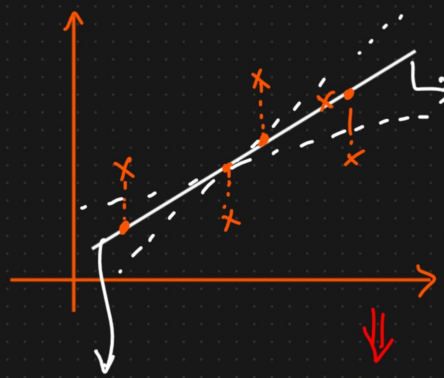
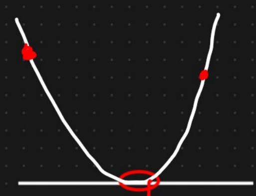
Linear Regression Using OLS {Ordinary Least Square}



↳ optimization

⇓

Loss function



↳ Global minima

⇓

→ OLS → Formula and Calculate

$$h_\theta(x) = \beta_0 + \beta_1 x_1 \quad \hookrightarrow \hat{y}$$

β_0 & β_1

Global Minima

Ordinary Least Square

Error ⇑ ⇐ $\boxed{MSE}$

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

$$y = x^2$$

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

⇓

Aim β_0, & β_1

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

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

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

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

$Eq \rightarrow 1$

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

$$\Downarrow$$

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

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

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

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

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

$\boxed{Eq\ 2 :}$

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

$$\Downarrow$$

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

$$\Downarrow$$

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

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

$$\Downarrow$$

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

Intercept

Ols

Coefficient

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

Coefficient

Intercept

x	y	$(y_i - \bar{y}) \div (x_i - \bar{x})$		β_1	$\beta_0 = (\bar{y} - \beta_1 \bar{x})$
—	—	—	—		
—	—	—	—		
—	—	—	—		
—	—	—	—		
$\bar{x}$	$\bar{y}$				

Ols $\approx$ Linear Regression (sklearn)