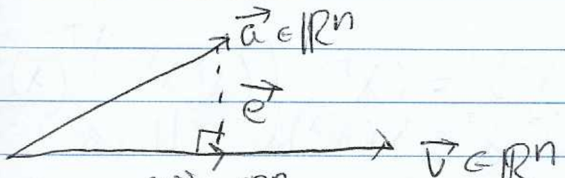
3/14/18

Lecture 12



$\text{proj}_{\vec{v}}(\vec{a}) \in \mathbb{R}^n$

$$\parallel$$

$$\underbrace{\frac{\vec{v}\,\vec{v}^T}{\|\vec{v}\|^2}}\,\vec{a} = H\vec{a}$$

$H \in \mathbb{R}^{n \times n}$

Notes: ① $\vec{a} = \text{proj}_{\vec{v}}(\vec{a}) + \vec{e}$

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

② $\text{proj}_{\vec{v}}(\vec{a}) \in \text{colsp}(\vec{v})$

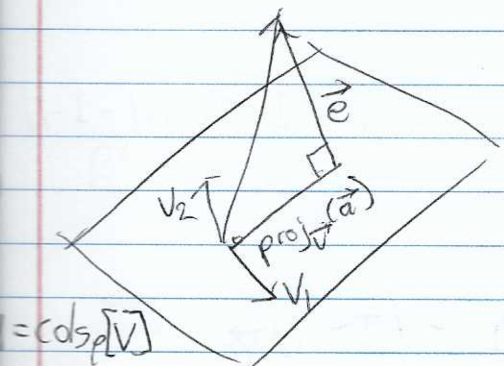
$\Rightarrow \text{proj}_{\vec{v}}(\vec{a}) = w\vec{v} = l\vec{v}_0 \rightarrow$ norm vector

↳ length

③ $\vec{e} \cdot \text{proj}_{\vec{v}}(\vec{a}) = 0$

$\Rightarrow \vec{e} \cdot \vec{v} = 0 \quad \Rightarrow \vec{v}^T\vec{e} = 0$

$$V = \big[\vec{V}_1 \,\vdots\, \vec{V}_2 \,\vdots\, \cdots \,\vdots\, \vec{V}_k\big]_{\vec{w}} \in \mathbb{R}^{n \times k}$$



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

$\Rightarrow \exists\, \vec{w}$ s.t. $\text{proj}_V(\vec{a}) = w_1\vec{V} + w_2\vec{V}_2 + \dots + w_k\vec{V}_k = V\vec{w}$

$\vec{e} = \vec{a} - V\vec{w} \in \mathbb{R}^n$

$\vec{e} \cdot \vec{V}_1 = 0\ ,\ \vec{e} \cdot \vec{V}_2 = 0, \dots, \vec{e} \cdot \vec{V}_k = 0$

$(\vec{a} - V\vec{w}) \cdot \vec{V}_1 = 0 \cdots \cdots$

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

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

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

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

$\Rightarrow V^T\vec{a} = \underbrace{V^TV}_{k \times k}\vec{w}$

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

Our goal is to solve for the projection of $\vec{a}$ onto $\vec{v}$.

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

Some Properties of proj. matrices

① Symmetric $H^T = H \Rightarrow \left(X(X^TX)^{-1}X^T\right)^T = (X^T)^T\left((X^TX)^{-1}\right)^T(X)^T$

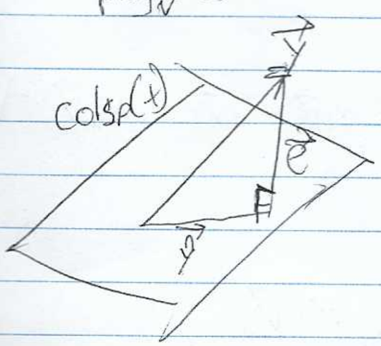
$$= X\left((X^TX)^T\right)^{-1}X^T = X(X^TX)^{-1}X^T = H \checkmark$$

$$\left((A^{-1})^T = (A^T)^{-1} : A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I \Rightarrow A^TB = I \Rightarrow B = (A^T)^{-1} \right)$$

② Idempotency

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

$$\underbrace{H\,H\,\vec{a}}_{proj_{V}(\vec{a})} = H_{proj_V(\vec{a})} = proj_V(proj_V(\vec{a}))$$



$$(I-H)(I-H) = II - 2HI + HH = I - 2H + H = I - H$$

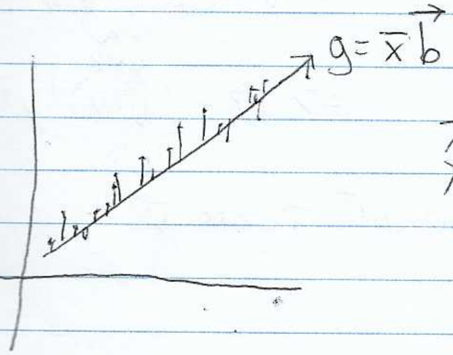
$$\vec{e} \cdot \vec{\hat{y}} = 0$$

$$\vec{e} = \vec{y} - \vec{\hat{y}} = \vec{y} - Hy = (I-H)\vec{y}$$

$I-H$ is the projection matrix onto $colsp(X)^\perp$ orthogonal space

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

(21 dots on graph)



$$\vec{\hat{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_{21} \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{21} \end{bmatrix} \qquad \vec{e} =$$

By pythagorean theorem $\|y\|^2 = \|\vec{\hat{y}}\|^2 + \|\vec{e}\|^2$

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2$$

Consider

$$\sum(y_i-\bar{y})^2 = \sum(y_i^2 - 2y_i\bar{y} + \bar{y}^2) = \sum y_i^2 - 2\bar{y}\sum y_i + \bar{y}^2\sum(1)$$
$$= \sum y_i^2 - 2n\bar{y}^2 + n\bar{y}^2 = \sum y_i^2 - n\bar{y}^2$$

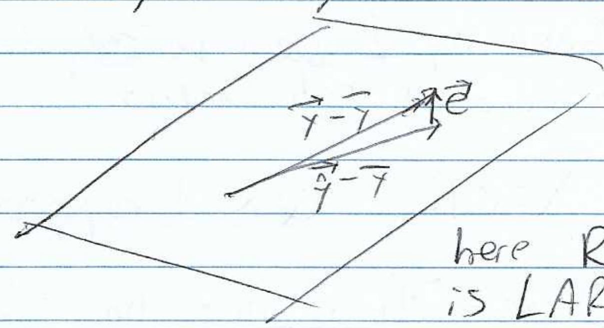
$$\sum(\hat{y}_i - \bar{y})^2 = \sum \hat{y}_i^2 - 2\bar{y}\sum\hat{y}_i + n\bar{y}^2 = \sum\hat{y}_i^2 - n\bar{y}^2$$

$$\sum y_i^2 - n\bar{y}^2 = \sum\hat{y}_i^2 - n\bar{y}^2 + \sum e_i^2$$
$$\Rightarrow \underbrace{\sum(y_i-\bar{y})^2}_{SST_{\text{total}}} = \underbrace{\sum(\hat{y}_i-\bar{y})^2}_{SSR} + \underbrace{\sum e_i^2}_{SSE}$$

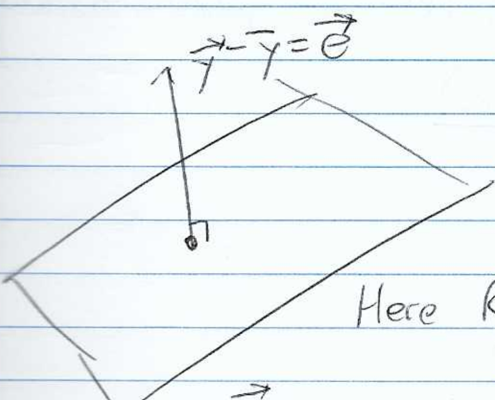
SST total = SSR regression or SSM model, SSE error

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\frac{1}{n-1}SSE}{\frac{1}{n-1}SST} = 1 - \frac{S^2_e}{S^2_y} = \frac{S_y^2 + S_e^2}{S^2_y}$$

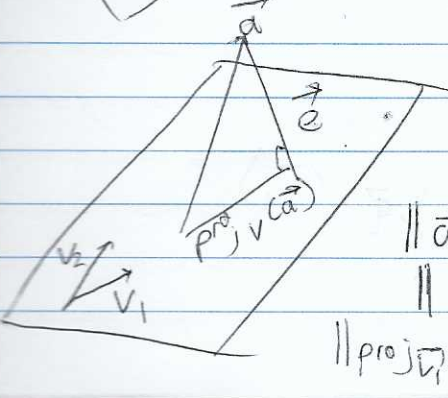
$$= \frac{SST}{SST} - \frac{SSE}{SST} = \frac{SSR}{SST}$$



here R^2 is LARGE



Here $R^2=0$



$$\|\vec{a}\|^2 = \|e\|^2 + \sum_{i=1}^{k}\|proj_{\vec{v}_i}(\vec{a})\|^2$$ if V is orthogonal

$$proj_V(\vec{a}) = proj_{\vec{v}_1}(\vec{a}) + proj_{\vec{v}_2}(a)$$

$$\|\vec{a}\|^2 \overset{?}{=} \underbrace{\|proj_{\vec{v}_1}(\vec{a})\|^2 + \|proj_{\vec{v}_2}(\vec{a})\|^2}_{proj_V(\vec{a})} + \|\vec{e}\|^2 \rightarrow \text{Only if } proj_{\vec{v}_1}(\vec{a}) \perp proj_{\vec{v}_2}(\vec{a})$$

$\|$

$\|proj_{\vec{v}_1}(\vec{a}) + proj_{\vec{v}_2}(\vec{a}) + \vec{e}\|^2$

& $proj_{\vec{v}_1}(a) \perp \vec{e}$ & $proj_{\vec{v}_2}(a) \perp \vec{e}$

$$\text{if } V \text{ orthog} \quad \underbrace{\text{proj}_V(\vec{a})}_{V(V^TV)^{-1}V^T} = \text{proj}_{\vec{V}_1}(\vec{a}) + \text{proj}_{\vec{V}_2}(\vec{a}) = \frac{\vec{V}_1\vec{V}_1^T}{\|\vec{V}_1\|^2} + \frac{\vec{V}_2\vec{V}_2^T}{\|\vec{V}_2\|^2}$$

$$= \left(\frac{\vec{V}_1\vec{V}_1^T}{\|\vec{V}_1\|^2} + \frac{\vec{V}_2\vec{V}_2^T}{\|\vec{V}_2\|^2}\right)\vec{a}$$

If V is orthonorma (orthogonal and each column is normalized to length 1)

$$= \left(\vec{V}_1\vec{V}_1^T + \vec{V}_2\vec{V}_2^T\right)\vec{a}$$

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

$Q \in \mathbb{R}^{n \times (p+1)}$ orthonormal $\Rightarrow$ full rank

$$\hat{y} = \overbrace{X(X^TX)^{-1}X^T}^{H}\vec{y} = \hat{y} = \overbrace{QQ^T}^{H}\vec{y}$$

Given X, I want to compute a Q with same colsp and orthonormal col. vectors.

$X = QR$ QR decomposition. (This is Gram-Schmidt)

Gram-Schmidt Algorithm (QR decomp)

Step 1: let $\vec{V}_{\cdot 1} = \vec{X}_{\cdot 1}$

Step 2: let $\vec{q}_{\cdot 1} = \dfrac{\vec{V}_{\cdot 1}}{\|\vec{V}_{\cdot 1}\|}$

Step 3: let $\vec{V}_{\cdot 2} = \vec{X}_{\cdot 2} - \text{proj}_{\vec{q}_{\cdot 1}}(\vec{X}_{\cdot 2})$

Step 4: let $\vec{q}_{\cdot 2} = \dfrac{\vec{V}_{\cdot 2}}{\|\vec{V}_{\cdot 2}\|}$

Step 5: let $\vec{V}_{\cdot 3} = \vec{X}_{\cdot 3} - \text{proj}_{\vec{q}_{\cdot 1}}(\vec{X}_{\cdot 3}) - \text{proj}_{\vec{q}_{\cdot 2}}(\vec{X}_{3})$

Step 6: let $\vec{q}_{\cdot 3} = \dfrac{\vec{V}_{\cdot 3}}{\|\vec{V}_{\cdot 3}\|}$

$\vdots$

$$\vec{x}_{\cdot 1} = \|\vec{v}_{\cdot 1}\| \vec{q}_{\cdot 1} = (\vec{q}_{\cdot 1} \cdot \vec{x}_{\cdot 1}) \vec{q}_{\cdot 1}$$

$$\vec{x}_{\cdot 2} = \|\vec{v}_{\cdot 1}\| \vec{q}_{\cdot 1} + \|\vec{v}_{\cdot 2}\| \vec{q}_{\cdot 2} = (\vec{x}_{\cdot 1} \cdot \vec{q}_{\cdot 1}) \vec{q}_{\cdot 1} + (\vec{x}_{\cdot 2} \cdot \vec{p}_{\cdot 2}) \vec{q}_{\cdot 2}$$

$$R = \begin{bmatrix} \vec{q}_{\cdot 1} \cdot \vec{x}_{\cdot 1} & \vec{q}_{\cdot 1} \cdot \vec{x}_{\cdot 2} & \cdots & \vec{q}_{\cdot 1} \vec{x}_{\cdot p+1} \\ 0 & \vec{q}_{\cdot 2} \cdot \vec{x}_{\cdot 2} & \cdots & \vec{q}_{\cdot 2} \cdot \vec{x}_{\cdot p+1} \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \vec{q}_{p+1} \cdot \vec{x}_{\cdot p+1} \end{bmatrix}$$

$$\boxed{X = QR}$$