Today’s Agenda

- Preliminaries on vector/matrix norms
- Error analysis
- Pivoting (motivation and basic idea)

Vector/matrix prelim

Vector norms

- Definition: A function $||\cdot||$ satisfies

  1. $\|\mathbf{x}\| \geq 0$ for any vector $\mathbf{x} \in \mathbb{R}^n$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$
  2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for any vector $\mathbf{x} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$
  3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for any vectors $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$.

- Popular vector norms:
  – L1 norm $\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$
  – L2 norm $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}^T\mathbf{x}}$
  – L infinity norm $\|\mathbf{x}\|_\infty = \max_{1\leq i\leq n} |x_i|$

Matrix norms

- Definition

$$\|A\| = 0 \text{ if } \mathrm{A} = 0 \text{ otherwise } \|A\| > 0;$$
$$\|kA\| = |k|\|A\| \quad (\text{the homogeneit condition});$$
$$\|A + B\| \leq \|A\| + \|B\|;$$
$$\|AB\| \leq \|A\|\|B\|.$$

- Any vector norm induces a matrix norm

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

- Important inequality

$$||Ax||_v \leq ||A||_M \cdot ||x||_v$$

Vector induced matrix norms

- L1 norm $\|A\|_1 = \max_{\|\mathbf{x}\|_1=1} \|A\mathbf{x}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|.$
  - Maximum column sum
- L infinity norm $\|A\|_\infty = \max_{\|\mathbf{x}\|_\infty=1} \|A\mathbf{x}\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|.$
  - Maximum row sum
- L2 norm $\|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2.$
  - Matrix spectral norm $\|A\|_2 = \max_{1\le i\le n} \sqrt{\lambda_i(A^T A)}.$

Frobenius norm

- The Frobenius norm

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2} .$$

- It is not an induced norm.

- It is equivalent to the vector norm when reshaping A into a vector

Equivalent norms

$$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n}\|\mathbf{x}\|_2,$$

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty,$$

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n\|\mathbf{x}\|_\infty.$$

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2,$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty,$$

$$\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty.$$

Error analysis

Error analysis

For a linear system $Ax = b$ having the true solution x and a computed solution $\tilde{x}$, we define

- Error vector: $e = \tilde{x} - x$
- Residual vector: $r = A\tilde{x} - b$

For two solutions, how do we evaluate which one is better?

- Look at the residual vector: smaller the better!

Condition number

- The condition number is defined by

$$\kappa(A) = ||A|| \cdot ||A^{-1}|| = \frac{\sigma_{max}(A)}{\sigma_{min}(A)}$$

- It indicates how close A is to being numerically singular.

- If $\kappa(A)$ is large, A is ill-conditioned; no expectation of a true solution or even close to it.

Pivoting

Pivoting

- Recall $a_{ij} \leftarrow a_{ij} - \left(\frac{a_{ik}}{a_{kk}}\right) a_{kj}$
- We must expect all quantities to be infected with roundoff errors.
- The roundoff error in a_{kj} is multiplied by $\left(\frac{a_{ik}}{a_{kk}}\right)$.
- The small pivot elements would lead to large multipliers and to worse roundoff errors.

Naïve Gaussian can fail

- Gaussian elimination would fail if $a_{11} = 0$.

$$\begin{cases} 0x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \qquad \begin{cases} \varepsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$$

- How about this for a small number $\epsilon \neq 0$?

$$\begin{cases} \varepsilon x_1 + x_2 = 1 \\ (1 - \varepsilon^{-1}) x_2 = 2 - \varepsilon^{-1} \end{cases}$$

On 8-digit decimal computer

- Consider $\epsilon = 10^{-9} \Rightarrow \epsilon^{-1} = 10^9$.
- To compute $2 - \epsilon^{-1}$, the computer must interpret the numbers as

$$
\begin{aligned}
\varepsilon^{-1} = 10^9 \quad &= \quad 0.10000\,000 \times 10^{10} \quad = \quad 0.10000\,00000\,00000\,0 \times 10^{10} \\
2 \quad &= \quad 0.20000\,000 \times 10^{1} \quad = \quad 0.00000\,00002\,00000\,0 \times 10^{10}
\end{aligned}
$$

- Thus $2 - \epsilon^{-1}$ is rounded to ϵ^{-1}.

Remedy

$$\begin{cases} 0x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \qquad \begin{cases} \varepsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$$

Switch the two rows

$$\begin{cases} x_1 + x_2 = 2 \\ 0x_1 + x_2 = 1 \end{cases} \qquad \begin{cases} x_1 + x_2 = 2 \\ \varepsilon x_1 + x_2 = 1 \end{cases}$$

Necessary to switch

- Given

$$\begin{cases} x_1 + x_2 = 2 \\ \varepsilon x_1 + x_2 = 1 \end{cases}$$

- After elimination

$$\begin{cases} x_1 + x_2 = 2 \\ (1-\varepsilon)x_2 = 1-2\varepsilon \end{cases}$$

- Solution

$$x_2 = 1-2\varepsilon/1-\varepsilon \approx 1$$

$$x_1 = 2 - x_2 \approx 1$$

- What if

$$\begin{cases} x_1 + \varepsilon^{-1}x_2 = \varepsilon^{-1} \\ x_1 + x_2 = 2 \end{cases}$$

- After elimination

$$\begin{cases} x_1 + \varepsilon^{-1}x_2 = \varepsilon^{-1} \\ (1-\varepsilon^{-1})x_2 = 2-\varepsilon^{-1} \end{cases}$$

- Solution

$$x_2 = (2-\varepsilon^{-1})/(1-\varepsilon^{-1}) \approx 1$$

$$x_1 = \varepsilon^{-1} - \varepsilon^{-1}x_2 \approx 0$$