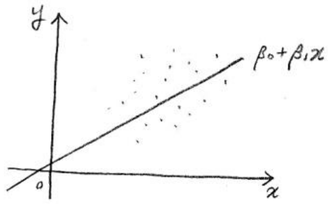
Multiple Regression model

* 관점의 차이 : 관점의 차이는 표현의 차이를 가져옴. (simple로 설명)

$\boxed{\text{Model 1}}$ $\quad y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma^2)$



각각의 자료 $(x_1, y_1)(x_2, y_2)\cdots(x_n, y_n)$에 대하여 임의의 y_i는 기댓값 $E(y_i) = \beta_0 + \beta_1 x_i$에 오차항 ε_i를 더해준 것이라 생각

⇒ $y_i = E(y_i) + \varepsilon_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ 이며, 좀더 잘 $E(y_i|x_i)$를 찾는 것이 목표.

$$\hat{y}_i = \widehat{E(y_i|x_i)} = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

⇒ y_i의 추정량 $\hat{y}_i$는 $x = x_i$에서의 평균반응으로 함!

$\boxed{\text{Model 2}}$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_0 \\ \vdots \\ \beta_0 \end{pmatrix} + \begin{pmatrix} \beta_1 x_1 \\ \beta_1 x_2 \\ \vdots \\ \beta_1 x_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \iff \begin{aligned} \mathbf{y} &= \beta_0 \mathbf{1} + \beta_1 \mathbf{x} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(0, I\sigma^2) \\ &= (\mathbf{1}\ \mathbf{x})\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \boldsymbol{\varepsilon} \end{aligned}$$

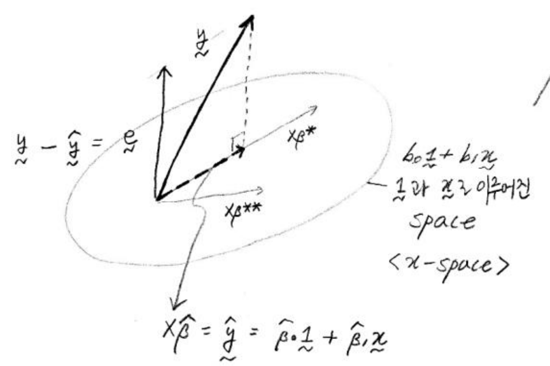
$\mathbf{y}$는 $\mathbf{x}$와 $\mathbf{1}$의 선형결합에 $\boldsymbol{\varepsilon}$를 더해준 것!

⇒ $\mathbf{y}$를 한 묶음으로 보고, $b_0\mathbf{1} + b_1\mathbf{x}$ 중에 가장 $\mathbf{y}$에 가까운 (설명을 잘하는) 것으로 추정 ⇒ $\hat{\mathbf{y}}$

i) 수식적,

$$S(b_0, b_1) = (\mathbf{y} - (b_0\mathbf{1} + b_1\mathbf{x}))'(\mathbf{y} - (b_0\mathbf{1} + b_1\mathbf{x}))$$ 을 최소화하는 $b_0, b_1 \Rightarrow \hat{\beta}_0, \hat{\beta}_1$

ii) 기하학적,



$\mathbf{y}$를 x-space로 projection 했을때의 '$b_0\mathbf{1} + b_1\mathbf{x}$'가 가장 $\mathbf{y}$와 가까움

∴ 이 $b_0\mathbf{1} + b_1\mathbf{x}$이 $\hat{\mathbf{y}}$!

Multiple Regression 으로 확장!

$\boxed{\text{Model 1}}$ $\quad y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \quad , \quad \epsilon_i \sim N(0, \sigma^2)$ $\quad i = 1, 2, \cdots, n$

$\boxed{\text{Model 2}}$ $\quad \underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\epsilon} \quad , \quad \underset{\sim}{\epsilon} \sim N(0, I\sigma^2)$

- $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ 자료가 n개 $n \times 1$

- $\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$ 상수항 + k개의 변수의 계수 $(k \times 1) \times 1$ or $p \times 1$

- $X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1 & x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{ik} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nj} & \cdots & x_{nk} \end{pmatrix} = (\underset{\sim}{1}, \underset{\sim}{x_1}, \underset{\sim}{x_2}, \cdots, \underset{\sim}{x_k}) \quad n \times p$

  * x_{ij} : j 번째 변수의
  i 번째 관측치.

- $\underset{\sim}{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$ $\underset{\sim}{\epsilon} \sim N(0, I\sigma^2)$ 각 ϵ_i는 독립 $n \times 1$

* LSE : $\hat{\beta}$ 유도하기

i) 수식적,

$$S(\underset{\sim}{b}) = (\underset{\sim}{y} - X\underset{\sim}{b})'(\underset{\sim}{y} - X\underset{\sim}{b})$$
$$= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{b}'X'\underset{\sim}{y} - \underset{\sim}{y}'X\underset{\sim}{b} + \underset{\sim}{b}'X'X\underset{\sim}{b}$$
$$= \underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{y}'X\underset{\sim}{b} + \underset{\sim}{b}'X'X\underset{\sim}{b} \quad (= \text{scalar})$$

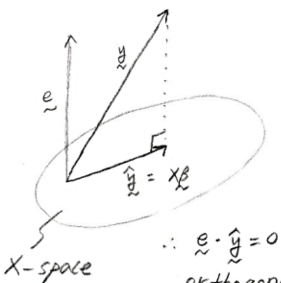
$\underset{\sim}{y}' = 1\times n$	$\underset{\sim}{b}' = 1\times p$
$X = n\times p$	$X' = p\times n$
$\underset{\sim}{b} = p\times 1$	$\underset{\sim}{y} = n\times 1$
$\underset{\sim}{y}'X\underset{\sim}{b} = 1\times 1$	$\underset{\sim}{b}'X'\underset{\sim}{y} = 1\times 1$

공식

$$\left.\frac{\partial S(\underset{\sim}{b})}{\partial \underset{\sim}{b}}\right|_{\underset{\sim}{b}=\hat{\underset{\sim}{\beta}}} = (-2\underset{\sim}{y}'X)' + \left.(X'X + (X'X)')\underset{\sim}{b}\right|_{\underset{\sim}{b}=\hat{\underset{\sim}{\beta}}} = \underset{\sim}{0}$$

$$\therefore -2X'\underset{\sim}{y} + 2X'X\hat{\underset{\sim}{\beta}} = \underset{\sim}{0} \quad (X'X : \text{symmetric})$$
$$X'\underset{\sim}{y} = X'X\hat{\underset{\sim}{\beta}}$$
$$\therefore \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

ii) 기하학적,



$\therefore \underset{\sim}{e} \cdot \hat{\underset{\sim}{y}} = 0$
orthogonal!

$$\underset{\sim}{e} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$$
$$(\underset{\sim}{y} - \hat{\underset{\sim}{y}})'\hat{\underset{\sim}{y}} = \underset{\sim}{0}$$
$$(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})'X\hat{\underset{\sim}{\beta}} = \underset{\sim}{0}$$
$$\underset{\sim}{y}'X\hat{\underset{\sim}{\beta}} - \hat{\underset{\sim}{\beta}}'X'X\hat{\underset{\sim}{\beta}} = \underset{\sim}{0}$$
$$\hat{\underset{\sim}{\beta}}'X'\underset{\sim}{y} = \hat{\underset{\sim}{\beta}}'X'X\hat{\underset{\sim}{\beta}}$$
$$X'\underset{\sim}{y} = X'X\hat{\underset{\sim}{\beta}}$$
$$\therefore \hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

$$\therefore \hat{\underset{\sim}{y}} = X\hat{\underset{\sim}{\beta}}$$
$$= X(X'X)^{-1}X'\underset{\sim}{y}$$
$$= H\underset{\sim}{y}$$

→ H는 $\underset{\sim}{y}$를 X-space로 projection

* think!

Model A : $y = \beta_0 + \beta_1 x_1 + \varepsilon \rightarrow e_A$

Model B : $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \rightarrow e_B$

$e_B \perp (x_1, x_2)$?

* Properties of $\hat{\beta}$ (distribution)

i) $E(\hat{\beta}) = \beta$

<u>proof</u> $\hat{\beta} = (X'X)^{-1}X'y$

$E(y) = X\beta$

$Var(y) = I\sigma^2$ ← ε만 확률변수 이므로. $\varepsilon \sim N(0, I\sigma^2)$

$$
\begin{aligned}
\therefore E(\hat{\beta}) &= E((X'X)^{-1}X'y) \\
&= (X'X)^{-1}X'E(y) \\
&= (X'X)^{-1}X'X\beta \\
&= \beta
\end{aligned}
$$

∴ unbiased estimator

ii)
$$
\begin{aligned}
Var(\hat{\beta}) &= Var((X'X)^{-1}X'y) \\
&= (X'X)^{-1}X'\,Var(y)\,((X'X)^{-1}X')' \\
&= (X'X)^{-1}X' \cdot I\sigma^2 \cdot X(X'X)^{-1} \\
&= (X'X)^{-1}\underline{(X'X)(X'X)^{-1}} \cdot \sigma^2 \\
&= (X'X)^{-1}\sigma^2
\end{aligned}
$$

$X'X$: symmetric

∴ Inverse도 symmetric

$\therefore \hat{\beta} \sim N(\beta, (X'X)^{-1}\sigma^2)$ $\because C = (X'X)^{-1}$

$\hat{\beta}_j \sim N(\beta_j, C_{jj}\sigma^2)$

* estimate of σ^2

$$\hat{\sigma}^2 = \frac{e'e}{n-p-1} = \frac{SSE}{n-p-1}$$

$(\because SSE = (y-\hat{y})'(y-\hat{y}) = e'e)$