Today: QR decomposition, least square problem, Q&A

Recap (Gram-Schmidt process) V: inner product space

$\{\vec{v}_1, \dots, \vec{v}_n\}$ l.i. set in V

$\Big\downarrow$ Gram-Schmidt process

$\{\vec{w}_1, \dots, \vec{w}_n\}$, where:

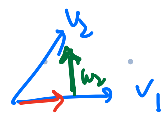
- they're l.i.,
- it's an orthogonal set.
- $\operatorname{Span}\{\vec{w}_1, \dots, \vec{w}_n\} = \operatorname{Span}\{\vec{v}_1, \dots, \vec{v}_n\}$.

$\Big\downarrow$ give an orthonormal set.

$$\left\{ \frac{\vec{w}_1}{\|\vec{w}_1\|}, \dots, \frac{\vec{w}_n}{\|\vec{w}_n\|} \right\}$$

$\vec{w}_1 = \vec{v}_1$

$\vec{w}_2 = \vec{v}_2 - \operatorname{proj}_{\operatorname{span}\{\vec{v}_1\}} \vec{v}_2$



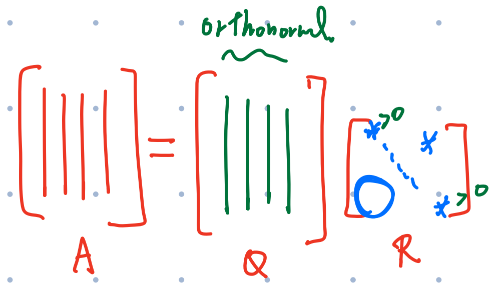
$\vec{w}_3 = \vec{v}_3 - \operatorname{proj}_{\operatorname{span}\{\vec{v}_1, \vec{w}_2\}} \vec{v}_3$

$\vdots$

Thm A: $m \times n$ with l.i. columns.

Then $\exists$ Q: $m \times n$, R: $n \times n$ s.t.

- $A = QR$
- Q has orthonormal columns. ($\Leftrightarrow Q^T Q = I_n$)
- R upper-triangular matrix, whose diagonal entries are positive.



Rmk: In HW you'll show that this decomposition is unique.

pf: Apply Gram-Schmidt on the column vectors of A:

Say $\{\vec{v}_1, \dots, \vec{v}_n\}$ are the columns of A.

(by assumption, this is a l.i. set)

$\vec{w}_1 = \vec{v}_1$

$$\vec{w}_2 = \vec{v}_2 - \underbrace{\operatorname{proj}_{\operatorname{span}\{\vec{v}_1\}} \vec{v}_2}_{= \, * \, \vec{v}_1}$$

$$\vdots$$

$$\vec{w}_3 = \vec{v}_3 - \underbrace{\operatorname{proj}_{\operatorname{span}\{v_1, w_2\}} \vec{v}_3}_{* \vec{v}_1 + * \vec{v}_2}$$

$$\underbrace{\begin{bmatrix} | & & | \\ \vec{w}_1 & \cdots & \vec{w}_n \\ | & & | \end{bmatrix}}_{\Phi} = \underbrace{\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix}}_{A} \begin{bmatrix} 1 & * & * & \\ 0 & 1 & * & \\ 0 & 0 & 1 & \\ \vdots & \vdots & 0 & \ddots \\ 0 & 0 & 0 & \end{bmatrix}$$

the columns form an orthogonal set,

inverse.

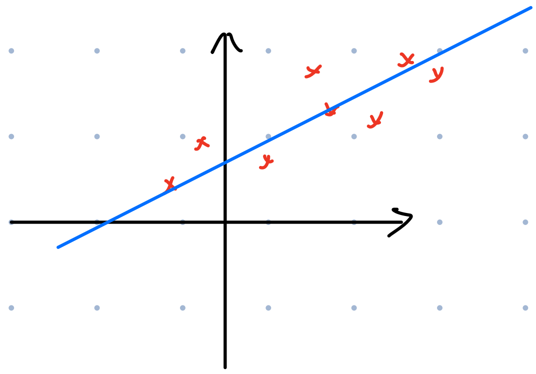
$$\Longrightarrow \underbrace{\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix}}_{A} = \begin{bmatrix} | & & | \\ \vec{w}_1 & \cdots & \vec{w}_n \\ | & & | \end{bmatrix} \begin{bmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} | & & | \\ \frac{\vec{w}_1}{\|\vec{w}_1\|} & \cdots & \frac{\vec{w}_n}{\|\vec{w}_n\|} \\ | & & | \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \|\vec{w}_1\| & & 0 \\ & \ddots & \\ 0 & & \|\vec{w}_n\| \end{bmatrix} \begin{bmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}_{R}$$

$$= \underbrace{\begin{bmatrix} | & & | \\ \frac{\vec{w}_1}{\|\vec{w}_1\|} & \cdots & \frac{\vec{w}_n}{\|\vec{w}_n\|} \\ | & & | \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \|\vec{w}_1\| & * & * \\ & \|\vec{w}_2\| & * \\ 0 & & \ddots \quad \|\vec{w}_n\| \end{bmatrix}}_{R}$$

Least square problem:

e.g.



Given data

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

We'd like to find

$y = ax + b$

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

that best approximate these points

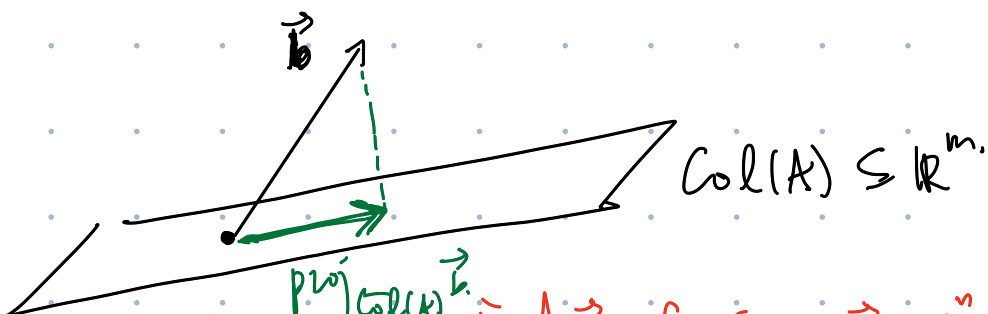
i.e. Find a, b s.t.

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \text{ is the closest to } \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

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

A least square solution of "$A\vec{x} = \vec{b}$" is a vector $\vec{x}_0 \in \mathbb{R}^n$ s.t.

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



$\text{proj}_{Col(A)} \vec{b} = A\vec{x}_0$ for some $\vec{x}_0 \in \mathbb{R}^n$,

and such $\vec{x}_0$ is a least square sol$^{\underline{n}}$ of "$A\vec{x} = \vec{b}$"

$\vec{x}_0$ is a least square sol$^{\underline{n}}$ of "$A\vec{x} = \vec{b}$"

$$\Longleftrightarrow \quad A\vec{x}_0 = \text{proj}_{Col(A)} \vec{b}$$

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

$\Updownarrow$

$A^T(\vec{b} - A\vec{x}_0) = \vec{0}$

$\Updownarrow$

$A^T A \vec{x}_0 = A^T \vec{b}$

"normal eq$^{\underline{n}}$" for "$A\vec{x} = \vec{b}$"

$\left(A = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}, \quad A^T = \begin{bmatrix} — v_1 — \\ \vdots \\ — v_n — \end{bmatrix} \right.$

$w \in \mathrm{Nul}(A^T) \Leftrightarrow A^T w = 0$

$\quad\quad\quad\quad \| $

$\begin{bmatrix} — v_1 — \\ \vdots \\ — v_n — \end{bmatrix} w$

$\Updownarrow$

$\langle v_1, w \rangle = \cdots = \langle v_n, w \rangle = 0$

$\Updownarrow$

$w \in \mathrm{Span}\{v_1, \cdots, v_n\}^\perp$

$\|$

$\left. \mathrm{Col}(A)^\perp \right)$

- the normal eq$^{\underline{n}}$ always has a sol$^{\underline{n}}$, and the solution(s) are the least square sol$^{\underline{n}}$'s to "$A\vec{x} = \vec{b}$".

<u>Thm</u>: A: $m \times n$ (real), the following are equivalent:

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

2) $A^T A$ is invertible. $(\Leftrightarrow \mathrm{Nul}(A^T A) = \{\vec{0}\})$

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

pf 1) $\Leftrightarrow$ 2) clear.

<u>2) $\Leftrightarrow$ 3)</u>:

<u>2) $\Rightarrow$ 3)</u>: $\mathrm{Nul}(A^T A) \supseteq \mathrm{Nul}(A)$ clear.

<u>3) $\Rightarrow$ 2)</u>: <u>Claim</u>: $\mathrm{Nul}(A^T A) = \mathrm{Nul}(A)$.

Suppose $\vec{x} \in \underline{\mathrm{Nul}(A^T A)}$, want to show: $\vec{x} \in \mathrm{Nul}(A)$.

$\|$

$$\Downarrow$$

$$A^T A \vec{x} = \vec{0}$$

$$(A\vec{x})^T A\vec{x}$$
$$\|$$
$$\vec{x}^T A^T A \vec{x} = \vec{0}$$
$$\|$$
$$\langle A\vec{x}, A\vec{x} \rangle$$

$$\Downarrow$$

$$A\vec{x} = \vec{0}$$

$$\boxed{A\vec{x} \in Col(A) \cap \underbrace{Nul(A^T)}_{\substack{\| \\ Col(A)^\perp}}} \Rightarrow A\vec{x} = \vec{0}$$

<u>Generalized eigenspace:</u> $A : n \times n$, λ is an eigenvalue of A.

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

<u>Thm:</u> Suppose A has distinct eigenvalues $\lambda_1, \ldots, \lambda_k$,

Then
$$\mathbb{C}^n = V_{\lambda_1}^{gen} \oplus \cdots \oplus V_{\lambda_k}^{gen}.$$

i.e. for any vector $\vec{v} \in \mathbb{C}^n$, $\exists!$ $\vec{v}_1, \ldots, \vec{v}_k$ s.t.

$$\vec{v} = \vec{v}_1 + \cdots + \vec{v}_k \quad \text{and} \quad \vec{v}_i \in V_{\lambda_i}^{gen}.$$

<u>Sketch of proof:</u> we'll be doing induction on the # of distinct eigenvalue of A.

Char. poly.

$$P_A(\lambda) = (\lambda_0 - \lambda)^n$$

$\Downarrow$ C-H thm.

$$(A - \lambda_0 I)^n = 0$$

<u>k = 1:</u> Suppose A has only one eigenvalue λ_0.

<u>Need to show:</u> $\mathbb{C}^n = V_{\lambda_0}^{gen}$ $\Longleftarrow$

i.e. $\forall \vec{v} \in \mathbb{C}^n$, $\exists k \geq 1$ s.t. $(A - \lambda_0 \mathbb{I})^k \vec{v} = \vec{0}$

Cayley-Hamilton thm: If p_A is the char. poly. of A,

$p_A(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0$

Then $p_A(A) = 0$ ← zero matrix.

$a_n A^n + a_{n-1} A^{n-1} + \cdots + a_0 \mathbb{I} = 0$

(In HW, you proved under an extra assumption that A is diagonalizable.)

Rmk: In HW, you proved that

C-H thm. ⇑ if A is nilpotent ($A^k = 0$ for some $k \geq 1$), then 0 is the only eigenvalue of A.

Sketch the proof of C-H thm.

Consider: $f : M_{n\times n}(\mathbb{C}) \longrightarrow M_{n\times n}(\mathbb{C})$

$A \longmapsto p_A(A)$

- In HW, you proved that $f(A) = 0$ when A is diagonalizable
- {diagonalizable matrices} $\subseteq M_{n\times n}(\mathbb{C})$ is "dense"
- f is a continuous map.

$\Rightarrow$ $f(A) = 0$ for any $A \in M_{n\times n}(\mathbb{C})$

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$$

↑

$$\begin{bmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & 1 & \\ & & \lambda_3 & 1 \\ & & & \lambda_4 \end{bmatrix} \qquad \lambda_i \to \lambda$$

$g : \mathbb{R} \to \mathbb{R}$ conti., $g(\mathbb{Q}) = 0$

$\cup$ dense

$\mathbb{Q}$

⇓

$g \equiv 0$