The probability that a slight perturbation of a numerical analysis problem is difficult

Peter Bürgisser

University of Paderborn

Joint work with Felipe Cucker and Martin Lotz

IMA Workshop on Complexity, Coding, and Communication

April 20, 2007

Outline

Introduction

Condition Number: Definition

Suppose we have a map f:

$$a \in \mathbb{R}^p \mapsto f(a) \in \mathbb{R}^m$$

Condition Number: Definition

Suppose we have a map f:

$$a \in \mathbb{R}^p \mapsto f(a) \in \mathbb{R}^m$$

The condition number $\kappa(f, a)$ of an input $a \in \mathbb{R}^p$ with respect to f measures the extent to which small perturbations Δa of the input alter the output.

Condition Number: Definition

Suppose we have a map f:

$$a \in \mathbb{R}^p \mapsto f(a) \in \mathbb{R}^m$$

The condition number $\kappa(f, a)$ of an input $a \in \mathbb{R}^p$ with respect to f measures the extent to which small perturbations Δa of the input alter the output.

More specifically, after fixing norms $\|\cdot\|$ on $\mathbb{R}^p$ and $\mathbb{R}^m$, we define

$$\kappa(f, a) := \limsup_{\Delta a \to 0} \frac{\|f(a+\Delta a) - f(a)\| / \|f(a)\|}{\|\Delta a\| / \|a\|}.$$

Condition Number: Definition

Suppose we have a map f:

$$a \in \mathbb{R}^p \mapsto f(a) \in \mathbb{R}^m$$

The condition number $\kappa(f,a)$ of an input $a \in \mathbb{R}^p$ with respect to f measures the extent to which small perturbations Δa of the input alter the output.

More specifically, after fixing norms $\|\cdot\|$ on $\mathbb{R}^p$ and $\mathbb{R}^m$, we define

$$\kappa(f,a) := \mathrm{limsup}_{\Delta a \to 0} \frac{\|f(a+\Delta a) - f(a)\|/\|f(a)\|}{\|\Delta a\|/\|a\|}.$$

If f is differentiable, we have

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

where $\|Df(a)\|$ denotes the operator norm of the Jacobian of f at a.

Condition Number for Matrix inversion

Fix $0 \neq b \in \mathbb{R}^n$ (for simplicity). Consider the map f

$$A \in \mathbb{R}^{n \times n} \mapsto x = A^{-1} b.$$

Condition Number for Matrix inversion

Fix $0 \neq b \in \mathbb{R}^n$ (for simplicity). Consider the map f

$$A \in \mathbb{R}^{n \times n} \mapsto x = A^{-1} b.$$

The classical condition number of a nonsingular matrix A is given by

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

where $\| \cdot \|$ denotes the operator norm.

Condition Number for Matrix inversion

Fix $0 \neq b \in \mathbb{R}^n$ (for simplicity). Consider the map f

$$A \in \mathbb{R}^{n \times n} \mapsto x = A^{-1} b.$$

The classical condition number of a nonsingular matrix A is given by

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

where $\| \cdot \|$ denotes the operator norm. In order to see this, suppose $Ax = b$ and $(A + \Delta A)(x + \Delta x) = b$. Then

$$\Delta x = -A^{-1} \Delta A \, x + \mathcal{O}(\|\Delta A\|^2),$$

Condition Number for Matrix inversion

Fix $0 \neq b \in \mathbb{R}^n$ (for simplicity). Consider the map f

$$A \in \mathbb{R}^{n \times n} \mapsto x = A^{-1} b.$$

The classical condition number of a nonsingular matrix A is given by

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

where $\| \cdot \|$ denotes the operator norm. In order to see this, suppose $Ax = b$ and $(A + \Delta A)(x + \Delta x) = b$. Then

$$\Delta x = -A^{-1} \Delta A \, x + \mathcal{O}(\|\Delta A\|^2),$$

hence

$$\frac{\|\Delta x\|}{\|x\|} \frac{\|A\|}{\|\Delta A\|} \leq \kappa(A) + \mathcal{O}(\|\Delta A\|).$$

Role of Condition Numbers

- Crucial issue for designing "numerically stable" algorithms:
  - want the output to be accurate when using finite precision arithmetic (round-off errors).

Role of Condition Numbers

- Crucial issue for designing “numerically stable” algorithms:
  - want the output to be accurate when using finite precision arithmetic (round-off errors).
- Even when assuming infinite precision arithmetic, the condition of an input often determines the running time of iterative algorithms:

Role of Condition Numbers

- Crucial issue for designing “numerically stable” algorithms:
  - want the output to be accurate when using finite precision arithmetic (round-off errors).
- Even when assuming infinite precision arithmetic, the condition of an input often determines the running time of iterative algorithms:
  - conjugate gradient method for solving linear equations

Role of Condition Numbers

- Crucial issue for designing "numerically stable" algorithms:
  - want the output to be accurate when using finite precision arithmetic (round-off errors).
- Even when assuming infinite precision arithmetic, the condition of an input often determines the running time of iterative algorithms:
  - conjugate gradient method for solving linear equations
  - Renegar's interior point method for linear optimization

Role of Condition Numbers

- Crucial issue for designing “numerically stable” algorithms:
  - want the output to be accurate when using finite precision arithmetic (round-off errors).
- Even when assuming infinite precision arithmetic, the condition of an input often determines the running time of iterative algorithms:
  - conjugate gradient method for solving linear equations
  - Renegar’s interior point method for linear optimization
  - Shub and Smale’s Newton homotopy method to solve systems of polynomial equations

Average-Case Analysis

- For ill-posed problems, the condition number and thus the running time are infinite.

Average-Case Analysis

- For ill-posed problems, the condition number and thus the running time are infinite.
- Hence worst-case analysis does not make sense in this context.

Average-Case Analysis

- For ill-posed problems, the condition number and thus the running time are infinite.
- Hence worst-case analysis does not make sense in this context.
- An average-case analysis of the running time of a numerical algorithm reduces to an analysis of the distribution (or expected value) of the condition number for random inputs a.
  Cf. Demmel, Edelman, Kostlan, Renegar, Shub, Smale,...

Smoothed Analysis

Smoothed analysis , proposed by D. Spielman and S.-H. Teng, is a new form of analysis, that arguably blends the best of both worst-case and average-case.

Smoothed Analysis

Smoothed analysis , proposed by D. Spielman and S.-H. Teng, is a new form of analysis, that arguably blends the best of both worst-case and average-case.

Let $f : \mathbb{R}^p \longrightarrow \mathbb{R}_+$ be a function (running time, condition number).

Instead of showing

> *"it is unlikely that $f(a)$ will be large"*

Smoothed Analysis

Smoothed analysis , proposed by D. Spielman and S.-H. Teng, is a new form of analysis, that arguably blends the best of both worst-case and average-case.

Let $f : \mathbb{R}^p \longrightarrow \mathbb{R}_+$ be a function (running time, condition number). Instead of showing

> *"it is unlikely that $f(a)$ will be large"*

one shows that

> *"for all a and all slight random perturbations Δa, it is unlikely that $f(a + \Delta a)$ will be large."*

Smoothed Analysis

Smoothed analysis , proposed by D. Spielman and S.-H. Teng, is a new form of analysis, that arguably blends the best of both worst-case and average-case.
Let $f : \mathbb{R}^p \to \mathbb{R}_+$ be a function (running time, condition number).
Instead of showing

"it is unlikely that $f(a)$ will be large"

one shows that

"for all a and all slight random perturbations Δa, it is unlikely that $f(a + \Delta a)$ will be large."

Worst case analysis	Average case analysis	Smoothed analysis
$\sup_{a \in \mathbb{R}^p} f(a)$	$\mathbf{E}_{a \in \Phi} f(a)$	$\sup_{a \in \mathbb{R}^p} \mathbf{E}_{z \in N(0,\sigma^2)} f(a+z)$

Smoothed Analysis

Smoothed analysis , proposed by D. Spielman and S.-H. Teng, is a new form of analysis, that arguably blends the best of both worst-case and average-case.
Let $f : \mathbb{R}^p \to \mathbb{R}_+$ be a function (running time, condition number).
Instead of showing

> *"it is unlikely that $f(a)$ will be large"*

one shows that

> *"for all a and all slight random perturbations Δa, it is unlikely that $f(a + \Delta a)$ will be large."*

Worst case analysis	Average case analysis	Smoothed analysis
$\sup_{a \in \mathbb{R}^p} f(a)$	$\mathbf{E}_{a \in \Phi} f(a)$	$\sup_{a \in \mathbb{R}^p} \mathbf{E}_{z \in N(0,\sigma^2)} f(a+z)$

Φ distribution on $\mathbb{R}^p$, $N(0, \sigma^2)$ Gaussian distribution.

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming
- Wschebor (2004): Smoothed analysis of the classical condition number

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming
- Wschebor (2004): Smoothed analysis of the classical condition number
- Cucker, Diao, and Wei (2005): Smoothed analysis of other condition numbers

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming
- Wschebor (2004): Smoothed analysis of the classical condition number
- Cucker, Diao, and Wei (2005): Smoothed analysis of other condition numbers
- Damerow, Meyer auf der Heyde, Raecke, Scheideler, Sohler (2003): Smoothed motion complexity

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming
- Wschebor (2004): Smoothed analysis of the classical condition number
- Cucker, Diao, and Wei (2005): Smoothed analysis of other condition numbers
- Damerow, Meyer auf der Heyde, Raecke, Scheideler, Sohler (2003): Smoothed motion complexity
- Bürgisser, Cucker, and Lotz (2006, 2007): Smoothed analysis of conic condition numbers

Previous Work on Smoothed Analysis

- Spielman and Teng (2001, 2004): Smoothed analysis of the Simplex algorithm
- Dunagan, Spielman, and Teng (2003): Smoothed analysis of Renegar's condition number for linear programming
- Wschebor (2004): Smoothed analysis of the classical condition number
- Cucker, Diao, and Wei (2005): Smoothed analysis of other condition numbers
- Damerow, Meyer auf der Heyde, Raecke, Scheideler, Sohler (2003): Smoothed motion complexity
- Bürgisser, Cucker, and Lotz (2006, 2007): Smoothed analysis of conic condition numbers
- and some more ...

Geometric Interpretation of Condition Numbers

The condition number of a square matrix A allows a geometric interpretation. The set of singular matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$ is the set of ill-posed problems.

Geometric Interpretation of Condition Numbers

The condition number of a square matrix A allows a geometric interpretation. The set of singular matrices $\Sigma \subseteq \mathbb{R}^{n\times n}$ is the set of ill-posed problems.

The Eckart-Young Theorem states that

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

where dist refers to the Frobenius norm $\|A\|^2 := \sum_{ij} a_{ij}^2$ (coming from the canonical scalar product on $\mathbb{R}^{n\times n}$).

Geometric Interpretation of Condition Numbers

The condition number of a square matrix A allows a geometric interpretation. The set of singular matrices $\Sigma \subseteq \mathbb{R}^{n\times n}$ is the set of ill-posed problems.

The Eckart-Young Theorem states that

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

where dist refers to the Frobenius norm $\|A\|^2 := \sum_{ij} a_{ij}^2$ (coming from the canonical scalar product on $\mathbb{R}^{n\times n}$).
The modified condition number

$$\kappa_F(A) := \|A\|_F \, \|A^{-1}\| = \frac{\|A\|_F}{\text{dist}(A, \Sigma)},$$

differs from $\kappa(A)$ at most by a factor of $\sqrt{n}$.

Conic Condition Numbers 1

- We consider an abstract setting with input space $\mathbb{R}^{p+1}$, together with a symmetric cone $\Sigma \subseteq \mathbb{R}^{p+1}$, the set of "ill-posed problems".

Conic Condition Numbers 1

- We consider an abstract setting with input space $\mathbb{R}^{p+1}$, together with a symmetric cone $\Sigma \subseteq \mathbb{R}^{p+1}$, the set of "ill-posed problems".
- We define the associated conic condition number $\mathscr{C}(a)$ of an input $a \in \mathbb{R}^p$ as

$$\mathscr{C}(a) := \frac{\|a\|}{\text{dist}(a, \Sigma)},$$

where the norm and distance are induced by the canonical inner product.

Conic Condition Numbers 1

- We consider an abstract setting with input space $\mathbb{R}^{p+1}$, together with a symmetric cone $\Sigma \subseteq \mathbb{R}^{p+1}$, the set of "ill-posed problems".
- We define the associated conic condition number $\mathscr{C}(a)$ of an input $a \in \mathbb{R}^p$ as

$$\mathscr{C}(a) := \frac{\|a\|}{\text{dist}(a, \Sigma)},$$

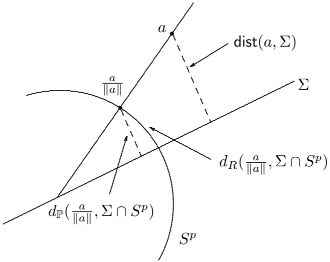
  where the norm and distance are induced by the canonical inner product.
- Let $\Sigma \subseteq \mathbb{R}^{n \times n}$ be the set of singular matrices A. The condition number $\kappa_F(A)$ is conic by the Eckart-Young Theorem:

$$\kappa_F(A) = \|A\|_F \|A^{-1}\| = \frac{\|A\|_F}{\text{dist}_F(A, \Sigma)}.$$

Conic Condition Numbers 2

$$\mathscr{C}(a) = \frac{\|a\|}{\mathsf{dist}(a,\Sigma)} = \frac{1}{d_{\mathbb{P}}(\frac{a}{\|a\|}, \Sigma \cap S^p)},$$

where $d_{\mathbb{P}}(x,y) = \sin d_R(x,y)$ denotes the projective distance on the sphere $S^p := \{x \in \mathbb{R}^{p+1} \mid \|x\| = 1\}$. Hence, we may restrict to data a lying on the sphere S^p.



Uniform Smoothed Analysis

- For $a \in S^p$ and $\sigma > 0$ let

$$B(a, \sigma) := \{z \in S^p \mid d_{\mathbb{P}^p}(z, a) \leq \sigma\}.$$

denote the ball of radius σ around a in S^p.

Uniform Smoothed Analysis

- For $a \in S^p$ and $\sigma > 0$ let
$$B(a,\sigma) := \{z \in S^p \mid d_{\mathbb{P}^p}(z,a) \le \sigma\}.$$
denote the ball of radius σ around a in S^p.
- Uniform smoothed analysis of a conic condition number $\mathscr{C}$ is the study of
$$\sup_{a \in S^p} \mathop{\mathbf{E}}_{z \in B(a,\sigma)} \ln \mathscr{C}(z),$$
where $z \in B(a,\sigma)$ means that z is uniformly distributed in $B(a,\sigma)$.

Uniform Smoothed Analysis

- For $a \in S^p$ and $\sigma > 0$ let

$$B(a,\sigma) := \{z \in S^p \mid d_{\mathbb{P}^p}(z,a) \leq \sigma\}.$$

denote the ball of radius σ around a in S^p.

- Uniform smoothed analysis of a conic condition number $\mathscr{C}$ is the study of

$$\sup_{a \in S^p} \mathop{\mathbf{E}}_{z \in B(a,\sigma)} \ln \mathscr{C}(z),$$

where $z \in B(a,\sigma)$ means that z is uniformly distributed in $B(a,\sigma)$.

  - $\sigma = 0$ yields worst-case analysis

Uniform Smoothed Analysis

- For $a \in S^p$ and $\sigma > 0$ let

$$B(a,\sigma) := \{z \in S^p \mid d_{\mathbb{P}^p}(z,a) \leq \sigma\}.$$

denote the ball of radius σ around a in S^p.

- Uniform smoothed analysis of a conic condition number $\mathscr{C}$ is the study of

$$\sup_{a \in S^p} \mathop{\mathbf{E}}_{z \in B(a,\sigma)} \ln \mathscr{C}(z),$$

where $z \in B(a,\sigma)$ means that z is uniformly distributed in $B(a,\sigma)$.

  - $\sigma = 0$ yields worst-case analysis
  - $\sigma = 1$ yields average-case analysis

Uniform Smoothed Analysis

- For $a \in S^p$ and $\sigma > 0$ let
$$B(a,\sigma) := \{z \in S^p \mid d_{\mathbb{P}^p}(z,a) \leq \sigma\}.$$
denote the ball of radius σ around a in S^p.
- Uniform smoothed analysis of a conic condition number $\mathscr{C}$ is the study of
$$\sup_{a \in S^p} \mathop{\mathbf{E}}_{z \in B(a,\sigma)} \ln \mathscr{C}(z),$$
where $z \in B(a,\sigma)$ means that z is uniformly distributed in $B(a,\sigma)$.
  - $\sigma = 0$ yields worst-case analysis
  - $\sigma = 1$ yields average-case analysis
- For any $a \in S^p$ and $\sigma > 0$, we are thus interested in studying the distribution of the random variables $\mathscr{C}(z)$ and $\ln \mathscr{C}(z)$ on $B(a,\sigma)$.

Main Result

Main Result

Theorem

Let $\mathscr{C}$ be a conic condition number with set of ill-posed inputs Σ. Assume that $\Sigma \cap S^p \subseteq W$ where $W \subset S^p$ is the zero set in S^p of homogeneous polynomials of degree at most d.
Then, for all $\sigma \in (0,1]$ and all $t \geq (2d+1)\frac{p}{\sigma}$,

$$\sup_{a \in S^p} \underset{z \in B(a,\sigma)}{\mathrm{Prob}} \{\mathscr{C}(z) \geq t\} \leq 26\, dp\, \frac{1}{\sigma t}.$$

and

$$\sup_{a \in S^p} \underset{z \in B(a,\sigma)}{\mathbf{E}} (\ln \mathscr{C}(z)) \leq 2 \ln p + 2 \ln d + 2 \ln \frac{1}{\sigma} + 4.7.$$

Main Result

Theorem

Let $\mathscr{C}$ be a conic condition number with set of ill-posed inputs Σ. Assume that $\Sigma \cap S^p \subseteq W$ where $W \subset S^p$ is the zero set in S^p of homogeneous polynomials of degree at most d.
Then, for all $\sigma \in (0,1]$ and all $t \geq (2d+1)\frac{p}{\sigma}$,

$$\sup_{a \in S^p} \underset{z \in B(a,\sigma)}{\mathrm{Prob}} \{\mathscr{C}(z) \geq t\} \leq 26\, dp \, \frac{1}{\sigma t}.$$

and

$$\sup_{a \in S^p} \underset{z \in B(a,\sigma)}{\mathbf{E}} (\ln \mathscr{C}(z)) \leq 2 \ln p + 2 \ln d + 2 \ln \frac{1}{\sigma} + 4.7.$$

Similar bounds on the tail of the distribution of $\mathscr{C}(z)$ in the framework of average case analysis ($\sigma = 1$) in terms of these parameters have been given by Demmel (1988) and by Beltrán-Pardo (2005) (over $\mathbb{C}$).

Applications

Linear Equation Solving

- Problem: Solving the system of equations $Ax = b$, $A \in \mathbb{R}^{n \times n}$.
- Set of ill-posed inputs: $\Sigma = \{A \in \mathbb{R}^{n \times n} \mid \det A = 0\}$
- Condition number: $\kappa_F(A) = \|A\|_F \|A^{-1}\|$.

Corollary

For all $A \in \mathbb{R}^{n \times n}$ of Frobenius norm one and $0 < \sigma \leq 1$ we have

$$\mathop{\mathbf{E}}_{Z \in B(A,\sigma)} (\ln \kappa_F(Z)) \leq 6 \ln n + 2 \ln \frac{1}{\sigma} + 4.7.$$

Linear Equation Solving

- Problem: Solving the system of equations $Ax = b$, $A \in \mathbb{R}^{n \times n}$.
- Set of ill-posed inputs: $\Sigma = \{A \in \mathbb{R}^{n \times n} \mid \det A = 0\}$
- Condition number: $\kappa_F(A) = \|A\|_F \|A^{-1}\|$.

Corollary

For all $A \in \mathbb{R}^{n \times n}$ of Frobenius norm one and $0 < \sigma \le 1$ we have

$$\mathop{\mathbf{E}}_{Z \in B(A,\sigma)} (\ln \kappa_F(Z)) \le 6 \ln n + 2 \ln \frac{1}{\sigma} + 4.7.$$

Proof: Apply the theorem with $p = n^2 - 1$, $W = \Sigma$, and $d = n$. □

Linear Equation Solving

- Problem: Solving the system of equations $Ax = b$, $A \in \mathbb{R}^{n \times n}$.
- Set of ill-posed inputs: $\Sigma = \{A \in \mathbb{R}^{n \times n} \mid \det A = 0\}$
- Condition number: $\kappa_F(A) = \|A\|_F \|A^{-1}\|$.

Corollary

For all $A \in \mathbb{R}^{n \times n}$ of Frobenius norm one and $0 < \sigma \leq 1$ we have

$$\mathop{\mathbf{E}}_{Z \in B(A,\sigma)} (\ln \kappa_F(Z)) \leq 6 \ln n + 2 \ln \frac{1}{\sigma} + 4.7.$$

Proof: Apply the theorem with $p = n^2 - 1$, $W = \Sigma$, and $d = n$. □

M. Wschebor derived similar bounds for Gaussian perturbations by direct methods (2004).

Eigenvalue Computations

- Problem: Compute the (complex) eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$.
- Set of ill-posed inputs: Matrices A having multiple eigenvalues.
- Condition number (Wilkinson): Satisfies $\kappa_{\text{eigen}}(A) \leq \frac{\sqrt{2}\,\|A\|_F}{\text{dist}(A,\Sigma)}$.

Corollary

For all $A \in \mathbb{R}^{n \times n}$ and $0 < \sigma \leq 1$ we have

$$\mathop{\mathbf{E}}_{Z \in B(A,\sigma)} (\ln \kappa_{\text{eigen}}(Z)) \leq 8 \ln n + 2 \ln \frac{1}{\sigma} + 5.1.$$

Eigenvalue Computations

- Problem: Compute the (complex) eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$.
- Set of ill-posed inputs: Matrices A having multiple eigenvalues.
- Condition number (Wilkinson): Satisfies $\kappa_{\text{eigen}}(A) \leq \frac{\sqrt{2}\,\|A\|_F}{\text{dist}(A,\Sigma)}$.

Corollary

For all $A \in \mathbb{R}^{n \times n}$ and $0 < \sigma \leq 1$ we have

$$\mathop{\mathbf{E}}_{Z \in B(A,\sigma)} (\ln \kappa_{\text{eigen}}(Z)) \leq 8 \ln n + 2 \ln \frac{1}{\sigma} + 5.1.$$

Proof: $W = \Sigma$ is the zeroset of the discriminant of the characteristic polynomial, which has degree $d = n^2 - n$. Apply the theorem. □

Complex Polynomial Systems

- Fix $d_1, \dots, d_n \in \mathbb{N} \setminus \{0\}$. We denote by $\mathcal{H}_{\mathbf{d}}$ the vector space of polynomial systems $f = (f_1, \dots, f_n)$ with $f_i \in \mathbb{C}[X_0, \dots, X_n]$ homogeneous of degree d_i. $\mathcal{H}_{\mathbf{d}}$ carries an invariant Hermitian product.

Complex Polynomial Systems

- Fix $d_1, \ldots, d_n \in \mathbb{N} \setminus \{0\}$. We denote by $\mathcal{H}_{\mathbf{d}}$ the vector space of polynomial systems $f = (f_1, \ldots, f_n)$ with $f_i \in \mathbb{C}[X_0, \ldots, X_n]$ homogeneous of degree d_i. $\mathcal{H}_{\mathbf{d}}$ carries an invariant Hermitian product.
- In a seminal series of papers, M. Shub and S. Smale studied the problem of, given $f \in \mathcal{H}_{\mathbf{d}}$, compute an approximation of a complex zero of f. They proposed an algorithm and studied its complexity in terms of a condition number $\mu_{\mathrm{norm}}(f)$ for f.

Complex Polynomial Systems

- Fix $d_1, \ldots, d_n \in \mathbb{N} \setminus \{0\}$. We denote by $\mathcal{H}_{\mathbf{d}}$ the vector space of polynomial systems $f = (f_1, \ldots, f_n)$ with $f_i \in \mathbb{C}[X_0, \ldots, X_n]$ homogeneous of degree d_i. $\mathcal{H}_{\mathbf{d}}$ carries an invariant Hermitian product.
- In a seminal series of papers, M. Shub and S. Smale studied the problem of, given $f \in \mathcal{H}_{\mathbf{d}}$, compute an approximation of a complex zero of f. They proposed an algorithm and studied its complexity in terms of a condition number $\mu_{\mathrm{norm}}(f)$ for f.

Corollary

For all $f \in \mathcal{H}_{\mathbf{d}}$ of norm one we have

$$\mathop{\mathbf{E}}_{g \in B(f,\sigma)} (\ln \mu_{\mathrm{norm}}(g)) \leq 2 \ln N + 4 \ln \mathcal{D} + 2 \ln n + 2 \ln \frac{1}{\sigma} + 6.1.$$

where $N = \dim \mathcal{H}_{\mathbf{d}} - 1$ and $\mathcal{D} = d_1 \cdots d_n$ is the Bézout number.

Complex Polynomial Systems

- Fix $d_1, \ldots, d_n \in \mathbb{N} \setminus \{0\}$. We denote by $\mathcal{H}_{\mathbf{d}}$ the vector space of polynomial systems $f = (f_1, \ldots, f_n)$ with $f_i \in \mathbb{C}[X_0, \ldots, X_n]$ homogeneous of degree d_i. $\mathcal{H}_{\mathbf{d}}$ carries an invariant Hermitian product.
- In a seminal series of papers, M. Shub and S. Smale studied the problem of, given $f \in \mathcal{H}_{\mathbf{d}}$, compute an approximation of a complex zero of f. They proposed an algorithm and studied its complexity in terms of a condition number $\mu_{\mathrm{norm}}(f)$ for f.

Corollary

For all $f \in \mathcal{H}_{\mathbf{d}}$ of norm one we have

$$\mathop{\mathbf{E}}_{g \in B(f,\sigma)} (\ln \mu_{\mathrm{norm}}(g)) \leq 2 \ln N + 4 \ln \mathcal{D} + 2 \ln n + 2 \ln \frac{1}{\sigma} + 6.1.$$

where $N = \dim \mathcal{H}_{\mathbf{d}} - 1$ and $\mathcal{D} = d_1 \cdots d_n$ is the Bézout number.

S. Smale and M. Shub obtained similar estimates for average complexity.

Outline of Proof of Main Theorem

Geometric Reformulation of the Problem

Given a conic condition number $\mathscr{C}$ with ill-posed set Σ. Let $W \subset S^p$ such that $\Sigma \cap S^p \subseteq W$. Then for $z \in S^p$

$$\mathscr{C}(z) \geq \frac{1}{\varepsilon} \Longleftrightarrow d_{\mathbb{P}^p}(z, \Sigma \cap S^p) \leq \varepsilon \Rightarrow d_{\mathbb{P}^p}(z, W) \leq \varepsilon.$$

Geometric Reformulation of the Problem

Given a conic condition number $\mathscr{C}$ with ill-posed set Σ. Let $W \subset S^p$ such that $\Sigma \cap S^p \subseteq W$. Then for $z \in S^p$

$$\mathscr{C}(z) \geq \frac{1}{\varepsilon} \Longleftrightarrow d_{\mathbb{P}^p}(z, \Sigma \cap S^p) \leq \varepsilon \Rightarrow d_{\mathbb{P}^p}(z, W) \leq \varepsilon.$$

We denote by $T(W, \varepsilon)$ the ε-neighborhood around W in S^p:

$$T(W, \varepsilon) := \{z \in S^p \mid d_{\mathbb{P}^p}(z, \Sigma) < \varepsilon\}.$$

Geometric Reformulation of the Problem

Given a conic condition number $\mathscr{C}$ with ill-posed set Σ. Let $W \subset S^p$ such that $\Sigma \cap S^p \subseteq W$. Then for $z \in S^p$

$$\mathscr{C}(z) \geq \frac{1}{\varepsilon} \Longleftrightarrow d_{\mathbb{P}^p}(z, \Sigma \cap S^p) \leq \varepsilon \Rightarrow d_{\mathbb{P}^p}(z, W) \leq \varepsilon.$$

We denote by $T(W, \varepsilon)$ the ε-neighborhood around W in S^p:

$$T(W, \varepsilon) := \{z \in S^p \mid d_{\mathbb{P}^p}(z, \Sigma) < \varepsilon\}.$$

Then:

$$\operatorname*{Prob}_{z \in B(a,\sigma)} \left\{ \mathscr{C}(z) \geq \frac{1}{\varepsilon} \right\} \leq \operatorname*{Prob}_{z \in B(a,\sigma)} \{ d_{\mathbb{P}^p}(z, W) \leq \varepsilon \} = \frac{\mathrm{vol}(T(W, \varepsilon) \cap B(a, \sigma))}{\mathrm{vol}(B(a, \sigma))}.$$

Geometric Version of Main Result

Let $\mathcal{O}_p := \mathrm{vol}(S^p) = \frac{2\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)}$ denote the volume of S^p.

Geometric Version of Main Result

Let $\mathcal{O}_p := \mathrm{vol}(S^p) = \frac{2\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)}$ denote the volume of S^p.

The main theorem follows from the following purely geometric statement on the volume of patches of tubes around subvarieties in spheres.

Geometric Version of Main Result

Let $\mathcal{O}_p := \mathrm{vol}(S^p) = \frac{2\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)}$ denote the volume of S^p.

The main theorem follows from the following purely geometric statement on the volume of patches of tubes around subvarieties in spheres.

Theorem'

Let $W \subset S^p$ be a real algebraic variety defined by homogeneous polynomials of degree at most $d \geq 1$. Then we have for $a \in S^p$ and $0 < \varepsilon, \sigma \leq 1$

$$\frac{\mathrm{vol}\left(T(W,\varepsilon) \cap B(a,\sigma)\right)}{\mathrm{vol} B(a,\sigma)} \leq 4 \sum_{k=1}^{p-1} \binom{p}{k} (2d)^k \left(1 + \frac{\varepsilon}{\sigma}\right)^{p-k} \left(\frac{\varepsilon}{\sigma}\right)^k + \frac{p\mathcal{O}_p}{\mathcal{O}_{p-1}} (2d)^p \left(\frac{\varepsilon}{\sigma}\right)^p.$$

The three main steps of the proof

- Upper bound on the volume of an ε-neighborhood of a smooth hypersurface in terms of integrals of absolute curvature.
  This is a variation of H. Weyl's exact formula for the volume of tubes, a formula which, however, only holds for sufficiently small ε.

The three main steps of the proof

I. Upper bound on the volume of an ε-neighborhood of a smooth hypersurface in terms of integrals of absolute curvature.
This is a variation of H. Weyl's exact formula for the volume of tubes, a formula which, however, only holds for sufficiently small ε.

II. Estimation of integrals of absolute curvature.
This is based on the kinematic formula of integral geometry and Bézout's theorem.

The three main steps of the proof

I Upper bound on the volume of an ε-neighborhood of a smooth hypersurface in terms of integrals of absolute curvature.
This is a variation of H. Weyl's exact formula for the volume of tubes, a formula which, however, only holds for sufficiently small ε.

II Estimation of integrals of absolute curvature.
This is based on the kinematic formula of integral geometry and Bézout's theorem.

III Remove the smoothness assumption by some perturbation argument.

Some differential geometry of hypersurfaces on spheres

- Let M be a compact oriented smooth hypersurface of S^p interpreted as a Riemannian submanifold. Denote by $\kappa_1(x), \ldots, \kappa_{p-1}(x)$ the principal curvatures at x of M.

Some differential geometry of hypersurfaces on spheres

- Let M be a compact oriented smooth hypersurface of S^p interpreted as a Riemannian submanifold. Denote by $\kappa_1(x), \ldots, \kappa_{p-1}(x)$ the principal curvatures at x of M.
- For $1 \leq i < p$ we define the ith curvature $K_{M,i}(x)$ of M at x as the ith elementary symmetric polynomial in $\kappa_1(x), \ldots, \kappa_{p-1}(x)$, and put $K_{M,0}(x) := 1$.

Some differential geometry of hypersurfaces on spheres

- Let M be a compact oriented smooth hypersurface of S^p interpreted as a Riemannian submanifold. Denote by $\kappa_1(x), \ldots, \kappa_{p-1}(x)$ the principal curvatures at x of M.
- For $1 \leq i < p$ we define the ith curvature $K_{M,i}(x)$ of M at x as the ith elementary symmetric polynomial in $\kappa_1(x), \ldots, \kappa_{p-1}(x)$, and put $K_{M,0}(x) := 1$.
- The integral of ith absolute curvature over the open subset U of M is defined as

$$|\mu_i|(U) := \int_U |K_{M,i}| \, dM.$$

Some differential geometry of hypersurfaces on spheres

- Let M be a compact oriented smooth hypersurface of S^p interpreted as a Riemannian submanifold. Denote by $\kappa_1(x), \ldots, \kappa_{p-1}(x)$ the principal curvatures at x of M.
- For $1 \leq i < p$ we define the ith curvature $K_{M,i}(x)$ of M at x as the ith elementary symmetric polynomial in $\kappa_1(x), \ldots, \kappa_{p-1}(x)$, and put $K_{M,0}(x) := 1$.
- The integral of ith absolute curvature over the open subset U of M is defined as

$$|\mu_i|(U) := \int_U |K_{M,i}| \, dM.$$

- Note $|\mu_0|(U) = \mathrm{vol}(U)$.

Bounding the volume of tubes

The ε-tube $T^{\perp}(M, \varepsilon)$ around M is a subset of $T(M, \varepsilon)$ obtained by "cutting off points close to the boundary" of M:

Bounding the volume of tubes

The ε-tube $T^{\perp}(M,\varepsilon)$ around M is a subset of $T(M,\varepsilon)$ obtained by "cutting off points close to the boundary" of M:

Variant of Weyl's Tube Formula

Let M be a compact, oriented, smooth hypersurface of S^p and U be an open subset of M. For all $0 < \varepsilon \leq 1$

$$\mathrm{vol}(T^{\perp}(U,\varepsilon)) \leq 2\sum_{i=0}^{p-1} J_{p,i+1}(\varepsilon)\cdot|\mu_i|(U),$$

where

$$J_{p,k}(\varepsilon) := \int_0^{\arcsin\varepsilon} (\sin\rho)^{k-1}(\cos\rho)^{p-k}\,d\rho.$$

Bounding the volume of tubes

The ε-tube $T^{\perp}(M, \varepsilon)$ around M is a subset of $T(M, \varepsilon)$ obtained by "cutting off points close to the boundary" of M:

Variant of Weyl's Tube Formula

Let M be a compact, oriented, smooth hypersurface of S^p and U be an open subset of M. For all $0 < \varepsilon \leq 1$

$$\mathrm{vol}(T^{\perp}(U, \varepsilon)) \leq 2 \sum_{i=0}^{p-1} J_{p,i+1}(\varepsilon) \cdot |\mu_i|(U),$$

where

$$J_{p,k}(\varepsilon) := \int_0^{\arcsin \varepsilon} (\sin \rho)^{k-1} (\cos \rho)^{p-k} d\rho.$$

Weyl's formula gives the exact volume of tubes, but only holds for sufficiently small radius.

Bounding the volume of tubes

The ε-tube $T^{\perp}(M,\varepsilon)$ around M is a subset of $T(M,\varepsilon)$ obtained by "cutting off points close to the boundary" of M:

Variant of Weyl's Tube Formula

Let M be a compact, oriented, smooth hypersurface of S^p and U be an open subset of M. For all $0 < \varepsilon \leq 1$

$$\mathrm{vol}(T^{\perp}(U,\varepsilon)) \leq 2\sum_{i=0}^{p-1} J_{p,i+1}(\varepsilon)\cdot|\mu_i|(U),$$

where

$$J_{p,k}(\varepsilon) := \int_0^{\arcsin\varepsilon} (\sin\rho)^{k-1}(\cos\rho)^{p-k}\,d\rho.$$

Weyl's formula gives the exact volume of tubes, but only holds for sufficiently small radius.

The leading term is $2J_{p,1}(\varepsilon)\cdot|\mu_0|(U) \approx 2\varepsilon\,\mathrm{vol}(U)$.

Bounding integrals of absolute curvature

Let $f \in \mathbb{R}[X_0, \ldots, X_p]$ be homogeneous of degree $d > 0$ with nonempty zero set $V \subseteq S^p$ such that the derivative of the restriction of f to S^p does not vanish on V. Then V is a compact smooth hypersurface of S^p.

Bounding integrals of absolute curvature

Let $f \in \mathbb{R}[X_0, \ldots, X_p]$ be homogeneous of degree $d > 0$ with nonempty zero set $V \subseteq S^p$ such that the derivative of the restriction of f to S^p does not vanish on V. Then V is a compact smooth hypersurface of S^p.

Proposition

For $a \in S^p$, $0 < \sigma \leq 1$, and $0 \leq i < p$ we have

$$|\mu_i|(V \cap B_{\mathbb{P}}(a,\sigma)) \;\leq\; 2\binom{p-1}{i}\mathcal{O}_{p-1}\, d^{i+1}\, \sigma^{p-i-1}.$$

Bounding integrals of absolute curvature

Let $f \in \mathbb{R}[X_0, \ldots, X_p]$ be homogeneous of degree $d > 0$ with nonempty zero set $V \subseteq S^p$ such that the derivative of the restriction of f to S^p does not vanish on V. Then V is a compact smooth hypersurface of S^p.

Proposition

For $a \in S^p$, $0 < \sigma \leq 1$, and $0 \leq i < p$ we have

$$|\mu_i|(V \cap B_{\mathbb{P}}(a, \sigma)) \;\leq\; 2\binom{p-1}{i}\mathcal{O}_{p-1}\, d^{i+1}\, \sigma^{p-i-1}.$$

Proof ingredients:

- Principal kinematic formula of integral geometry for spheres
- Bézout's theorem

Crofton's formula from integral geometry

We denote by dG the volume element on the orthogonal group $G = O(p+1)$ (compact Lie group), normalized such that the volume of G equals one. G operates on S^p in the natural way.

Crofton's formula from integral geometry

We denote by dG the volume element on the orthogonal group $G = O(p+1)$ (compact Lie group), normalized such that the volume of G equals one. G operates on S^p in the natural way.

Crofton's formula

Let T be a submanifold of S^p with dim $T = p - 1$. Then

$$\frac{\mathrm{vol}_{p-1}(T)}{\mathcal{O}_{p-1}} = \frac{1}{2} \int_{g \in G} \#(T \cap gS^1)\, dG(g).$$

Crofton's formula from integral geometry

We denote by dG the volume element on the orthogonal group $G = O(p+1)$ (compact Lie group), normalized such that the volume of G equals one. G operates on S^p in the natural way.

Crofton's formula

Let T be a submanifold of S^p with $\dim T = p-1$. Then

$$\frac{\mathrm{vol}_{p-1}(T)}{\mathcal{O}_{p-1}} = \frac{1}{2}\int_{g\in G} \#(T \cap gS^1)\, dG(g).$$

- This allows to bound $\mathrm{vol}(V \cap B_{\mathbb{P}}(a,\sigma))$ via Bézout's theorem.

Crofton's formula from integral geometry

We denote by dG the volume element on the orthogonal group $G = O(p+1)$ (compact Lie group), normalized such that the volume of G equals one. G operates on S^p in the natural way.

Crofton's formula

Let T be a submanifold of S^p with $\dim T = p-1$. Then

$$\frac{\mathrm{vol}_{p-1}(T)}{\mathcal{O}_{p-1}} = \frac{1}{2}\int_{g\in G} \#(T \cap gS^1)\, dG(g).$$

- This allows to bound $\mathrm{vol}(V \cap B_{\mathbb{P}}(a,\sigma))$ via Bézout's theorem.
- To bound μ_i for $i > 0$ we need a more sophisticated tool.

Principal kinematic formula from integral geometry

A far reaching generalization of Crofton's formula was obtained by Chern for Euclidean space.

Principal kinematic formula from integral geometry

A far reaching generalization of Crofton's formula was obtained by Chern for Euclidean space.
We will need the following version of this formula for spheres (cf. Santaló, Howard).

Principal kinematic formula

Let U be an open subset of a compact oriented smooth hypersurface M of S^p and $0 \leq i < p-1$. Then we have

$$\mu_i(U) = \mathcal{C}(p,i) \int_{g \in G} \mu_i(gU \cap S^{i+1})\, dG(g),$$

where $\mathcal{C}(p,i) = (p-i-1)\binom{p-1}{i} \frac{\mathcal{O}_{p-1}\mathcal{O}_p}{\mathcal{O}_i \mathcal{O}_{i+1} \mathcal{O}_{p-i-2}}$.

Principal kinematic formula from integral geometry

A far reaching generalization of Crofton's formula was obtained by Chern for Euclidean space.
We will need the following version of this formula for spheres (cf. Santaló, Howard).

Principal kinematic formula

Let U be an open subset of a compact oriented smooth hypersurface M of S^p and $0 \le i < p-1$. Then we have

$$\mu_i(U) = \mathcal{C}(p,i) \int_{g \in G} \mu_i(gU \cap S^{i+1})\, dG(g),$$

where $\mathcal{C}(p,i) = (p-i-1)\binom{p-1}{i} \frac{\mathcal{O}_{p-1}\mathcal{O}_p}{\mathcal{O}_i \mathcal{O}_{i+1} \mathcal{O}_{p-i-2}}$.

This allows to reduce the estimation of μ_i to the case of codimension one, i.e., Gaussian curvature, which can be treated directly.

References

References for the results presented and some relevant previous work:

- C. Beltrán and L.M. Pardo. Upper bounds on the distribution of the condition number of singular matrices. To appear in *Found. Comput. Math.*.
- P. Bürgisser, F. Cucker, and M. Lotz. The probability that a slight perturbation of a numerical analysis problem is difficult. Accepted for *Mathematics of Computation*
- P. Bürgisser, F. Cucker, and M. Lotz. General formulas for the smoothed analysis of condition numbers. *Comptes rendus de l'Académie des sciences Paris*, Ser. I 343, 145-150 (2006).
- P. Bürgisser, F. Cucker, and M. Lotz. Smoothed analysis of complex conic condition numbers *Journal de Mathématiques Pures et Appliquées* 86: 293-309 (2006).
- J. Demmel. The probability that a numerical analysis problem is difficult. *Mathematics of Computation* 50: 449-480 (1988).
- A. Gray. *Tubes*. Progress in Mathematics, vol. 221, 2004.
- A. Ocneanu (1984, lecture at MSRI, unpublished).

- J. Renegar. On the efficiency of Newton's method in approximating all zeros of systems of complex polynomials. *Math. of Oper. Research* 12: 121–148 (1987).
- L. A. Santaló. *Integral Geometry and Geometric Probability*. Encyclopedia of Mathematics and its Applications, vol. 1, 1976.
- M. Shub and S. Smale. Complexity of Bézout's theorem II: volumes and probabilities. In *Computational Algebraic Geometry*, pp. 267–285. Progress in Mathematics 109, Birkhäuser, 1993
- S. Smale. The fundamental theorem of algebra and complexity theory. *Bulletin of the AMS* 4: 1–36 (1981).
- D.A. Spielman and S.-H. Teng. Smoothed Analysis: Why The Simplex Algorithm Usually Takes Polynomial Time. *Journal of the ACM* 51: 385–463 (2004).
- H. Weyl. On the volume of tubes. *Amer. J. Math.* 61: 461-472 (1939).
- R. Wongkew. Volumes of tubular neighbourhoods of real algebraic varieties. *Pacific J. of Mathematics* 159: 177–184 (2003).
- M. Wschebor. Smoothed Analysis of $\kappa(A)$. *J. of Complexity* 20: 97–107 (2004).