Ordinary Least Squares Estimation

Minji Jung

2025.1.19

Introduction

Ordinary Least Squares (OLS) estimation is a method for estimating the parameters of a linear regression model by minimizing the sum of the squared residuals. This document explains the derivation of the Least Squares Estimator (LSE) for both Simple Linear Regression (SLR) and Multiple Linear Regression (MLR).

1. Simple Linear Regression (SLR)

The Simple Linear Regression model is expressed as:

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

where:

- Y is the response variable,
- X is the predictor variable,
- β_0 is the intercept,
- β_1 is the slope coefficient,
- ϵ is the error term.

Objective

The goal is to minimize the sum of squared residuals:

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

Derivation of Normal Equations

1. Take the partial derivative with respect to β_0:

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

Set this to zero:

$$\sum_{i=1}^{n} Y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} X_i.$$

2. Take the partial derivative with respect to β_1:

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

Set this to zero:

$$\sum_{i=1}^{n} X_i Y_i = \beta_0 \sum_{i=1}^{n} X_i + \beta_1 \sum_{i=1}^{n} X_i^2.$$

3. Using the two equations derived, solve for β_0 and β_1:

$$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}, \quad \beta_0 = \bar{Y} - \beta_1 \bar{X}.$$

Thus, the Least Squares Estimators for β_0 and β_1 are explicitly given by these equations.

2. Multiple Linear Regression (MLR)

The Multiple Linear Regression model is expressed as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where:

- $\mathbf{Y}$ is an $n \times 1$ vector of responses,
- $\mathbf{X}$ is an $n \times p$ matrix of predictors (including a column of ones for the intercept),
- $\boldsymbol{\beta}$ is a $p \times 1$ vector of coefficients,
- $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of errors.

Objective

The goal is to minimize the sum of squared residuals:

$$S(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

Derivation of Normal Equation

1. Expand the objective function:

$$S(\boldsymbol{\beta}) = \mathbf{Y}^T\mathbf{Y} - 2\mathbf{Y}^T\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}.$$

2. Take the derivative with respect to $\boldsymbol{\beta}$:

$$\frac{\partial S}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^T\mathbf{Y} + 2\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}.$$

3. Set the derivative to zero:

$$\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^T\mathbf{Y}.$$

4. Solve for $\boldsymbol{\beta}$:

$$\boldsymbol{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}.$$

Thus, the Least Squares Estimator for $\boldsymbol{\beta}$ is explicitly given by this equation.

Conclusion

The OLS method provides a framework to estimate regression coefficients by minimizing the sum of squared residuals. The derivation involves setting up the objective function, differentiating, and solving the resulting normal equations. These principles are fundamental in regression analysis and have wide applications in predictive modeling.