중회귀모형에서 $\boldsymbol{\beta}$의 LSE는 오차제곱합

$$S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{Z} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Z} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{Z}'\mathbf{Z} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Z} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}$$

를 최소로 하는 $\boldsymbol{\beta}$이다. $S(\boldsymbol{\beta})$를 $\boldsymbol{\beta}$에 관하여 미분하여 0으로 놓으면 다음의 정규방정식

$$\frac{S(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{Z} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$

으로부터 $\boldsymbol{\beta}$의 LSE는

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z} \tag{2-8}$$

가 된다.

2.1.3 최소제곱추정량의 성질들

중회귀모형에서 최소제곱추정법에 의하여 구한 모수 $\boldsymbol{\beta}$의 LSE $\hat{\boldsymbol{\beta}}$의 중요한 성질들을 요약 · 정리하면 다음과 같다.

① $\hat{\boldsymbol{\beta}}$은 평균 $\boldsymbol{\beta}$, 공분산 $(\mathbf{X}'\mathbf{X})^{-1}\sigma_\varepsilon^2$를 갖는 다변량 정규분포를 따른다. 즉, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}$는 정규확률 벡터 $\mathbf{Z}$의 선형결합이고

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{Z}) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}) \\ &= \boldsymbol{\beta} \end{aligned}$$

이다. 또한 $A = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$라 놓으면 $\hat{\boldsymbol{\beta}}$의 공분산 행렬은

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol\beta}) &= \mathrm{Cov}(AZ) = A\mathrm{Cov}(Z)A' = A\sigma_\varepsilon^2 IA' \\ &= \sigma_\varepsilon^2(X'X)^{-1}X'X(X'X)^{-1} = \sigma_\varepsilon^2(X'X)^{-1}\end{aligned}$$

과 같다. 예를 들어, 단순회귀모형 $Z_t = \beta_0 + \beta_1 X_t + \varepsilon_t$의 경우 LSE $\hat\beta_0$과 $\hat\beta_1$의 공분산행렬은 아래와 같다.

$$\mathrm{Cov}(\hat{\boldsymbol\beta}) = \begin{pmatrix}\mathrm{Var}(\hat\beta_0) & \mathrm{Cov}(\hat\beta_0, \hat\beta_1) \\ \mathrm{Cov}(\hat\beta_0, \hat\beta_1) & \mathrm{Var}(\hat\beta_1)\end{pmatrix}$$

$$= \sigma_\varepsilon^2\left(\begin{pmatrix}1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n\end{pmatrix}\begin{pmatrix}1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n\end{pmatrix}\right)^{-1}$$

$$= \sigma_\varepsilon^2\begin{pmatrix}n & \sum_{t=1}^{n} X_t \\ \sum_{t=1}^{n} X_t & \sum_{t=1}^{n} X_t^2\end{pmatrix}^{-1}$$

$$= \sigma_\varepsilon^2\begin{pmatrix}\frac{1}{n} + \frac{\overline{X}^2}{\sum_{t=1}^{n}(X_t - \overline{X})^2} & -\frac{\overline{X}}{\sum_{t=1}^{n}(X_t - \overline{X})^2} \\ -\frac{\overline{X}}{\sum_{t=1}^{n}(X_t - \overline{X})^2} & \frac{1}{\sum_{t=1}^{n}(X_t - \overline{X})^2}\end{pmatrix}$$

② Z의 **적합값**(fitted value)은 $\hat{\mathbf{Z}} = X\hat{\boldsymbol\beta}$이고 오차항의 추정값으로 사용할 수 있는 **잔차**(residual)는

$$\boldsymbol{e} = \mathbf{Z} - \hat{\mathbf{Z}} = \mathbf{Z} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z} = [\boldsymbol{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Z}$$

이다. 또한 잔차벡터 $\boldsymbol{e}$의 평균과 공분산행렬은 다음과 같다.

$$E(\boldsymbol{e}) = [\boldsymbol{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']E(\mathbf{Z}) = [\boldsymbol{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{X}\boldsymbol\beta = 0$$

과

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{e}) &= [\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\mathrm{Cov}(\mathbf{Z})[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'] \\ &= [\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'][\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\sigma_\varepsilon^2 \\ &= [\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}']\sigma_\varepsilon^2\end{aligned}$$

③ 종속변수의 변동이 회귀모형에 있는 설명변수들에 의해서 얼마만큼 설명되는지를 나타내기 위해 다음과 같이 제곱합들을 정의하기로 한다. SST는 종속변수의 총 제곱합(total sum of squares)이라 하고 종속변수의 평균에 의하여 수정된 변동합으로서

$$\mathrm{SST} = \sum_{t=1}^{n}(\mathrm{Z}_t - \overline{\mathrm{Z}})^2 = \sum_{t=1}^{n}\mathrm{Z}_t^2 - n\overline{\mathrm{Z}}^2 = \mathbf{Z}'\mathbf{Z} - n\overline{\mathrm{Z}}^2$$

이다. SST는 설명변수들에 의해 설명되는 변동합, 즉 회귀모형에 의해 설명되는 변동합인 회귀제곱합(sum of squares due to regression ; SSR)과 회귀모형에 의해 설명되지 못하고 남아 있는 오차변동합인 오차제곱합(sum of squares due to error ; SSE)의 합이며, 각 제곱합은 다음과 같다.

$$\mathrm{SSR} = \sum_{t=1}^{n}(\hat{\mathrm{Z}}_t - \overline{\mathrm{Z}})^2 = \sum_{t=1}^{n}\hat{\mathrm{Z}}_t^2 - n\overline{\mathrm{Z}}^2 = \hat{\boldsymbol{\beta}}'\mathrm{X}'\mathrm{Z} - n\overline{\mathrm{Z}}^2$$

$$\mathrm{SSE} = \sum_{t=1}^{n}(\mathrm{Z}_t - \hat{\mathrm{Z}}_t)^2 = \boldsymbol{e}'\boldsymbol{e}$$

$$\begin{aligned}\mathrm{SST} &= \sum_{t=1}^{n}(\mathrm{Z}_t - \overline{\mathrm{Z}})^2 = \sum_{t=1}^{n}[(\mathrm{Z}_t - \hat{\mathrm{Z}}_t) + (\hat{\mathrm{Z}}_t - \overline{\mathrm{Z}})]^2 \\ &= \sum_{t=1}^{n}(\hat{\mathrm{Z}}_t - \overline{\mathrm{Z}})^2 + \sum_{t=1}^{n}(\mathrm{Z}_t - \hat{\mathrm{Z}}_t)^2 = \mathrm{SSE} + \mathrm{SSR}\end{aligned}$$

회귀모형의 설명력을 나타내는 측도인 **결정계수**(coefficient of determination) R^2는

$$R^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}}$$

와 같이 정의된다. 결정계수 R^2가 가질 수 있는 값의 범위는 $0 \le R^2 \le 1$이며, R^2가 1에 가까울수록 종속변수가 회귀모형에 의해 잘 설명된다고 판단된다.

④ 다음과 같은 **분산분석**(analysis of variance ; ANOVA)표는 회귀모형이 통계적으로 유의한지를 검정하기 위하여 작성된다.

분산분석표

원인	제곱합	자유도	평균제곱합	F비
모형	SSR	p	$\mathrm{MSR}=\dfrac{\mathrm{SSR}}{p}$	$F=\dfrac{\mathrm{MSR}}{\mathrm{MSE}}$
오차	SSE	$n-p-1$	$\mathrm{MSE}=\dfrac{\mathrm{SSE}}{(n-p-1)}$	
수정합	SST	$n-1$		

⑤ 오차분산 $\hat{\sigma}_\varepsilon^2$는 평균오차제곱합인 MSE에 의해 추정된다. 즉,

$$\hat{\sigma}_\varepsilon^2=\mathrm{MSE}=\frac{\mathrm{SSE}}{n-p-1}\equiv s^2$$

이며 s^2는 σ_ε^2의 불편추정량이다. 즉, $\dfrac{\mathrm{SSE}}{\sigma_\varepsilon^2}$는 자유도 $n-p-1$을 갖는 χ^2분포를 따르므로

$$E(s^2)=E\left(\frac{\mathrm{SSE}}{n-p-1}\right)=\frac{\sigma_\varepsilon^2}{n-p-1}E\left(\frac{\mathrm{SSE}}{\sigma_\varepsilon^2}\right)=\sigma_\varepsilon^2$$

가 된다.

⑥ LSE $\hat{\boldsymbol{\beta}}$의 추정된 공분산행렬은

$$\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}})=\hat{\sigma}_\varepsilon^2(\mathbf{X}'\mathbf{X})^{-1}=s^2(\mathbf{X}'\mathbf{X})^{-1}$$

이다. 예를 들어 $\hat{\beta}_i$의 추정된 표준편차는 $s_{\hat{\beta}_i}=s\sqrt{c_{ii}}$이다. 여기서 c_{ii}는 행렬 $(\mathbf{X}'\mathbf{X})^{-1}$의 i 번째 대각원소값이다.

⑦ LSE $\hat{\boldsymbol{\beta}}$과 s^2은 서로 독립이다.