MA385 Part 4: Linear Algebra 2

4.3: Condition Numbers

Dr Niall Madden

November 2025

1. Outline Section 4.3

For more, see Section 2.7 of Suli and Mayers:
https://ebookcentral.proquest.com/lib/nuig/reader.action?docID=221072&ppg=51&c=UERG

2. Motivation

Numerical solutions to some linear systems are adversely affected by round-off errors.

This phenomenon is due to the *matrices* in the linear systems. Those matrices for which the issue is particularly prevalent are referred to as being **ill-conditioned**.

For any matrix, we can assign a numerical *score* that gives an indication of whether it is ill-conditioned. That score, called the **condition number**, is defined in terms of matrix norms, and is the subject of these section.

3. Consistency of matrix norms

Suppose we have a vector norm, $\|\cdot\|$ and associated subordinate matrix norm. It is not hard to see that

$$\|A\mathbf{u}\| \leq \|A\|\|\mathbf{u}\| \quad \text{for } \textit{any } \mathbf{u} \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}.$$

Here is why:

By definition

$$\|A\| := \max_{u \in \mathbb{R}^n/\{0\}} \frac{\|Au\|}{\|u\|}$$

So, for any vector u,

$$\|A\| \geq \frac{\|Au\|}{\|u\|}$$

$$\Rightarrow \quad \|A\| \cdot \|u\| \geq \|Au\|.$$

3. Consistency of matrix norms

There is an analogous statement for the product of two matrices:

Definition 4.3.1 (Consistent matrix norm)

A matrix norm $\|\cdot\|$ is **consistent** (or "*sub-multiplicative*") if

$$\|AB\| \leq \|A\|\|B\|, \qquad \text{for all } A, B \in \mathbb{R}^{n\times n}.$$

Theorem 4.3.1

Any subordinate matrix norm is consistent.

The proof is left to an exercise. That exercises also demonstrates that there are matrix norms which are *not* consistent.

4. Computer representation of numbers

[Please read this slide in your own time!]

Modern computers don't store numbers in decimal (base 10), but in binary (base 2) "floating point numbers" of the form :

$$x = \pm a \times 2^{b-M}.$$

Most use *double precision*, where 8 bytes (64 bits or *binary digits*) are used to store

- the sign (1 bit),
- a, called the "significand" or "mantissa" (52 bits)
- and the exponent, $b - 1023$ (11 bits)

Note that a has roughly 16 decimal digits.
(Some older computer systems sometimes use *single precision* where a has 23 bits — giving 8 decimal digits — and b has 7; so too do many new GPU-based systems).

4. Computer representation of numbers

When we try to store a real number x on a computer, we actually store the nearest floating-point number. That is, we end up storing $x + \delta x$, where δx is the "round-off" error.

But the quantity we are mainly interested in is the **relative error:** $|\delta x|/|x|$.

Since this is not a course on computer architecture, we'll simplify a little and just take it that single and double precision systems lead to a relative error of 10^{-8} and 10^{-16} respectively.

5. Condition Numbers

(*See p68–70 of Süli and Mayers for a full development of the concept of a condition number*).

Suppose we use, say, LU-factorization and back-substitution on a computer to solve

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

Because of the "round-off error" we actually solve

$$A(\mathbf{x} + \delta\mathbf{x}) = (\mathbf{b} + \delta\mathbf{b}).$$

Our problem now is, for a given A, if we know the (relative) error in $\mathbf{b}$, can we find an upper-bound on the relative error in $\mathbf{x}$?

Definition 4.3.2

The *condition number* of a matrix, with respect to a particular matrix norm $\|\cdot\|_\star$ is

$$\kappa_\star(A) = \|A\|_\star \|A^{-1}\|_\star.$$

If $\kappa_\star(A) \gg 1$ then we say A is *ill-conditioned*.

Example: Find the condition number κ_∞ of

$$A = \begin{pmatrix} 10 & 12 \\ 0.08 & 0.1 \end{pmatrix}. \qquad A^{-1} = \begin{pmatrix} 2.5 & 300 \\ -2 & 250 \end{pmatrix}$$

Then $\|A\|_\infty = 22$ $\qquad \|A^{-1}\|_\infty = 302.5$

So $\kappa_\infty(A) = (22)(302.5) = 6{,}655.$

Theorem 4.3.2

Suppose that $A \in \mathbb{R}^{n\times n}$ is nonsingular and that $\mathbf{b}, \mathbf{x} \in \mathbb{R}^n$ are non-zero vectors. If $A\mathbf{x} = \mathbf{b}$ and $A(\mathbf{x} + \delta\mathbf{x}) = (\mathbf{b} + \delta\mathbf{b})$ then

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A)\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}.$$

Since $Ax = b$

And $A(x + \delta x) = b + \delta b$, so $Ax + A\delta x = b + \delta b$,

we get $A\delta x = \delta b$

Next $Ax = b \Rightarrow \|Ax\| = \|b\|$.

$\Rightarrow \|b\| \leq \|A\| \cdot \|x\|$

Theorem 4.3.2

Suppose that $A \in \mathbb{R}^{n\times n}$ is nonsingular and that $\mathbf{b}, \mathbf{x} \in \mathbb{R}^n$ are non-zero vectors. If $A\mathbf{x} = \mathbf{b}$ and $A(\mathbf{x} + \delta\mathbf{x}) = (\mathbf{b} + \delta\mathbf{b})$ then

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \kappa(A)\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}.$$

Similarly $A\delta x = \delta b \implies \delta x = A^{-1}\delta b.$

$\implies \|\delta x\| \le \|A^{-1}\| \cdot \|\delta b\|$

Since we had $\|b\| \le \|A\| \cdot \|x\|$

$\implies \|\delta x\| \cdot \|b\| \le \underbrace{\|A\| \cdot \|A^{-1}\|}_{\kappa(A)} \cdot \|\delta b\| \cdot \|x\|$

$\implies \dfrac{\|\delta x\|}{\|x\|} \le \kappa(A)\dfrac{\|\delta b\|}{\|b\|}.$

Example 4.3.1

Suppose we are using a computer to solve $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{pmatrix} 10 & 12 \\ 0.08 & 0.1 \end{pmatrix} \quad \text{and } \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

But, due to round-off error, right-hand side has a relative error (in the ∞-norm) of 10^{-6}. Give a bound for the relative error in $\mathbf{x}$ in the ∞-norm.

We had that $\kappa_\infty(A) = 6655$

Since $\frac{\|\delta b\|}{\|b\|} = 10^{-6}$

So $\frac{\|\delta x\|}{\|x\|} \leq 6.655 \times 10^{-3}$

6. Calculating κ_∞ and κ_1

For every matrix norm we get a different condition number.

Example 4.3.2

Let A be the $n \times n$ matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 1 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

What are $\kappa_1(A)$, and $\kappa_\infty(A)$?

First we compute $\|A\|_1$ and $\|A\|_\infty$.

$\|A\|_1 = 2 \qquad \|A\|_\infty = n.$

6. Calculating κ_∞ and κ_1

For this very special example, it is easy to write down the inverse of A:

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ -1 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

$\|A^{-1}\|_1 = 2$ $\qquad$ $\|A^{-1}\|_\infty = n$

$\Rightarrow \quad \kappa_1(A) = 4$ $\qquad$ $\kappa_\infty(A) = n^2$

To compute $\kappa_1(A)$ and $\kappa_\infty(A)$, we need to know A^{-1}, which is usually not practical. However, for κ_2, we are able to *estimate* the condition number of A without knowing A^{-1}.
Recall that $\|A\|_2 = \sqrt{\lambda_n}$ where λ_n is the largest eigenvalue of $B = A^T A$.
We can also show that $\|A^{-1}\|_2 = \dfrac{1}{\sqrt{\lambda_1}}$ where λ_1 is the smallest eigenvalue of B

- $A^T A$ and AA^T have the same eigenvalues (they are "*similar*"):
  $A^T A x = \lambda x \Rightarrow A(A^T A)x = \lambda (Ax)$
  $\Rightarrow (AA^T)(Ax) = \lambda (Ax) \Rightarrow \lambda$ is an ε'val of AA^T with ε'vec Ax
- For any non-singular matrix X, we have that $(X^T)^{-1} = (X^{-1})^T$.
- $(A^T A)^{-1} = (AA^T)^{-1}$

So

$$\kappa_2(A) = \left(\lambda_n/\lambda_1\right)^{1/2}.$$

. . . .

Motivated by this, we'll finish MA385, by studying an easy way of estimating the eigenvalues of a matrix.

Finished here

Exercise 4.3.1

(i) Prove that, if $\|\cdot\|$ is a subordinate matrix norm, then it is *consistent*, i.e., for any pair of $n \times n$ matrices, A and B, we have $\|AB\| \leq \|A\|\|B\|$.

(ii) One might think it intuitive to define the "max" norm of a matrix as follows:

$$\|A\|_{\widetilde{\infty}} = \max_{i,j} |a_{ij}|.$$

Show that this is indeed a norm on $\mathbb{R}^{n\times n}$. Show that, however, it is not consistent.

Exercise 4.3.2

Let A be the matrix

$$A = \begin{pmatrix} 0.1 & 0 & 0 \\ 10 & 0.1 & 10 \\ 0 & 0 & 0.1 \end{pmatrix}$$

Compute $\kappa_\infty(A)$. Suppose we wish to solve the system of equations $A\mathbf{x} = \mathbf{b}$ on *single precision* computer system (i.e., the relative error in any stored number is approximately 10^{-8}). Give an upper bound on the relative error in the computed solution $\mathbf{x}$.