김은영

제3장 중선형회귀모형

3.1 중선형회귀모형

3.1.1 모집단 중회귀모형

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{p-1} x_{i,p-1} + \epsilon_i \qquad (i = 1, \cdots, n) \quad \epsilon_i \sim i.i.d.\, N(0, \sigma^2)$$

3.1.2 선형모형과 비선형모형

회귀모형 $y = f(x, \beta) + \epsilon$

회귀식 $f(x, \beta)$이 모든 β_j들에 대해 선형모형 → '선형모형'

[예] $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$

$y = \beta_0 + \beta_1 \log x_1 + \beta_2 \log x_2 + \epsilon$

3.1.3 행렬을 이용한 모형식

행렬 표현: $y = X\beta + \epsilon$ 여기서 X: 디자인행렬(design matrix)

$$\begin{aligned} y_1 &= \beta_0 1 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_{p-1} x_{1,p-1} + \epsilon_1 \\ y_2 &= \beta_0 1 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_{p-1} x_{2,p-1} + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 1 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_{p-1} x_{n,p-1} + \epsilon_n \end{aligned}$$

$$y_{(n\times 1)} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\ X_{(n\times p)} = \begin{bmatrix} 1 & x_{11} & x_{12} \cdots x_{1,p-1} \\ 1 & x_{21} & x_{22} \cdots x_{2,p-1} \\ \vdots & & \vdots \ \vdots \ \vdots \\ 1 & x_{n1} & x_{n2} \cdots x_{n,p-1} \end{bmatrix},\ \beta_{(p\times 1)} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix},\ \epsilon_{(n\times 1)} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$$y = X\beta + \epsilon \qquad \epsilon \sim N_n(0, \sigma^2 I)$$

3.2 회귀계수의 추정

3.2.1 최소제곱법

$$Q = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon^t \epsilon = (y - X\beta)^t (y - X\beta)$$

$$\hat{\beta}_{LSE} = \underset{\beta}{\arg\min}\ (y - X\beta)^t (y - X\beta)$$

$$= \underset{\beta}{\arg\min}\ (y^t - \beta^t X^t)(y - X\beta)$$

$$= \underset{\beta}{\arg\min}\ (y^t y - \beta^t X^t y - y^t X\beta + \beta^t X^t X\beta)$$

$$= \underset{\beta}{\arg\min}\ (y^t y - 2\beta^t X^t y + \beta^t X^t X\beta)$$

$\dfrac{\partial Q}{\partial \beta} = -2X^t y + 2X^t X\beta = 0 \ \rightarrow \ X^t X\beta = X^t y$: 정규방정식

만약 $X^t X$ 가 정칙행렬이면, $(X^t X)^{-1} X^t X\hat{\beta} = (X^t X)^{-1} X^t y \ \Rightarrow \ \hat{\beta}_{LSE} = (X^t X)^{-1} X^t y$

$\rightarrow$ linear in y

3.2.2 최소제곱추정량의 성질

1. $E(\hat{\beta}) = E\left[(X^t X)^{-1} X^t y\right] = (X^t X)^{-1} X^t E(y) = (X^t X)^{-1} X^t X\beta = \beta$

   $\rightarrow$ $\hat{\beta}$: β의 불편추정량

2. $Cov(\hat{\beta}) = Cov\left[(X^t X)^{-1} X^t y\right] = (X^t X)^{-1} X^t Cov(y) X (X^t X)^{-1}$

   $= (X^t X)^{-1} X^t X (X^t X)^{-1} Cov(y) = (X^t X)^{-1} Cov(y) = (X^t X)^{-1} \sigma^2$

3. 잔차벡터 $e = y - \hat{y} = y - Hy = (I - H)y$

   $\hat{y} = X\hat{\beta} = X(X^t X)^{-1} X^t y = Hy$

   여기서, $H = X(X^t X)^{-1} X^t$ $\rightarrow$ projection[hat] matrix

   $H^2 = X(X^t X)^{-1} X^t X (X^t X)^{-1} X^t = X(X^t X)^{-1} X^t = H$

   $\rightarrow$ H : 멱등행렬

4. 잔차벡터의 성질

(1) $e = y - \hat{y} = y - Hy = (I - H)y$

여기서, $I - H$: 멱등행렬

(2) $\hat{y} \perp e$

[증명] $\hat{y}^t e = (Hy)^t (I-H)y = y^t H(I-H)y = y^t (H-H^2)y = y^t (H-H)y = 0$

(3) $X^t e = 0$

[증명] $X^t(I-H)y = X^t[I - X(X^t X)^{-1} X^t]y = [X^t - X^t X(X^t X)^{-1} X^t]y = [X^t - X^t]y = 0$

[정리 3.1] Gauss-Markov Theorem

회귀모형에서 $E[\epsilon]=0,\ Cov(\epsilon)=I_n\sigma^2$일 때 최소제곱추정량 $\hat{\beta}$은 y_i들의 선형함수로 주어지는 β의 선형불편추정량들 중에서 가장 작은 분산을 갖는다. 즉, $\hat{\beta}$은 β의 최량선형불편추정량(BLUE)이다.

[증명]

$\hat{\beta} = \left(\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_{p-1}\right)^t$

$\hat{\beta}^* = \left(\hat{\beta}_0^*, \hat{\beta}_1^*, \cdots, \hat{\beta}_{p-1}^*\right)^t$: β의 선형불편추정량

$$\boxed{Var(\hat{\beta}_j) \le Var(\hat{\beta}_j^*) \qquad (j = 0, 1, \cdots, p-1)}$$

$\hat{\beta}^* = Ay = \left[(X^t X)^{-1} X^t + D\right]y$

여기서, 적당한 $(p \times n)$ 상수행렬 D에 대해 $A = (X^t X)^{-1} X^t + D$

$\hat{\beta}^*$: β의 선형불편추정량 $\rightarrow$ $\beta = E(\hat{\beta}^*) = E\left[\left\{(X^t X)^{-1} X^t + D\right\}y\right]$

$$= (X^t X)^{-1} X^t X\beta + DX\beta = \beta + DX\beta$$

$$\Rightarrow\ DX = 0$$

$$Cov(\hat{\beta}^*) = Cov\left[\left\{(X^t X)^{-1} X^t + D\right\}y\right] = \left\{(X^t X)^{-1} X^t + D\right\} Cov(y) \left\{X(X^t X)^{-1} + D^t\right\}$$

$$= \left\{(X^t X)^{-1} X^t + D\right\}\left\{X(X^t X)^{-1} + D^t\right\} Cov(y)$$

$$= \left[(X^t X)^{-1} X^t X(X^t X)^{-1} + (X^t X)^{-1}(DX)^t + DX(X^t X)^{-1} + DD^t\right]\sigma^2$$

$$= \left[(X^t X)^{-1} + DD^t\right]\sigma^2 = (X^t X)^{-1}\sigma^2 + DD^t\sigma^2$$

$$= Cov(\hat{\beta}) + DD^t\sigma^2$$

$\Rightarrow\ Cov(\hat{\beta}^*) = Cov(\hat{\beta}) + DD^t\sigma^2$

j번째 회귀계수추정량의 분산은 위 행렬의 j번째 대각원소이므로

$Var(\hat{\beta}_j^*) = Var(\hat{\beta}_j) + \sigma^2 \sum_{i=1}^{n} d_{ji}^2$ 여기서, $\sum_{i=1}^{n} d_{ji}^2$: 항상 양수

$\ge Var(\hat{\beta}_j)$