Condition Numbers
Tutorial Talk

Peter Bürgisser

Technical University of Berlin

Perspectives on Matrix Computations:

Theoretical Computer Science Meets Numerical Analysis

Banff, March 7, 2023

Outline

Turing's condition number of a matrix

A. Turing, 1948

J. von Neumann and H. Goldstine, 1947

General definition of condition number

- Suppose we have a numerical computation problem

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

  and input x has small error Δx.
- Size of errors can be measured in different ways: absolute or relative errors, componentwise or normwise ...
- Will focus on *normwise relative error* $\|\Delta x\|/\|x\|$, which depends on choice of norm.
- Want to bound relative error $\|\Delta y\|/\|y\|$ of output in terms of relative error $\|\Delta x\|/\|x\|$ of input.
- This is done by the normwise relative condition number $\kappa(f, x)$ at x:

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

- Formal definition for differentiable f:

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

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

Turing's condition number

- Number inversion $f\colon \mathbb{R}\setminus\{0\}\to\mathbb{R}, x\mapsto x^{-1}$ has condition number

$$\kappa(f,x) \;=\; |f'(x)|\,\frac{|x|}{|f(x)|} = 1.$$

- Matrix inversion

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

has derivative $Df(A)(\dot{A}) = -A^{-1}\dot{A}A^{-1}$, hence using spectral norm, $\|Df(A)\| = \|A^{-1}\|^2$.

- Obtain classical condition number of A:

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

- Note that $\kappa(\lambda A) = \kappa(A)$ for $\lambda\in\mathbb{R}^*$.
- $\kappa(A)$ was introduced by A. Turing in 1948.

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

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

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

- Hence

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

- $\mathrm{dist}(A, \Sigma)$ equals the *smallest singular value* of A.

Finite precision

- Digital computers operate with floating-point numbers, and every arithmetic operations produces a round-off error.
- Let ϵ_{mach} denote the round-off unit (e.g., 10^{-12}).
- Suppose we compute the approximation $\widetilde{x}$ of $x \in \mathbb{R}$ with relative error δ, i.e. $\widetilde{x} = x(1+\delta)$.
- The best we can hope for is $\delta \leq \frac{1}{2}\epsilon_{\text{mach}}$.
- One calls $\log_{10}\left(\frac{\delta}{\epsilon_{\text{mach}}}\right)$ the loss of precision in decimal digits.
- Turing's condition number is relevant for finite precision analysis of linear algebra

Backward-error analysis and condition

- Suppose $\mathcal{A}$ is a finite-precision algorithm approximately computing the function $f\colon \mathbb{R}^p \to \mathbb{R}^q$.
- Suppose we can show that for all inputs x there exists $e \in \mathbb{R}^p$ such that $A(x) = f(x+e)$ with small e (called "backward-error").
- Can bound "forward-error" by

$$\|\mathcal{A}(x) - f(x)\| = \|f(x+e) - f(x)\| \lesssim \; \kappa(f,x) \; \|e\|$$

- Example: The *Householder QR factorization algorithm* is one of the main engines in numerical linear algebra.
- N. Higham: On input an invertible $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$, this algorithm computes $\widetilde{x}$ close to $x = A^{-1}b$ such that there exist $\widetilde{A} \in \mathbb{R}^{n\times n}$ and $\widetilde{b} \in \mathbb{R}^n$ satisfying

$$\|\widetilde{A} - A\|_F \leq n\gamma_{cn}\|A\|_F, \quad \|\widetilde{b} - b\| \leq n\gamma_{cn}\|b\|,$$

where $c > 0$ is a small constant and $\gamma_k := \frac{k\epsilon_{\text{mach}}}{1-k\epsilon_{\text{mach}}}$ for $k < \epsilon_{\text{mach}}^{-1}$
- Loss of precision is bounded by

$$\log\left(\frac{\|\widetilde{x}-x\|}{\epsilon_{\text{mach}}\|x\|}\right) \leq \log\kappa(A) + 2\log n + O(1)$$

Condition as complexity parameter

- Many numerical algorithms are iterative. Often, the number of iterations to achieve a certain precision ε can be bounded in terms of the condition of the input.
- Famous example: *method of conjugate gradients*. On input a positive definite $S \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, and start value $x_0 \in \mathbb{R}^n$, this algorithm computes a sequence $x_1, x_2, \ldots$ converging to $A^{-1}b$.
- In order to achieve relative error ε, it suffices to execute

$$\frac{1}{2}\sqrt{\kappa(S)}\ \ln\left(\frac{1}{\varepsilon}\right)$$

  iterations (Hestenes and Stiefel, 1952).
- There are many results in this spirit of condition based analysis.
- linear programming (Renegar, ...)
- polynomial equation solving (Shub & Smale,...)
- ...

Probabilistic analysis of condition number

- Typical values of condition of an instance?
- Suppose $A \in \mathbb{R}^{n \times n}$ is a random matrix with independent standard Gaussian entries.
- Random matrix theory provides the joint probability density of the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ of $A^T A$.
- From this one can derive for the expectation (Edelman 1988):

$$\mathbb{E}(\log \kappa(A)) = \log n + O(1)$$

- Hence Householder solving $Ax = b$ via QR factorization algorithm has average loss of precision $O(\log n)$.

Smoothed analysis of condition number

- Smoothed analysis is a more refined form of probabilistic analysis.
- Fix any $\overline{A} \in \mathbb{R}^{n \times n}$ with $\|\overline{A}\| \leq 1$ and assume A is isotropic Gaussian with mean $\overline{A}$ and variance σ^2. Wschebor proved (2004):

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

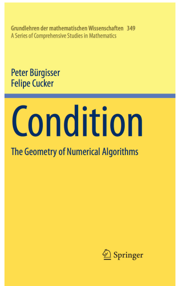
- This implies *for all* $\overline{A}$ with $\|\overline{A}\| \leq 1$

$$\mathbb{E}_{A \sim N(\overline{A}, \sigma^2 I)} \log(\kappa(A)) = \log n + \log \frac{1}{\sigma} + O(1)$$

For all $\overline{A}$ and all slight random perturbations A of $\overline{A}$, it is unlikely that $\kappa(A)$ will be large.

Smoothed analysis of numerical algorithms

- Smoothed analysis was proposed as a new form of analysis of algorithms by Spielman and Teng that blends the best of both worst-case and average-case.
- They carried out a smoothed analysis of the running time of the simplex algorithm (2001).
- For many numerical algorithms, a smoothed analysis of their running time can be reduced to a smoothed analysis of condition numbers.
  - linear programming (Renegar, ...)
  - polynomial equation solving (Shub & Smale,...)
  - ...
- See my monograph "Condition" with Felipe Cucker (Springer 2013).



Variants of Condition Numbers: Structured Data

CN for structured pertubations

- We generally defined normwise condition of $f : \mathbb{R}^p \to \mathbb{R}^q$ at x by

$$\operatorname{cond}^f(x) := \lim_{\delta \to 0} \sup_{\operatorname{RelError}(x) \leq \delta} \frac{\operatorname{RelError}(f(x))}{\operatorname{RelError}(x)}$$

- For matrix inversion $f : A \mapsto A^{-1}$ this gives $\operatorname{cond}^f(A) = \kappa(A)$
- However, for structured data, one should only allow structured perturbations in the definition
- E.g., when focusing on matrices L of certain sparsity pattern, $\operatorname{cond}^f_{sparse}(L) \leq \kappa(L)$, and the upper bound may be pessimistic

Triangular matrices

- Suppose $L \in \mathbb{R}^{n \times n}$ is lower triangular with independent standard Gaussian entries ℓ_{ij} for $i \geq j$. Viswanathan and Trefethen (1998):

$$\mathbb{E}(\log \kappa(L)) = \Omega(n)$$

- Would the loss of precision in the solution of triangular systems conform to this bound, we would not be able to accurately find these solutions!
- But practitioners observed since long that triangular systems of equations are generally solved to high accuracy.

Explanation?

Componentwise relative errors

- Classical condition number (matrix inversion)

$$\kappa(A) = \lim_{\delta \to 0} \sup_{\mathrm{RelError}(A) \le \delta} \frac{\mathrm{RelError}(A^{-1})}{\mathrm{RelError}(A)}$$

is defined w.r.t. normwise relative error (with spectral norm $\| \ \|$)

$$\mathrm{RelError}(A) := \frac{\|\widetilde{A} - A\|}{\|A\|}$$

- Instead we may use the componentwise relative error (respects sparsity)

$$\mathrm{CwRelError}(A) := \max_{i,j} \frac{|\widetilde{a_{ij}} - a_{ij}|}{|a_{ij}|}$$

- Componentwise condition number of matrix inversion defined as

$$\mathrm{Cw}^{\dagger}(A) := \lim_{\delta \to 0} \sup_{\mathrm{CwRelError}(A) \le \delta} \frac{\mathrm{CwRelError}(A^{-1})}{\mathrm{CwRelError}(A)}$$

Backward substitution is componentwise stable

- Backward substitution is the obvious algorithm for solving a triangular linear system $Lx = b$.
- The loss of precision of backward substitution can be shown to be bounded by $\mathcal{O}(\log \mathrm{Cw}^{\dagger}(L) + \log n)$
- Cheung & Cucker (2009):

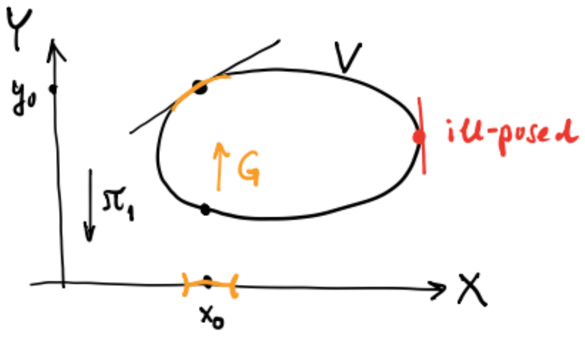
$$\mathbb{E}(\log \mathrm{Cw}^{\dagger}(L)) = \mathcal{O}(\log n)$$

  for a random lower-triangular matrix $L \in \mathbb{R}^{n \times n}$ with independent standard Gaussian random entries ℓ_{ij}
- This explains why linear triangular systems can be solved by backward substitution with high accuracy.

General geometric framework for condition numbers

J.R. Rice 1966

Differential geometric setting

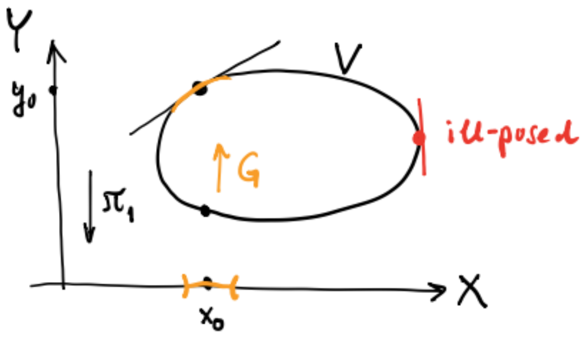


- X (smooth) manifold of inputs, Y manifold of outputs, $V \subseteq X \times Y$ submanifold, $n := \dim X = \dim V$
- $(x, y) \in V$ expresses that y is a solution for input x
- Implicit Function Thm: The projection $\pi_1 \colon V \to X, (x, y) \mapsto x$ can be locally inverted around $(x_0, y_0) \in V$ if derivative $D\pi_1(x_0, y_0)$ has full rank n. (Otherwise, call (x_0, y_0) ill-posed.)
- The local inverse $x \mapsto (x, G(x))$ of π_1 is given by the solution map G, where $G(x_0) = y_0$. Its derivative

$$DG(x_0) \colon T_{x_0}X \to T_{y_0}Y$$

is called the condition map.

General definition of condition



- Assume tangent spaces $T_x X$ and $T_y Y$ are normed vector spaces.
- This is the case if X and Y are Riemannian manifolds: they have smoothly varying inner products on tangent spaces $T_x X$ and $T_y Y$.
- Define (absolute) normwise condition number as operator norm

$$\|DG(x_0)\| := \max_{\|\dot{x}\|=1} \|DG(x_0)(\dot{x})\|$$

Condition of eigenpairs of matrices

Geometric framework for eigenpairs

- Problem: Compute eigenvectors and eigenvalues of given matrix
- Input manifold $X = \mathbb{C}^{n\times n}$, output manifold $Y = \mathbb{C} \times \mathbb{P}(\mathbb{C}^n)$
- Submanifold $V := \{(A, \lambda, v) \in X \times Y \mid Av = \lambda v\}$
- Endow X and Y with standard Riemannian metrics (on $\mathbb{P}(\mathbb{C}^n)$ take Fubini-Study metric).
- (A, λ, v) well-posed iff λ is simple eigenvalue of A
- Denote by $A_{\lambda,v}$ the linear iso of $v^\perp \simeq \mathbb{C}^{n-1}$ induced by $A - \lambda I$.
- Components of solution map $G = (G_{evector}, G_{evalue})$ give

  $$\|DG_{evector}\| = \|A_{\lambda,v}^{-1}\|, \quad \|DG_{evalue}\| = \frac{\|u\|\|v\|}{|\langle u, v\rangle|} = \frac{1}{\cos\theta}$$

  where u is left-eigenvector of A: $A^* u = \overline{\lambda} u$.
- If A is hermitian, $\|DG_{evalue}\| = 1$ and $\|DG_{evector}\|^{-1}$ equals distance of λ to closest eigenvalue.

Condition number for eigenpairs

- Denote by Σ' the set of ill-posed triples (A, λ, v), i.e., λ is multiple eigenvalue of A.
- For $(A, \lambda, v) \notin \Sigma'$ define scale-invariant condition number

$$\mu(A, \lambda, v) := \|A\|_F \cdot \|DG_{evector}\| = \|A\|_F \cdot \|A_{\lambda,v}^{-1}\|$$

  Reason: $\|DG_{evector}\|$ dominates $\|DG_{evalue}\|$
- Armentano 2014: Condition number theorem in spirit of Eckart-Young:

$$\mu(A, \lambda, v) \leq \frac{const}{\mathrm{dist}\big((A, \lambda, v), \Sigma'_v\big)}$$

  here Σ'_v denotes the fibre over v of the projection $\Sigma' \to \mathbb{P}(\mathbb{C}^n)$
- Generalizes earlier result by Wilkinson (1965)

Stable and efficient algorithms for eigenpairs

- Bezout series by Shub and Smale (1993–1996): development of rigorous geometric framework for numerically solving systems of polynomial equations (Smale's 17th problem)
- Algorithms based on homotopy continuation with stepsizes controlled by condition numbers
- Underlying principles are widely applicable
- Armentano, Beltran, B, Cucker, Shub 2018 elaborated on this to develop a numerically stable and theoretically efficient algorithm for computation of eigenpairs.
- Their algorithm runs in average (and smoothed) poly time, but it is not competitive with the algorithms used in practice.
- Motivation: the algorithms for eigenpair computation sucessfully used in practice (Hessenberg QR with shifts) were not analyzed
- Exciting recent progress by Banks, Vargas, Shrivastava on global Convergence of the Hessenberg QR algorithm with shifts.

Homotopy continuation algorithm for eigenpairs

- Recall solution variety

$$V := \{(A, \lambda, v) \in X \times Y \mid Av = \lambda v\} \subseteq \mathbb{C}^{n\times n} \times \mathbb{C} \times \mathbb{P}(\mathbb{C}^n)$$

with subvariety Σ' of ill-posed triples

- Use a well-conditioned start triple $(A_0, \lambda_0, v_0) \in V$
- On input $A \in \mathbb{C}^{n\times n}$ consider the line segment $[A_0, A]$, consisting of

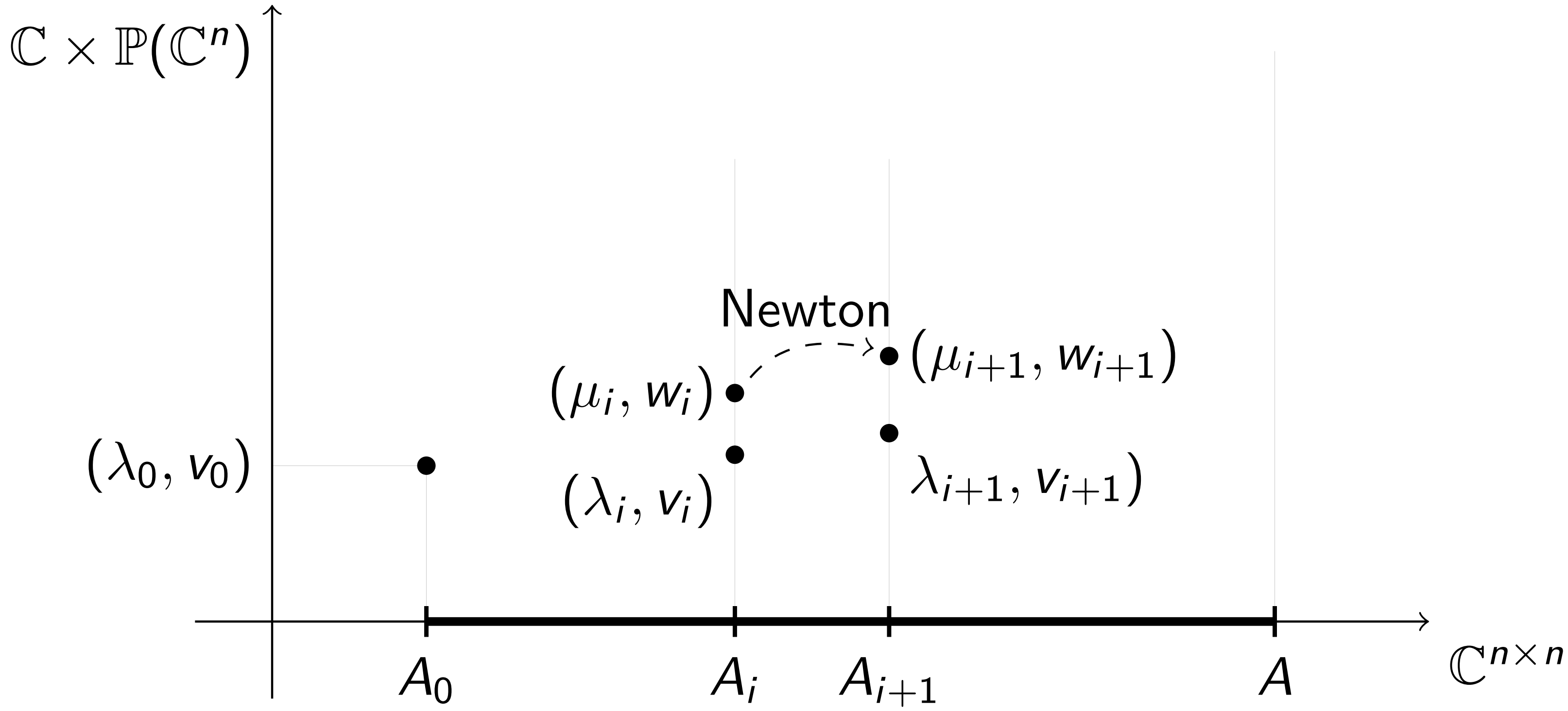
$$A_t := (1-t)A_0 + tA \quad \text{for } t \in [0,1]$$

- If $[A_0, A]$ does not meet the discriminant variety (i.e., none of the A_t has a multiple eigenvalue), then there exists a unique lifting to V,

$$\gamma\colon [0,1] \to V,\ t \mapsto (A_t, \lambda_t, v_t),$$

called solution curve.

Adaptive linear homotopy continuation



- The idea is to follow the solution curve γ numerically: partition $[0, 1]$ into $t_0 = 0, \ldots, t_k = 1$. Writing $A_i := A_{t_i}$, $\lambda_i := \lambda_{t_i}$, $v_i := v_{t_i}$, successively compute approximations (μ_i, w_i) of (λ_i, v_i) by Newton's method starting with $(\mu_0, v_0) = (\lambda_0, v_0)$.

Stepsize and condition length

- How to choose the stepsize $t_{i+1} - t_i$?

Essential theorem: The radius of quadratic attraction of Newton iteration can be upper bounded by inverse condition number.

- We choose the step size $t_{i+1} - t_i$ as an appropriate function of the current condition number $\mu(A_i, \lambda_i, v_i)$.
- One can prove that the number of Newton steps can be upper bounded by the condition length of the solution curve γ:

$$\int_0^1 \mu(\gamma(t)) \|\dot{\gamma}(t)\| \, dt$$

Probabilistic analysis

- We fix a well-conditioned start triple $(A_0, \lambda_0, v_0) \in V$.
- For a standard Gaussian input matrix $A \in \mathbb{C}^{n \times n}$ we show that the average number of Newton iterations is bounded by

$$O\big(n^4 \mu(A_0, \lambda_0, v_0)^2\big)$$

- Smoothed analysis: let $\overline{A}$ satisfy $\|\overline{A}\|_F = 1$ and assume $A \sim N(\overline{A}, \sigma^2 I)$. We can bound the smoothed average number of Newton iterations by

$$O\big(n^4 \mu(A_0, \lambda_0, v_0)^2 \sigma^{-2}\big)$$

- Can achieve average number of Newton iterations $O\big(n^4\big)$ for computing one eigenpair; each iteration costs $O(n^3)$ arithmetic operations
- Algorithm is provably numerically stable and strongly accurate (can produce approximation a la Smale and hence ε-forward approximation)

Thank you for your attention!