Multiple Regression

설명변수 2개인 경우 - Data Structure

■ 자료구조

반응변수 response	설명변수 explanatory	
y_1	x_{11}	x_{12}
y_2	x_{21}	x_{22}
$\vdots$	$\vdots$	$\vdots$
y_n	x_{n1}	x_{n2}

i번째 (2)

x_{ij}

$j = 1, 2$ (설명변수 의미)

몇번째 데이터인지 case의값의미

Example

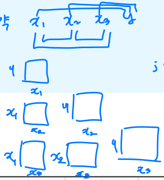
Table: 총판매액의 자료

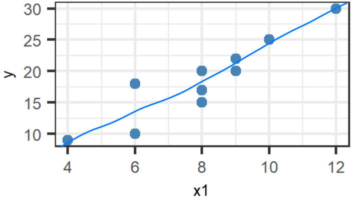
상점번호	1	2	3	4	5	6	7	8	9	10
광고료 x_1(단위 : 10만 원)	4	8	9	8	8	12	6	10	6	9
상점의 크기 x_2(단위 : 평)	4	10	8	5	10	15	8	13	5	12
총판매액 y(단위 : 100만 원)	9	20	22	15	17	30	18	25	10	20

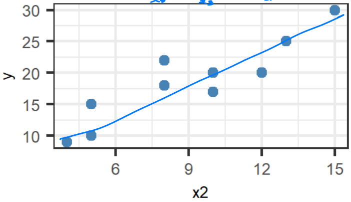
Example

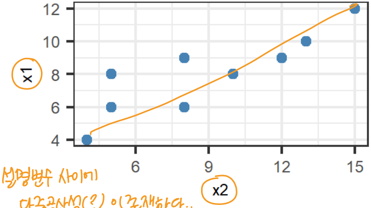
만약 x_1 x_2 x_3 상

; 산점도행렬









설명변수 사이에
다중공산성(?) 이 존재한다..

중회귀모형(Multiple Linear Regression)

- 설명변수가 2개인 다중(선형)회귀모형

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \quad i = 1, 2, \ldots, n$$

  - 회귀모수 : $\beta_0, \beta_1, \beta_2$
  - 설명변수(독립변수) :

$$X_1 = (x_{11}, \ldots, x_{n1})^\top, X_2 = (x_{12}, \ldots, x_{n2})^\top$$

  - 반응변수(종속변수) : $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$
  - 오차항 : $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^\top, \ \epsilon_i \sim_{i.i.d} N(0, \sigma^2)$

Least Square Estimation

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2}$

- 최소제곱추정량 :

$$\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right) = \underset{(\beta_0,\beta_1,\beta_2)\in\mathbb{R}^3}{\arg\min} \sum_{i=1}^{n} \{y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})\}^2$$

제약조건 3개

$\Rightarrow \sum e_i = 0$

$$\frac{\partial S}{\partial \beta_0} = \cancel{2} \sum_{i=1}^{n} \Big(y_i - \overbrace{(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2})}^{\hat{y}_i}\Big)(+1) = 0$$

$\Rightarrow \sum e_i x_{i1} = 0$

② $$\frac{\partial S}{\partial \beta_1} = 2 \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2}\right)(-x_{i1}) = 0$$

$\Rightarrow \sum e_i x_{i2} = 0$

③ $$\frac{\partial S}{\partial \beta_2} = \cancel{2} \sum_{i=1}^{n} \left(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2})\right)(+x_{i2}) = 0$$

Least Square Estimation

- 정규방정식 (normal equations)

1

$$
\begin{aligned}
\sum_{i=1}^{n} y_i &= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} \\
\sum_{i=1}^{n} x_{i1} y_i &= \hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1} x_{i2} \\
\sum_{i=1}^{n} x_{i2} y_i &= \hat{\beta}_0 \sum_{i=1}^{n} x_{i2} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} x_{i2} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}^2
\end{aligned}
$$

$$① \sum_{i=1}^{n} y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}$$

① : $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2$

② $\sum e_i x_{i1} = 0$

$$\sum e_i x_{i1} = \sum e_i (x_{i1} - \bar{x}_1 + \bar{x}_1)$$
$$= \sum e_i (x_{i1} - \bar{x}_1) + \bar{x}_1 \underbrace{\sum e_i}_{=0}$$

$$e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2})$$
$$= y_i - \bar{y} + \hat{\beta}_1 \bar{x}_1 + \hat{\beta}_2 \bar{x}_2 - \cdots$$
$$= (y_i - \bar{y}) - \hat{\beta}_1 (x_{i1} - \bar{x}_1) - \hat{\beta}_2 (x_{i2} - \bar{x}_2)$$
$$= \sum (y_i - \bar{y})(x_{i1} - \bar{x}_1) - \hat{\beta}_1 \sum (x_{i1} - \bar{x}_1)^2 - \hat{\beta}_2 \sum (x_{i2} - \bar{x}_2)(x_{i1} - \bar{x}_1)$$
$$= S_{(1y)} - \hat{\beta}_1 S_{(11)} - \hat{\beta}_2 S_{(12)} = 0$$

$$③ \begin{cases} \hat{\beta}_1 S_{(11)} + \hat{\beta}_2 S_{(12)} = S_{(1y)} \\ \hat{\beta}_1 S_{(12)} + \hat{\beta}_2 S_{(22)} = S_{(2y)} \end{cases}$$

$S_{(12)}$
1번째와 2번째의
공분산 느낌

$S_{(12)}$가 0에가까워지면..

$$\hat{\beta}_1 = \frac{S_{(1y)}}{S_{(11)}} \fallingdotseq \frac{S_{(xy)}}{S_{(xx)}}$$

$$\begin{cases} \hat{\beta}_1 S_{(11)} + \hat{\beta}_2 S_{(12)} = S_{(1y)} \\ \hat{\beta}_1 S_{(12)} + \hat{\beta}_2 S_{(22)} = S_{(2y)} \end{cases}$$

$$\begin{pmatrix} S_{(11)} & S_{(12)} \\ S_{(12)} & S_{(22)} \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} S_{(1y)} \\ S_{(2y)} \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} S_{(11)} & S_{(12)} \\ S_{(12)} & S_{(22)} \end{pmatrix}^{-1} \begin{pmatrix} S_{(1y)} \\ S_{(2y)} \end{pmatrix}$$

단순 선형 회귀모형

행렬의 사용

- j 번째 관측치

$$y_j = \beta_0 + \beta_1 x_j + \epsilon_j = \begin{pmatrix} 1 & x_j \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \epsilon_j$$

- n 개의 관측치

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

단순 선형 회귀모형

- $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{n\times 1} = \underbrace{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}}_{n\times 2} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}}_{2\times 1} + \underbrace{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}}_{n\times 1}$$

단순 선형 회귀모형

Example

$$
\boldsymbol{y} = \begin{pmatrix} 9 \\ 20 \\ 22 \\ 15 \\ 17 \\ 30 \\ 18 \\ 25 \\ 10 \\ 20 \end{pmatrix}, \quad
\boldsymbol{X} = \begin{array}{l} \;\text{절편} \\ \begin{pmatrix} 1 & 4 \\ 1 & 8 \\ 1 & 9 \\ 1 & 8 \\ 1 & 8 \\ 1 & 12 \\ 1 & 6 \\ 1 & 10 \\ 1 & 6 \\ 1 & 9 \end{pmatrix} \end{array}, \quad
\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \end{pmatrix}, \quad
\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}
$$

단순 선형 회귀모형

$\sum (y_i - \beta_0 - \beta_1 x_i)^2$

의 최소값 찾기 위해서..

$y = X\beta + \epsilon$

- 오차제곱합

$$\sum_{j=1}^{n} \epsilon_j^2 = \underbrace{\underset{1\times n}{\epsilon^\top}\,\underset{n\times 1}{\epsilon}}_{1\times 1} = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

$(y^T - \beta^T X^T)(y - X\beta) = y^T y - \underbrace{\underset{1\times n}{y^T}\,\underset{n\times 2}{X}\,\underset{2\times 1}{\beta}}_{=1\times 1} - \underbrace{\beta^T X^T \beta}_{=1\times 1} + \beta^T X^T X \beta$

상수이므로..

$$= \boldsymbol{y}^\top \boldsymbol{y} - 2\boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{y} + \boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta}$$

미분해서 0으로 만드는 β를 찾자!

☆
$(A+B)^T = A^T + B^T$
$(AB)^T = B^T A^T$
$(A^{-1})^T = (A^T)^{-1}$

미분법

constant vector

- $\mathbf{c}$, $\mathbf{x}$: $n \times 1$ 벡터

$$\underset{1\times n}{\mathbf{c}^\top}\underset{n\times 1}{\mathbf{x}} = (c_1, \ldots, c_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} c_i x_i$$

- $\mathbf{c}^\top\mathbf{x}$의 $\mathbf{x}$에 대한 편도함수(partial derivative)

$\frac{\partial(ax)}{\partial x} = a$ 와 같은 미분개념

벡터미분
: 각 element로 미분하여
열벡터 값을 가진다

$$\frac{\partial(\mathbf{c}^\top\mathbf{x})}{\partial\mathbf{x}} = \begin{pmatrix} \frac{\partial(\mathbf{c}^\top\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial(\mathbf{c}^\top\mathbf{x})}{\partial x_n} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \mathbf{c}$$

미분법

- $$\frac{\partial(\mathbf{c}^\top \mathbf{x})}{\partial \mathbf{x}} = \mathbf{c}$$ $\quad \frac{\partial(x^T c)}{\partial x} = c$

- 이차형식 $\underbrace{\underset{1\times n}{\boldsymbol{y}^\top}\,\underset{n\times n}{A}\,\underset{n\times 1}{\boldsymbol{y}}}_{1\times 1}$

$$A:\ \text{symmetric} \Rightarrow \frac{\partial(\boldsymbol{y}^\top A\boldsymbol{y})}{\partial \underset{n\times 1}{\boldsymbol{y}}} = 2\overbrace{\underset{n\times n}{A}\,\underset{n\times 1}{\boldsymbol{y}}}^{n\times 1}$$

대칭행렬

: 미분결과도 분모 형식과 같은. dimension 나와야 함

$$A:\ \text{non-symmetric} \Rightarrow \frac{\partial(\boldsymbol{y}^\top A\boldsymbol{y})}{\partial \boldsymbol{y}} = (A + A^\top)\boldsymbol{y}$$

단순 선형 회귀모형

- 오차제곱합

$\hat{\beta} = \arg\min \varepsilon^T \varepsilon$

$$\sum_{j=1}^{n} \epsilon_j^2 = \boldsymbol{\epsilon}^\top \boldsymbol{\epsilon} = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$$

$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$

$$= \boldsymbol{y}^\top \boldsymbol{y} - 2\underbrace{\boldsymbol{\beta}^\top}_{2\times1} \underbrace{\boldsymbol{X}^\top}_{2\times n} \underbrace{\boldsymbol{y}}_{n\times1} + \boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta}$$

$\dfrac{\partial S}{\partial \underset{(2\times1)}{\beta}} = 0 - 2X^T y + 2X^T X \beta = 0$

β^T 빼고 계수안남음

- 정규방정식

$$\boldsymbol{X}^\top \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}^\top \boldsymbol{y},$$

- 최소제곱합 (LSE)

$$\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)^\top = \hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}$$

단순 선형 회귀모형

$$
\boldsymbol{X}^\top \boldsymbol{y} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum y_j \\ \sum x_j y_j \end{pmatrix}
$$

$$
\boldsymbol{X}^\top \boldsymbol{X} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix}
$$

단순 선형 회귀모형

$$(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{\sum x_j^2}{n\sum x_j^2 - \left(\sum x_j^2\right)^2} & \dfrac{-\sum x_j}{n\sum x_j^2 - \left(\sum x_j^2\right)^2} \\ \dfrac{-\sum x_j}{n\sum x_j^2 - \left(\sum x_j^2\right)^2} & \dfrac{n}{n\sum x_j^2 - \left(\sum x_j^2\right)^2} \end{pmatrix}$$

$$\Rightarrow \hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat\beta_0 \\ \hat\beta_1 \end{pmatrix} = \begin{pmatrix} \dfrac{(\sum x_j^2)(\sum y_j) - (\sum x_j)(\sum x_j y_j)}{n\sum x_j^2 - (\sum x_j)^2} \\ \dfrac{n\sum x_j y_j - (\sum x_j)(\sum y_j)}{n\sum x_j^2 - (\sum x_j)^2} \end{pmatrix}$$

$\Rightarrow S_{(xy)}$

↗ n으로 나누면

$\sum x_j^2 - \frac{1}{n}(\sum x_j)^2$

$= \sum(x_j - \bar{x})^2 = S_{(xx)}$

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

$\hat\beta_1 = \dfrac{S_{(xy)}}{S_{(xx)}}$

단순 선형 회귀모형

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \dfrac{(\sum x_j^2)(\sum y_j) - (\sum x_j)(\sum x_j y_j)}{n \sum x_j^2 - (\sum x_j)^2} \\ \\ \dfrac{n \sum x_j y_j - (\sum x_j)(\sum y_j)}{n \sum x_j^2 - (\sum x_j)^2} \end{pmatrix}$$

Data Structure and Model

- 자료구조 (설명변수 p개)

$$
\begin{array}{ccccc}
\text{response} & & & \text{explanatory} & \\
y_1 & & x_{11} & \cdots & x_{1p} \\
y_2 & & x_{21} & \cdots & x_{2p} \\
\vdots & & \vdots & \ddots & \vdots \\
y_n & & x_{n1} & \cdots & x_{np}
\end{array}
$$

중회귀모형(Multiple Linear Regression)

- 설명변수가 p개인 다중(선형)회귀모형

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i \quad i = 1, 2, \ldots, n$$

- 회귀모수 : $\beta_0, \beta_1, \ldots, \beta_p$ ← p+1개
- 설명변수(독립변수) :

$$X_1 = (x_{11}, \ldots, x_{n1})^T, \ldots, X_p = (x_{1p}, \ldots, x_{np})^T$$

- 반응변수(종속변수) : $Y = (y_1, \ldots, y_n)^T$
- 오차항 : $\epsilon_1, \ldots, \epsilon_n$, $(\sim_{i.i.d}\ N(0, \sigma^2))$

중회귀모형(Multiple Linear Regression)

- 자료구조 (설명변수 p개)

$y = X\beta + \varepsilon$

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}}_{\boldsymbol{X}} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}}_{\boldsymbol{\epsilon}}$$

중회귀모형(Multiple Linear Regression)

- 자료구조

rank(X) = p+1

설명변수간 상관관계가 없음

선형적으로 독립

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{n\times 1} = \underbrace{\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}}_{n\times(p+1)} \underbrace{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}}_{(p+1)\times 1} + \underbrace{\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}}_{n\times 1}$$

1 X_1 X_2 X_p

$n \geq p+1$ (∵ 연립방정식 풀이.)

중회귀모형(Multiple Linear Regression)

- 설명변수가 p개인 다중(선형)회귀모형: 행렬형식

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- 회귀모수 : $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$, $(p+1) \times 1$ vector
- 설명변수(독립변수) :

$X : X \times (p+1)$

$$\boldsymbol{X} = (\mathbf{1}, X_1, \ldots, X_p), \; rank(\boldsymbol{X}) = p+1, \; n > p$$

- 반응변수(종속변수) : $\boldsymbol{y} = (y_1, \ldots, y_n)^\top$

$I_n = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 1 \end{pmatrix}$

- 오차항 : $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^\top \sim \underline{N(0, \sigma^2 \mathbf{I}_n)}$

← 항등행렬

다변량정규분포

Least Square Estimation

- 최소제곱추정량 :

$$\left(\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p\right) = \underset{(\beta_0,\ldots,\beta_p)\in\mathbb{R}^{p+1}}{\arg\min} \sum_{i=1}^{n} \{y_i - (\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})\}^2$$

또는

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{p+1}}{\arg\min} \underbrace{||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}||^2}_{S(\boldsymbol{\beta})}$$

Least Square Estimation

$$= \sum \varepsilon_i^2 = \varepsilon^\top \varepsilon = (y - X\beta)^\top (y - X\beta)$$

$$\begin{aligned}
S(\boldsymbol{\beta}) &= ||\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}||^2 \\
&= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\
&= \boldsymbol{y}^\top \boldsymbol{y} - \boldsymbol{\beta}^\top \boldsymbol{X} \boldsymbol{y} - \boldsymbol{y}^\top \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{y}^\top \boldsymbol{y} - 2\boldsymbol{y}^\top \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta}
\end{aligned}$$

- 정규방정식

$$\frac{\partial S}{\partial \boldsymbol{\beta}}\Big|_{\hat{\boldsymbol{\beta}}} = -2\boldsymbol{X}^\top \boldsymbol{y} + 2\boldsymbol{X}^\top \boldsymbol{X}\hat{\boldsymbol{\beta}} = 0 \quad \Rightarrow \boxed{\boldsymbol{X}^\top \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}^\top \boldsymbol{y}}$$

Least Square Estimation

- 최소제곱추정량 :

rank(X) = p+1

⟺ $(X^T X)^{-1}$ 존재함.

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}$$

  단, $(\boldsymbol{X}^\top \boldsymbol{X})^{-1}$ 이 존재하여야 함.

- 추정된 회귀직선 : $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y}$

Example

$$
\boldsymbol{y} = \begin{pmatrix} 9 \\ 20 \\ 22 \\ 15 \\ 17 \\ 30 \\ 18 \\ 25 \\ 10 \\ 20 \end{pmatrix}, \quad
\boldsymbol{X} = \begin{array}{c} \begin{matrix} 1 & X_1 & X_2 \end{matrix} \\ \begin{pmatrix} 1 & 4 & 4 \\ 1 & 8 & 10 \\ 1 & 9 & 8 \\ 1 & 8 & 5 \\ 1 & 8 & 10 \\ 1 & 12 & 15 \\ 1 & 6 & 8 \\ 1 & 10 & 13 \\ 1 & 6 & 5 \\ 1 & 9 & 12 \end{pmatrix} \end{array}, \quad
\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad
\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \end{pmatrix}
$$

Example

$$\boldsymbol{X}^\top \boldsymbol{X} = \begin{pmatrix} 10 & 80 & 90 \\ 80 & 686 & 784 \\ 90 & 784 & 932 \end{pmatrix}, \quad \boldsymbol{X}^\top \boldsymbol{y} = \begin{pmatrix} 186 \\ 1608 \\ 1866 \end{pmatrix}$$

R에서 solve함수?

$$(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = \begin{pmatrix} 1.62902 & -0.26385 & 0.06464 \\ -0.26385 & 0.08047 & -0.04222 \\ 0.06464 & -0.04222 & 0.03034 \end{pmatrix}$$

Example

단순 선형에서

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \underline{x_i}$ x_i 1증가시 $\hat{\beta}_1$ 만큼증가..

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y} = \begin{pmatrix} -0.651 \\ 1.551 \\ 0.760 \end{pmatrix}$$

- 추정된 회귀직선

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 = -0.651 + 1.551 x_1 + 0.760 x_2$$

이론 해석상..

$\hat{\beta}_1$: x_2가 변하지 않을때 x_1이 1만큼변하면 y는 $\hat{\beta}_1$만큼 증가 (1.5)

$\hat{\beta}_2$: x_1가 변하지않을때 x_2가 1단위 증가하면 y는 $\hat{\beta}_2$만큼증가 (0.7)

but 실제로는..!? 상관관계있음

Decomposition of sum of squares

- 제곱합의 분해 : $SST = SSE + SSR$

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$

제곱합의 종류	정의 및 기호	자유도
총제곱합	$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$	$n-1$
잔차제곱합	$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$n-(\underbrace{p+1}_{\text{제약조건}})$
회귀제곱합	$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	p (설명변수 개수)

Decomposition of sum of squares

- 총변동 : SST (총제곱합)

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n(\bar{y})^2 = \boldsymbol{y}^\top \boldsymbol{y} - n(\bar{y})^2$$

또는

$$n\left(\bar{y}\right)^2 = \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 = \boldsymbol{y}^\top \mathbf{1}\mathbf{1}^\top \boldsymbol{y}/n$$

Decomposition of sum of squares

- 회귀모형에 의해 설명되지 않는 변동 : SSE (잔차제곱합)

$$
\begin{aligned}
SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 &= (\boldsymbol{y} - \hat{\boldsymbol{y}})^\top(\boldsymbol{y} - \hat{\boldsymbol{y}}) \\
&= (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \boldsymbol{y}^\top\boldsymbol{y} - 2\hat{\boldsymbol{\beta}}^\top\boldsymbol{X}^\top\boldsymbol{y} + \hat{\boldsymbol{\beta}}^\top\boldsymbol{X}^\top\boldsymbol{X}\hat{\boldsymbol{\beta}} \\
&= \boldsymbol{y}^\top\boldsymbol{y} - \boldsymbol{y}^\top(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} \\
&= \boldsymbol{y}^\top\left[\mathbf{I}_n - \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\right]\boldsymbol{y}
\end{aligned}
$$

또는 $SSE = \boldsymbol{y}^\top\boldsymbol{y} - \hat{\boldsymbol{\beta}}^\top\boldsymbol{X}^\top\boldsymbol{y}$

$$SSE = \sum (y_i - \hat{y}_i)^2 = (\underset{\sim}{y} - \underset{\sim}{\hat{y}})^T (\underset{\sim}{y} - \underset{\sim}{\hat{y}})$$

$$\hat{y} = X\hat{\beta} = X(X^TX)^{-1}X^Ty$$

비중심카이제곱분포!

$$e = (y - \hat{y}) = y - X(X^TX)^{-1}X^Ty$$

$$= \underbrace{(I_n - X(X^TX)^{-1}X^T)}_{A}\, y$$

$$e = Ay$$

$$SSE = (Ay)^T(Ay) = y^TA^TAy$$

$$A^TA = (I_n - X(X^TX)^{-1}X^T)(I_n - X(X^TX)^{-1}X^T)$$

$$= I_n - 2X(X^TX)^{-1}X^T + X(X^TX)^{-1}X^T \cdot X(X^TX)^{-1}X^T$$

$$= I_n - X(X^TX)^{-1}X^T$$

$$A^TA = A \longrightarrow \therefore \underline{SSE = y^TAy}$$

$$\underset{\sim}{y} \sim N(\mu, \sigma^2 I_n)$$

$$\underset{\sim}{\mu} = X\underset{\sim}{\beta}$$

$$\boxed{y \sim N(X\beta, I_n\sigma^2)}$$

Decomposition of sum of squares

$= y^T(\cancel{I_n} - \frac{1}{n}J_n)\, y - y^T(\cancel{I_n} - X(X^TX)^{-1}X^T)\, y$

- 회귀모형에 의해 설명되는 변동 : SSR (회귀제곱합)

$$
\begin{aligned}
SSR &= SST - SSE \\
&= \left[\boldsymbol{y}^\top \boldsymbol{y} - n(\bar{y})^2\right] - \left[\boldsymbol{y}^\top \boldsymbol{y} - \hat{\boldsymbol{\beta}}^\top \boldsymbol{X}^\top \boldsymbol{y}\right] \\
&= \hat{\boldsymbol{\beta}}^\top \boldsymbol{X}^\top \boldsymbol{y} - n(\bar{y})^2 \\
&= \boldsymbol{y}^\top \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y} - n(\bar{y})^2
\end{aligned}
$$

분산분석표

■ 중회귀모형의 분산분석표

χ^2

요인	제곱합	자유도(df)	평균제곱(MS)	F_0	유의확률
회귀	SSR	p	$MSR = \dfrac{SSR}{p}$	$\dfrac{MSR}{MSE}$	$P(F \geq f)$
잔차	SSE	$n-(p+1)$	$MSE = \dfrac{SSE}{n-p-1}$	χ^2 분포를 따른다.	
계	SST	$n-1$			

회귀직선의 유의성 검정

■ 회귀직선의 유의성 검정 (F-test) 설명변수가 다 의미가 없다.

- 가설 : $H_0 : \beta_1 = \cdots = \beta_p = 0 \ vs. \ H_1 : \text{not } H_0$

↳ 최소한 한개는 의미가 있다.

- 검정통계량

$$F = \frac{MSR}{MSE} = \frac{SSR/p}{SSE/(n-(p+1)))} \sim_{H_0} F(p, n-p-1)$$

- 검정통계량의 관측값 : F_0
- 유의수준 α에서의 기각역 : $F_0 > F_\alpha(p, n-p-1)$
- 유의확률 $= P(F > F_0)$

Example

$$
\begin{aligned}
SST &= \mathbf{y}^\top \mathbf{y} - n(\bar{y})^2 \\
&= 9^2 + 20^2 + \cdots + 20^2 - 10(18.6)^2 = 368.40 \\
SSR &= \hat{\boldsymbol{\beta}}^\top \boldsymbol{X}^\top \mathbf{y} - n(\bar{y})^2 \\
&= \begin{pmatrix} -0.651 & 1.551 & 0.760 \end{pmatrix} \begin{pmatrix} 186 \\ 1608 \\ 1866 \end{pmatrix} - 10(18.6)^2 = 332.12 \\
SSE &= SST - SSR = 36.28
\end{aligned}
$$

Example

가설검정 할때 꼭써주기.

- 회귀직선의 유의성 검정 (F-test)
  - 가설 : $H_0 : \beta_1 = \beta_2 = 0 \ vs. \ H_1$: not H_0
  - 분산분석표

요 인	제곱합	자유도	평균제곱	F_0	$F_{0.05}(2,7)$
회 귀	332.12	2	166.06	32.04**	4.74
잔 차	36.28	7	5.18		
계	368.40	9			

  - $F_0 = 32.04 > F_\alpha(p, n-p-1) = F_{0.05}(2,7) = 4.74$ 이므로 귀무가설 기각

회귀모형의 정도

$\left\{\begin{array}{l} MSE \downarrow \\ F_0 \uparrow \\ R^2 \uparrow \end{array}\right.$

- 회귀모형의 적합도
  - MSE 평균오차제곱합: 작을수록 적합
  - $F_0 = MSR/MSE$ $F_0 \uparrow \rightarrow SSR \uparrow$: 모형 적합
  - $R^2 = SSR/SST$ 1에 가까울수록 적합
  - 회귀계수 $\boldsymbol{\beta}^\top = (\beta_0, \beta_1, \beta_2, \cdots, \beta_p)$의 정확한 추정

↳ 오차가 작다.

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$

$s.e.(\hat{\beta}_1) = \dfrac{\sigma^2}{\sqrt{S_{(xx)}}}$

$S_{(xx)} = \sum (x_i - \bar{x})^2$

Estimation of error variance

MSE = $\hat{\sigma}^2$: 불편추정량

HW

SSE = $y^T A y$

정리 5.1 통해서

불편추정량임을

보이기

- 오차분산 (σ^2)의 추정:
  - 잔차 (residual) : $e_i = y_i - \hat{y}_i$
  - 잔차(오차) 제곱합 (residual (or error) sum of squares) :

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$$

  - 평균제곱오차 (mean squared error) : $MSE = \dfrac{SSE}{n-(p+1)}$
  - 오차분산의 추정값 : $\hat{\sigma}^2 = MSE$

$$\boxed{\mathrm{E}(MSE) = \sigma^2}$$

$SSE = \Sigma(y_i - \hat{y})^2$
$(\beta_0, \beta_1, \cdots, \beta_p) =$

$MSE = \frac{SSE}{n-(p+1)}$

- 결정계수 (Coefficient of determination)

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

- 수정된 결정계수 (Adjusted multiple correlation coefficient)

$$R^2_{adj} = 1 - \frac{SSE/(n-p-1)}{SST/(n-1)}$$

자유도로 나누어줌 → 평균생각

ex) $\frac{(\beta_1)}{M_1}$ $\frac{(\beta_1, \beta_2)}{M_2}$

$(\hat{\beta}_0, \hat{\beta}_1) = \text{minimize} \ (y_i - \hat{y}_i)^2$

$\Rightarrow \underset{\beta_0, \beta_1}{\text{minimize}} \sum (y_i - \beta_0 - \beta_1 x_1)^2 = S_1$ SSE_1

$(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) \Rightarrow \underset{\beta_0, \beta_1, \beta_2}{\text{minimize}} \sum (y_i - \beta_0 - \beta_1 x_1 \underbrace{- \beta_2 x_2}_{\text{설명변수 하나 더 추가..}})^2 = S_2$ SSE_2

S_1은.. S_2에서 $\beta_2 = 0$이라고 생각한 case

$\Rightarrow \boxed{S_1 > S_2}$ 무조건..

설명변수가 추가되면

SSE는 작아지는

성질이 있다.

$SSE_1 > SSE_2$

$\Rightarrow -SSE_1 < -SSE_2$

$\Rightarrow 1 - \frac{SSE_1}{SST} < 1 - \frac{SSE_2}{SST}$

$\Rightarrow \boxed{R_1^2 < R_2^2}$

LSE

$\mathrm{Var}(Ay) = A\,\mathrm{Var}(y)\,A^T$

- $\hat{\beta}$

$$
\begin{aligned}
\mathrm{E}(\hat{\boldsymbol{\beta}}) &= \mathrm{E}\left[\underbrace{(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top}_{A:\text{상수}}\mathbf{y}\right] \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathrm{E}(\mathbf{y}) = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta} \quad \text{불편추정량} \\
\mathrm{Var}(\hat{\boldsymbol{\beta}}) &= \mathrm{Var}\left[(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \mathbf{y}\right] \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \left[\mathrm{Var}(\mathbf{y})\right] \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \left(I\sigma^2\right) \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\sigma^2
\end{aligned}
$$

LSE

- $\hat{\boldsymbol{\beta}}$

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \mathrm{Var}\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_p \end{pmatrix} = \begin{pmatrix} \mathrm{Var}(\hat{\beta}_0) & \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_2) & \cdots & \mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_p) \\ & \mathrm{Var}(\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \cdots & \mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_p) \\ & & \mathrm{Var}(\hat{\beta}_2) & \cdots & \mathrm{Cov}(\hat{\beta}_2, \hat{\beta}_p) \\ & & & & \vdots \\ & & & & \mathrm{Var}(\hat{\beta}_p) \end{pmatrix}$$

$$= (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\sigma^2 := (c_{ij})\sigma^2$$

평균반응 및 개별 y 의 추정량

$E(\hat{y}) = x^T E(\hat{\beta}) = x^T \beta$

$(1 \quad 1 \quad x_{01} \cdots x_{0p}) \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}$

- $\hat{y} = E(\widehat{y|\boldsymbol{X} = \boldsymbol{x}}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta} x_p = \boldsymbol{x}^\top \hat{\boldsymbol{\beta}}$

하나의관측값..

$$\mathrm{Var}(\hat{y}) = \mathrm{Var}(\boldsymbol{x}^\top \hat{\boldsymbol{\beta}}) = \boldsymbol{x}^\top \mathrm{Var}(\hat{\boldsymbol{\beta}}) \boldsymbol{x}$$

$$= \underbrace{\underset{(1\times(p+1))}{\boldsymbol{x}^\top} \underset{(p+1)\times(p+1)}{(\boldsymbol{X}^\top \boldsymbol{X})^{-1}} \underset{(p+1)\times 1}{\boldsymbol{x}}}_{1\times 1} \sigma^2$$

기댓값 : 똑같음

- 개별관측값 $\hat{y}_s$

$$\mathrm{Var}(\hat{y}_s) = \left[1 + \boldsymbol{x}^\top (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{x} \right] \sigma^2$$

Example

- $\mathrm{Var}(\hat{\boldsymbol{\beta}})$

$$\mathrm{Var}\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\end{pmatrix} = (X^\top X)^{-1}\sigma^2 = \begin{pmatrix}1.63 & -0.26 & 0.06\\ -0.26 & 0.08 & -0.04\\ 0.06 & -0.04 & 0.03\end{pmatrix}\sigma^2$$

$$= \begin{pmatrix}\mathrm{Var}(\hat{\beta}_0) & \mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_2)\\ \mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_1) & \mathrm{Var}(\hat{\beta}_1) & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2)\\ \mathrm{Cov}(\hat{\beta}_0,\hat{\beta}_2) & \mathrm{Cov}(\hat{\beta}_1,\hat{\beta}_2) & \mathrm{Var}(\hat{\beta}_2)\end{pmatrix}$$

Example

$$\begin{array}{ccc} & x_1 & x_2 \end{array}$$
$$x = (1, 10, 10)$$

- $\mathrm{Var}(\hat{y})$

$$\begin{aligned} \mathrm{Var}(\hat{y}) &= \boldsymbol{x}^\top (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{x}\ \sigma^2 \\ &= \begin{pmatrix} 1 & 10 & 10 \end{pmatrix} \begin{pmatrix} 1.63 & -0.26 & 0.06 \\ -0.26 & 0.08 & -0.04 \\ 0.06 & -0.04 & 0.03 \end{pmatrix} \begin{pmatrix} 1 \\ 10 \\ 10 \end{pmatrix} \sigma^2 \\ &= 0.28182\sigma^2 \end{aligned}$$

제곱합의 분포

- 중회귀모형

$$\boldsymbol{y} = \overbrace{X\boldsymbol{\beta}}^{\text{y의 기댓값}} + \boldsymbol{\epsilon},\ \ \boldsymbol{\epsilon} \sim N(\boldsymbol{0}, I_n\sigma^2)$$

$$SST = \boldsymbol{y}^\top \boldsymbol{y} - n(\bar{y})^2 = \boldsymbol{y}^\top \left[I_n - \frac{J_n}{n} \right] \boldsymbol{y}$$

$$SSR = \boldsymbol{y}^\top \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y} - n(\bar{y})^2$$

$$= \boldsymbol{y}^\top \left[\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top - \frac{J_n}{n} \right] \boldsymbol{y}$$

$$SSE = \boldsymbol{y}^\top \left[I_n - \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \right] \boldsymbol{y}$$

제곱합의 분포

$$\frac{SST}{\sigma^2} \sim \chi^2 \left\{ n-1,\ \boldsymbol{\beta}^\top \boldsymbol{X}^\top \left(I_n - \frac{J_n}{n} \right) \boldsymbol{X}\boldsymbol{\beta}/2\sigma^2 \right\}$$

$$\frac{SSR}{\sigma^2} \sim \chi^2 \left\{ p,\ \boldsymbol{\beta}^\top \boldsymbol{X}^\top \left(I_n - \frac{J_n}{n} \right) \boldsymbol{X}\boldsymbol{\beta}/2\sigma^2 \right\}$$

$$\frac{SSE}{\sigma^2} \sim \chi^2 (n - p - 1)$$

제곱합의 분포

$$F_0 = \frac{MSR}{MSE} = \frac{\left(\frac{SSR}{\sigma^2}\right)/k}{\left(\frac{SSE}{\sigma^2}\right)/(n-k-1)} \sim F'(k,\ n-k-1,\ \lambda)$$

where

$$\lambda = \boldsymbol{\beta}^\top \boldsymbol{X}^\top \left(I_n - \frac{J_n}{n}\right) \boldsymbol{X}\boldsymbol{\beta}/2\sigma^2$$