Derivation of Least Squares Estimator in Simple and Multiple Linear Regression

1. Simple Linear Regression

2020118029 Hyeongjin Kim

Model Setup

The model for Simple Linear Regression is given as:

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

where y_i is the dependent variable, x_i is the independent variable, β_0 is the intercept, β_1 is the slope (regression coefficient), and ϵ_i represents the error term.

Objective

Using the method of Least Squares, we aim to find $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize the Sum of Squared Errors (SSE). The SSE is defined as:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_i)\right)^2.$$

Minimization Process

1. Compute the partial derivatives of $S(\beta_0, \beta_1)$ with respect to β_0 and β_1:

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

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

2. Set the derivatives to zero to find the critical points:

$$\sum_{i=1}^{n} \left(y_i - (\beta_0 + \beta_1 x_i)\right) = 0, \quad \sum_{i=1}^{n} x_i \left(y_i - (\beta_0 + \beta_1 x_i)\right) = 0.$$

3. Solve the above equations for $\hat{\beta}_0$ and $\hat{\beta}_1$: - Slope ($\hat{\beta}_1$):

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

where $\bar{x}$ and $\bar{y}$ are the means of x and y, respectively. - Intercept ($\hat{\beta}_0$):

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

—

2. Multiple Linear Regression

Model Setup

The model for Multiple Linear Regression is:

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

where:

- $\mathbf{y} \in \mathbb{R}^n$: Vector of dependent variables,
- $\mathbf{X} \in \mathbb{R}^{n \times p}$: Matrix of independent variables (n samples and p predictors),
- $\boldsymbol{\beta} \in \mathbb{R}^p$: Vector of regression coefficients,
- $\boldsymbol{\epsilon} \in \mathbb{R}^n$: Vector of errors.

Objective

The goal is to minimize the SSE:

$$S(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Minimization Process

1. Compute the gradient of $S(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$:

$$\frac{\partial S}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^\top (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

2. Set the gradient to zero:

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

3. Solve for $\hat{\boldsymbol{\beta}}$:

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

—

Conclusion

- For **Simple Linear Regression**, the coefficients $\hat{\beta}_1$ and $\hat{\beta}_0$ are derived using the means, variances, and covariances of x and y. - For **Multiple Linear Regression**, the coefficients are obtained through matrix operations, provided that $(\mathbf{X}^\top\mathbf{X})$ is invertible. - Both methods rely on minimizing the sum of squared residuals to find the best-fitting regression line or hyperplane.