Basics in Applied Mathematics – Part I: Numerics

Lecture 3: Norms and conditioning – the proofs

JProf. Dr. Diyora Salimova

Department of Applied Mathematics, University of Freiburg

27 October 2026

Last time, and today

Last time: conditioning versus stability; norms; $\mathrm{cond}(A) = \|A\|\,\|A^{-1}\|$; Gaussian elimination as $A = LU$, and Cholesky as statements.

After this lecture you can

(i) prove that $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ are norms and that they are equivalent, with the constants $\sqrt{n}$ and n,

(ii) derive $\|A\|_1$, $\|A\|_\infty$ and $\|A\|_2$ from the definition of the induced norm,

(iii) prove the perturbation bound for $Ax = b$ when A and b are perturbed, and the residual–error inequalities,

(iv) explain why a pivot that is small compared to the entries below it is harmful, and why SPD matrices need no pivoting.

1.4 Norms · 2.2 Conditioning of linear systems · 2.4.3 Pivoting · 2.5 SPD matrices

The programme

Norms (Section 1.4)

- Proposition 1.15: $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ are norms
- Theorem 1.16: equivalence of the three norms
- Proposition 1.20: induced norms
- Remark 1.21: Frobenius norm
- Theorem 1.22: $\|A\|_1$, $\|A\|_\infty$, $\|A\|_2$ explicitly

Linear systems (Sections 2.2–2.5)

- Lemma 2.4, Theorem 2.5: perturbation of A and b
- Example 2.15: why we pivot
- Proposition 2.19: SPD matrices

Definition 1.14. A **norm** on $\mathbb{R}^n$: $\|\cdot\|\colon \mathbb{R}^n \to [0,\infty)$ with
(i) $\|x\| = 0 \iff x = 0$, (ii) $\|\alpha x\| = |\alpha|\,\|x\|$, (iii) $\|x+y\| \le \|x\| + \|y\|$.

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|, \qquad \|x\|_2 = \Big(\sum_{j=1}^{n} |x_j|^2\Big)^{1/2} = \sqrt{x^\top x}, \qquad \|x\|_\infty = \max_{1\le j\le n} |x_j|$$

Proposition 1.15

$\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are norms on $\mathbb{R}^n$.

Equivalence of the three norms

Theorem 1.16 (equivalence of the 1-, 2- and ∞-norm)

For all $x \in \mathbb{R}^n$,

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty, \qquad \|x\|_\infty \le \|x\|_1 \le n\,\|x\|_\infty, \qquad \|x\|_2 \le \|x\|_1 \le \sqrt{n}\,\|x\|_2.$$

All six inequalities are sharp: equality on the left for $x = e_1$, on the right for $x = (1, \dots, 1)^\top$.

Theorem 1.17. Any two norms on $\mathbb{R}^n$ are equivalent: $c\,\|x\|_a \le \|x\|_b \le C\,\|x\|_a$ $(0 < c \le C$; proof in the notes).

In $\mathbb{R}^n$, convergence does not depend on the norm; the constants do, and they may depend on n—the factors $\sqrt{n}$ and n above.

Definition 1.19 (induced norm)

One vector norm is fixed on each space ($\mathbb{R}^n$, $\mathbb{R}^m$) and used throughout, both written $\|\cdot\|$; for $A \in \mathbb{R}^{m\times n}$

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|.$$

$$\|Ax\| \le \|A\|\,\|x\| \quad \text{for all } x, \tag{1.4}$$

and $\|A\|$ is the smallest such constant.

Proposition 1.20 (properties of induced norms)

(i) The induced norm is a norm on $\mathbb{R}^{m\times n}$.

(ii) $\|I\| = 1$ for the identity matrix $I \in \mathbb{R}^{n\times n}$.

(iii) Induced norms are submultiplicative: $\|AB\| \le \|A\|\,\|B\|$ for all $A \in \mathbb{R}^{l\times m}$, $B \in \mathbb{R}^{m\times n}$.

The Frobenius norm

Remark 1.21. The Frobenius norm $\|A\|_F = \big(\sum_{i,j} a_{ij}^2\big)^{1/2}$ is submultiplicative as well: by Cauchy–Schwarz

$$|(AB)_{ij}|^2 = \Big|\sum_k a_{ik} b_{kj}\Big|^2 \le \sum_k a_{ik}^2 \sum_k b_{kj}^2,$$

and summing over i and j gives $\|AB\|_F^2 \le \|A\|_F^2 \, \|B\|_F^2$.

It is *not* induced by any vector norm: $\|I\|_F = \sqrt{n} \neq 1$ for $n \ge 2$.

Explicit formulas

Theorem 1.22 (explicit formulas for induced norms)

For $A \in \mathbb{R}^{m\times n}$: $\|A\|_1 = \max_k \sum_i |a_{ik}|$ (maximal absolute column sum), $\|A\|_\infty = \max_i \sum_k |a_{ik}|$ (maximal absolute row sum), $\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$ (square root of the largest eigenvalue of the symmetric positive semidefinite $A^\top A$).

$\|A\|_1$ and $\|A\|_\infty$ cost $O(mn)$ flops; $\|A\|_2$ is an eigenvalue problem we consider later.

Memory aid: ∞—rows (the symbol lies flat), 1—columns.

Throughout: $\|\cdot\|$ a vector norm on $\mathbb{R}^n$ and the matrix norm it induces; $A \in \mathbb{R}^{n\times n}$ regular (invertible).

Definition 2.1 and Theorem 2.2 (Thursday)

$\operatorname{cond}(A) = \|A\|\,\|A^{-1}\|$. If $Ax = b \neq 0$ and $A(x + \Delta x) = b + \Delta b$, then

$$\frac{\|\Delta x\|}{\|x\|} \;\leq\; \operatorname{cond}(A)\,\frac{\|\Delta b\|}{\|b\|}, \qquad \text{sharp for suitable } b \text{ and } \Delta b.$$

The *relative* data error in b is amplified by at most $\operatorname{cond}(A)$ (absolute errors by at most $\|A^{-1}\|$).

Question: what if the matrix A is perturbed as well—is $A + \Delta A$ still regular, and how large is Δx?

The perturbation lemma

Lemma 2.4 (perturbation lemma)

$B \in \mathbb{R}^{n \times n}$ with $\|B\| < 1$ for an induced matrix norm. Then $I + B$ is regular and

$$\|(I + B)^{-1}\| \leq \frac{1}{1 - \|B\|}.$$

A small perturbation of the identity stays regular, and its inverse is controlled.

Perturbation of matrix and right-hand side

Theorem 2.5 (perturbation of matrix and right-hand side)

A regular, $Ax = b \neq 0$, $\|A^{-1}\|\,\|\Delta A\| < 1$. Then $A + \Delta A$ is regular, and for every Δb the solution $x + \Delta x$ of $(A + \Delta A)(x + \Delta x) = b + \Delta b$ satisfies

$$\frac{\|\Delta x\|}{\|x\|} \;\leq\; \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)\,\dfrac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).$$

Both relative data errors enter with the factor $\operatorname{cond}(A)$ (the denominator is close to 1 when $\operatorname{cond}(A)\,\|\Delta A\|/\|A\| \ll 1$).

What Theorem 2.5 means

- in short, with $\kappa := \operatorname{cond}(A)$, $e_A := \|\Delta A\|/\|A\|$, $e_b := \|\Delta b\|/\|b\|$: $\dfrac{\|\Delta x\|}{\|x\|} \le \dfrac{\kappa}{1-\kappa e_A}\,(e_A + e_b)$; so for $e_A, e_b \le \varepsilon$ with $\kappa\varepsilon < 1$

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{2\kappa\varepsilon}{1-\kappa\varepsilon} \approx 2\kappa\varepsilon \qquad (\kappa\varepsilon \ll 1)$$

- storing A and b in $\mathbb{F}$ produces relative data errors of the order u (at most; none for representable data)
- **data rounding alone can produce solution errors of order** $\operatorname{cond}(A)\,u$, whatever the algorithm—even if the stored system is then solved exactly

Example 2.15 (small pivot)

$$A = \begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}, \qquad 0 < \varepsilon \ll 1.$$

A is regular and well-conditioned: $\operatorname{cond}_2(A) \approx 2.6$ for small ε.

Gaussian elimination without pivoting takes the pivot $a_{11} = \varepsilon$.

Predict: with $\varepsilon = 10^{-17}$ in double precision, what do the computed factors $\tilde{L}\tilde{U}$ reproduce—A, or something else?

Partial pivoting

Partial pivoting: in step k exchange rows so that the pivot is the entry of largest modulus among $a_{ik}^{(k)}$, $i = k, \dots, n$ $\Rightarrow$ all $|l_{ik}| \le 1$.

Theorem 2.16 (LU decomposition with partial pivoting)

For every regular $A \in \mathbb{R}^{n \times n}$ there are a permutation matrix P, a unit lower triangular L with $|l_{ij}| \le 1$ and a regular upper triangular U with $PA = LU$. Solve $Ax = b$: $Ly = Pb$, then $Ux = y$.

Proof by induction on n (in the notes, marked *). `numpy.linalg.solve` uses LAPACK's LU solver with partial pivoting (`gesv`).

Definition 2.18

$A \in \mathbb{R}^{n\times n}$ is **symmetric positive definite (SPD)** if $A = A^\top$ and $x^\top A x > 0$ for all $x \neq 0$.

Proposition 2.19 (properties of SPD matrices)

A SPD. Then

(i) A is regular;

(ii) all leading principal submatrices $A_k = (a_{ij})_{i,j=1}^{k}$, $k = 1, \ldots, n$, are SPD;

(iii) all eigenvalues of A are positive, and $\det A > 0$;

(iv) $a_{ii} > 0$ for all i, and $\max_{i,j} |a_{ij}| = \max_i a_{ii}$.

(ii) + Theorem 2.13 $\Rightarrow$ LU without pivoting; symmetry halves the work: Cholesky, $A = LL^\top$ with L lower triangular with positive diagonal—not the unit lower triangular LU factor (Theorem 2.20).

Summary

- $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ are norms and are equivalent, with constants $\sqrt{n}$ and n (Proposition 1.15, Theorem 1.16)
- induced norms: $\|Ax\| \le \|A\|\,\|x\|$, $\|AB\| \le \|A\|\,\|B\|$ (Proposition 1.20); $\|A\|_1$ max. absolute column sum, $\|A\|_\infty$ max. absolute row sum, $\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$ (Theorem 1.22)
- perturbing A and b (small relative perturbations): relative error of $x \lesssim \operatorname{cond}(A)\times$ (sum of the relative data errors) (Lemma 2.4, Theorem 2.5); data rounding alone can cost errors of order $\operatorname{cond}(A)\,u$, whatever the algorithm
- pivoting avoids large multipliers (Example 2.15, Theorem 2.16); SPD matrices: regular, positive eigenvalues, SPD leading blocks (Proposition 2.19)