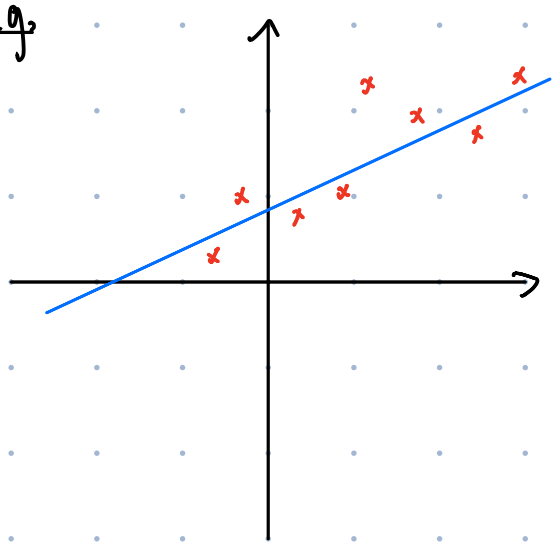
Office hour: ~~Thur 12-1:30~~ $\rightarrow$ Fri. 12-1:30 PDT.

Least square problem:

e.g.



Given data $\{(x_1,y_1),\dots,(x_n,y_n)\}$.

Want to find $y = ax+b$ approximating the data.

$y_1 \sim ax_1+b$
$y_2 \sim ax_2+b$
$\vdots$

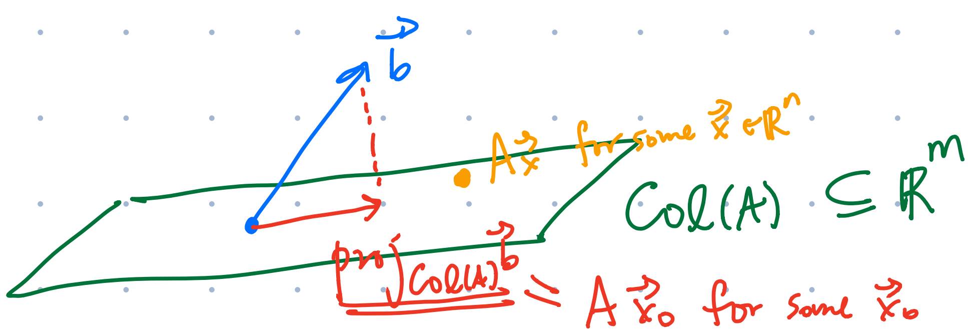
i.e. find a, b s.t.

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \text{ "closest" to } \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Def: A: $m\times n$ matrix, $\vec b \in \mathbb{R}^m$

A <u>least square sol^n</u> of "$A\vec x = \vec b$" is an $\vec x_0 \in \mathbb{R}^n$ s.t.

$$\|\vec b - A\vec x_0\| \le \|\vec b - A\vec x\| \quad \forall \vec x \in \mathbb{R}^n.$$



For any $\vec{x}_0 \in \mathbb{R}^n$ s.t. $A\vec{x}_0 = \text{proj}_{Col(A)}\vec{b}$,

we have $\|\vec{b} - A\vec{x}_0\| \le \|\vec{b} - A\vec{x}\| \quad \forall \vec{x} \in \mathbb{R}^n$.

$\Rightarrow \quad b - A\vec{x}_0 = b - \text{proj}_{Col(A)}\vec{b} \in Col(A)^\perp$

$\Leftrightarrow \quad \underline{A^T(b - A\vec{x}_0) = 0} \quad (\because Col(A)^\perp = Nul(A^T))$

$\Rightarrow \quad \underline{A^TA\vec{x}_0 = A^T\vec{b}}$.

"normal eq'n" for $A\vec{x} = \vec{b}$.

Thm: $A^TA\vec{x}_0 = A^T\vec{b}$ always has a sol^n, and the sol^n(s) are the least square sol^n of $A\vec{x} = \vec{b}$.

Thm A: $\overset{\text{real}}{m \times n}$. The following are equivalent:

1) $A\vec{x} = \vec{b}$ has a unique least square sol^n

2) A^TA is invertible

3) columns of A are l.i. $\Leftrightarrow Nul(A) = \{\vec{0}\}$

pf. $\underline{1) \Leftrightarrow 2)}$ clear.

$\underline{2) \Leftrightarrow 3)}$: $Nul(A^TA) \supseteq Nul(A) \Rightarrow$ "2) $\Rightarrow$ 3)"

"$Nul(A^TA) \subseteq Nul(A)$"

$\vec{v} \in Nul(A^TA)$

$A^TA\vec{v} = \vec{0}$

$\Downarrow$

$0 = \boxed{\vec{v}^TA^T}\boxed{A\vec{v}} = (A\vec{v})^T(A\vec{v}) \Rightarrow A\vec{v} = \vec{0}$ $\square$

$\langle A\vec{v}, A\vec{v} \rangle$ $\|$

$A\vec{v} \in \mathbb{R}^n$

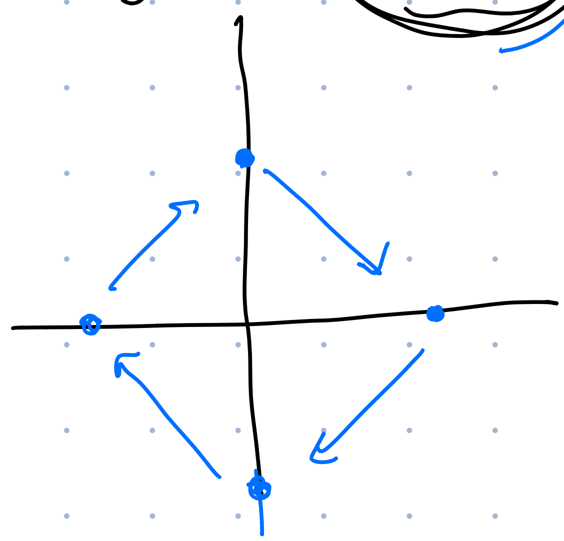
Rmk When A has l.i. columns, $A = QR$
the unique least square sol^n of $A\vec{x} = \vec{b}$ is given by:

$$\boxed{\vec{x}_0 = R^{-1} Q^T \vec{b}.}$$

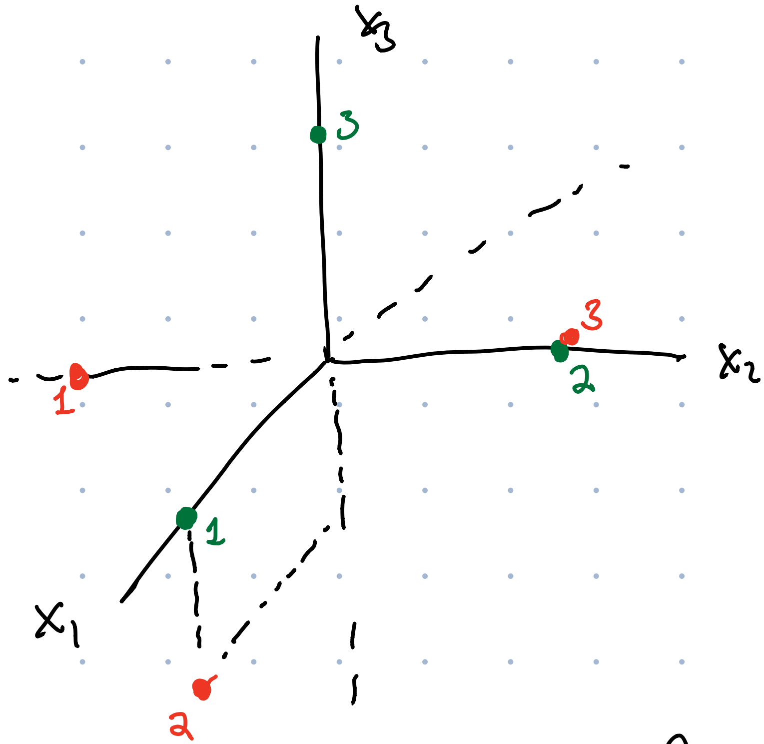
Check: $A^T A \vec{x}_0 \overset{??}{=} A^T \vec{b}$

$$A^T A \vec{x}_0 = R^T Q^T Q \cancel{R} \cancel{R^{-1}} Q^T \vec{b} = R^T \underbrace{Q^T Q}_{= I} Q^T \vec{b} = R^T Q^T \vec{b} = A^T \vec{b}$$

$$\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} = b \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$



$\lambda_1, \dots, \lambda_k$ are distinct eigvalues of A.

generalized eigenspace $\quad$ A: $n\times n$ $\quad$ λ eigenvalue of A

$$V^g_\lambda := \{\vec{v} \in \mathbb{C}^n \mid (A-\lambda I)^k \vec{v} = \vec{0} \text{ for some } k \geq 1\}.$$

Thm: $$\mathbb{C}^n = V^g_{\lambda_1} \oplus \cdots \oplus V^g_{\lambda_k}$$

Sketch of proof:

$k=1$: A has only one eigenvalue λ_0 $\Rightarrow$ char. poly $p(\lambda) = (\lambda_0 - \lambda)^n$

$\boxed{\text{Need to show: } \mathbb{C}^n = V^g_{\lambda_0} \quad \text{i.e. } \forall \vec{v} \in \mathbb{C}^n, \exists k \geq 1 \text{ s.t. } (A-\lambda_0 I)^k \vec{v} = \vec{0}}$

Cayley-Hamilton thm: If p is the char. poly. of A, then $p(A) = 0$

$p = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0$

0 = zero matrix

$$\boxed{A^n + b_{n-1}A^{n-1} + \cdots + b_0 I = 0}$$

(In HW5 #6, you proved C-H for diag. A)

char. poly. of A is $(\lambda_0 - \lambda)^n$

By C-H $\Rightarrow$ $\boxed{(\lambda_0 I - A)^n = 0}$

$\Rightarrow$ $(A - \lambda_0 I)^n \vec{v} = \vec{0}$

$\Rightarrow$ $\mathbb{C}^n = V^g_{\lambda_0}$. □

Rmk A is nilpotent ($A^k = 0$ for some $k \geq 1$)

In HW, $\Rightarrow$ 0 is the only eigenvalue of A.

C-H

Suppose the statement is true for any linear transf. w/ $\leq k-1$ eigenvalues

$\boxed{\text{Suppose } A \text{ has } k \text{ eigenvalues}}$ $\boxed{\text{WTS: } A \text{ also has such decomp.}}$

HW5 extra credit: $T : V \to V$ finite dim'l v.s. V.

$0 \subseteq \ker(T) \subseteq \ker(T^2) \subseteq \cdots \subseteq \underbrace{\ker(T^k) = \ker(T^{k+1}) = \cdots}_{= \ker^s(T)}$

$V \supseteq \operatorname{Im}(T) \supseteq \operatorname{Im}(T^2) \supseteq \cdots \supseteq \underbrace{\operatorname{Im}(T^k) = \operatorname{Im}(T^{k+1}) = \cdots}_{= \operatorname{Im}^s(T)}$

- $V = \ker^s(T) \oplus \operatorname{Im}^s(T)$
- T preserves this decomp.: $T(\ker^s(T)) \subseteq \ker^s(T)$
  $T(\operatorname{Im}^s(T)) \subseteq \operatorname{Im}^s(T)$
- $T|_{\ker^s(T)}$: nilpotent, $T|_{\operatorname{Im}^s(T)}$: invertible

---

Apply to: $T = T_{(A - \lambda_k \mathbb{I})}$

$$\mathbb{C}^n = \boxed{\ker^s(A - \lambda_k \mathbb{I})} \oplus \boxed{\operatorname{Im}^s(A - \lambda_k \mathbb{I})}$$

$$\ker^s(A - \lambda_k \mathbb{I}) = V^g_{\lambda_k} = \{\vec{v} \mid (A - \lambda_k \mathbb{I})^{\ell} \vec{v} = \vec{0} \text{ for some } \ell \geq 1\}$$

Claim: $T_{A - \lambda_k \mathbb{I}}|_{\operatorname{Im}^s(A - \lambda_k \mathbb{I})}$ only has eigenvalues $\lambda_1, \ldots, \lambda_{k-1}$.

pf $T_{A - \lambda_k \mathbb{I}}|_{\operatorname{Im}^s(A - \lambda_k \mathbb{I})} : \operatorname{Im}^s(A - \lambda_k \mathbb{I}) \longrightarrow \operatorname{Im}^s(A - \lambda_k \mathbb{I})$

invertible.

$\forall \vec{v} \in \operatorname{Im}^s(A - \lambda_k \mathbb{I}) \setminus \{\vec{0}\}$,

$$(A - \lambda_k \mathbb{I}) \vec{v} \neq \vec{0}$$

$\Rightarrow$ λ_k is not an eigenvalue of $T_{A - \lambda_k \mathbb{I}}|_{\operatorname{Im}^s(A - \lambda_k \mathbb{I})}$

By inductive hypothesis,

$$\operatorname{Im}^s(A - \lambda_k \mathbb{I}) = \left(\operatorname{Im}^s(A - \lambda_k \mathbb{I}) \cap V^g_{\lambda_1}\right) \oplus \cdots \oplus \left(\operatorname{Im}^s(A - \lambda_k \mathbb{I}) \cap V^g_{\lambda_{k-1}}\right)$$

$$\operatorname{Im}^s(A - \lambda_k \mathbb{I}) \cap V^g_{\lambda_1} = V^g_{\lambda_1}, \qquad \operatorname{Im}^s(A - \lambda_k \mathbb{I}) \cap V^g_{\lambda_{k-1}} = V^g_{\lambda_{k-1}}$$

Obs: $(A-\lambda_1 I)$ preserves $\ker^s(A-\lambda_k I) \oplus \operatorname{Im}^s(A-\lambda_k I)$

$$V^g_{\lambda_1} = \ker^s(A-\lambda_1 I) = \underbrace{(\ker^s(A-\lambda_1 I) \cap \ker^s(A-\lambda_k I))}_{\{\vec{0}\}} \oplus (\ker^s(A-\lambda_1 I) \cap \operatorname{Im}^s(A-\lambda_k I))$$

Prop $S: V \to V$, $V = U \oplus W$, $S(U) \subseteq U$, $S(W) \subseteq W$,
then $\ker^s(S) = (\ker^s(S) \cap U) \oplus (\ker^s(S) \cap W)$

The only thing left: $\boxed{\ker^s(A-\lambda_1 I) \cap \ker^s(A-\lambda_k I)} = \{0\}$

$\forall \vec{v} \in W \subseteq \mathbb{C}^n$, $W \neq \{\vec{0}\}$

$\forall \vec{v}$, $(A-\lambda_1 I)^k \vec{v} = \vec{0}$ for some $k \geq 1$.

Claim: The only eigenvalue of A on W is λ_1

$Aw = \lambda' w$ $\quad \vec{w} \neq \vec{0}$

$(A-\lambda_1 I)^k w = 0$

$\Rightarrow (\lambda'-\lambda_1)^k w = 0$

$\Rightarrow \lambda' = \lambda_1$

Cayley-Hamilton. (sketch of proof)

$$f: M_{n\times n}(\mathbb{C}) \longrightarrow M_{n\times n}(\mathbb{C})$$
$$A \longmapsto p_A(A)$$

where p_A is the char. poly. of A.

- A is diagonizable, $f(A) = p_A(A) = 0$
- $\{$diagonalizable matrices$\} \subseteq M_{n \times n}(\mathbb{C})$ is "dense"
- f is continuous.

$\Longrightarrow$ $f(A) = 0$ for any A.