Math 390.4
lecture 8; 2/20

$\mathcal{Y} = \mathbb{R}$, $p = 2$

Linear model : $\mathcal{H} = \{ w_0 + w_1 x_1 + w_2 x_2 : \vec{w} \in \mathbb{R}^3 \}$

$\mathbb{D} = \langle X, y \rangle$; $X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{41} \end{bmatrix}$

Algorithm will return one $\vec{w}$
and all $\hat{y}$'s can be compacted in

$\hat{\vec{y}} = X\vec{w}$
$\vec{e} = \vec{y} - \hat{\vec{y}}$
A : OLS

$SSE = \sum e_i^2 = \vec{e}^T \vec{e} = \vec{y}^T \vec{y} - 2\vec{w}^T X^T + \vec{w}^T X^T X \vec{w}$

$A : \vec{b} = \underset{\vec{w} \in \mathbb{R}}{\text{argmin}} \{ \vec{e}^T \vec{e} \}$

$\frac{d}{dw}[SSE] := \begin{bmatrix} \frac{d}{dw_0}[SSE] \\ \frac{d}{dw_1}[SSE] \\ \frac{d}{dw_2}[SSE] \end{bmatrix} = \vec{0}_3$

$\frac{d}{dx}[SSE] = \frac{d}{d\vec{w}}[\vec{y}^T\vec{y} - 2\vec{w}^T X^T \vec{y} + \vec{w}^T X^T X \vec{w}]$

$\vec{0}_3 - 2X^T\vec{y} + 2X^T X\vec{w} = \vec{0}_3$

$2X^T X \vec{w} = 2X^T \vec{y}$

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

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

Assume X is full rank ($X^T X$ full rank)
linearly independent $(p+1)$

$X = \begin{bmatrix} 1 & \vec{x}_1 & \vec{x}_2 & \dots & \vec{x}_p \\ \vdots & \downarrow & \downarrow & & \downarrow \\ 1 & & & & \end{bmatrix}$

X_* is new observation $\in 1 \times \mathbb{R}^{p+1}$

$g(\vec{x}_*) = \vec{x}_* \vec{b} = \hat{y}_*$

$g(X_*) = X_* \vec{b} = \hat{\vec{y}}_*$

$\hat{\vec{y}} = g(X) = X\vec{b} = \underbrace{X(X^T X)^{-1} X^T}_{H \in \mathbb{R}^{n \times n}} \vec{y} = \hat{\vec{y}} = H\vec{y} = T(\vec{y})$

$\underset{p+1}{\text{dim(rank)}} + \underset{n-(p+1)}{\text{dim(nullity)}} = n$

$\vec{y} = \hat{\vec{y}} + e$

$I\vec{y} = \underbrace{H\hat{\vec{y}}}_{T(\vec{y})} + \underbrace{(I-H)\vec{y}}_{u(\vec{y})}$

$SSE = \vec{e}^T e$

$MSE = \frac{1}{n-(p+1)} SSE$

$RMSE = \sqrt{MSE}$

$R^2 = 1 - \frac{SSE}{SST}$

231 Review :
let $\vec{x} \in \mathbb{R}^n$; a: scalar constant w.r.t all x_j's

$\frac{d}{d\vec{x}}[a] = \vec{0}_n$

$\vec{a}$: column vector

$\frac{d}{dx}[\vec{a}^T \vec{x}] = \vec{a}$

f, g are both functions $\mathbb{R}^4 \to \mathbb{R}^4$; a & b scalars

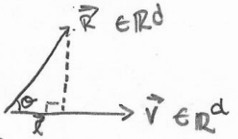
$\frac{d}{dx}[a f(\vec{x}) + b g(\vec{x})]$

$a \frac{d}{dx} f(\vec{x}) + b \frac{d}{dx} g(\vec{x})$

let $A \in \mathbb{R}^{n \times m}$ be a symmetrical matrix of constants w.r.t x_j's

$\frac{d}{dx}[\vec{x}^T A \vec{x}] \to \frac{d}{dx}\left[[x_1 x_2 \dots x_n] \begin{bmatrix} a_1 \vec{x} \\ a_2 \vec{x} \\ \vdots \\ a_n \vec{x} \end{bmatrix} \right] \to$

$\to 2\vec{a}_1 \cdot \vec{x} \to 2A\vec{x}$



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

orthogonal projection

$(\perp)$

we want a formula for $\vec{l}$ as a function of inputs $\vec{a}, \vec{v}$

Recall Law of cosines

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

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

$$\vec{l} = \|\vec{l}\| \frac{\vec{v}}{\|\vec{v}\|} = \frac{\vec{a} \cdot \vec{v}\,\vec{v}}{\|\vec{v}\|^2} = \frac{\overbrace{\vec{v}\vec{v}^T}^{d \times d}}{\|\vec{v}\|^2}\vec{a} = H\vec{a} = \text{proj}_{\vec{v}}(\vec{a})$$

$\vec{v}$ straight line in $\mathbb{R}^d$