Ill-conditioning means 'nearly singular'

Using any subordinate matrix norm

$$\frac{1}{\kappa_p(A)} = \min \frac{\|\mathbf{\Delta A}\|_p}{\|\mathbf{A}\|_p}$$

where $\mathbf{A} + \mathbf{\Delta A}$ is singular

if $\kappa(A)$ is large, $\mathbf{A}$ is close to a singular matrix

Example

The determinant does not help

Note: $\det \mathbf{A}$ is NOT a good measure of how close to singular a matrix is. It gives false negatives

Example

and false positives

Example

Geometric interpretation

An ill-conditioned matrix $\Longrightarrow$ a system of hyperplanes that are **almost parallel**

$\Longrightarrow$ wouldn't have to tilt them much to make them parallel and hence have no or infinitely many solutions.

In MATLAB

$\kappa(A)$ is computed in MATLAB with the `cond` command.
Not cheap to compute $\kappa(A)$
instead usually estimate its value.

In MATLAB one can use either

- `condest` to estimate $\kappa_1(A)$
- or `rcond` which estimates its reciprocal.

Backwards error and residuals

Suppose all backward error is in $\mathbf{A}$ not $\mathbf{b}$. Then

$$(\mathbf{A} + \Delta\mathbf{A})\hat{\mathbf{x}} = \mathbf{b}$$

so that

$$\Delta\mathbf{A}\,\hat{\mathbf{x}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} = \mathbf{r}$$

where $\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ is the **residual**.

Then $\|\mathbf{r}\| \leq \|\Delta\mathbf{A}\|\ \|\hat{\mathbf{x}}\| \implies \|\Delta\mathbf{A}\| \geq \frac{\|\mathbf{r}\|}{\|\hat{\mathbf{x}}\|}$
so the relative backward error

$$\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \geq \frac{\|\mathbf{r}\|}{\|\mathbf{A}\|\ \|\hat{\mathbf{x}}\|} \equiv \eta$$

is measured by the **relative residual** η.

Hence

- large relative residuals $\implies$ large (normwise) relative backward error

A stable method (small backward error) produces small relative residuals.

- the bound can be attained, so that small relative residuals signal small (normwise) relative backward error

BUT small residuals do NOT $\implies$ small errors in $\mathbf{x}$
(it depends on the size of the condition number of $\mathbf{A}$)

Backward error analysis: triangular systems

Under standard model for Floating point arithmetic:

Theorem

If the $n \times n$ triangular system

$$\mathbf{T}\mathbf{x} = \mathbf{b}$$

is solved by substitution, the computed solution $\hat{\mathbf{x}}$ satisfies

$$(\mathbf{T} + \Delta\mathbf{T})\hat{\mathbf{x}} = \mathbf{b}$$

where (componentwise bound)

$$|\ \Delta T_{ij}\ | \leq \gamma_n\ |\ T_{ij}\ | \approx nu\ |\ T_{ij}\ | \quad since\ \gamma_n \equiv \frac{nu}{1 - nu} \approx nu$$

i.e. it solves a system with a nearby matrix (and the correct **b**!).

Solving triangular systems by forward/backsubstitution is backward stable

In general, componentwise stability is a more stringent requirement than normwise stability, since for any monotone norm

$$| \Delta T_{ij} | \leq \gamma_n | T_{ij} | \Rightarrow \|\Delta \mathbf{T}\| \leq \gamma_n \|\mathbf{T}\|$$

Solving triangular systems by forward/backsubstitution is normwise backward stable

which is the usual usage of the term **backward stable**.

$\Longrightarrow$ final stages of solving a linear system are **no problem**; any problems must arise in the factorization stage!

Gauss elimination: any pivoting strategy

Theorem

If the $n \times n$ system

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

is solved by Gauss elimination, the computed solution $\hat{\mathbf{x}}$ satisfies

$$(\mathbf{A} + \Delta\mathbf{A})\hat{\mathbf{x}} = \mathbf{b}$$

where (componentwise again)

$$| \Delta A_{ij} | \leq \gamma_{3n} \sum_k | \hat{L}_{ik} || \hat{U}_{kj} |$$

which is **not what we're after**. Want a bound for $\Delta\mathbf{A}$ involving $| \mathbf{A} |$.

Gauss elimination with partial pivoting (GEPP)

The problem with Gauss elimination is that $\sum_k \mid \hat{L}_{ik} \mid\mid \hat{U}_{kj} \mid$ can be much bigger than $\mid A_{ij} \mid = \mid \sum_k \hat{L}_{ik} \hat{U}_{kj} \mid$
i.e. $\mathbf{A}$ must have been produced from $\mathbf{L}, \mathbf{U}$ using lots of cancellation.

Partial pivoting ensures that

$$\mid \hat{L}_{ik} \mid \leq 1$$

but what can we say about $\mid \hat{U}_{kj} \mid$?
To progress, go to a normwise analysis:

$$\|\Delta \mathbf{A}\| \leq \gamma_{3n} \|\hat{\mathbf{L}}\| \; \|\hat{\mathbf{U}}\|$$

for a monotone norm.

The growth factor

Definition

The **growth factor** (using the max-norm):

$$\rho := \|\hat{\mathbf{L}}\|_M \|\hat{\mathbf{U}}\|_M / \|\mathbf{A}\|_M$$

Using the definition $\implies$ for the 1 or ∞ norms

$$\|\Delta \mathbf{A}\| \leq n^2 \gamma_{3n} \rho \|\mathbf{A}\|$$

normwise backward error is determined by the growth factor.

Without pivoting...

there is **no bound for the growth factor**.

Example

The matrix $\begin{bmatrix} \delta & 1 \\ 1 & 1 \end{bmatrix}$ for $\delta \ll 1$ has $\rho \approx \delta^{-2}$ which can be arbitrarily large.

even for 2×2 system!

For Gauss elimination with partial pivoting

- There are matrices that have $\rho = 2^{n-1}$ — for them GEPP is **not backward stable**.
- These matrices appear to be **very rare** in practise. For most matrices, ρ is very modest in size e.g < 10
- For random matrices e.g. matrices with components drawn from a standard normal distribution, $\rho \sim n^{1/2}$

Hence experts say that, **for practical purposes, GEPP is (normwise) backward stable** — the default direct method for solving general dense linear systems.

SUMMARY

1. Gauss elimination without pivoting is **not backward stable**.
2. GEPP is, for practical purposes, (normwise) backward stable.
3. Cholesky factorization is (normwise) backward stable
4. There are other factorizations (QR) that are (normwise) backward stable for general matrices but more expensive than GEPP
5. A backward stable method produces small (relative) residuals.
6. This does not imply small (forward) errors if $\kappa(\mathbf{A}) >> 1$ i.e your problem is ill-conditioned.

What if **A** is ill-conditioned?

Remedy:

1. do you really need $\mathbf{x}$ or just $f(\mathbf{x})$ or $\mathbf{r}$?
2. do the problem another way i.e. change $\mathbf{A}$
3. live with the reduced accuracy

End of Week 7