Changing the Matrix

So far, only discussed changing the RHS, i.e. \( Ax = b \rightarrow A\hat{x} = \hat{b} \).

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

\[
Ax = b \rightarrow \hat{A}\hat{x} = \hat{b}.
\]

What can we say about the error due to an approximate matrix?

\[
\hat{x} = x + \delta x \quad \hat{A} = A + \delta A
\]

\[
\frac{\|\delta x\|}{\|x\|} \leq K \cdot \frac{\|\delta A\|}{\|A\|}
\]

\[
L \rightarrow \delta x = \hat{x} - x = \hat{x} - A^{-1}\hat{b} = A^{-1}(A\hat{x} - \hat{A}\hat{x}) = -A^{-1}\delta A\hat{x}
\]

\[
\frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\| \cdot ||\delta A|| \cdot \|\hat{x}\|
\]

\[
\|\delta x\| \leq \|\delta A\| \cdot \|A\|^{-1} \|x\|
\]

\[
k = \|A\| \cdot \|A^{-1}\|
\]
Changing Condition Numbers

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

$$DAx = Db, \quad \|DA\| \cdot \|A^{-1}D^{-1}\| \geq \|I\| = 1$$

$$AD\bar{x} = b, \quad \|AD\| \cdot \|D^{-1}A^{-1}\|$$

What is this called as a general concept?

- Left preconditioning $\pi = (DA)^{-1}(Db)$
- Right preconditioning $\bar{x} = (AD)^{-1}b, \quad x = D\bar{x}$
In-class activity: Matrix Norms and Conditioning
Singular Value Decomposition (SVD)

What is the *Singular Value Decomposition* of an $m \times n$ matrix?

\[
A = U \Sigma V^T
\]

- col. of $U$: left sing. vec.
- col. of $V$: righ sing. vec.
- sing. values: $\sigma_1, \ldots, \sigma_m$

\[
\begin{align*}
U^T U &= I \\
V^T V &= I \\
\Sigma &= \begin{pmatrix} \sigma_1 & \ldots & 0 \\ 0 & \ddots & \vdots \\ 0 & \ldots & \sigma_m \end{pmatrix}
\end{align*}
\]
Computing the 2-Norm

Using the SVD of $A$, identify the 2-norm.

$$A = U \Sigma V^T$$

- $\|A\|_2 = \|\Sigma\|_2 = \sigma_{\text{max}}$
- $\|QB\|_2 = \|B\|_2 = \|BQ\|_2$, $Q$ oth.

Express the matrix condition number $\text{cond}_2(A)$ in terms of the SVD:

$$\text{cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \sigma_{\text{max}} / \sigma_{\text{min}}$$
Not a matrix norm: Frobenius

The 2-norm is very costly to compute. Can we make something simpler?

\[ \|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} \]

Frobenius norm.

What about its properties?

- \( \|A\|_F \geq 0 \), \( \|A\|_F = 0 \iff A = 0 \)
- \( \|\lambda A\|_F = \lambda \|A\|_F \)
- \( \|A + B\|_F \leq \|A\|_F + \|B\|_F \).
Frobenius Norm: Properties

Is the Frobenius norm induced by any vector norm?

\[ \|A\|_F = \max_{\|x\|_2=1} \|Ax\|_2 \quad \|I\|_F = \sqrt{n} \]

How does it relate to the SVD?

\[ \|A\|_F = \sqrt{\sum_{i=1}^{\min(n,m)} \sigma_i^2} \]
\[ \|A\|_F = \sqrt{\text{tr}(A^T A)} \]

* Errata: the square root was missing from the original scribbles
Solving Systems: Simple cases

Solve $Dx = b$ if $D$ is diagonal. (Computational cost?)

$$x_i = b_i / D_{ii} \quad \mathcal{O}(n)$$

Solve $Qx = b$ if $Q$ is orthogonal. (Computational cost?)

$$x = Q^T b \quad \mathcal{O}(n^2)$$

Given SVD $A = U\Sigma V^T$, solve $Ax = b$. (Computational cost?)

$$U \Sigma V^T x = b \quad y = V^T x$$

$$\Sigma V^T x = U^T b$$

$$V y = x$$

$$\mathcal{O}(n^2) \rightarrow \mathcal{O}(n^3) \text{ to find } x$$
Solving Systems: Triangular matrices

Solve

\[
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} & a_{14} \\
  a_{22} & a_{23} & a_{24} \\
  a_{33} & a_{34} \\
  a_{44}
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  z \\
  w
\end{bmatrix}
= 
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3 \\
  b_4
\end{bmatrix}.
\]

Demo: Coding back-substitution [cleared]

What about non-triangular matrices?

\[
\begin{align*}
a_{37} x + a_{34} w &= b_3 \\
a_{44} w &= b_4
\end{align*}
\]
Gaussian Elimination

**Demo:** Vanilla Gaussian Elimination [cleared]

What do we get by doing Gaussian Elimination?

\[ \text{REF} \]

How is that different from being upper triangular?

What if we do not just eliminate downward but also upward?

\[ \text{Gauss-Jordan elimination} \]
What is the **LU factorization**?

\[ A = LU \]

- \( L \) - lower \( \Delta \), diag of all ones
- \( U \) - upper

\[ a_{11} = \ell_{11} \cdot u_{11} \]
Solving $Ax = b$?

Does LU help solve $Ax = b$?

$A = LU$

$Ax = b$

$Ly = b \Leftrightarrow \text{forward subst.} \Rightarrow O(n^2)$

$Ux = y \Leftrightarrow \text{back subst.} \Rightarrow O(n^2)$
Demonstrating an LU factorization

\[ A = \begin{bmatrix} l_{11} & \tilde{a}_{12} \\ \tilde{a}_{12} & \begin{bmatrix} a_{11} & a_{12}^T \end{bmatrix} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ \tilde{a}_{12} & A_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ U_{22} \end{bmatrix} \]

\[ u_{11} = a_{11} \]
\[ u_{12} = a_{12}^T \]
\[ \tilde{a}_{12} = u_{11} \cdot \tilde{q}_{12} \]
\[ \tilde{q}_{12} = \tilde{a}_{12} / u_{11} \]
\[ A_{22} = \tilde{q}_{12} U_{22} \]
\[ \tilde{q}_{22} U_{22} = A_{22} - l_{11} u_{12}^T \]

**Demo:** LU Factorization [cleared]
Computational Cost

What is the computational cost of multiplying two $n \times n$ matrices?

$$(AB)_{ij} = \sum_k a_{ik} b_{kj} \quad O(n^3)$$

- $u_{11} = a_{11}, \quad u_{12}^T = a_{12}^T$.
- $l_{21} = a_{21}/u_{11}$.
- $L_{22} U_{22} = A_{22} - l_{21} u_{12}^T$.

What is the computational cost of carrying out LU factorization on an $n \times n$ matrix?

$O(n^3)$

**Demo:** Complexity of Mat-Mat multiplication and LU [cleared]