Ordinary Least Squares Estimation

1. 단순 선형 회귀

단순 선형 회귀의 모델은 다음과 같이 정의됩니다:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$

목표는 SSE 최소화하는 β_0와 β_1를 찾는 것입니다:

$$SSE = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

2. SSE를 편미분하여 방정식 도출

(a) β_0에 대한 미분

$$\frac{\partial SSE}{\partial \beta_0} = \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

$$\frac{\partial SSE}{\partial \beta_0} = -2 \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i)) = 0$$

(b) β_1에 대한 미분

$$\frac{\partial SSE}{\partial \beta_1} = \frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2$$

$$\frac{\partial SSE}{\partial \beta_1} = -2 \sum_{i=1}^{n} x_i (y_i - (\beta_0 + \beta_1 x_i)) = 0$$

3. β_0와 β_1 구하기

(a) β_0를 β_1로 표현

첫 번째 방정식을 β_0에 대해 정리합니다:

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i}{n} - \beta_1 \frac{\sum_{i=1}^{n} x_i}{n}$$

즉:

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

(b) β_1 계산

위 식을 두 번째 방정식에 대입합니다:

$$\sum_{i=1}^{n} x_i y_i = (\bar{y} - \beta_1 \bar{x}) \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$

(c) 식 정리

$(\bar{y} - \beta_1 \bar{x}) \sum_{i=1}^{n} x_i$를 분배하여 정리하면:

$$\sum_{i=1}^{n} x_i y_i = \bar{y} \sum_{i=1}^{n} x_i - \beta_1 \bar{x} \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$

$$\sum_{i=1}^{n} x_i y_i - \bar{y} \sum_{i=1}^{n} x_i = \beta_1 \left(\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i \right)$$

4. 식 정리

두 번째 방정식에서 최종적으로 β_1를 구하기 위해 식을 정리합니다. 방정식은 다음과 같습니다:

$$\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - \bar{y} \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}$$

이를 다시 쓰면:

$$\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - \bar{y} \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i\right)^2}$$

위 식은 x와 y의 평균값, 분산, 공분산을 활용하여 간소화될 수 있습니다.

5. 공분산과 분산으로 식 정리

(a) 공분산의 정의

공분산 $\mathrm{Cov}(x, y)$는 다음과 같이 정의됩니다:

$$\mathrm{Cov}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

위 식을 풀어서 계산하면 다음과 같은 형태로 바꿔쓸 수 있습니다:

$$\mathrm{Cov}(x, y) = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \left(\sum_{i=1}^{n} x_i\right) \left(\sum_{i=1}^{n} y_i\right)}{n}$$

(b) 분산의 정의

분산 $\mathrm{Var}(x)$는 다음과 같이 정의됩니다:

$$\mathrm{Var}(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

이를 계산하면 다음과 같이 쓸 수 있습니다:

$$\mathrm{Var}(x) = \frac{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

(c) 공분산과 분산을 이용한 β_1 계산

공분산과 분산을 위 정의에 따라 계산하면 β_1는 다음과 같이 표현됩니다:

$$\beta_1 = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)}$$

따라서:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

6. 결론

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

β_1를 구한 후, $\beta_0 = \bar{y} - \beta_1\bar{x}$를 이용하여 β_0를 계산할 수 있습니다.

7. 다중 선형 회귀

다중 선형 회귀 모델은 다음과 같이 정의됩니다:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 I)$$

여기서:

- $\mathbf{y}$: $n \times 1$ 벡터
- $\mathbf{X}$: $n \times p$ 행렬 (p는 설명 변수의 개수)
- $\boldsymbol{\beta}$: $p \times 1$ 벡터
- $\boldsymbol{\epsilon}$: $n \times 1$ 벡터

(a) 오차 제곱합 (SSE) 정의

최소 제곱법은 SSE를 최소화하는 $\boldsymbol{\beta}$를 찾는 것입니다:

$$SSE = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2$$

이를 행렬 곱으로 표현하면:

$$SSE = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

(b) SSE를 $\boldsymbol{\beta}$에 대해 미분

SSE를 최소화하기 위해 $\boldsymbol{\beta}$에 대해 편미분합니다. 미분 과정은 다음과 같습니다:

$$\frac{\partial SSE}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \left[\mathbf{y}^\top \mathbf{y} - 2\mathbf{y}^\top \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} \right]$$

미분 결과는:

$$\frac{\partial SSE}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^\top \mathbf{y} + 2\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta}$$

$\frac{\partial SSE}{\partial \boldsymbol{\beta}} = 0$으로 두면:

$$-2\mathbf{X}^\top \mathbf{y} + 2\mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} = 0$$

이를 정리하면:

$$\mathbf{X}^\top \mathbf{y} = \mathbf{X}^\top \mathbf{X}\boldsymbol{\beta}$$

(c) $\boldsymbol{\beta}$ 계산 (Normal Equation)

위 식을 정리하여 $\boldsymbol{\beta}$를 구합니다:

$$\boldsymbol{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$