Solving a Linear System

- FP
- norm (vector)
- conditioning

Given:

- $m \times n$ matrix A
- m-vector $\boldsymbol{b}$

What are we looking for here, and when are we allowed to ask the question?

$$\underset{m\times n}{A}\ \underset{n\times 1}{x} = \underset{m\times 1}{b}$$

- solution may not exist
- solution may not be unique.

$m=n$, A^{-1} exists.

Next: Want to talk about conditioning of this operation. Need to measure distances of matrices.

Matrix Norms

What norms would we apply to matrices?

- $\|A\| \geq 0$, $\|A\| = 0$ iff $A = 0$
- $\|\lambda A\| = |\lambda| \, \|A\|$, $\lambda \in \mathbb{R}$
- triangle inequality.

- "flatten" norm.
- "compatible" norm

Intuition for Matrix Norms

Provide some intuition for the matrix norm.

$$y = \frac{x}{\|x\|}$$

$$\max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \neq 0} \frac{\left\| A \frac{x}{\|x\|} \cdot \|x\| \right\|}{\left\| \frac{x}{\|x\|} \cdot \|x\| \right\|} = \max_{\|y\|=1} \frac{\|Ay\|}{\|y\|}$$

Matrix Norm Properties

What is $\|A\|_1$? $\|A\|_\infty$?

$$\|A\|_1 = \max_{\|x\|=1} \frac{\|Ax\|_1}{\|x\|_1} = \max_j \sum_i |A_{ij}| \quad , \quad \|A\|_\infty$$

How do matrix and vector norms relate for $n \times 1$ matrices?

$$\max_{\|x\|=1} \|Ax\| = \max_{x \in \{-1,1\}} \|Ax\| = \|A[:,1]\|$$

Demo: Matrix norms [cleared]

Properties of Matrix Norms

$\max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$

Matrix norms inherit the vector norm properties:

- $\|A\| > 0 \Leftrightarrow A \neq 0$.
- $\|\gamma A\| = |\gamma| \, \|A\|$ for all scalars γ.
- Obeys triangle inequality $\|A + B\| \leqslant \|A\| + \|B\|$

$\|Ax\| \le \|A\| \, \|x\| \quad | : \|x\|$

$\frac{\|Ax\|}{\|x\|} \le \|A\|$

But also some more properties that stem from our definition:

- $\|Ax\| \le \|A\| \, \|x\|$
- $\|AB\| \le \|A\| \, \|B\|$

"Submultiplicativity"

Conditioning

What is the condition number of solving a linear system $A\boldsymbol{x} = \boldsymbol{b}$?

Inputs: b with error Δb

Output: x with error Δx

$$\frac{\text{rel. err. in out}}{\text{rel. err. in in}} \approx \frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|} = \frac{\|\Delta x\| \cdot \|b\|}{\|\Delta b\| \cdot \|x\|}$$

$Ax=b$

$$A(x+\Delta x) = b + \Delta b \Rightarrow A\Delta x = \Delta b \Leftrightarrow \Delta x = A^{-1}\Delta b$$

$$= \frac{\|A^{-1}\Delta b\| \, \|Ax\|}{\|\Delta b\| \cdot \|x\|} \le \frac{\|A^{-1}\| \, \|A\| \, \cancel{\|\Delta b\|} \, \|x\|}{\cancel{\|\Delta b\| \, \|x\|}}$$

Conditioning of Linear Systems: Observations

Showed $\kappa(\text{Solve } A\boldsymbol{x} = \boldsymbol{b}) \leq \|A^{-1}\| \, \|A\|$.
I.e. found an *upper bound* on the condition number. With a little bit of fiddling, it's not too hard to find examples that achieve this bound, i.e. that it is *sharp*.

So we've found the *condition number of linear system solving*, also called the condition number of the matrix A:

$$\operatorname{cond}(A) = \kappa(A) = \|A\| \left\|A^{-1}\right\|.$$

$$\operatorname{cond}_2(A) = \|A\|_2 \, \|A^{-1}\|_2$$

$$\operatorname{cond}(I) = 1$$

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \, \|A^{-1}\| = \operatorname{cond}(A)$$

Conditioning of Linear Systems: More properties

- cond is relative to a given norm. So, to be precise, use

$$\mathrm{cond}_2 \quad \text{or} \quad \mathrm{cond}_\infty .$$

$Ax = b$ solve

- If A^{-1} does not exist: $\mathrm{cond}(A) = \infty$ by convention.

$Ay = c$ matvec

$y = A^{-1}c$

What is $\kappa(A^{-1})$?

$\kappa(A)$

What is the condition number of matrix-vector multiplication?

$\kappa(A)$

Demo: Condition number visualized [cleared]
Demo: Conditioning of 2x2 Matrices [cleared]

Residual Vector

What is the residual vector of solving the linear system

$$\boldsymbol{b} = A\boldsymbol{x}?$$

$$\hat{x} = x + \Delta x$$

$$b - A\hat{x} := r$$

Residual and Error: Relationship

How do the (norms of the) residual vector $\boldsymbol{r}$ and the error $\Delta\boldsymbol{x} = \boldsymbol{x} - \widehat{\boldsymbol{x}}$ relate to one another?

$$\frac{\|\Delta x\|}{\|\hat{x}\|} = \frac{\|A^{-1} r\|}{\|\hat{x}\|} \le \frac{\|A^{-1}\|\cdot\|r\|}{\|\hat{x}\|}$$

$$\|\Delta x\| = \|x - \hat{x}\| = \|A^{-1} r\|$$

$$\kappa = \|A\|\cdot\|A^{-1}\|$$

$$\|A\hat{x}\| \le \|A\|\,\|\hat{x}\|$$

$$\frac{\kappa}{\|A\|}\cdot\frac{\|r\|}{\|\hat{x}\|} = \kappa\cdot\frac{\|r\|}{\|A\|\,\|\hat{x}\|}$$

$$\le \kappa\cdot\frac{\|r\|}{\|A\hat{x}\|}$$

Changing the Matrix

So far, all our discussion was based on changing the right-hand side, i.e.

$$A\boldsymbol{x} = \boldsymbol{b} \quad \rightarrow \quad A\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{b}}.$$

The matrix consists of FP numbers, too–it, too, is approximate. I.e.

$$A\boldsymbol{x} = \boldsymbol{b} \quad \rightarrow \quad \hat{A}\widehat{\boldsymbol{x}} = \boldsymbol{b}.$$

What can we say about the error now?

Changing Condition Numbers

Once we have a matrix A in a linear system $A\boldsymbol{x} = \boldsymbol{b}$, are we stuck with its condition number? Or could we improve it?

What is this called as a general concept?

In-Class Activity: Matrix Norms and Conditioning

In-class activity: Matrix Norms and Conditioning