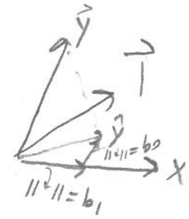
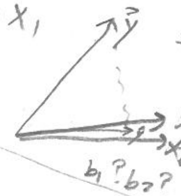
Lec10 Math 390.4 3/12/19

Back to Linear Algebra... demos after...

What is $\vec{b}$



Collinearity?

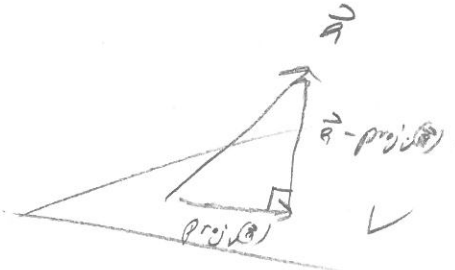


$\Rightarrow$ Instability in $\vec{b}$ but not $\hat{y}$.

Let $V = \text{span}\{v_1, v_2\} = \text{colsp}[V]$

If $V = [v_1 \mid v_2]$

$$\text{Does } \underset{\substack{\uparrow \\ \text{Orthogonal proj}}}{\text{proj}_V(\vec{a})} \overset{?}{=} \overbrace{\text{proj}_{\vec{v}_1}(\vec{a})}^{=H_1\vec{a}} + \overbrace{\text{proj}_{\vec{v}_2}(\vec{a})}^{=H_2\vec{a}} = c_1\vec{v}_1 + c_2\vec{v}_2$$



① $\text{proj}_V(\vec{a}) \in \text{colsp}(V)$

② $\text{proj}_V(\vec{a}) \perp (\vec{a} - \text{proj}_V(\vec{a}))$

since $c_1\vec{v}_1 + c_2\vec{v}_2 \in \text{span}\{\vec{v}_1, \vec{v}_2\}$

$$\Rightarrow \|\vec{a}\|^2 = \|\text{proj}_V(\vec{a})\|^2 + \|\vec{a} - \text{proj}_V(\vec{a})\|^2$$

$$= \|\text{proj}_{\vec{v}_1}(\vec{a}) + \text{proj}_{\vec{v}_2}(\vec{a})\|^2 + \|\vec{a} - \text{proj}_{\vec{v}_1}(\vec{a}) - \text{proj}_{\vec{v}_2}(\vec{a})\|^2$$

$$= \|c_1\vec{v}_1 + c_2\vec{v}_2\|^2 + \|\vec{a} - c_1\vec{v}_1 - c_2\vec{v}_2\|^2$$

$$\|\vec{a}\|^2 = \|c_1\vec{v}_1\|^2 + \|c_2\vec{v}_2\|^2 + 2\|c_1\vec{v}_1\|\|c_2\vec{v}_2\|\cos(c_1v_1, c_2v_2)$$

$$+ \|\vec{a}\|^2 + \|c_1\vec{v}_1\|^2 + \|c_2\vec{v}_2\|^2 - 2\|\vec{a}\|\|c_1v_1\|\cos(\vec{a}, c_1\vec{v}_1) - 2\|\vec{a}\|\|c_2v_2\|\cos(\vec{a}, c_2\vec{v}_2) + 2\|c_1\vec{v}_1\|\|c_2v_2\|\cos(c_1v_1, c_2v_2)$$

$$\Rightarrow \|a\|\left(\|c_1v_1\|\cos(\vec{a}, \vec{v}_1) + \|c_2v_2\|\cos(\vec{a}, \vec{v}_2)\right) = c_1^2\|\vec{v}_1\|^2 + c_2\|\vec{v}_2\|^2$$

$$0 = \text{proj}_V(\vec{a})^T(\vec{a} - \text{proj}_V(\vec{a})) = \text{proj}_V(\vec{a})^T\vec{a} - \|\text{proj}_V(\vec{a})\|^2$$

$\|H\vec{a}\|^2 = (H\vec{a})^T(H\vec{a}) = \vec{a}^T H^T H\vec{a} = \vec{a}^T H\vec{a}$

$\|\vec{a}\|^2 = \vec{v}^T\vec{v}$

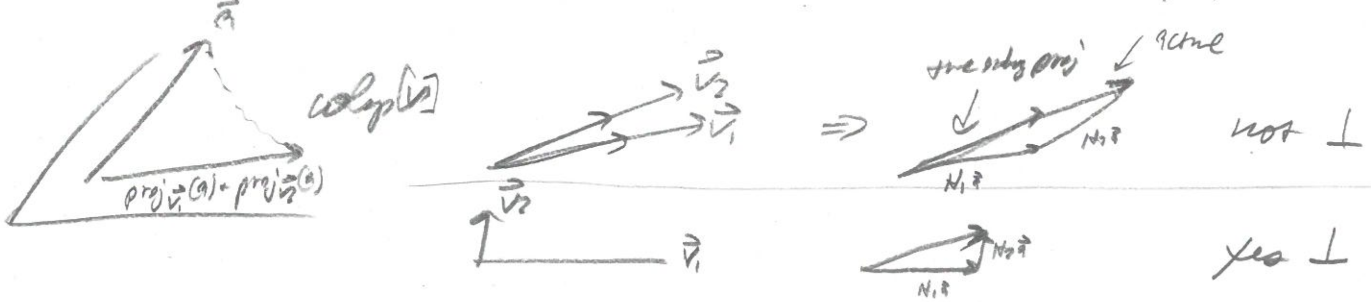
$$= (H_1\vec{a} + H_2\vec{a})^T\vec{a} - \|H_1\vec{a} + H_2\vec{a}\|^2$$

$$= (\vec{a}^T H_1 + \vec{a}^T H_2)\vec{a} - \|H_1\vec{a}\|^2 - \|H_2\vec{a}\|^2 + \|H_1\vec{a}\|\|H_2\vec{a}\|\cos(H_1\vec{a}, H_2\vec{a})$$

$$= \vec{a}^T H_1\vec{a} + \vec{a}^T H_2\vec{a} - \vec{a}^T H_1\vec{a} - \vec{a}^T H_2\vec{a} + \underbrace{\|H_1\vec{a}\|}_{>0}\underbrace{\|H_2\vec{a}\|}_{>0}\underbrace{\cos(\vec{v}_1, \vec{v}_2)}_{\substack{\| \\ 0 \\ \text{only if } \vec{v}_1 \perp \vec{v}_2}}$$

if $\vec{a}$ not orthogonal to V

What just happened? If $\vec{v}_1 \not\perp \vec{v}_2$ then the projection is too long or too short to be orthogonal but if $\perp$, it is precisely the orthogonal projection.



Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_d$ be orthogonal $\quad V = [\vec{v}_1 \mid \vec{v}_2 \mid \ldots \mid \vec{v}_d]$

$$\text{Proj}_V(\vec{a}) = \text{proj}_{\vec{v}_1}(\vec{a}) + \ldots + \text{proj}_{\vec{v}_d}(\vec{a}) = \frac{\vec{v}_1 \vec{v}_1^T}{\|\vec{v}_1\|^2}\vec{a} + \ldots \frac{\vec{v}_d \vec{v}_d^T}{\|\vec{v}_d\|^2}\vec{a}$$

$$= \left(\frac{\vec{v}_1 \vec{v}_1^T}{\|\vec{v}_1\|^2} + \ldots + \frac{\vec{v}_d \vec{v}_d^T}{\|\vec{v}_d\|^2}\right)\vec{a}$$

If $\|\vec{v}_1\|^2 = \ldots = \|\vec{v}_d\|^2 = 1$ i.e. scaled to unit vector, $Q = [\vec{v}_1 \mid \ldots \mid \vec{v}_d]$ is called an orthonormal matrix

$$\text{Proj}_V(\vec{a}) = \left(\vec{v}_1 \vec{v}_1^T + \ldots + \vec{v}_d \vec{v}_d^T\right)\vec{a}$$

$$= \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ v_1 & v_2 & \cdots & \vec{v}_d \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \leftarrow & v_1^T & \rightarrow \\ \leftarrow & v_2^T & \rightarrow \\ & \vdots & \\ \leftarrow & \vec{v}_d^T & \rightarrow \end{bmatrix} \vec{a}$$

$$[A \; B \; C] \begin{bmatrix} D \\ E \\ F \end{bmatrix} = AD + BE + CF$$

Block matrix multiplication

$$= \underbrace{Q\,Q^T}_{H}\vec{a}$$

Another proof

all 0's except 1 in jth place

$$QQ^T\vec{a} = Q\sum_{j=1}^{d} \vec{v}_j^T \vec{a}\, \vec{e}_j = \sum_{j=1}^{d} \vec{v}_j^T \vec{a}\, Q\vec{e}_j$$

$$= \sum_{j=1}^{d} \vec{v}_j^T \vec{a}\, \vec{v}_j = \sum_{j=1}^{d} \vec{v}_j \vec{v}_j^T \vec{a}$$

since true for all $\vec{a}$

$$\Rightarrow QQ^T = \sum_{j=1}^{d} \vec{v}_j \vec{v}_j^T$$

$$\Rightarrow QQ^T = V(V^TV)^{-1}V^T$$

Why? $\text{colsp}[Q] = \text{colsp}[V]$ represents same subspace.

$$\text{proj}_Q(\vec{a}) = Q(Q^TQ)^{-1}Q^T\vec{a} = QQ^T\vec{a} = \sum_{j=1}^{d} \vec{q}_j \vec{q}_j^{\,T}\vec{a}$$

$$Q^TQ = \begin{bmatrix} \leftarrow q_1 \rightarrow \\ \leftarrow q_2 \rightarrow \\ \vdots \\ \leftarrow q_d \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ q_1 & q_2 & \cdots & q_d \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{bmatrix} = I_d, \quad I_d^{-1} = I_d$$

$\vec{q}_i \cdot \vec{q}_j = 0$ if $i \neq j$ since all cols orthogonal

$\vec{q}_i \cdot \vec{q}_i = 1$ since $\|\vec{q}_i\|^2 = 1$ since all cols are normalized

---

$X \longrightarrow Q$ via Gram-Schmidt

$\text{colsp}[X] = \text{colsp}[Q]$

$$\underbrace{X}_{n \times (p+1)} = \underbrace{Q}_{n \times (p+1)} \underbrace{R}_{(p+1)\times(p+1)}$$ ← the "remainder" from the algorithm

"Q-R decomposition"

What does R look like? R is upper Δ, R square and R is full rank otherwise impossible X and Q could both be same rank

How does this help for OLS?

$$\vec{b} = (X^TX)^{-1}X^T\vec{y}$$

$$\Rightarrow X^TX\vec{b} = X^T\vec{y}$$

$$(QR)^T(QR)\vec{b} = (QR)^T\vec{y}$$

$$R^T\underbrace{Q^TQ}_{I}R\vec{b} = R^T\underbrace{Q^T\vec{y}}_{\vec{z}}$$

$$\Rightarrow \cancel{R^T}R\vec{b} = \cancel{R^T}\vec{z} \quad \text{since } R^T \text{ invertible}$$

$$\Rightarrow R\vec{b} = \vec{z}$$ easily solvable by back-substitution

$$\begin{bmatrix} c & d & e \\ 0 & f & g \\ 0 & 0 & h \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

$$b_2 = \frac{z_3}{h}$$

$$fb_1 + gb_2 = z_2 \Rightarrow b_1 = \frac{z_2 - g\frac{z_3}{h}}{f}$$

etc...