To derive the formula for the coefficients in **multiple linear regression** in matrix form, let's begin with the general framework of the problem.

1. Multiple Linear Regression Model

The general form of the multiple regression model can be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- $\mathbf{y}$ is an $n \times 1$ vector of the response variable.
- $\mathbf{X}$ is an $n \times (p+1)$ matrix of the features (covariates) with an intercept term (where p is the number of predictors).
- $\boldsymbol{\beta}$ is a $(p+1) \times 1$ vector of the coefficients $\beta_0, \beta_1, \ldots, \beta_p$.
- $\boldsymbol{\epsilon}$ is an $n \times 1$ vector of errors (residuals).

2. Least Squares Estimation

The goal is to minimize the **Residual Sum of Squares** (RSS), which is the sum of squared residuals:

$$RSS = \sum_{i=1}^{n} e_i^2 = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2$$