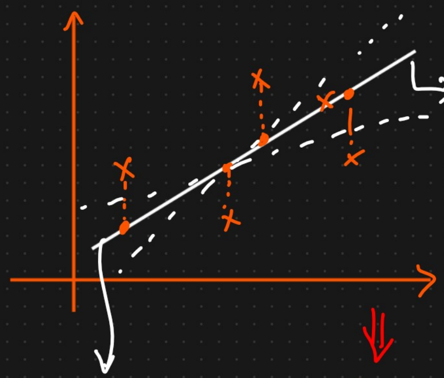
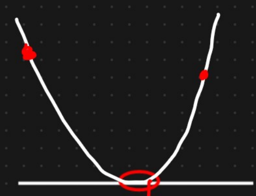
Linear Regression Using OLS {Ordinary Least Square}



→ optimization

⇓

Loss function



Global Minima

⇓

→ $\underline{\underline{\text{OLS}}}$ → Formula and Calculate

$\underline{\beta_0}$ & $\underline{\beta_1}$

$h_\theta(x) = \beta_0 + \beta_1 x_1$

→ $\hat{y}$

Global Minima

Ordinary Least Square

Error ⇑ =

$\boxed{\text{MSE}}$

$$S(\beta_0, \beta_1) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1)^2$$

⇓

Aim: $\underline{\beta_0}$, & $\underline{\beta_1}$

$$y = x^2$$

$$\frac{\partial y}{\partial x} = 2(x)^{2-1} \frac{\partial (x)}{\partial x}$$

$$= 2x$$

$$\frac{\partial S}{\partial \beta_0}(\beta_0, \beta_1) = \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(0 - 1 - 0)$$

$$= \frac{-2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0 \longrightarrow (1)$$

$$\frac{\partial S}{\partial \beta_1}(\beta_0, \beta_1) = \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1)(-x_i) = 0 \longrightarrow (2)$$

$$= \frac{-2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1)(x_1) = 0 \longrightarrow (2)$$

Eq $\rightarrow$ 1

$$\frac{-2}{n}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)=0$$

$$\Downarrow$$

$$-\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)=0$$

$$-\sum_{i=1}^{n}y_i+n*\beta_0+\beta_1\sum_{i=1}^{n}x_i=0$$

$$n\beta_0+\beta_1\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i$$

$$\beta_0=\frac{\sum_{i=1}^{n}y_i}{n}-\frac{\beta_1\sum_{i=1}^{n}x_i}{n}$$

Intercept $\rightarrow$ $\boxed{\beta_0=\bar{y}-\beta_1\bar{x}}$

$\boxed{\text{Eq 2 :}}$

$$\frac{-2}{n}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)(x_i)=0$$

$$\Downarrow$$

$$\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)(x_i)=0$$

$$\Downarrow$$

$$\sum_{i=1}^{n}x_i y_i-\beta_0\sum_{i=1}^{n}(x_i)-\beta_1\sum_{i=1}^{n}(x_i)^2=0$$

$$\sum_{i=1}^{n} \left(x_i y_i - \beta_0 x_i - \beta_1 x_i^2 \right) = 0$$

$$\Downarrow$$

Replace $\beta_0 = \bar{y} - \beta_1 \bar{x}$

$$\sum_{i=1}^{n} \left(x_i y_i - (\bar{y} - \beta_1 \bar{x}) x_i - \beta_1 x_i^2 \right) = 0$$

$$\sum_{i=1}^{n} \left(x_i y_i - x_i \bar{y} + \beta_1 \bar{x} x_i - \beta_1 x_i^2 \right) = 0$$

$$\sum_{i=1}^{n} \left(y_i - \bar{y} + \beta_1 \bar{x} - \beta_1 x_i \right) x_i = 0$$

$$\sum_{i=1}^{n} \left[(y_i - \bar{y}) + \beta_1 (\bar{x} - x_i) \right] = 0$$

$$\sum_{i=1}^{n} (y_i - \bar{y}) + \sum_{i=1}^{n} \beta_1 (\bar{x} - x_i) = 0$$

$$\sum_{i=1}^{n} \beta_1 (\bar{x} - x_i) = - \sum_{i=1}^{n} (y_i - \bar{y})$$

$$\beta_1 = \frac{- \sum_{i=1}^{n} (y_i - \bar{y})}{\sum_{i=1}^{n} (\bar{x} - x_i)}$$

$$\beta_1 = \frac{\cancel{-} \sum_{i=1}^{n} (y_i - \bar{y})}{\cancel{-} \sum_{i=1}^{n} (x_i - \bar{x})}$$

Coefficient $\Rightarrow$ $\boxed{\beta_1 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})}}$

Intercept $\Downarrow$

OLS

Coefficient

$$\boxed{\beta_0 = \bar{y} - \beta_1 \bar{x}}$$

X	y	$(y_i - \bar{y}) \div (x_i - \bar{x})$		Coefficient $\downarrow$ β_1	Intercept $\downarrow$ $\beta_0 = (\bar{y} - \beta_1 \bar{x})$
–	–	–	–		
–	–	–	–		
–	–	–	–		
–	–	–	–		
$\bar{x}$	$\bar{y}$				

OLS $\approx$ Linear Regression (sklearn)