Lecture - 06 02/13/2020

Review for lab-03

$$g(x^{-1}) = \mathbb{1}_{\vec{w}.\vec{x} \geq 0} = \hat{y}$$

$$SAE \overset{(\vec{w})}{=} \sum_{i=1}^{n} \mathbb{1}_{g(\vec{x}_i) \neq y_i}$$

$$\underset{\hat{y}_i}{\overset{\|}{g(\vec{x}_i)}}$$

$$H = \{\mathbb{1}_{\vec{w}.\vec{x} \geq 0} : \vec{w} \in \mathbb{R}^{p+1}\}$$

Recall:

$$y \leq \mathbb{R}$$

$$p = 1$$

$$H = \{\underset{w_0 + w_1 x_1}{\overset{\|}{\vec{w}.\vec{x}}} : \vec{w} \in \mathbb{R}^{p+1}\}$$ all linear models

$$\vec{x}' = [1 \ \vec{x}]$$

SSE => Sum of sqd. error

ordinary least square (OLS) regression $\leftarrow A : b_0, b_1 := \underset{\vec{w} \in \mathbb{R}^{p+1}}{\text{avgmax}} \left\{ \sum_{i=1}^{n} (y_i - \vec{w}.\vec{x}_i')^2 \right\}$

$\uparrow$ least square

using calculus....

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

Var[X] is estimated by $S^2 x$

$$S^2 x := \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i^2 - 2\bar{x}x_i + \bar{x}^2)$$

$$= \frac{1}{n-1} \left(\Sigma x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \right)$$

$$= \frac{1}{n-1} \left(\Sigma x_i^2 - n\bar{x}^2 \right)$$

For 2 r.v's X, Y

Correlation : $\text{corr}[X,Y] := \dfrac{\text{Corr}[X,Y]}{\sqrt{\text{var}[X]\,\text{var}[Y]}}$

Covariance : $\text{cov}[X,Y] := E[(X-\mu_x)(Y-\mu_y)]$

cov[X,Y] estimated by S_{xy}

$$S_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} (x_i y_i - x_i \bar{y} - y_i \bar{x} + \bar{x}\bar{y})$$

$$= \frac{1}{n-1} \left(\Sigma x_i y_i - n\bar{x}\bar{y} - \cancel{n\bar{x}}\bar{y} + \cancel{n\bar{x}}\bar{y} \right)$$

$$= \frac{1}{n-1} \left(\Sigma x_i y_i - n\bar{x}\bar{y} \right)$$

Correlation

$Corr[x, y]$ is estimated by

$$r := \frac{S_{xy}}{\sqrt{S^2_x S^2_y}} = \frac{S_{xy}}{S_x S_y}$$

Using calculus

$$b_0 = \bar{y} - b_1 \bar{x} = \bar{y} - r \, {S_y}/{S_x} \, \bar{x}$$

$$b_1 = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{\sum x_i^2 - n \bar{x}^2}$$

$$= \frac{(n-1) S_{xy}}{(n-1) S^2_x} = \frac{S_{xy}}{S^2_x} = \frac{r S_x S_y}{S^2_x} = r \frac{S_y}{S_x}$$

$\Longrightarrow$ $g(x) = b_0 + b_1 x$ (how to predict)

A: data