Square Linear Systems

Goal: solve n linear eq.'s in n unknowns.

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n
\end{aligned}
$$

$$
\begin{matrix} n \\ \text{equations} \end{matrix}
\overset{A}{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}
\underset{\begin{matrix} n \\ \text{unknowns} \end{matrix}}{\overset{x}{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}}
=
\overset{b}{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}
$$

Computational task: "solve" $Ax = b$

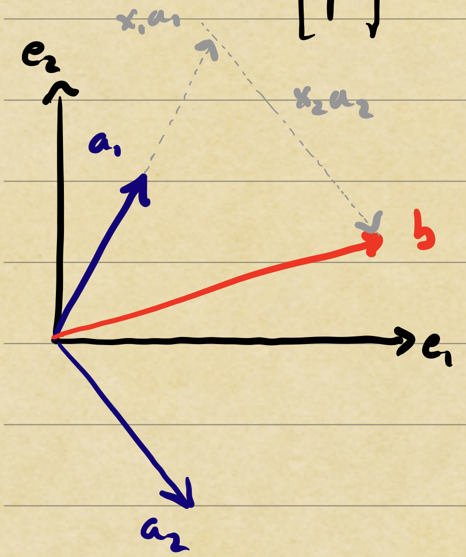
In high-level languages like Julia/Matlab

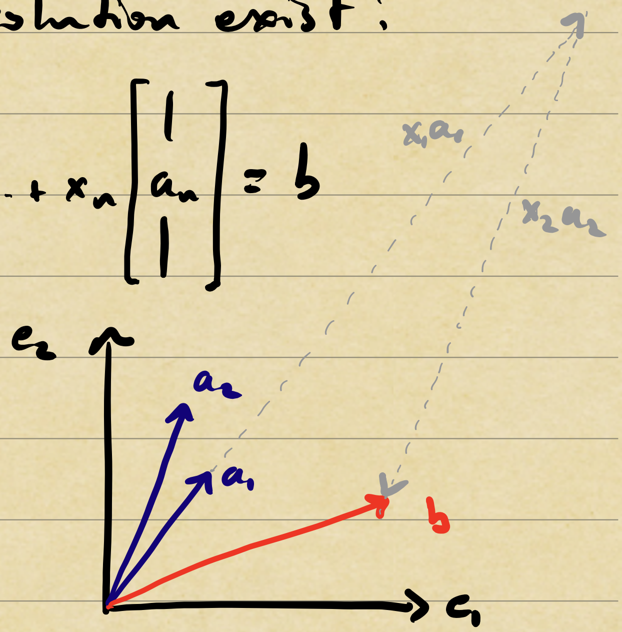
$$x = A\backslash b$$

↑ "backslash" = mldivide

Q1: When does a solution exist?

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





If columns of A are linearly independent, A is invertible and unique soln is $x = A^{-1}b$.

See Trefethen Lecture 1 Thm 1.3 for list of equivalent conditions.

Q2: When is $Ax = b$ well-conditioned?

(Solution sensitive to perturbations)

Sensitivity to perturbations in b (not A)

$$\underset{\substack{\uparrow \\ \text{output}}}{x} = A^{-1} \underset{\substack{\uparrow \\ \text{input}}}{b}$$

how much can A^{-1} amplify b?

From Lecture 3 Def.

$$\kappa_{A^{-1}}(b) = \frac{\|J_{A^{-1}}(b)\|}{\|A^{-1}b\|/\|b\|} = \frac{\|A^{-1}\|\,\|b\|}{\|A^{-1}b\|}$$

how much can A^{-1} shrink b?

We can bound $\kappa_{A^{-1}}(b)$ independent of b

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

$$\Rightarrow \frac{\|b\|}{\|A^{-1}b\|} \leq \|A\|$$

$$\kappa_{A^{-1}}(b) \leq \underbrace{\|A^{-1}\|\,\|A\|}_{\text{"Condition \# of A"}} = \kappa(A)$$

Condition # of A is also condition #

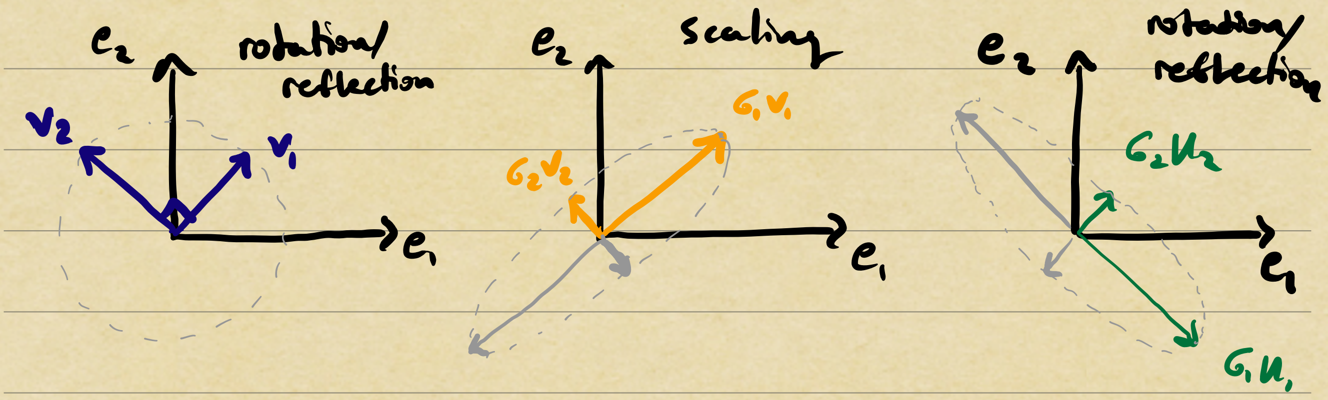
of $x = A^{-1}b$ w.r.t. perturbations in A.

The Singular Value Decomposition

A powerful way to understand how A shrinks & stretches different vectors is through the singular value decomposition (SVD).

$$A = \underset{\substack{U \\ \text{unitary}}}{\overset{\text{left singular vecs}}{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}}} \underset{\substack{\Sigma \\ \text{diagonal} \\ \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n}}{\overset{\text{singular vals}}{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}}} \underset{\substack{V^* \\ \text{unitary}}}{\overset{\text{right singular vecs}}{\begin{bmatrix} - & v_1^* & - \\ & \vdots & \\ - & v_n^* & - \end{bmatrix}}}$$

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



Unitary / Orthogonal matrices
(complex) (real)

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

equivalently $Q^*Q = I$ $\quad (Q^{-1} = Q^*!)$

Multiplying by Q preserves length

$$\|Qx\|_2 = \|x\|_2$$

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

Key Idea: Express $\kappa(A)$ using singular vals of A.

$$Ax = U \underset{\uparrow \text{ all amplification/stretching happen here}}{\Sigma} V^* x$$

(1) $$\|A\|_2 = \sup_{x \in \mathbb{R}^n} \frac{\|U\Sigma V^* x\|_2}{\|x\|_2} = \sup_{x \in \mathbb{R}^n} \frac{\|\Sigma V^* x\|_2}{\|x\|_2}$$

$$\underset{\nearrow \; y = V^*x}{=} \sup_{y \in \mathbb{R}^n} \frac{\|\Sigma y\|_2}{\|Vy\|_2} = \sup_{y \in \mathbb{R}^n} \frac{\|\Sigma y\|_2}{\|y\|_2}$$

$$= \|\Sigma\|_2$$

2-norm is unitarily invariant

(2) $$\|\Sigma\|_2 = \sup_{x \in \mathbb{R}^n} \frac{\|\Sigma x\|_2}{\|x\|_2} = \sigma_1$$

Similarly $$A^{-1} = V \Sigma^{-1} U^*$$

$$\hookrightarrow = \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_n^{-1} \end{bmatrix}$$

$$\|A^{-1}\|_2 = \|\Sigma^{-1}\|_2 = 1/\sigma_n^{-1}$$

↑ largest entry in Σ^{-1}

$$\Rightarrow \quad \kappa(A) = \|A\| \|A^{-1}\|$$

$$= \sigma_1/\sigma_n$$

← ratio of extreme singular values of A

σ_1 - amplifying action

σ_n - shrinking action