3/19/18

Lecture 13

Projections onto colsp(V), s.t. $V = [\vec{V}_1 \; \cdots \; \vec{V}_k]$

$$\text{proj}_V(\vec{a}) = \sum_{j=1}^{k} \frac{\vec{V}_j \vec{V}_j^T}{\|\vec{V}_j\|} = \sum_{j=1}^{k} \vec{V}_j \vec{V}_j^T \vec{a}$$

→ if cols of V are normalized (i.e., length is 1)

If V is orthogonal, then:

$$\|\text{proj}_V(\vec{a})\|^2 = \left\| \sum_{j=1}^{k} \frac{\vec{V}_j \vec{V}_j^T \vec{a}}{\|\vec{V}_j\|^2} \right\|^2 = \sum_{j=1}^{k} \left\| \frac{\vec{V}_j \vec{V}_j^T \vec{a}}{\|\vec{V}_j\|^2} \right\|^2 = \sum_{j=1}^{k} \|\vec{V}_j \vec{V}_j^T \vec{a}\|^2$$

$X = QR \leftarrow \mathbb{R}^{(p\times 1)\times(p\times 1)}$

X: $\mathbb{R}^{n\times(p\times 1)}$ full rank; Q: $\mathbb{R}^{n\times(p\times 1)}$ full rank

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

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

$$\hat{\vec{y}} = \underbrace{QQ^T}_{H} \vec{y}$$

$$QQ^T = X(X^TX)^{-1}X^T = (QR)\left((QR)^T(QR)\right)^{-1}(QR)^T = QR(R^T\underbrace{Q^TQ}_{I_{p+1}}R)^{-1}R^TQ^T$$

$$\begin{bmatrix} \leftarrow q_1^T \rightarrow \\ \leftarrow q_2^T \rightarrow \\ \vdots \\ \leftarrow q_{p+1}^T \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ q_1 & q_2 & \cdots & q_{p+1} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} = I_{p+1}$$

$$QQ^T = \underbrace{QRR^{-1}}_{I_{p+1}}\underbrace{(R^T)^{-1}R^T}_{I_{p+1}}Q^T = QQ^T$$

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

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

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

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

$$(R^T)^{-1}R^TR\vec{b} = (R^T)^{-1}R^T\vec{z}$$

$\Rightarrow R\vec{b} = \vec{z}$ solve for $\vec{b}$

→

$$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \Rightarrow f b_2 = z_3 \rightarrow b_2 = \frac{z_3}{f}$$

$$\Rightarrow d b_1 = e b_2 = z_2$$

$$\Rightarrow b_1 = \frac{z_2 - e b_2}{d} = \frac{z_2 - e\frac{z_3}{f}}{d}$$

If R^2 is 0 then correlation is also 0

correlation 0 means no slope

$$\hat{\vec{y}} = \text{proj}_Q(\vec{y}) = H\vec{y} = QQ^T\vec{y} = \sum_{j=1}^{p+1} \text{proj}_{q_j}(\vec{y})$$

$$\sum \hat{y}_i^2 = \|\hat{\vec{y}}\|^2 = \sum_{j=1}^{p+1} \|\text{proj}_{q_j}(\vec{y})\|^2$$

$$SSR := \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum \hat{y}_i^2 - n\bar{y}^2 = \sum_{j=1}^{p+1} \|\text{proj}_{q_j}(\vec{y})\|^2 - n\bar{y}^2$$

$SST = SSR + SSE$

function of $\vec{y}$, not x's

$$SSR \uparrow \Rightarrow R^2 \uparrow,\ SSE \downarrow,\ RMSE \downarrow$$

$X = QR$

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

full rank

$$SSR \Rightarrow R^2 \downarrow,\ SSE \uparrow,\ RMSE \uparrow$$

$X_{new} = [X \,\vdots\, \vec{X}_{new}]$

adding a new predictor s.t. X_{new} is full rank

$\text{rank}[X_{new}] = (p+1) + 1$

$$\hat{y}_{new} = H_{new}\vec{y} = Q_{new} Q_{new}^T \vec{y} = \sum_{j=1}^{p+1} \text{proj}_{q_j}(\vec{y}) + \text{proj}_{q_{new}}(\vec{y})$$

$$SR := \sum_{j=1}^{p+1} \|\text{proj}_{q_j}(\vec{y})\|^2 + \|\text{proj}_{q_{new}}(\vec{y})\|^2 \Rightarrow SSR_{new} \geq SSR \Rightarrow R^2_{new} \geq R^2$$

$$\Rightarrow SSE_{new} \leq SSE$$

A full rank of a projection is an image of itself

What if $p+1=n \Rightarrow X \in \mathbb{R}^{n \times n}$ square and full rank

$Colsp(X) = \mathbb{R}^n$ the whole space

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

$$\vec{\hat{y}} = H\vec{y} = I\vec{y} = \vec{y} \Rightarrow \vec{e} = 0 \Rightarrow SSE = 0 \Rightarrow R^2 = 1$$

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

$$\vec{\hat{y}} = X\vec{b} \Rightarrow \vec{y} \approx X\vec{b} \qquad \vec{y} = X\vec{b} \Rightarrow \vec{b} = X^{-1}(\vec{y})$$