Linear Regression Model

Artur Zhdan

February 2025

1 Overview

We have:

- A target variable y (e.g., column 10 of our data).
- Two fixed predictors: FTP $= X_1$ and WE $= X_9$.
- Candidate predictors X_j for $j \in \{2, 3, 4, 5, 6, 7, 8\}$.

We want to find which X_j provides the best model when added to the two fixed predictors.

2 Design Matrix

For each candidate X_j, we form a design matrix $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} 1 & X_1^{(1)} & X_9^{(1)} & X_j^{(1)} \\ 1 & X_1^{(2)} & X_9^{(2)} & X_j^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_1^{(n)} & X_9^{(n)} & X_j^{(n)} \end{bmatrix}.$$

The first column of 1's is for the intercept term.

3 Parameter Estimation

We estimate the regression coefficients $\boldsymbol{\beta}$ using the Normal Equation:

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

The predictions then become:

$$\hat{\mathbf{y}} = \mathbf{X} \, \boldsymbol{\beta}.$$

4 Model Comparison

To measure accuracy, we use the Mean Squared Error (MSE):

$$\mathrm{MSE}_j = \frac{1}{n}\sum_{i=1}^{n}\bigl(y_i - \hat{y}_i\bigr)^2.$$

We compute MSE_j for each candidate X_j. Then we choose the best feature $X_{\hat{j}}$ by

$$\hat{j} = \arg\min_j \ \mathrm{MSE}_j.$$

This feature produces the lowest MSE and thus gives the "best" linear fit with our target variable.