15. Problem condition

- condition of a mathematical problem
- matrix norm
- condition number
Sources of error in numerical computation

**Example:** evaluate a function $f : \mathbb{R} \to \mathbb{R}$ at a given $x$

sources of error in the result:

- $x$ is not exactly known
  - measurement errors
  - errors in previous computations
  $\rightarrow$ how sensitive is $f(x)$ to errors in $x$?

- the algorithm for computing $f(x)$ is not exact
  - discretization (e.g., algorithm uses a table to look up function values)
  - truncation (e.g., function is evaluated by truncating a Taylor series)
  - rounding error during the computation
  $\rightarrow$ how large is the error introduced by the algorithm?
Condition (conditioning) of a problem
describes sensitivity of the solution with respect to errors in the data

- **well-conditioned problem:**
  small changes in the data produce small changes in the solution

- **ill-conditioned (badly conditioned) problem:**
  small changes in the data can produce large changes in the solution

A rigorous definition depends on what ‘large error’ means

- absolute or relative error, which norm is used, . . .
- the informal definition is sufficient for our purposes
Example: function evaluation

here the problem is: given $x$, evaluate $y = f(x)$

• if $x$ is changed to $x + \Delta x$, solution changes to

$$y + \Delta y = f(x + \Delta x)$$

• condition with respect to absolute error in $x$ and $y$

$$|\Delta y| \approx |f'(x)||\Delta x|$$

problem is ill-conditioned w.r.t. absolute error if $|f'(x)|$ is very large

• condition with respect to relative errors in $x$ and $y$

$$\frac{|\Delta y|}{|y|} \approx \frac{|f'(x)||x||\Delta x|}{|f(x)||x|}$$

ill-conditioned w.r.t. relative error if $|f'(x)||x|/|f(x)|$ is very large
Roots of a polynomial

\[ p(x) = (x - 1)(x - 2) \cdots (x - 10) + \delta \cdot x^{10} \]

roots of \( p \) computed by MATLAB for two values of \( \delta \)

\[ \delta = 10^{-5} \]

\[ \delta = 10^{-3} \]

roots are very sensitive to errors in the coefficients
Condition of a set of linear equations

• assume $A$ is nonsingular and $Ax = b$

• if we change $b$ to $b + \Delta b$, the new solution is $x + \Delta x$ with

$$A(x + \Delta x) = b + \Delta b$$

• the change in $x$ is

$$\Delta x = A^{-1} \Delta b$$

Condition

• the equations are well-conditioned if small $\Delta b$ results in small $\Delta x$
• the equations are ill-conditioned if small $\Delta b$ can result in large $\Delta x$
Example of ill-conditioned equations

\[ A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 + 10^{-10} & 1 - 10^{-10} \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 1 - 10^{10} & 10^{10} \\ 1 + 10^{10} & -10^{10} \end{bmatrix} \]

• solution for \( b = (1, 1) \) is \( x = (1, 1) \)

• change in \( x \) if we change \( b \) to \( b + \Delta b \):

\[ \Delta x = A^{-1} \Delta b = \begin{bmatrix} \Delta b_1 - 10^{10}(\Delta b_1 - \Delta b_2) \\ \Delta b_1 + 10^{10}(\Delta b_1 - \Delta b_2) \end{bmatrix} \]

small \( \Delta b \) can lead to extremely large \( \Delta x \)
Outline

• condition of a mathematical problem

• matrix norm

• condition number
Norm of a matrix

$m \times n$ matrix $A$ defines a linear function $f(x) = Ax$

\[ x \xrightarrow{A} y = f(x) = Ax \]

- $\|Ax\|/\|x\|$ gives the amplification factor or gain in the direction $x$
- we define the norm of $A$ as the maximum gain over all directions:

\[
\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|
\]

this norm is also called the spectral norm or 2-norm

- in MATLAB: $\text{norm}(A)$
Computing the norm of a matrix

**Simple matrices:** sometimes it is easy to maximize $\|Ax\|/\|x\|$

- zero matrix: $\|0\| = 0$
- identity matrix: $\|I\| = 1$
- diagonal matrix:

\[
A = \begin{bmatrix}
A_{11} & 0 & \cdots & 0 \\
0 & A_{22} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_{nn}
\end{bmatrix}, \quad \|A\| = \max_{i=1,\ldots,n} |A_{ii}|
\]

- matrix with orthonormal columns: $\|A\| = 1$

**General matrices:** $\|A\|$ must be computed by numerical algorithms
Other matrix norms

many other definitions of matrix norms exist, e.g., the *Frobenius* norm

\[ \|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2} \]

- has a simple explicit expression
- in MATLAB: use `norm(A, ’fro’)`

in this course: \(\|A\|\) stands for the norm on defined on page 15-8

- no simple explicit expression (except for special \(A\))
- readily computed numerically
Properties of the matrix norm

- *nonnegative definiteness:* \( \|A\| \geq 0 \) for all \( A \); \( \|A\| = 0 \) iff \( A = 0 \)

- *homogeneity:* \( \|\alpha A\| = |\alpha| \|A\| \)

- *triangle inequality:* \( \|A + B\| \leq \|A\| + \|B\| \)

- \( \|Ax\| \leq \|A\| \|x\| \) if the product \( Ax \) exists

- \( \|AB\| \leq \|A\| \|B\| \) if the product \( AB \) exists

- if \( A \) is nonsingular: \( \|A\| \|A^{-1}\| \geq 1 \)

- if \( A \) is nonsingular: \( 1/\|A^{-1}\| = \min_{x \neq 0} (\|Ax\|/\|x\|) \)

- \( \|A^T\| = \|A\| \)
Outline

• condition of a mathematical problem

• matrix norm

• condition number
Bound on absolute error

suppose $A$ is nonsingular

$$x = A^{-1}b, \quad \Delta x = A^{-1} \Delta b$$

**Upper bound** on $\|\Delta x\|$\footnote{follows from property 4 on page 15-11}

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\|$$

• small $\|A^{-1}\|$ means that $\|\Delta x\|$ is small when $\|\Delta b\|$ is small
• large $\|A^{-1}\|$ means that $\|\Delta x\|$ can be large, even when $\|\Delta b\|$ is small
• for every $A$, there exists $\Delta b$ such that $\|\Delta x\| = \|A^{-1}\| \|\Delta b\|$ (no proof)
Bound on relative error

suppose in addition that $b \neq 0$; hence $x \neq 0$

Upper bound on $\|\Delta x\| / \|x\|$

$$\frac{\|\Delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}$$

(follows from $\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\|$ and $\|b\| \leq \|A\| \|x\|$)

• $\|A\| \|A^{-1}\|$ small means $\|\Delta x\| / \|x\|$ is small when $\|\Delta b\| / \|b\|$ is small
• $\|A\| \|A^{-1}\|$ large means $\|\Delta x\| / \|x\|$ can be much larger than $\|\Delta b\| / \|b\|$
• for every $A$, there exist $b$, $\Delta b$ such that equality holds in (1) (no proof)
Condition number

**Definition:** the condition number of a nonsingular matrix $A$ is

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

**Properties**

- $\kappa(A) \geq 1$ for all $A$ (last property on page page 15-11)

- $A$ is a *well-conditioned* matrix if $\kappa(A)$ is small (close to 1):
  - the relative error in $x$ is not much larger than the relative error in $b$

- $A$ is *badly conditioned* or *ill-conditioned* if $\kappa(A)$ is large:
  - the relative error in $x$ can be much larger than the relative error in $b$
Example

- $A$ is blurring matrix, nonsingular with condition number $\approx 10^9$
- we apply $A$ to image $x$

\[
y_1 = Ax
\]

blurred image

\[
y_2 = Ax + \text{small noise}
\]

blurred and noisy image
Example

we solve $Ax = y$ for the two blurred images

- illustrates ill conditioning of $A$
- explains need for regularization in deblurring algorithms