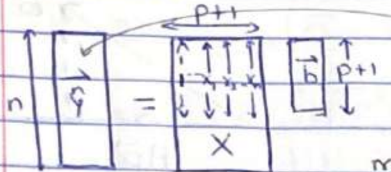
Lecture - 09. 02/25/2020.

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

Full Rank i.e rank[X] = p+1

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

$\vec{\hat{y}} = X\vec{b}$

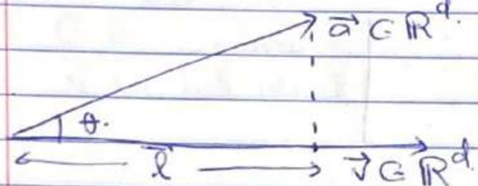


$b_0 + b_1 x_{11} + \dots + b_p x_{pp}$

$\vec{\hat{y}} \in \mathbb{R}^n$

more specifically $\vec{\hat{y}} \in$ colsp[X]

$\vec{\hat{y}} = b_0 \vec{1} + b_1 \vec{x}_1 + \dots + b_p \vec{x}_p$



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

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

Othonogal projection.

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

unit direction.

$= \|\vec{l}\| \underbrace{\frac{\vec{v}}{\|\vec{v}\|}}_{\text{direction of } \vec{v} \text{ with length of } 1}$

$= \frac{\vec{a} \cdot \vec{v}}{\|v\|} \cdot \frac{\vec{v}}{\|\vec{v}\|}$

$= \underbrace{\left(\frac{1}{\|v\|^2} \underset{d\times 1}{\vec{v}} \, \underset{1\times d}{\vec{v}^T} \right)}_{\substack{d \times d \\ H}} \vec{a} = H\vec{a}$

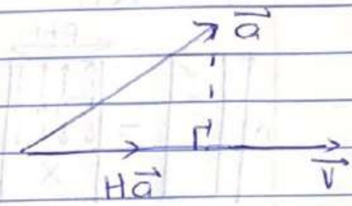
rank $[H] = 1$

$$H = \frac{\vec{v}}{\|\vec{v}\|^2} [v_1, v_2, \ldots, v_d] = \begin{bmatrix} \frac{v_1}{\|\vec{v}\|^2}\vec{v} & \frac{v_2}{\|\vec{v}\|^2}\vec{v} & \cdots & \frac{v_d}{\|\vec{v}\|^2}\vec{v} \end{bmatrix}$$

If $\vec{x}$ is an eiganvector, there exists a scalar λ . . .

$$H\vec{x} = \lambda \vec{x}$$

the only eiganvector is $\vec{v}$



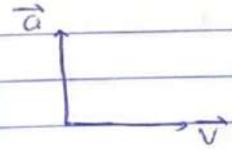
$$H\vec{v} = \underbrace{(1)}_{\lambda}\vec{v}$$

$$HV = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{bmatrix} V$$

$$= \begin{bmatrix} 1 & & 0 \\ & 0 & \\ 0 & & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \cdots \\ \vdots & \end{bmatrix}$$

rank $[A]$ = # of non zero.

Consider $\vec{a} \perp \vec{v}$. What is the projection$_{\vec{v}}(\vec{a}) = \vec{0}$



$$\text{proj}_{v}(\vec{a}) = \frac{1}{\|\vec{v}\|^2}\vec{v}\underbrace{\vec{v}^T\vec{a}}_{0}$$

$$= \vec{0}_d$$

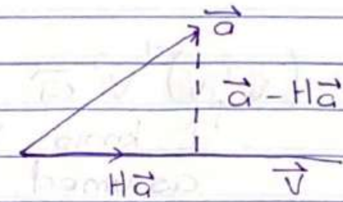
$$\text{proj}_{\vec{v}}(\text{proj}_{\vec{v}}(\vec{a})) = \text{proj}_{\vec{v}}(\vec{a})$$

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

$$HH = \left(\frac{1}{\|\vec{V}\|^2}\vec{V}\vec{V}^T\right)\left(\frac{1}{\|\vec{V}\|^2}\vec{V}\vec{V}^T\right)$$

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

$$\vec{O}_d = \text{proj}_{\vec{V}}(\vec{a} - H\vec{a})$$
$$= H(\vec{a} - H\vec{a})$$
$$= H\vec{a} - HH\vec{a}$$
$$= H\vec{a} - H\vec{a} = \vec{0}d$$



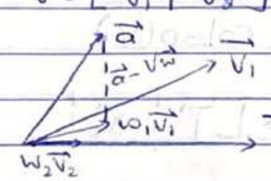
$V = [\,\vec{V}_1 \mid \vec{V}_2 \mid \dots \mid \vec{V}_k\,]$

$\in \mathbb{R}^{d \times k}$ where $k < d$

V in full rank

$k = 2$

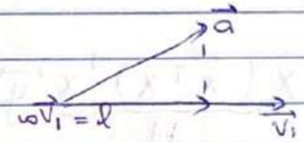
$V = [\,\vec{V}_1 \mid \vec{V}_2\,]$



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

$V \; \| \; \text{colsp}(V)$

Two Facts



I $\text{Proj}_V(\vec{a}) = V\vec{w}$ when $\vec{w} \in \mathbb{R}^k$

$V \approx \text{colsp}$ $\quad //$

II $\vec{a} - V\vec{w} \perp \vec{V}_1 \qquad \vec{V}_1^T(a - V\vec{w}) = 0$

$\vec{a} - V\vec{w} \perp \vec{V}_2 \qquad \vec{V}_2^T(a - V\vec{w}) = 0$

$\vdots \qquad \Longrightarrow \qquad \vdots \qquad \Longrightarrow$

$\vec{a} - V\vec{w} \perp \vec{V}_k \qquad \vec{V}_k^T(a - V\vec{w}) = 0$

$$\begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_k^T \end{bmatrix} (\vec{a} - V\vec{w}) = \vec{0}_k \implies V^T(\vec{a} - V\vec{w}) = \vec{0}_k \implies$$

$$V^T\vec{a} - V^TV\vec{w} = \vec{0}_v \implies V^T\vec{a} = V^TV\vec{w} \implies$$

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

I know V^TV is invertible since V was assumed full rank.

$$\underset{\substack{V \\ \| \\ colsp(V)}}{proj}(\vec{a}) = \underbrace{V(V^TV)^{-1}V^T}_{H}\vec{a}$$

$X = [\vec{1} \mid \vec{x}_1 \mid \cdots \mid \vec{x}_p]$, full rank

$$\vec{\hat{Y}} = proj_X(\vec{Y}) = \underbrace{X(X^TX)^{-1}X^T}_{H}\vec{y} = X\underset{\substack{\| \\ (X^TX)^{-1}X^T\vec{y}}}{\vec{b}}$$

$\implies \vec{\hat{y}}$ is the orthogonal projection of $\vec{y}$

$$HH = X(X^TX)^{-1}(X^T$$
$$= X(X^TX)^{-1}X^T = H$$

Thm

If matrix A is symmetric & $\cdot \implies A$ is an orthogonal proj matrix.

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

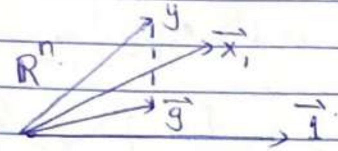
$$= X((X^TX)^T)^{-1}X^T = X(X^TX)^{-1}X^T = H$$

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

If A is symmetric.

$\text{proj}_{X\perp}(\vec{y}) = (I-H)\vec{y}$ orthogonal

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



$\hat{y} \in \text{colsp}([\vec{1} \,;\, \vec{x}])$
$= b_0\vec{1} + b_1\vec{x}$

$\text{colsp}[X] \subset \mathbb{R}^n$

$n \; [\, \underbrace{X}_{p+1} \mid \underbrace{X^\perp}_{n-(p+1)} \,]$ orthogonal confident of $X \in \mathbb{R}^{n\times n}$ and full rank.

$(I-H)^T = I^T - H^T = I - H$

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

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

wts $\implies \hat{y}^T\vec{e} = 0 \implies (H\vec{y})^T((I-H)\vec{y}) = 0$

$\implies \vec{y}^T H^T (I-H)\vec{y} = 0 \implies \vec{y}^T H(I-H)\vec{y} = 0$

$\implies \vec{y}^T(HI - HH)\vec{y} = 0 \implies \vec{y}^T(H-H)\vec{y} = 0$

$\implies \vec{y}^T(0_{n\times n})\vec{y} = 0 \implies \vec{y}^T\vec{0} = 0$

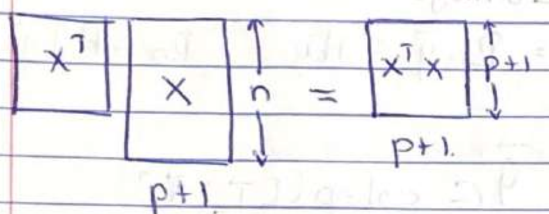
$h^*(x) = X\vec{\beta} + \vec{\varepsilon}$
iid normals

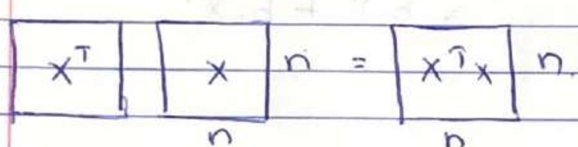
$\text{rank}[H] = p+1$

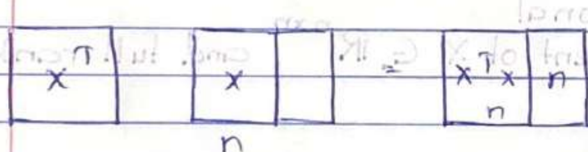
$\vec{\hat{y}} = H\vec{y}$

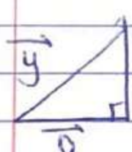
$\vec{\hat{y}} \in \text{colsp}[X] \quad \dim[\text{colsp}[X]] = p+1$

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





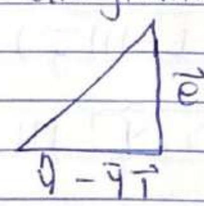




$\Rightarrow \vec{y} \cdot \vec{y} = \vec{\hat{y}}^T \vec{\hat{y}} + \vec{e}^T \vec{e}$

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

Is this diagram accurate? yes.



$$\begin{bmatrix} \uparrow \\ \vec{y} \\ \downarrow \end{bmatrix} - \bar{y}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \\ \vec{y} \\ \\ \end{bmatrix} - \begin{bmatrix} \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \end{bmatrix}$$

$\vec{e} = (\vec{y} - \bar{y}\vec{1}) - (\vec{\hat{y}} - \bar{y}\vec{1}) = \vec{y} - \vec{\hat{y}}$

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

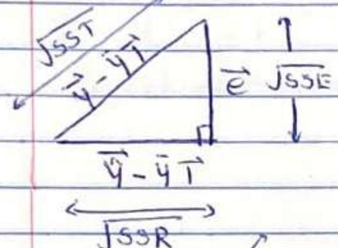
$$= \vec{\hat{y}} - \bar{y}\vec{1}$$

$$X = \begin{bmatrix} \uparrow & \uparrow & \uparrow & & \uparrow \\ \vec{1} & \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_p \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{bmatrix} \implies \vec{1} \in \text{colsp}[X]$$

sum of square identity

$$\implies \underbrace{\|\vec{y} - \bar{y}\vec{1}\|^2}_{SST} = \underbrace{\|\vec{\hat{y}} - \bar{y}\vec{1}\|^2}_{SSR} + \underbrace{\|\vec{e}\|^2}_{SSE}$$

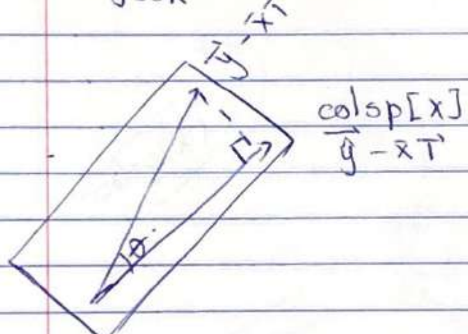
$$SST = \Sigma(y_i - \bar{y})^2 \qquad SSR = \Sigma(\hat{y}_i - \bar{y})^2 \qquad SSE = \Sigma e_i^2$$



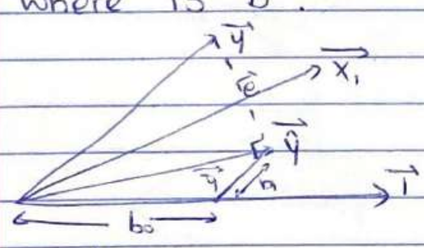
$$R^2 = 1 - \frac{SSE}{SST} = \frac{SST - SSE}{SST} = \frac{SSR}{SST} = \left(\frac{\sqrt{SSR}}{\sqrt{SST}}\right)^2$$

$$= \cos^2(\theta(\vec{y} - \bar{x}\vec{1}, \hat{y} - \bar{y}1)) = [0, 1]$$

$$\theta(\vec{y} - \bar{x}\vec{1}, \hat{y} - \bar{y}1) = \theta[\vec{y}, \vec{\hat{y}}]$$



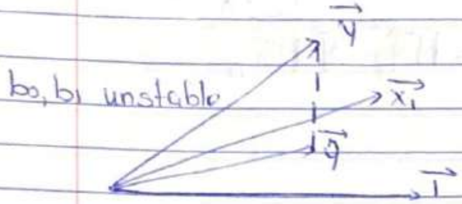
Where is $\vec{b}$?



$$\vec{\hat{y}} = X\vec{b} = H\vec{y}$$

$$\vec{\hat{y}} \in \text{colsp}[X]$$

$$\implies \vec{\hat{y}} = b_0\vec{1} + b_1\vec{x}_1$$



$\vec{x}_1 = \vec{1} + \vec{n}$ 'noise'

$\text{colsp}\left[\,[\vec{1} \mid \vec{x}_1]\,\right] = \text{colsp}\left[\,[\vec{1} \mid \vec{x}_1']\,\right]$

$\Longrightarrow$ $\hat{\vec{y}}$ is the same in both situations.

Multicolinearity :-

When the column of X are very similar in direction $\hat{\vec{y}}$ being change but $\vec{b}$ is unstable.