Today

- FP
- num. LA
  - ↳ norms for matrices
  - ↳ conditioning
  - ↳ solving

Announcements

- HW1 out
- Exercise 0 ongoing
- Quiz deadlines today
  - ↳ next Wed.

Implementing Arithmetic

How is floating point addition implemented?
Consider adding $a = (1.101)_2 \cdot 2^1$ and $b = (1.001)_2 \cdot 2^{-1}$ in a system with three bits in the significand.

$$a = (1.101)_2 \cdot 2^1$$

$$b = 0.01001)_2 \cdot 2^1$$

$$1.11101 \cdot 2^1$$

round $\quad 1.111 \cdot 2^1$

Problems with FP Addition

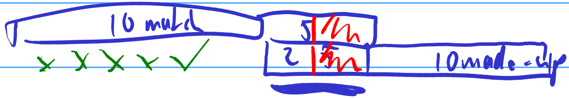
What happens if you subtract two numbers of very similar magnitude?
As an example, consider $a = (1.1011)_2 \cdot 2^0$ and $b = (1.1010)_2 \cdot 2^0$.

$$
\begin{array}{rl}
a = & 1.1011 \\
b = & 1.1010 \\
\hline
a-b = & 0.0001 \cdot 2^0 \\
 & \quad\;\; \underline{1}.\,????\ \cdot 2^{-4}
\end{array}
$$

Demo: Catastrophic Cancellation

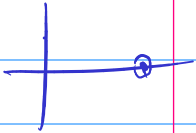
Suppose a, b are known to 12 digits and stored in DP. → 52 bits in the stored fraction

Their 10 leading digits match.



↳ 5 digits left because of cancellation

↳ 2 left because of rel. error



$a \cdot 2^c \cdot b \cdot 2^f$

Supplementary Material

- Josh Haberman, Floating Point Demystified, Part 1
- David Goldberg, What every computer programmer should know about floating point

Outline

Solving a Linear System

Given:

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

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

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

↳ lin. comb. of the col. of A to yield $\vec{b}$

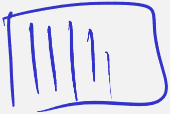
↳ $m = n$ for now

↳ sol. may not exist.

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

Matrix Norms

What norms would we apply to matrices?



submultiplicativity

$$\|Ax\| \le \underbrace{\|A\|}_{\text{defining now}} \|x\| \quad \text{for all } x \neq 0$$

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

$$\|A\| := \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

$$= \max_{x \neq 0} \Big\| A \underbrace{\frac{x}{\|x\|}}_{\text{norm } 1} \Big\|$$

$$= \max_{\|y\|=1} \|Ay\|$$

Matrix Norm Properties

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

$$\|A\|_1 = \max_{\text{col } j} \sum_{\text{row } i} |A_{ij}|$$

$$\|A\|_\infty = \max_{\text{row } i} \sum_{\text{col } j} |A_{ij}|$$

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

$[\ ]_{n\times 1}$ $\underbrace{[x]}_{\|\ \|=1}$ $\le \|Ax\|$? "they agree"

Demo: Matrix norms

Properties of Matrix Norms

Matrix norms inherit the vector norm properties:

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

But also some more properties that stem from our definition:

$$\|Ax\| \leq \|A\| \, \|x\|$$

$$\|AB\| \leq \|A\| \, \|B\|$$

Conditioning

output $\quad$ input

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

Δb

Δx

$$\frac{\text{rel. error in output}}{\text{rel. error in input}} = \frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|} = \frac{\|\Delta x\|\,\|b\|}{\|\Delta b\|\,\|x\|}$$

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

↳ shows an upper bound

↳ need to show that bound is sharp

↳ find an example that reaches the bound

Conditioning of Linear Systems: Observations

Showed $\kappa(\text{Solve } A\mathbf{x} = \mathbf{b}) \le \left\|A^{-1}\right\| \|A\|$.
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:

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

Conditioning of Linear Systems: More properties

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

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

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

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

$\kappa(A)$ $\qquad \|\underline{A}\underline{v}\| \le \|\underline{A}\|\|\underline{v}\|$?

What is the condition number of matrix-vector multiplication?

$x \to Ax$

$x \to A^{-1}y = x \qquad \to \qquad y = Ax$

Demo: Condition number visualized

Demo: Conditioning of 2x2 Matrices

Residual Vector

What is the residual vector of solving the linear system

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

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \, \|A^{-1}\|$$