

Last time: SVD of matrix  $A \in \mathbb{C}^{m \times n}$

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

$\uparrow \quad \quad \uparrow \quad \quad \swarrow$

$U \in \mathbb{C}^{m \times m} \quad \Sigma \in \mathbb{R}_{\geq 0}^{m \times n} \quad V \in \mathbb{C}^{n \times n}$   
 $U^* U = I_m \quad \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \quad V^* V = I_n$   
 $p = \min\{m, n\}$

• Existence: ①  $A^* A = V \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_p^2 \\ & & & 0 \end{bmatrix} V^*$   
 $\Sigma^2$

②  $U \Sigma = A V$

• Practical algorithm: Solve eigenvalue problem for

$$H = \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} \leftarrow \text{Hermitian}$$

use Golub-Kahan bidiagonalization: operation count  $\approx 4mn^2$  FLOPs  
(will go into detail when we talk about eigenvalue problems)

- Application 1: Point cloud registration
- Application 2: Overdetermined least-squares

Given  $b(a) \in L^2([0, 1])$

Find  $n$ -degree polynomial  $p(a) \in \mathcal{P}_n([0, 1])$

such that  $\min_{p \in \mathcal{P}_n} \int_0^1 |p(a) - b(a)|^2 da$

Discretize the problem:

quadrature:  $\int_0^1 f(a) da \approx \frac{1}{m} \sum_{i=1}^m f(a_i)$ ,  $a_i = \frac{i-1}{m-1}$ ,  $i = 1, \dots, m$

$$p(x) = \sum_{k=0}^{n-1} x_k a^k$$

$$\Rightarrow \min_{c \in \mathbb{R}^n} \frac{1}{m} \sum_{i=1}^m \left| \sum_{k=0}^{n-1} x_k a_i^k - b(x_i) \right|^2$$

$$\text{Let } A = \begin{bmatrix} 1 & a_1 & \dots & a_1^{n-1} \\ 1 & a_2 & \dots & a_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_m & \dots & a_m^{n-1} \end{bmatrix} \quad b = (b(x_1), \dots, b(x_m))^T$$

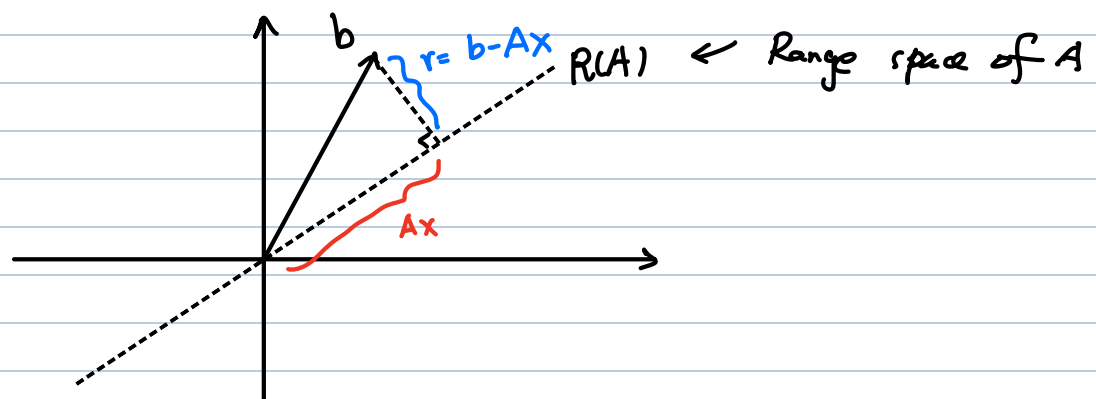
$$x = (x_1, \dots, x_n)^T$$

$$\Rightarrow \min_x \|Ax - b\|_2^2 \quad \leftarrow \text{least-squares problem}$$

$m \text{ equations} > n \text{ variables}$

$$\begin{matrix} & n \\ m & \begin{bmatrix} \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \end{bmatrix} \end{matrix} \begin{matrix} \begin{bmatrix} \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \end{bmatrix} \end{matrix} = \begin{matrix} \begin{bmatrix} \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \end{bmatrix} \\ b \end{matrix}$$

$A \quad x \quad b$



• Solve least-squares via normal equations

$$\hat{x} \in \arg \min_x \|Ax - b\|_2^2$$

$$\Leftrightarrow r = b - A\hat{x} \perp R(A)$$

$$\Leftrightarrow b - A\hat{x} \in N(A^*)$$

$$\Leftrightarrow A^*A\hat{x} = A^*b \quad \leftarrow \text{Normal equations}$$

assume  $\text{rank}(A) = n$  so  $\text{rank}(A^*A) = \text{rank}(A) = n$

and  $A^*A$  invertible

- Solving least-squares via normal equations

Step 1: compute  $C = A^*A$ ,  $d = A^*b$

Step 2: Solve  $C\hat{x} = d$  operation count  $\approx 2mn^2$  FLOPs ( $m \gg 1$ )

- Solving normal equations directly can suffer from ill-conditioning

For the least square problem using monomials,

$$\text{as } m \rightarrow +\infty, \frac{1}{m} (A^*A)_{ij} = \frac{1}{m} \sum_{k=1}^m a_k^{i+j-2}$$

$$= \frac{1}{m} \sum_{k=1}^m \left( \frac{k-1}{m-1} \right)^{i+j-2}$$

$$\rightarrow \int_0^1 a^{i+j-2} da = \frac{1}{i+j-1}$$

So  $\frac{1}{m} A^*A \rightarrow H_n$  Hilbert matrix - ill-conditioned!

Beyond NLA: monomials are "bad" basis,  
choose basis that are orthogonal to each other

$$\text{In general, } \kappa_2(A^*A) = \frac{\sigma_1^2}{\sigma_n^2} \gg \frac{\sigma_1}{\sigma_n} (= \kappa_2(A))$$

- Solve least-squares via SVD

$$\text{Let } A = U \Sigma V^*$$

$$\min_x \|U \Sigma V^* x - b\|_2^2$$

$$\text{Let } d = U^*b, \quad y = V^*x$$

$$\min_y \|\Sigma y - d\|_2^2 = \min_{y \in \mathbb{C}^n} \left\| \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} \right\|_2^2$$

$$\text{Let } \Sigma^+ = \begin{bmatrix} d_1 & & & \\ & \ddots & & \\ & & d_n & \\ & & & 0 \end{bmatrix}$$

$$\hat{y} := \Sigma^+ d = \arg \min \|\Sigma y - d\|_2^2$$

$$\text{hence } \hat{x} = V \hat{y} = \underbrace{V \Sigma^+ U^*}_{A^+ \text{ pseudoinverse of } A} b \in \arg \min \|Ax - b\|^2$$

Actually, from normal equation,

$$\begin{aligned} \hat{x} &= \underbrace{(A^* A)^{-1} A^*}_{= A^+} b \\ &= (V \Sigma^* \Sigma V^*)^{-1} V \Sigma^* U^* b \\ &= V (\Sigma^* \Sigma)^{-1} V^* V \Sigma^* U^* b \\ &= V \Sigma^+ U^* b \end{aligned}$$

Algorithm:

Step 1:	Compute SVD : $A = U \Sigma V^*$	operation count $\approx 4mn^2$ FLOPs
Step 2:	$\hat{x} = V \Sigma^+ U^* b$	

- Stability of least-squares

$$\text{Define } \kappa_2(A) = \|A\|_2 \|A^+\|_2 = \sqrt{\kappa_2(A^* A)}$$

Thm (Wedin) Let  $A \in \mathbb{R}^{m \times n}$ ,  $A + \Delta A$  both be of full rank

$$\text{Let } \|b - Ax\|_2 = \min \quad , \quad r = b - Ax$$

$$\|(b + \Delta b) - (A + \Delta A) y\|_2 = \min \quad , \quad s = b + \Delta b - (A + \Delta A) y$$

$$\frac{\|\Delta A\|_2}{\|A\|_2} \quad , \quad \frac{\|\Delta b\|_2}{\|b\|_2} \leq \epsilon_{\text{mach}}$$

Then provided that  $\kappa_2(A) \epsilon_{\text{mach}} < 1$

we have

$$\frac{\|x-y\|_2}{\|x\|_2} \leq \frac{K_2(A) \epsilon_{mach}}{1 - K_2(A) \epsilon_{mach}} \left( 2 + (K_2(A) + 1) \frac{\|r\|_2}{\|A\|_2 \|x\|_2} \right)$$

$$\frac{\|r-s\|_2}{\|b\|_2} \leq (1 + 2K_2(A)) \epsilon_{mach}$$

These bounds are approximately attainable

Pf: See Higham Thm 20.1  $\square$

Remark: The first bound is usually interpreted as

saying that the sensitivity of least-squares is measured by  $K_2(A)$  when  $\|r\|$  is small or zero

and by  $K_2^2(A)$  otherwise

Thm Solving least-squares via SVD is backward stable.

The computed  $\hat{x}$  satisfies

$$\hat{x} = \underset{x}{\operatorname{argmin}} \| (A + \Delta A)x + b + \Delta b \|_2$$

$$\text{and } \frac{\|\Delta A\|}{\|A\|} \cdot \frac{\|\Delta b\|}{\|b\|} \leq C_{min} \epsilon_{mach}$$

Summary:	method	normal eqn	SVD
	conditioning	$\approx K_2^2(A)$	$\approx K_2(A)$ ( $\ r\ $ small)
	operation count	$\approx 2mn^2$ FLOPs	$\approx 4mn^2$ FLOPs

- Solving least-squares via QR factorization

Computing SVD is expensive. can we work with orthogonal transform but lower cost?

• (reduced) QR factorization

$$A = Q R$$

$\uparrow$   
 $Q \in \mathbb{C}^{m \times n}$   
 $Q^* Q = I_n$   
 orthonormal columns (  $m \geq n$  )

$\leftarrow$   
 $R \in \mathbb{C}^{n \times n}$   
 upper triangular matrix

Not  $Q Q^* = I_m$

when  $\text{rank}(A) = n$ ,  $\Rightarrow \text{rank}(R) = n$

hence  $R(A) = R(Q)$

Let  $Q = [q_1 \dots q_n]$ ,  $q_i \in \mathbb{C}^m$

projection  
onto  $R(A)$

Then  $P_A x = P_Q x = \sum_{i=1}^n q_i (q_i^* x) = Q Q^* x$

Hence  $\hat{x} \in \arg \min_x \|Ax - b\|_2^2$

$$\Leftrightarrow r = b - A\hat{x} \perp R(A)$$

$$\Leftrightarrow P_Q r = 0$$

$$\Leftrightarrow Q R \hat{x} = Q Q^* b$$

$$\Leftrightarrow R \hat{x} = Q^* b$$

$$\Leftrightarrow \hat{x} = R^{-1} Q^* b$$

Actually, from normal equations,

$$\hat{x} = (A^* A)^{-1} A^* b = (R^* R)^{-1} R^* Q^* b = R^{-1} Q^* b$$