Ordinary Least Squares Estimation

JAEYOUNG LEE

January 2025

1 오차 최소화

우리는 데이터의 실제 값인 y를 잘 예측하는 모델을 찾고 싶어한다. 그럼 실제 값 y_i와 예측 값 $\hat{y}_i$의 차이를 최대한 줄이는 선을 찾는 게 목표. 우리는 오차 최소화 방법을 Least Squares Method를 통해 푼다. 이는 $[y_i$ - $\hat{y}_i]$의 제곱을 최소화 시키는 것이다.

2 Simple Linear Regression

2.1 SSE

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

여기서, β_0는 절편 그리고 β_1는 기울기이다. 우리는 이 두 변수를 조절해서 SSE를 제일 작게 만들어야 한다!

2.2 미분으로 최적화

오차 제곱합(SSE)을 최소화하려면 변수를 하나씩 미분해서 기울기가 0이 되는 지점을 찾는다.

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

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

이 두 식을 풀면:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \quad \beta_0 = \bar{y} - \beta_1 \bar{x}$$

3 Multiple Linear Regression

3.1 행렬 형태로 일반화

다중 선형 회귀의 SSE를 구하는 아이디어는 단순 선형 회귀와 일치한다. 하지만 이를 표현하는 방식으로 matrix와 vector를 활용한다고 생각하면 편하다.

$$\mathbf{y} = \mathbf{X}\beta$$

- $\mathbf{X}$: 독립 변수 행렬 $(n \times p)$
- β: 회귀 계수 벡터 $(p \times 1)$
- $\mathbf{y}$: 종속 변수 벡터 $(n \times 1)$

3.2 SSE

오차 제곱합을 행렬 형태로 정의하면:

$$J(\beta) = \|\mathbf{y} - \mathbf{X}\beta\|^2$$

3.3 미분으로 최적화

오차를 최소화하려고 $J(\beta)$를 β로 미분한다:

$$\frac{\partial J}{\partial \beta} = -2\mathbf{X}^\top(\mathbf{y} - \mathbf{X}\beta) = 0$$

이걸 정리하면:

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

양변을 $\mathbf{X}^\top\mathbf{X}$로 나누면:

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

위와 같이 우리는 normal equation을 구할 수 있다.