linear regression.

④

target variable → numerical variable → regression.

Assumption: 독립변수와 종속변수 간의 선형적인 관계를 가정.

① 선형성
② 독립성
③ 등분산성 분산이 같다.
④ 정규성. 정규분포 인지?

1. 독립변수와 종속변수 간의 선형관계를 가정.
2. 오차항은 정규분포를 따른다고 가정.
(Normality)
3. 오차항은 등분산성을 가정
(Homoscedasticity)
4. 공분산성이 없다는 가정.

$$y_1 = \beta_0 \cdot 1 + \beta_1 x_{11} + \beta_2 x_{21} + \cdots + \beta_p x_{p1}$$

$$y_2 = \beta_0 \cdot 1 + \beta_1 x_{12} + \beta_2 x_{22} + \cdots + \beta_p x_{p2}$$

$$\vdots$$

$$y_n = \beta_0 \cdot 1 + \beta_1 x_{1n} + \beta_2 x_{2n} + \cdots + \beta_p x_{pn}.$$

n: samples.

p: input variables

$$\Downarrow$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} \qquad \begin{bmatrix} \boxed{\begin{matrix} 1 \\ \vdots \\ 1 \end{matrix}} & \begin{matrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{matrix} & \begin{matrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{matrix} & \cdots & \begin{matrix} x_{p1} \\ x_{p2} \\ \vdots \\ x_{pn} \end{matrix} \end{bmatrix}$$

intercept, β_0 표현을 위한 값.

$$\beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1) \times 1} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

추정할 β

Least square metod

(3)

$$RSS(\beta) = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & & \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Np} \end{bmatrix} \text{ and } y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}.$$

$$y - X\beta = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & & & \\ 1 & x_{N1} & \cdots & & x_{Np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$= \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} \\ \\ \beta_0 + \beta_1 x_{N1} + \cdots + \beta_p x_{Np} \end{bmatrix}$$

$$= \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{bmatrix}$$

$$= \begin{bmatrix} y_1 - f(x_1) \\ y_2 - f(x_2) \\ \vdots \\ y_N - f(x_N) \end{bmatrix}$$

$$(y - X\beta)^T (y - X\beta) = \sum_{i=1}^{N} (y - f(x))^2 = RSS(\beta)$$

$$RSS(\beta) = (y - X\beta)^T (y - X\beta)$$

Partial differential

$$RSS(\beta) = \begin{bmatrix} \frac{\partial RSS(\beta)}{\partial \beta_0} \\ \vdots \\ \frac{\partial RSS(\beta)}{\partial \beta_p} \end{bmatrix}$$

$$\min E = \min \|y - X\beta\|$$

$$\frac{\partial E}{\partial \beta} \rightarrow \left(\frac{\partial RSS(\beta)}{\partial \beta_0}, \frac{\partial RSS(\beta)}{\partial \beta_1} \cdots \frac{\partial RSS(\beta)}{\partial \beta_n} \right)$$