Math 390.4
lecture 11 3/3

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

colsp$[V] :=$ span$[\{\vec{v}_1, \vec{v}_2\}]$

$\text{proj}_V(\bar{a}) \neq \text{proj}_{\vec{v}_1}(\bar{a}) + \text{proj}_{\vec{v}_2}(\bar{a})$

$H\vec{a} = H_1\vec{a} + H_2\vec{a}$

is $H_1 + H_2$ orthogonal

1) $\text{proj}_V(\vec{a}) = c\vec{v}_1 + c_2\vec{v}_2 \in$ colsp$[V]$

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

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

$\hookrightarrow 2\|H_1\vec{a}\|\|H_2\vec{a}\| *$
$\cos(\theta(H_1\vec{a}, H_2\vec{a})) = 0$

holds if...

1) $\theta = -90^\circ$ or 90°

2) $H_1 a \perp H_2 a$

$\vec{v}_1 \perp v_2$

$\text{proj}_V(\vec{a}) = \text{proj}_{\vec{v}_1}(a) + \dots + \text{proj}_{\vec{v}_j}(a)$

$H\vec{a} = H_1\vec{a} + \dots + H_d\vec{a}$

$= \dfrac{\vec{v}_1\vec{v}_1^T}{\|v_1\|}\vec{a} + \dots + \dfrac{\vec{v}_d\vec{v}_d^T}{\|v_d\|}\vec{a}$

$\left(\dfrac{\vec{v}_1\vec{v}_1^T}{\|v_1\|} + \dots + \dfrac{\vec{v}_d\vec{v}_d^T}{\|v_d\|}\right)\vec{a}$

let $\|\vec{v}_1\| = \dots = \|\vec{v}_d\| = 1$ $H\vec{a}$

$Q = [\vec{v}_1 \mid \dots \mid \vec{v}_d]$

↓ orthogonal basis "orthonormal"

↓ unit length & orthogonal

$\text{proj}_V(\vec{a}) = (\underbrace{\vec{v}_1\vec{v}_1^T + \dots + \vec{v}_d\vec{v}_d^T}_{H})\vec{a}$

iff colsp$[V]$ = colsp$[Q]$

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

$= [v_1\ v_2 \dots v_d]\begin{bmatrix} v_1^T \\ v_2^T \\ v_d^T \end{bmatrix} = QQ^T = H = V\underbrace{(V^TV)^{-1}}V^T$

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

eliminates double counting (orthonormalizing)

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

change of basis

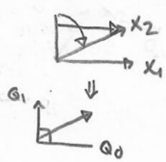
$X \Rightarrow Q$

$\underset{n \times p+1}{X} = \underset{n \times p+1}{Q}\underset{p+1 \times p+1}{R}$ $\quad Q = XR^{-1}$

full rank $\quad$ full rank

R is upper triangular

$X = QR = Q\begin{bmatrix} c & d & e \\ 0 & f & g \\ 0 & 0 & h \end{bmatrix}$



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

$(X^TX)\vec{b} = X^Ty$

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

$R^TQ^TQR\vec{b} = R^TQ\vec{y}$

$R^TR\vec{b} = R^T\vec{z}$

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

$R\vec{b} = \vec{z}$

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

$b_1 = \dfrac{\vec{z} - g\frac{z}{h}}{f}$

Review

$SST = SSR + SSE$

$\sum(y_i - \bar{y})^2$

$SST = SSR + SSE$

↑ $\quad$ ↓

$R^2 \uparrow$ RMSE $\downarrow$

Fixed function of y (does not change)

$$SSR = \sum_{i=1}^{n} (\hat{y} - \bar{y})^2 = \sum_{i=1}^{n} \hat{y}_i - 2\bar{y}\sum_{i=1}^{n} \hat{y}_i + \sum_{i=1}^{n} \bar{y}^2 = \|\vec{\hat{y}}\|^2 - 2n\bar{y}^2 + n\bar{y}^2 = \|\vec{\hat{y}}\|^2 - n\bar{y}^2 =$$

$$= \|H\vec{y}\|^2 - n\bar{y}^2$$

$$\sum \hat{y}_i = \vec{\hat{y}}^T \vec{1}_n = (H\vec{y})^T \vec{1}_n = \vec{y}^T H \vec{1}_n = \vec{y}^T \vec{1}_n = \sum y_i - n\bar{y}$$

$$= \|(\sum_{j=0}^{p} proj_j(\vec{y}))\|^2 - n\bar{y}^2$$

$$X = \begin{bmatrix} | & \uparrow & & \uparrow \\ | & x_1 & \cdots & x_p \\ | & \downarrow & & \downarrow \end{bmatrix} \xrightarrow{G.S} Q = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ q_0 & q_1 & \cdots & q_p \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

note: $\|\vec{a} + \vec{b}\|^2 = \|a\|^2 + \|b\|^2$

add new column to X

$$= \|proj_{\vec{q}_0}(\vec{y})\|^2 + \|proj_{\vec{q}_1}(\vec{y})\|^2 \cdots + \|proj_{\vec{q}_p}(\vec{y})\|^2 - n\bar{y}^2$$

≥ 0 $\qquad \geq 0$

$$X = \begin{bmatrix} | & \uparrow & & \uparrow & \uparrow \\ | & x_1 & \cdots & x_p & \vec{x}_* \\ | & \downarrow & & \downarrow & \downarrow \end{bmatrix} \xrightarrow{G.S} Q = \begin{bmatrix} q_0 & q_1 & \cdots & q_p & \vec{q}_* \end{bmatrix}$$

$SST = SSR_* + SSE_*$ $\qquad H = QQ^T$

$SSR_* > SSR$ $\qquad H_\perp = I - H = Q_\perp Q_\perp^T$

$SSE_* < SSE$

lets add all $n-(p+1)$ columns to get

$X = [\ _{p+1}\ ][\ _{n-p+1}\ ] \in \mathbb{R}^{n \times n}$ full rank

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

Is I orthogonal proj. matrix?

$I^T = I$

$II = I$

$$SSE = \sum_{i=1}^{n} (y - y_i)^2 = \sum 0^2 = 0$$

* new columns are not related to $\vec{z}_j$'s, the true causal inputs

$R^2 = 1$; $RMSE = 0$

"overfitting" (core concept in data science ...

predictive performance on ID

plague of data science in the world)

generalization error

affects future performance

$$y = g(x) + (\underbrace{\underbrace{h^*(\vec{x}) - g(\vec{x})}_{\text{estimation}}) + \underbrace{\underbrace{(h^*(\vec{x}) - f(\vec{x}))}_{\text{mispecification}} + \underbrace{(t(\vec{z}) - f(x))}_{\text{ignorance } \delta}}_{\varepsilon}}_{e}$$

Assume overfit

$\Rightarrow f(\vec{x})$ & $h(\vec{x})$ is fixed

$h^*(x) = \vec{x}\vec{\beta} \Rightarrow \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$; $f(x) = \sin(x_1 x_2) + x_3 + \ldots$

overfitting is estimation error

$g(x) = \vec{x}\vec{b}$

asymmetric overfit

$\vec{b}$ & $\vec{\beta}$ diverge

$h^*(\vec{x}) - g(\vec{x})$: estimation error

$\vec{x}\vec{\beta} - x\vec{b}$

$x(\vec{\beta} - \vec{b})$