Math 390.4
lecture 9 2/25

$X \in \mathbb{R}^{n \times (p+1)}$

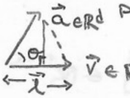
Full rank: $\text{rank}(X) = p+1$
no redundancy $\triangleq$

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

$\hat{\vec{y}} = Xb$

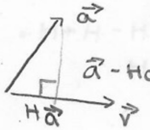
$$n\left\{\begin{bmatrix} \hat{\vec{y}} \end{bmatrix} = \underbrace{\begin{bmatrix} & X & \end{bmatrix}}_{p+1} \underset{\text{weights}}{\underset{\uparrow}{[b]}}\right\}p+1 \quad \vec{y} \in \mathbb{R}^n$$

$\hat{\vec{y}} \in \text{columnspace}(X)$
$\in b_0 \vec{1} + b_1 \vec{x}_1 + \ldots + b_p \vec{x}_p$



$\vec{l} :=$ projection $(\text{proj})_{\vec{v}}(\vec{a})$
orthogonal projection

$(\vec{a} - \vec{l}) \perp \vec{l}$



$$\frac{\vec{a} \cdot \vec{v}}{\|a\|\|v\|} = \cos(\theta) = \frac{\|\vec{l}\|}{\|\vec{a}\|}$$

$$= \|l\| \underbrace{\frac{\vec{v}}{\|\vec{v}\|}}_{\text{direction of unit length}} = \frac{\vec{a} \cdot \vec{v}}{\|v\|} \cdot \frac{\vec{v}}{\|v\|} = \underbrace{\left(\frac{1}{\|v\|^2} \overset{d \times 1}{\vec{v}} \overset{1 \times d}{\vec{v}^T}\right)}_{H} \vec{a} = H\vec{a}$$

$\text{rank}(H) = 1$

$$H = \frac{\vec{v}}{\|v\|^2}[v_1, v_2 \ldots v_d] = \left[\frac{v_1}{\|v\|^2}\vec{v} \quad \frac{v_2}{\|v\|^2}\vec{v} \quad \ldots \frac{v_d}{\|v\|^2}\vec{v}\right]$$

all values are linearally dependent

if $\vec{x}$ is an eigenvector, there exists a scalar λ
$H\vec{x} = \lambda\vec{x}$ the only eigenvalue is $\vec{v}$ $\quad H\vec{v} = \underset{\lambda}{\underset{\uparrow}{(1)}}\vec{v}$

$HV = \begin{bmatrix} \lambda & 0 \\ 0 & \ddots & \lambda_d \end{bmatrix} V = \begin{bmatrix} \lambda_0 & 0 \\ 0 & \ddots & 0 \end{bmatrix} [\vec{v}_1 \ldots \vec{v}_d]$

$\text{rank}(A) = \#$ of non zero eigenvalues

consider $\vec{a} \perp \vec{v}$. what is $\text{proj}_{\vec{v}}(\vec{a}) = \vec{0}_d$

$\text{proj}_{\vec{v}}(\text{proj}_{\vec{v}}(\vec{a})) = \text{proj}_{\vec{v}}(a)$

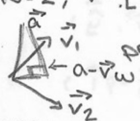
$H(H\vec{a}) = H\vec{a} = HH = H$ independent

$$HH = \left(\frac{1}{\|v\|^2}\vec{v}\vec{v}^T\right)\left(\frac{1}{\|v\|^2}\vec{v}\vec{v}^T\right) = \frac{1}{\|v\|^4}\underbrace{\vec{v}\vec{v}^T\vec{v}\vec{v}^T}_{\|v\|^2} = \frac{1}{\|\vec{v}\|^2}\vec{v}\vec{v}^T = H$$

$\vec{0}_d = \text{proj}_{\vec{v}}(\vec{a} - Ha) = H(\vec{a} - Ha) = Ha - HH\vec{a} = H\vec{a} - H\vec{a} = \vec{0}_d$

$V = [\vec{v}_1 \mid \vec{v}_2 \mid \cdots \mid \vec{v}_K] \in \mathbb{R}^{d \times K}$; full rank

ex) $K = 2$; $V = [v_1 \mid v_2]$



$\text{proj}_V(\vec{a}) = P(V, \vec{a})$

Two Facts

1) $\underset{\text{colsp}(V)}{\text{proj}_V(\vec{a}) = \underset{\uparrow}{V}\vec{w}}$ where $\vec{w} \in \mathbb{R}^K$

2) $\vec{a} - V\vec{w} \perp \vec{v}_1 \qquad \vec{v}_1^T(\vec{a} - V\vec{w}) = 0$
$\vec{a} - V\vec{w} \perp \vec{v}_2 \quad = \quad \vec{v}_2^T(\vec{a} - V\vec{w}) = 0 \Rightarrow \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_K^T \end{bmatrix}(\vec{a} - V\vec{w}) = \vec{0}_K \Rightarrow V^T(\vec{a} - V\vec{w}) = \vec{0}_K$
$\vdots$
$\vec{a} - V\vec{w} \perp \vec{v}_K \qquad \vec{v}_K^T(\vec{a} - V\vec{w}) = 0$

$V^T\vec{a} - V^T V\vec{w} = \vec{0}_K$
$V^T\vec{a} = V^T V\vec{w}$
$(V^T V\vec{w})^{-1} V^T = \vec{a}$

since $V^T V$ is invertible since V has assumed full rank K

$\text{proj}_V(\vec{a}) = \underbrace{V(V^T V)^{-1} V^T}_{H}\vec{a}$

$X = [\vec{1} \mid \vec{x}_1 \mid \cdots \mid \vec{x}_p]$, full rank $\quad (X^T X)^{-1} X^T \vec{y}$

$\text{proj}_X(\vec{y}) = \underbrace{X(X^T X)^{-1} X^T}_{H} \vec{y} = X\vec{b} \Rightarrow \hat{\vec{y}}$ is an orthogonal projection of $\vec{y}$ onto colsp(X)

$$HH = (X(X^TX)^{-1}\underbrace{X^T)(X(X^TX)^{-1}}_{I}X^T)$$
$$= X(X^TX)^{-1}X^T$$
$$= H$$

$$\begin{bmatrix} \bar{y} - r\frac{s_y}{s_x}\bar{x} \\ r\frac{s_y}{s_x} \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \vec{b} = (X^TX)^{-1}X^T\vec{y}$$

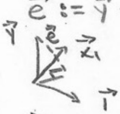
Thm:
If matrix A is symmetric & independent $\Rightarrow$
A is an orthogonal projection matrix

assuming A is symmetric ...

$$(A^{-1})^T = (A^T)^{-1}$$

$$H^T = (X(X^TX)^{-1})X^T)^T = X^{T^T}\underbrace{((X^TX)^{-1})^T}_{} = X(\underbrace{(X^TX)^T}_{})^{-1}X^T = X(X^TX)^{-1}X^T = H$$

$$\vec{e} := \vec{y} - \hat{\vec{y}} = \vec{y} - H\vec{y} \Rightarrow (I_n - H)\vec{y}$$



$\vec{e} := \text{projection}_{X_\perp}(\vec{y}) = (I-H)\vec{y}$ orthogonal complement projection matrix

$$n\updownarrow\Big[\underbrace{X}_{p+1} \,\Big|\, \underbrace{X_\perp}_{n-p+1}\Big] \in \mathbb{R}^{n\times n}$$

$$\hat{\vec{y}} \in \text{colsp}(\vec{1} \mid \vec{x}) = b_0\vec{1} + b_1\vec{x}$$

$$(I-H)^T = I^T - H^T = I - H$$
$$(I-H)(I-H) = II - HI - IH - -HH = I - H - H + H =$$
$$= I - H$$

prove $\hat{\vec{y}} \perp \vec{e}$

show $\hat{\vec{y}}^T \cdot \vec{e} = 0$

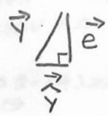
$$(H\vec{y})^T((I-H)\vec{y}) = 0$$
$$\vec{y}^T H^T (I-H)\vec{y} = 0$$
$$\vec{y}^T H (I-H)\vec{y} = 0$$
$$\vec{y}^T(H - HH)\vec{y} = 0$$
$$\vec{y}^T(H - H)\vec{y} = 0$$
$$\vec{y}^T(0_{n\times n})\vec{y} = 0$$
$$0 = 0$$

$$[X^T]\begin{bmatrix} X \end{bmatrix}\Big\}n = \begin{bmatrix} X^TX \end{bmatrix} p+1$$
$$\quad p+1 \qquad\qquad p+1$$

$\text{rank}(H) = p+1 \quad \text{rank}(I-H) = n-(p+1)$

$\hat{\vec{y}} = H\vec{y} \quad \hat{\vec{y}} \in \text{colsp}(X)$

$\dim(\text{colsp}(X)) = p+1$



$$\vec{y}^T\vec{y} = \hat{\vec{y}}^T\hat{\vec{y}} + \vec{e}^T\vec{e}$$
$$\|\vec{y}\|^2 = \|\hat{\vec{y}}\|^2 + \underbrace{\|\vec{e}\|^2}_{SSE}$$

$$X = \begin{bmatrix} | & \uparrow & & \uparrow \\ 1 & \vec{x}_1 & \cdots & \vec{x}_p \\ | & \downarrow & & \downarrow \end{bmatrix}$$

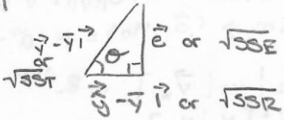
is this diagram accurate? (yes)

$$\text{proj}_X(\vec{y} - \bar{y}\vec{1}) = H(\vec{y} - \bar{y}\vec{1}_n) = \underbrace{H\vec{y}}_{\hat{\vec{y}}} - \bar{y}\overbrace{H\vec{1}_n}^{\text{proj}_X(\vec{1})}$$
$$\hat{\vec{y}} - \bar{y}\vec{1}$$

$\vec{1} \in \text{colsp}(X)$

Pythag. holds

$$\underbrace{\|\vec{y} - \bar{y}\vec{1}\|^2}_{SST} = \underbrace{\|\hat{\vec{y}} - \bar{y}\vec{1}\|^2}_{SSR} + \underbrace{\|\vec{e}\|^2}_{SSE}$$
$$\Sigma(y - \bar{y})^2 \qquad \Sigma(\hat{y} - \bar{y})^2 \qquad \Sigma e_i^2$$



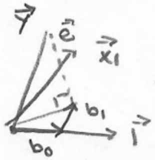
$\theta(\vec{y}, \hat{\vec{y}})$

small θ; larger R^2
$\vec{y}$ more similar to $\hat{\vec{y}}$

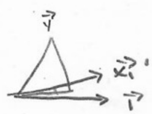
$$R^2 = 1 - \frac{SSE}{SST} = \frac{SST - SSE}{SST} = \left(\frac{\sqrt{SSR}}{\sqrt{SST}}\right)^2 = \cos^2\theta(\vec{y} - \bar{y}\vec{1}, \hat{\vec{y}} - \bar{y}\vec{1})$$
$$\in (0,1)$$

where is $\vec{b}$?

$$\hat{\vec{y}} = X\vec{b} = H\vec{y}$$
$$\hat{\vec{y}} \in \text{colsp}(X)$$
$$\Rightarrow \hat{\vec{y}} = b_0\vec{1} + b_1\vec{x}_1$$



change of basis



$x_1 = \vec{1} + \vec{\eta}$
(noise)

$$\text{colsp}([\vec{1} \mid \vec{x}_1]) = \text{colsp}([\vec{1} \mid \vec{x}_1{}'])$$

$\hat{y}$ is the same in both subspaces

multicollinearity

columns of X are very similar in direction

$\hat{\vec{y}}$ doesn't change

(R^2 doesn't change b/c $\hat{\vec{y}}$ is same)

- but $\vec{b}$ is unstable