Mathematical Solution to Linear Regression

1. Simple Linear Regression

Model

The model equation for simple linear regression is:

$$y = \theta_0 + \theta_1 x$$

where:

- y: Dependent variable (output),
- x: Independent variable (input),
- θ_0: Intercept,
- θ_1: Slope.

Objective

The objective is to minimize the **Sum of Squared Errors (SSE)**:

$$SSE = \sum_{i=1}^{n} \left(y_i - (\theta_0 + \theta_1 x_i)\right)^2$$

Solution

The formulas for the coefficients are:

$$\theta_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$
$$\theta_0 = \bar{y} - \theta_1 \bar{x}$$

where $\bar{x} = \frac{1}{n} \sum x_i$ and $\bar{y} = \frac{1}{n} \sum y_i$ are the means of x and y, respectively.

2. Multiple Linear Regression

Model

The model equation for multiple linear regression is:

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_p x_p$$

or equivalently, in matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta}$$

where:

- $\mathbf{y}$: $n \times 1$ vector of observed values,
- $\mathbf{X}$: $n \times (p+1)$ matrix of predictors (first column is ones for θ_0),
- $\boldsymbol{\theta}$: $(p+1) \times 1$ vector of coefficients.

Objective

The objective is to minimize the **SSE**:

$$SSE = (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\theta})$$

Solution

The formula for the coefficients is:

$$\boldsymbol{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

3. Summary Table

The following table summarizes the key formulas for simple and multiple linear regression:

Aspect	**Simple Linear Regression**	**Multiple Linear Regression**
Model Equation	$y = \theta_0 + \theta_1 x$	$y = \theta_0 + \theta_1 x_1 + \cdots + \theta_p x_p$
Objective	$\min \sum_{i=1}^{n} (y_i - (\theta_0 + \theta_1 x_i))^2$	$\min(\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\theta})$
Solution for Coefficients	$\theta_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$ $\theta_0 = \bar{y} - \theta_1 \bar{x}$	$\boldsymbol{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$