrices $\Sigma \subseteq \mathbb{R}^{m\times m}$ the set of ill-posed instances for matrix inversion. Clearly, $A \in \Sigma \Leftrightarrow \det A = 0$.
- The Eckart-Young Theorem from 1936 states that

$$\kappa(A) = \|A\|\,\|A^{-1}\| = \frac{\|A\|}{\operatorname{dist}(A, \Sigma)}.$$

  where dist either refers to operator norm or to Frobenius norm (Euclidean norm on $\mathbb{R}^{n\times n}$).

Distance to ill-posedness

- We call the set of singular matrices $\Sigma \subseteq \mathbb{R}^{m\times m}$ the set of ill-posed instances for matrix inversion. Clearly, $A \in \Sigma \Leftrightarrow \det A = 0$.
- The Eckart-Young Theorem from 1936 states that

$$\kappa(A) = \|A\| \, \|A^{-1}\| = \frac{\|A\|}{\mathrm{dist}(A, \Sigma)}.$$

  where dist either refers to operator norm or to Frobenius norm (Euclidean norm on $\mathbb{R}^{n\times n}$).
- This is a prototype of a Condition Number Theorem; see Demmel.

Role of condition numbers

Obvious: Condition numbers are a crucial issue for designing "numerically stable" algorithms.

Role of condition numbers

Obvious: Condition numbers are a crucial issue for designing "numerically stable" algorithms.

Less obvious, but true: Even when assuming infinite precision arithmetic, the condition of an input often dominates the running time of iterative algorithms.

Role of condition numbers

Obvious: Condition numbers are a crucial issue for designing "numerically stable" algorithms.

Less obvious, but true: Even when assuming infinite precision arithmetic, the condition of an input often dominates the running time of iterative algorithms.

Three important examples for this phenomenon:

- conjugate gradient method for solving linear equations (Hestenes and Stiefel)

Role of condition numbers

Obvious: Condition numbers are a crucial issue for designing "numerically stable" algorithms.

Less obvious, but true: Even when assuming infinite precision arithmetic, the condition of an input often dominates the running time of iterative algorithms.

Three important examples for this phenomenon:

- conjugate gradient method for solving linear equations (Hestenes and Stiefel)
- interior point methods for linear optimization (Renegar)

Role of condition numbers

Obvious: Condition numbers are a crucial issue for designing "numerically stable" algorithms.

Less obvious, but true: Even when assuming infinite precision arithmetic, the condition of an input often dominates the running time of iterative algorithms.

Three important examples for this phenomenon:

- conjugate gradient method for solving linear equations (Hestenes and Stiefel)
- interior point methods for linear optimization (Renegar)
- Newton homotopy methods to solve systems of polynomial equations (Shub and Smale)

Probabilistic analysis of Turing's condition number

- Assume slight perturbation of A due to noise, round-off, etc.

Probabilistic analysis of Turing's condition number

- Assume slight perturbation of A due to noise, round-off, etc.
- We model this with an isotropic Gaussian distribution with mean $\overline{A} \in \mathbb{R}^{n \times n}$ and covariance matrix $\sigma^2 I$, which has the density:

$$\rho(A) = \frac{1}{(\sigma\sqrt{2\pi})^{n^2}} \exp\left(-\frac{\|A - \overline{A}\|_F^2}{2\sigma^2}\right).$$

If $\overline{A} = 0$ and $\sigma = 1$: Ginibre ensemble.

Probabilistic analysis of Turing's condition number

- Assume slight perturbation of A due to noise, round-off, etc.
- We model this with an isotropic Gaussian distribution with mean $\overline{A} \in \mathbb{R}^{n \times n}$ and covariance matrix $\sigma^2 I$, which has the density:

$$\rho(A) = \frac{1}{(\sigma\sqrt{2\pi})^{n^2}} \exp\left(-\frac{\|A - \overline{A}\|_F^2}{2\sigma^2}\right).$$

  If $\overline{A} = 0$ and $\sigma = 1$: Ginibre ensemble.
- **Theorem.** (Wschebor, improving Sankar, Spielman, and Teng)

$$\sup_{\|\overline{A}\|=1} \operatorname*{Prob}_{A \sim N(\overline{A}, \sigma^2 I)} \{\kappa(A) \geq t\} = \mathcal{O}\left(\frac{n}{\sigma t}\right).$$

  This is a prototype of a smoothed analysis.

Probabilistic analysis of Turing's condition number

- Assume slight perturbation of A due to noise, round-off, etc.
- We model this with an isotropic Gaussian distribution with mean $\overline{A} \in \mathbb{R}^{n\times n}$ and covariance matrix $\sigma^2 I$, which has the density:

$$\rho(A) = \frac{1}{(\sigma\sqrt{2\pi})^{n^2}} \exp\Big(-\frac{\|A-\overline{A}\|_F^2}{2\sigma^2}\Big).$$

  If $\overline{A} = 0$ and $\sigma = 1$: Ginibre ensemble.
- **Theorem.** (Wschebor, improving Sankar, Spielman, and Teng)

$$\sup_{\|\overline{A}\|=1} \operatorname*{Prob}_{A\sim N(\overline{A},\sigma^2 I)} \{\kappa(A) \geq t\} = \mathcal{O}\Big(\frac{n}{\sigma t}\Big).$$

  This is a prototype of a smoothed analysis.
- Tao and Vu (2010) have results for general distributions.

Part II: Polynomial Equations

Complexity of Bezout's Theorem

(Shub and Smale 1993–1996)

Smale's 17th problem

Guiding question:

> *Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?*

Smale's 17th problem

Guiding question:

> *Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?*

The answer is yes!

Smale's 17th problem

Guiding question:

> *Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?*

The answer is yes!

Very recently, the problem was completely solved by Pierre Lairez at TU Berlin (2015), building on work by Shub, Smale, Beltrán, Pardo, Bürgisser, Cucker.

The framework

- For a degree vector $d = (d_1, \ldots, d_n)$ we define

  $$\mathcal{H}_d := \{f = (f_1, \ldots, f_n) \mid f_i \in \mathbb{C}[X_0, \ldots, X_n] \text{ homogeneous of degree } d_i\}.$$

The framework

- For a degree vector $d = (d_1, \ldots, d_n)$ we define

  $$\mathcal{H}_d := \{f = (f_1, \ldots, f_n) \mid f_i \in \mathbb{C}[X_0, \ldots, X_n] \text{ homogeneous of degree } d_i\}.$$

- The input size is $N := \dim_{\mathbb{C}} \mathcal{H}_d$.

The framework

- For a degree vector $d = (d_1, \ldots, d_n)$ we define

  $$\mathcal{H}_d := \{f = (f_1, \ldots, f_n) \mid f_i \in \mathbb{C}[X_0, \ldots, X_n] \text{ homogeneous of degree } d_i\}.$$

- The input size is $N := \dim_{\mathbb{C}} \mathcal{H}_d$.
- We look for zeros ζ of f in complex projective space $\mathbb{P}^n$: $f(\zeta) = 0$.

The framework

- For a degree vector $d = (d_1, \dots, d_n)$ we define

  $$\mathcal{H}_d := \{f = (f_1, \dots, f_n) \mid f_i \in \mathbb{C}[X_0, \dots, X_n] \text{ homogeneous of degree } d_i\}.$$

- The input size is $N := \dim_{\mathbb{C}} \mathcal{H}_d$.
- We look for zeros ζ of f in complex projective space $\mathbb{P}^n$: $f(\zeta) = 0$.
- The Bombieri-Weyl hermitian inner product $\langle\ \rangle$ on $\mathcal{H}_d$ is invariant under the natural action of the unitary group $U(n+1)$ on $\mathcal{H}_d$ and allows to define $\|f\| := \langle f, f\rangle^{1/2}$.

The framework

- For a degree vector $d = (d_1, \dots, d_n)$ we define

$$\mathcal{H}_d := \{f = (f_1, \dots, f_n) \mid f_i \in \mathbb{C}[X_0, \dots, X_n] \text{ homogeneous of degree } d_i\}.$$

- The input size is $N := \dim_{\mathbb{C}} \mathcal{H}_d$.
- We look for zeros ζ of f in complex projective space $\mathbb{P}^n$: $f(\zeta) = 0$.
- The Bombieri-Weyl hermitian inner product $\langle\ \rangle$ on $\mathcal{H}_d$ is invariant under the natural action of the unitary group $U(n+1)$ on $\mathcal{H}_d$ and allows to define $\|f\| := \langle f, f\rangle^{1/2}$.
- We have a standard Gaussian distribution on $\mathcal{H}_d$ with density

$$\rho(f) = \frac{1}{\sqrt{2\pi}^{2N}} \exp\Big(-\frac{1}{2}\|f\|^2\Big).$$

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \left\{(f, \zeta) \mid f(\zeta) = 0\right\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \left\{(f, \zeta) \mid f(\zeta) = 0\right\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold
- We have a solution map

$$G : \mathcal{H}_d \to \mathbb{P}^n,\ \widetilde{f} \mapsto \widetilde{\zeta}$$

defined locally around the simple zero ζ (implicit function theorem).

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \left\{(f, \zeta) \mid f(\zeta) = 0\right\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold
- We have a solution map

$$G\colon \mathcal{H}_d \to \mathbb{P}^n,\ \widetilde{f} \mapsto \widetilde{\zeta}$$

  defined locally around the simple zero ζ (implicit function theorem).
- We consider the derivative $D_f G\colon T_f\mathcal{H}_d \to T_\zeta\mathbb{P}^n$.

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \{(f, \zeta) \mid f(\zeta) = 0\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold
- We have a solution map

$$G \colon \mathcal{H}_d \to \mathbb{P}^n, \ \widetilde{f} \mapsto \widetilde{\zeta}$$

  defined locally around the simple zero ζ (implicit function theorem).
- We consider the derivative $D_f G \colon T_f\mathcal{H}_d \to T_\zeta\mathbb{P}^n$.
- According to our general principles, we define by

$$\mu'(f, \zeta) := \|f\| \cdot \|D_f G\|,$$

  the condition number at (f, ζ).

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \left\{(f,\zeta) \mid f(\zeta) = 0\right\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold
- We have a solution map

$$G\colon \mathcal{H}_d \to \mathbb{P}^n,\ \widetilde{f} \mapsto \widetilde{\zeta}$$

  defined locally around the simple zero ζ (implicit function theorem).
- We consider the derivative $D_f G\colon T_f\mathcal{H}_d \to T_\zeta\mathbb{P}^n$.
- According to our general principles, we define by

$$\mu'(f,\zeta) := \|f\| \cdot \|D_f G\|,$$

  the condition number at (f,ζ).
- One can show $\|D_f G\| = \|Df(\zeta)^\dagger\|$ if $\|z\| = 1$.

Condition number

- Let $f(\zeta) = 0$. How much does ζ change when we perturb f a little?
- Consider the solution variety $V := \Big\{(f, \zeta) \mid f(\zeta) = 0\Big\} \subseteq \mathcal{H}_d \times \mathbb{P}^n$, which is a smooth Riemannian submanifold
- We have a solution map

$$G \colon \mathcal{H}_d \to \mathbb{P}^n,\ \widetilde{f} \mapsto \widetilde{\zeta}$$

  defined locally around the simple zero ζ (implicit function theorem).
- We consider the derivative $D_f G \colon T_f \mathcal{H}_d \to T_\zeta \mathbb{P}^n$.
- According to our general principles, we define by

$$\mu'(f, \zeta) := \|f\| \cdot \|D_f G\|,$$

  the condition number at (f, ζ).
- One can show $\|D_f G\| = \|Df(\zeta)^\dagger\|$ if $\|z\| = 1$.
- For the sake of elegance: rather use $\|Df(\zeta)^\dagger \operatorname{diag}(\sqrt{d_i})\|$ resulting in the condition number $\mu(f, \zeta)$ for polynomial equation solving.

Newton iteration and approximate zeros

- Have a projective Newton iteration

$$z_{k+1} = N_f(z_k)$$

with Newton operator $N_f : \mathbb{P}^n \to \mathbb{P}^n$ and starting point z_0.

Newton iteration and approximate zeros

- Have a projective Newton iteration

$$z_{k+1} = N_f(z_k)$$

with Newton operator $N_f : \mathbb{P}^n \to \mathbb{P}^n$ and starting point z_0.

- Let d refer to the geodesic distance on the Riemannian manifold $\mathbb{P}^n$ (Fubini-Study metric). Think of d as an angle.

Newton iteration and approximate zeros

- Have a projective Newton iteration

$$z_{k+1} = N_f(z_k)$$

  with Newton operator $N_f \colon \mathbb{P}^n \to \mathbb{P}^n$ and starting point z_0.
- Let d refer to the geodesic distance on the Riemannian manifold $\mathbb{P}^n$ (Fubini-Study metric). Think of d as an angle.

Theorem. (Smale) The condition controls the radius of the basin of attraction of the Newton iteration. Let $(f, \zeta) \in V$ and put $D := \max_i d_i$. If $z \in \mathbb{P}^n$ satisfies

$$d(z, \zeta) \leq \frac{0.3}{D^{3/2}\, \mu(f, \zeta)},$$

then for all $i \in \mathbb{N}$

$$d(z_i, \zeta) \leq \frac{1}{2^{2^i - 1}}\, d(z_0, \zeta).$$

We call z an approximate zero of f associated with the zero ζ.

Homotopy path continuation: a topological idea

- Given a start system $(g, \zeta) \in V$ and an input $f \in \mathcal{H}_d$.

Homotopy path continuation: a topological idea

- Given a start system $(g, \zeta) \in V$ and an input $f \in \mathcal{H}_d$.
- Connect g and f by line segment $[g, f]$ consisting of

$$q_t := (1-t)g + tf \qquad \text{for } t \in [0,1].$$

Homotopy path continuation: a topological idea

- Given a start system $(g, \zeta) \in V$ and an input $f \in \mathcal{H}_d$.
- Connect g and f by line segment $[g, f]$ consisting of

$$q_t := (1-t)g + tf \quad \text{for } t \in [0, 1].$$

- If $[g, f]$ does not meet the discriminant variety (none of the q_t has a multiple zero), then there exists a unique lifting to a solution path γ

$$\gamma\colon [0, 1] \to V, t \mapsto (f_t, \zeta_t)$$

such that $(f_0, \zeta_0) = (g, \zeta)$.

Homotopy path continuation: a topological idea

- Given a start system $(g, \zeta) \in V$ and an input $f \in \mathcal{H}_d$.
- Connect g and f by line segment $[g, f]$ consisting of

$$q_t := (1 - t)g + tf \quad \text{for } t \in [0, 1].$$

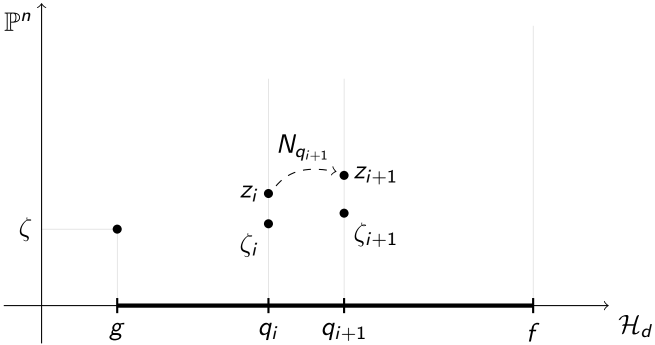
- If $[g, f]$ does not meet the discriminant variety (none of the q_t has a multiple zero), then there exists a unique lifting to a solution path γ

$$\gamma \colon [0, 1] \to V, t \mapsto (f_t, \zeta_t)$$

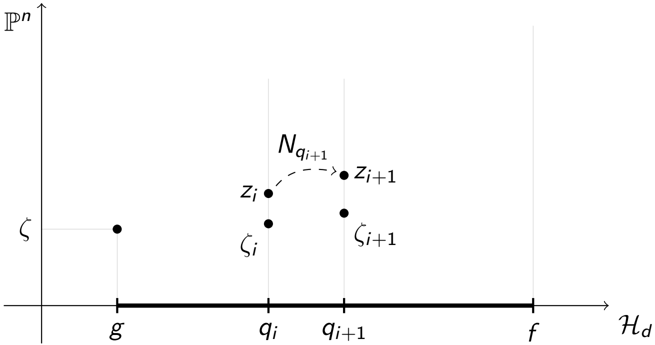
  such that $(f_0, \zeta_0) = (g, \zeta)$.

- Choosing adaptive step sizes is dictated by a condition number theorem that characterizes $\mu(f, \zeta)^{-1}$ as the inverse distance of f to the set Σ_ζ of systems $\widetilde{f}$ that have a multiple zero at ζ.

Discretization



Discretization



- Follow γ numerically: Let $t_0 = 0, \ldots, t_k = 1$ and write $q_i := q_{t_i}$.
- Successively compute approximations z_i of $\zeta_i := \zeta_{t_i}$ by Newton's method

$$z_{i+1} := N_{q_{i+1}}(z_i)$$

starting with $z_0 := \zeta$.

Complexity

- Compute t_{i+1} adaptively from t_i such that

$$d(q_{i+1}, q_i) = \frac{c}{D^{3/2}\mu(q_i, z_i)^2}.$$

This defines the Adaptive Linear Homotopy algorithm.

Complexity

- Compute t_{i+1} adaptively from t_i such that

$$d(q_{i+1}, q_i) = \frac{c}{D^{3/2}\mu(q_i, z_i)^2}.$$

  This defines the Adaptive Linear Homotopy algorithm.

- We denote by $K(f, g, \zeta)$ the number k of Newton continuation steps that are needed to follow the homotopy.

Complexity

- Compute t_{i+1} adaptively from t_i such that

$$d(q_{i+1}, q_i) = \frac{c}{D^{3/2}\mu(q_i, z_i)^2}.$$

  This defines the Adaptive Linear Homotopy algorithm.

- We denote by $K(f, g, \zeta)$ the number k of Newton continuation steps that are needed to follow the homotopy.

Theorem. (Shub & Smale). z_i is an approximate zero of ζ_i for all i. Moreover,

$$K(f, g, \zeta) \leq 217\, D^{3/2} \int_0^1 \mu(\gamma(t))^2 \, \|\dot{\gamma}(t)\| \, dt.$$

Average expected polynomial time

- Suppose the start system $(g, \zeta) \in V$ is chosen "at random":

  choose $g \in \mathcal{H}_d$ from standard Gaussian, choose one of the $d_1 \cdots d_n$ many zeros ζ of g uniformly at random.

  Efficient sampling of (g, ζ) is possible (Beltrán & Pardo).

Average expected polynomial time

- Suppose the start system $(g, \zeta) \in V$ is chosen "at random":

  choose $g \in \mathcal{H}_d$ from standard Gaussian, choose one of the $d_1 \cdots d_n$ many zeros ζ of g uniformly at random.

  Efficient sampling of (g, ζ) is possible (Beltrán & Pardo).
- Further assume that the input system $f \in \mathcal{H}_d$ is standard Gaussian.

Average expected polynomial time

- Suppose the start system $(g, \zeta) \in V$ is chosen "at random":

  choose $g \in \mathcal{H}_d$ from standard Gaussian, choose one of the $d_1 \cdots d_n$ many zeros ζ of g uniformly at random.

  Efficient sampling of (g, ζ) is possible (Beltrán & Pardo).
- Further assume that the input system $f \in \mathcal{H}_d$ is standard Gaussian.

- **Theorem.** (Beltrán and Pardo) The average number of homotopy steps is bounded by

$$\mathbb{E}_{(g,\zeta),f} K(f, g, \zeta) \leq \mathcal{O}\big(D^{3/2} N n\big),$$

  where $D := \max_i d_i$ and $N = \dim {}_{\_}C\mathcal{H}_d$.

Average expected polynomial time

- Suppose the start system $(g, \zeta) \in V$ is chosen "at random":

  choose $g \in \mathcal{H}_d$ from standard Gaussian, choose one of the $d_1 \cdots d_n$ many zeros ζ of g uniformly at random.

  Efficient sampling of (g, ζ) is possible (Beltrán & Pardo).
- Further assume that the input system $f \in \mathcal{H}_d$ is standard Gaussian.

- **Theorem.** (Beltrán and Pardo) The average number of homotopy steps is bounded by

$$\mathbb{E}_{(g,\zeta),f} K(f, g, \zeta) \leq \mathcal{O}\big(D^{3/2} N n\big),$$

  where $D := \max_i d_i$ and $N = \dim {}_{\mathbb{C}}\mathcal{H}_d$.

- When allowing randomized algorithms, this is a solution to Smale's 17th problem.

Average expected polynomial time

- Suppose the start system $(g, \zeta) \in V$ is chosen "at random":

  choose $g \in \mathcal{H}_d$ from standard Gaussian, choose one of the $d_1 \cdots d_n$ many zeros ζ of g uniformly at random.

  Efficient sampling of (g, ζ) is possible (Beltrán & Pardo).
- Further assume that the input system $f \in \mathcal{H}_d$ is standard Gaussian.

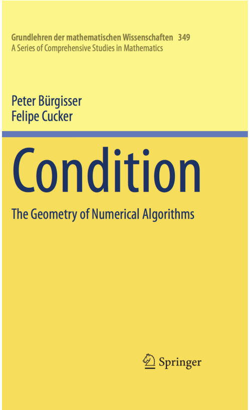
- **Theorem.** (Beltrán and Pardo) The average number of homotopy steps is bounded by

$$\mathbb{E}_{(g,\zeta),f} K(f, g, \zeta) \leq \mathcal{O}\big(D^{3/2} N n\big),$$

  where $D := \max_i d_i$ and $N = \dim_{\mathbb{C}} \mathcal{H}_d$.

- When allowing randomized algorithms, this is a solution to Smale's 17th problem.
- "Derandomizations": B & Cucker, Lairez

For more details I refer to my new monograph (Springer 2013) with Felipe Cucker:



Part III: Condition of intersecting a projective variety with a varying linear subspace

Setting of intersection problem

- $Z \subseteq \mathbb{P}^n$ fixed m-dimensional irreducible projective variety
- $L \subseteq \mathbb{P}^n$ varying linear subspace of complementary dimension $s = n - m$
- Bezout: $\#(Z \cap L) = \deg Z$ if L is in sufficiently general position

Setting of intersection problem

- $Z \subseteq \mathbb{P}^n$ fixed m-dimensional irreducible projective variety
- $L \subseteq \mathbb{P}^n$ varying linear subspace of complementary dimension $s = n - m$
- Bezout: $\#(Z \cap L) = \deg Z$ if L is in sufficiently general position
- Goal: Define and study the condition to numerically compute elements of $Z \cap L$
- Motivated by a problem in algebraic vision: S. Agarwal, H. L. Lee, R. Thomas, and B. Sturmfels

The kernel intersection condition number

- Let L be given as the kernel of a full rank matrix $A \in \mathbb{C}^{m\times(n+1)}$
- Suppose L has a transversal intersection with Z at the smooth point z of Z. For $\widetilde{A} \approx A$, let $G(\widetilde{A})$ be the intersection point of $\ker \widetilde{A} \cap Z$ close to L

The kernel intersection condition number

- Let L be given as the kernel of a full rank matrix $A \in \mathbb{C}^{m\times(n+1)}$
- Suppose L has a transversal intersection with Z at the smooth point z of Z. For $\widetilde{A} \approx A$, let $G(\widetilde{A})$ be the intersection point of $\ker \widetilde{A} \cap Z$ close to L
- Consider the derivative $D_A G\colon T_A\mathbb{C}^{m\times(n+1)} \to T_z\mathbb{P}^n$ of the solution map G and its operator norm

$$\|D_A G\| := \sup_{\|\dot{A}\|_F=1} \|D_A G(\dot{A})\|,$$

  defined with respect to the Frobenius norm $\|\dot{A}\|_F$.
- The kernel intersection condition number of A at z (with respect to the variety Z) is defined as

$$\mathrm{kercond}_Z(A,z) := \|A\| \cdot \|D_A G\|$$

The intrinsic intersection condition number

- The complex Grassmann manifold $\mathbb{G} := \mathbb{G}(\mathbb{P}^n, s)$ is the set of s-dimensional projective linear subspaces L of $\mathbb{P}^n$. One can define a unitary invariant hermitian metric on $\mathbb{G}$.

The intrinsic intersection condition number

- The complex Grassmann manifold $\mathbb{G} := \mathbb{G}(\mathbb{P}^n, s)$ is the set of s-dimensional projective linear subspaces L of $\mathbb{P}^n$. One can define a unitary invariant hermitian metric on $\mathbb{G}$.
- Have now a solution map $\widetilde{L} \mapsto \gamma(\widetilde{L}) \in \mathbb{P}^n$ with the derivative $D_L\gamma \colon T_L\mathbb{G} \to T_z\mathbb{P}^n$ and define the (intrinsic) intersection condition number of L at z (with respect to the variety Z) as

$$\kappa_Z(L, z) := \|D_L\gamma\|.$$

The intrinsic intersection condition number

- The complex Grassmann manifold $\mathbb{G} := \mathbb{G}(\mathbb{P}^n, s)$ is the set of s-dimensional projective linear subspaces L of $\mathbb{P}^n$. One can define a unitary invariant hermitian metric on $\mathbb{G}$.
- Have now a solution map $\widetilde{L} \mapsto \gamma(\widetilde{L}) \in \mathbb{P}^n$ with the derivative $D_L\gamma \colon T_L\mathbb{G} \to T_z\mathbb{P}^n$ and define the (intrinsic) intersection condition number of L at z (with respect to the variety Z) as

$$\kappa_Z(L, z) := \|D_L\gamma\|.$$

- **Theorem.** If $L \in \mathbb{G}$ is the kernel of $A \in \mathbb{C}^{m\times(n+1)}$ and $z \in L$:

$$\kappa_Z(L, z) \;\leq\; \mathrm{kercond}_Z(A, z) \;\leq\; \kappa(A) \cdot \kappa_Z(L, z),$$

where $\kappa(A) := \|A\| \cdot \|A^\dagger\|$.

Geometric characterizations

- Let $z \in Z \cap L$ be a smooth point of Z.
- A small minimum angle between L and Z at z means a "glancing intersection" of L and Z, which should result in a large intersection condition. This can be made formal:

Geometric characterizations

- Let $z \in Z \cap L$ be a smooth point of Z.
- A small minimum angle between L and Z at z means a "glancing intersection" of L and Z, which should result in a large intersection condition. This can be made formal:
- **Theorem.** We have $\kappa_Z(L, z) = 1/\sin\alpha$, where α be the minimum angle between $T_z Z$ and $T_z L$.

Geometric characterizations

- Let $z \in Z \cap L$ be a smooth point of Z.
- A small minimum angle between L and Z at z means a "glancing intersection" of L and Z, which should result in a large intersection condition. This can be made formal:
- **Theorem.** We have $\kappa_Z(L, z) = 1/\sin\alpha$, where α be the minimum angle between T_zZ and T_zL.
- The Schubert variety $\Sigma_z(Z)$ of Z at z consists of the $L \in \mathbb{G}_z$ having a nontransversal intersection with Z at z. The geodesic distance d_g and projection distance d_p both define a metric on the subspace $\mathbb{G}_z$ of $\mathbb{G}$.

Geometric characterizations

- Let $z \in Z \cap L$ be a smooth point of Z.
- A small minimum angle between L and Z at z means a "glancing intersection" of L and Z, which should result in a large intersection condition. This can be made formal:
- **Theorem.** We have $\kappa_Z(L, z) = 1/\sin\alpha$, where α be the minimum angle between T_zZ and T_zL.
- The Schubert variety $\Sigma_z(Z)$ of Z at z consists of the $L \in \mathbb{G}_z$ having a nontransversal intersection with Z at z. The geodesic distance d_g and projection distance d_p both define a metric on the subspace $\mathbb{G}_z$ of $\mathbb{G}$.
- **Condition Number Theorem.** We have $d_p(L, \Sigma_z) = \sin d_g(L, \Sigma_z)$ and

$$\kappa_Z(L, z) = \frac{1}{d_p(L, \Sigma_z)}.$$

Towards a probabilistic analysis

- The Hurwitz variety $\Sigma(Z)$ of Z is defined as the Zariski closure of the union of the local Schubert varieties $\Sigma_z(Z)$, taken over all regular points z of Z. It consists of all L touching Z (and the limits thereof).
- $\Sigma(Z)$ is a hypersurface in $\mathbb{G}$ if Z is not linear (Sturmfels 2014).

Towards a probabilistic analysis

- The Hurwitz variety $\Sigma(Z)$ of Z is defined as the Zariski closure of the union of the local Schubert varieties $\Sigma_z(Z)$, taken over all regular points z of Z. It consists of all L touching Z (and the limits thereof).
- $\Sigma(Z)$ is a hypersurface in $\mathbb{G}$ if Z is not linear (Sturmfels 2014).

- Define $\kappa_Z(L) := \max_{z \in Z \cap L} \kappa_Z(L, z)$. Then

$$\kappa_Z(L) \geq \epsilon^{-1} \Longleftrightarrow d_p(L, \Sigma(Z)) \leq \epsilon.$$

- For uniformly random $L \in \mathbb{G}$, bounding the probability that $\kappa_Z(L)$ is large means bounding the volume of the tube around $\Sigma(Z)$.

Towards a probabilistic analysis

- The Hurwitz variety $\Sigma(Z)$ of Z is defined as the Zariski closure of the union of the local Schubert varieties $\Sigma_z(Z)$, taken over all regular points z of Z. It consists of all L touching Z (and the limits thereof).
- $\Sigma(Z)$ is a hypersurface in $\mathbb{G}$ if Z is not linear (Sturmfels 2014).
- Define $\kappa_Z(L) := \max_{z \in Z \cap L} \kappa_Z(L, z)$. Then

$$\kappa_Z(L) \geq \epsilon^{-1} \iff d_p(L, \Sigma(Z)) \leq \epsilon.$$

- For uniformly random $L \in \mathbb{G}$, bounding the probability that $\kappa_Z(L)$ is large means bounding the volume of the tube around $\Sigma(Z)$.
- Can express the volume of $\Sigma(Z)$ in terms of its degree.

Thank you