Linear Systems of Eq.'s

Part 1: The singular value decomposition (SVD)

m linear Eqn's in n unknowns:

$$
\begin{aligned}
a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n &= b_1 \\
a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n &= b_2 \\
&\vdots \\
a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n &= b_m
\end{aligned}
$$

Matrix notation $\Rightarrow$

$$
m \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{n} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{n \text{ unknowns}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{m \text{ eqns}}
$$

Fundamental task: "solve" $Ax = b$

In high-level languages like Julia/Matlab

$$x = A \backslash b$$

"backslash"

$\Rightarrow$ Julia example: square systems ; under/overdetermined systems

What does it mean to solve $Ax = b$?

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

columns of A

$$x_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} + \dots + x_n \begin{bmatrix} | \\ a_n \\ | \end{bmatrix} = \begin{bmatrix} | \\ b \\ | \end{bmatrix}$$

$\Rightarrow$ "Find the coordinates of b in the basis $\{a_1, a_2, \dots, a_n\}$"

$b = b_1 e_1 + b_2 e_2 + \dots + b_n e_n$

$\Downarrow$

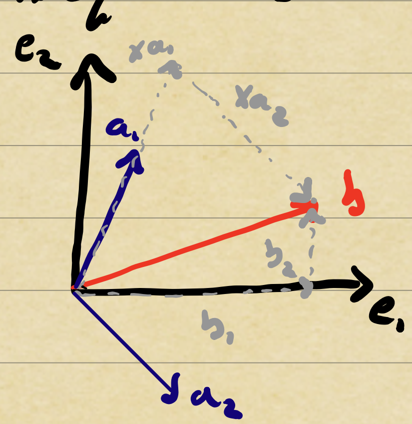
$b = x_1 a_1 + x_2 a_2 + \dots + x_n a_n$

$e_j = [0, \dots, 0, \underset{j^{th}\text{ entry}}{1}, 0, \dots, 0]^T$

canonical basis

When can we perform this change of basis?

Case 1: $n = m$ #unknowns = #equations



If A is square and invertible,

$\Rightarrow x = A^{-1} b$

unique solution

Thm 1.3 (Characterizations of invertibility)

These conditions are equivalent for $A \in \mathbb{R}^{m \times m}$

(a) A is invertible

(b) $\text{rank}(A) = m$

(c) $\text{range}(A) = \mathbb{R}^m$

(d) $\text{null}(A) = \{0\}$

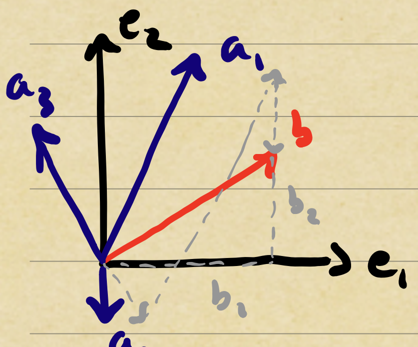
(e) $0 \notin \lambda(A)$ (no zero eigenvalues)

(f) $0 \notin \sigma(A)$ (no zero singular vals)

(g) $\det(A) \neq 0$

If any of these conditions conditions is unfamiliar, please read Lecture 1 in Trefethen carefully and review 18.06 mat. as necessary.

Case 2: $n > m$ # unknowns > # equations



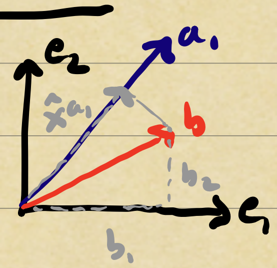
If columns of A are linearly independent, then there is no <u>unique</u> solution.

$\Rightarrow$ many ways to write

$x_1 a_1 + x_2 a_2 + x_3 a_3 = b$

$\Rightarrow$ We need to specify additional constraints, like a minimal norm soln.

Case 3: $n < m$ # unknowns < # equations

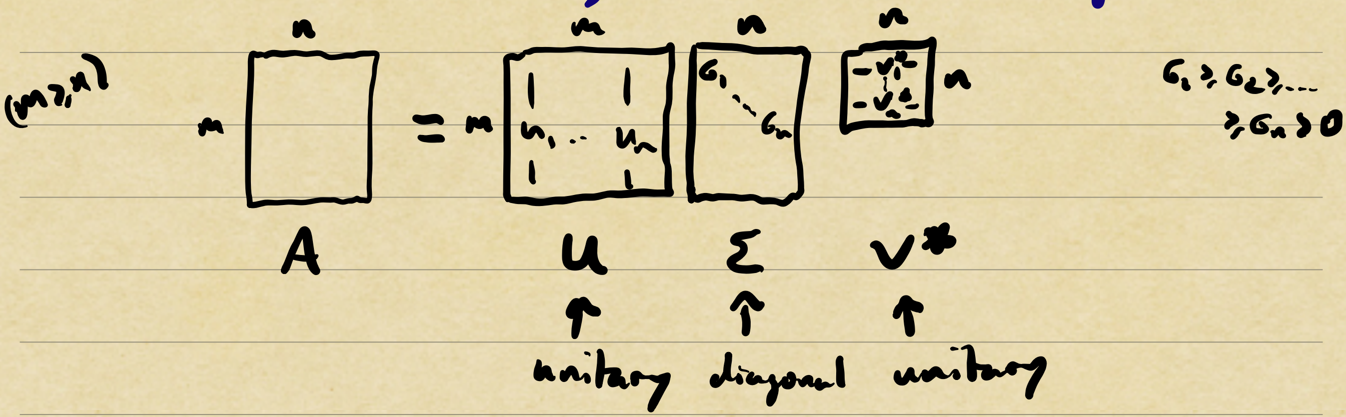


$\hat{x}$ = least squares soln.

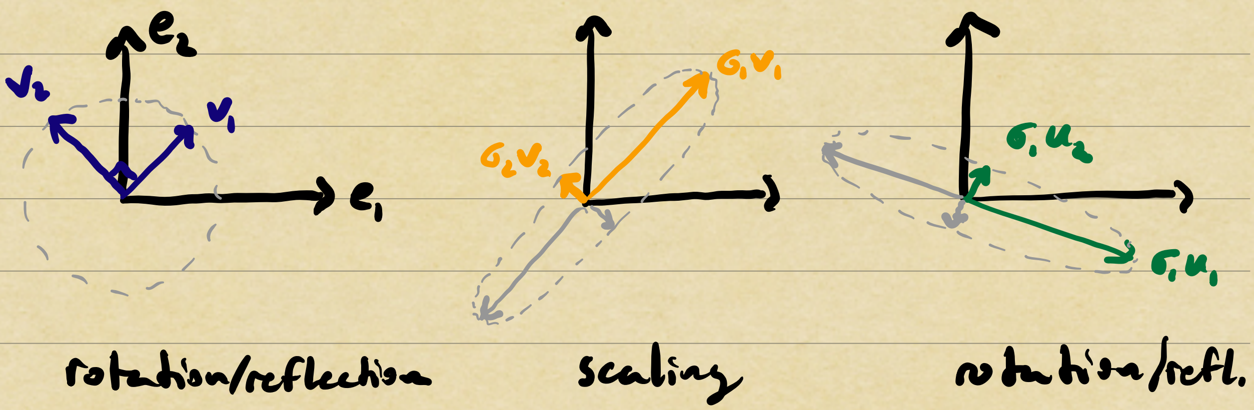
minimizes $\|Ax - b\|_2$

$= \sqrt{\text{"sum of squares"}}$

A powerful way to understand all 3 cases (and many other topics in NLA) is with the singular value decomposition.



The SVD is one of many LA factorizations that decompose a matrix into a product of highly structured matrices, whose individual action is easier to understand.



Unitary / orthogonal matrices
(complex) (real)

$$\tilde{Q} = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_m \\ | & & | \end{bmatrix} \quad \text{if} \quad q_i^* q_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$$Q^*Q = \begin{bmatrix} - & q_1^* & - \\ & \vdots & \\ - & q_m^* & - \end{bmatrix} \begin{bmatrix} | & & | \\ q_1 & \cdots & q_m \\ | & & | \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & & \\ 0 & 1 & 0 & \\ & 0 & \ddots & 0 \\ & & 0 & 1 \end{bmatrix}}_{I}$$

$Q^{-1} = Q^*$ for a unitary matrix!

Multiplying by Q means changing to a new orthonormal basis.

$\Rightarrow$ Length $\|Qx\|_2 = \|x\|$!

b/c $\|Qx\|_2^2 = (Qx)^*(Qx) = x^*Q^*Qx = x^*x$
$= \|x\|_2^2$

key idea: Since U and V are both unitary, the SVD usually allows us to replace

$$Ax = b \Leftrightarrow U\Sigma V^*x = b \Leftrightarrow \Sigma \underbrace{V^*x}_{y} = \underbrace{U^*b}_{d}$$

with a diag matrix $\Rightarrow \Sigma y = d$

Thm Every $A \in \mathbb{R}^{m \times n}$ has an SVD.
(See Trefethen Lecture 5 for uniqueness*)

Case 1 $m = n$

$$\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$$

invertible IFF $\sigma_n > 0$.

If $\sigma_{k+1} = 0$, then $\text{rank}(A) = k$.

reduced SVD: drop zero σ's

$$A = \begin{array}{c} \\ m \end{array} \begin{array}{c} k \\ \begin{bmatrix} u_1 & \cdots & u_k \end{bmatrix} \end{array} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} - v_1^* - \\ \vdots \\ - v_k^* - \end{bmatrix}$$

Case 2 $n > m$

$$\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{array}\right] \begin{bmatrix} y_1 \\ \vdots \\ y_m \\ y_{m+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}$$

Solution if $\sigma_m > 0$! We say A has full row rank if $\text{rank}(A) = m$

Case 3 $n < m$

$$\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ \vdots \\ d_m \end{bmatrix}$$

Exact soln possible only if

$d_{n+1}, \dots, d_m = 0$ and $\sigma_n > 0$

We say A has full column rank if $\operatorname{rank}(A) = n$

In general, least squares residual is

$$\|Ax - b\| = \sqrt{\sum_{i=n+1}^{m} |d_i|^2}$$