MULTIPLE LINEAR REGRESSION

A Short Theoretical Tutorial

@S. S. Roy , Date: 27 July 2026

1. What It Is

Multiple linear regression extends the straight-line model to **two or more predictor variables**. A single response Y is expressed as a linear combination of k predictors X_1, X_2, …, X_k plus a random error:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

β_0 is the intercept; each β_j is a **partial regression coefficient**, giving the change in Y for a one-unit change in X_j when all other predictors are held fixed. ε is the random error, assumed to have mean zero, constant variance, and independence — and normality when tests are required. "Linear" refers to linearity in the parameters β, not in the predictors.

2. Matrix Form

With n observations the model is written compactly using matrices:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon$$

where Y is the n×1 vector of responses, X is the n×(k+1) design matrix (a leading column of 1s for the intercept, then one column per predictor), β is the (k+1)×1 vector of unknown parameters, and ε is the n×1 error vector.

3. Least-Squares Estimation

The parameters are estimated by least squares, minimising the sum of squared residuals. The solution is the **normal-equations** result in matrix form:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

This single formula generalises the straight-line estimates b_1 = S_XY/S_XX and $b_0 = \bar{Y} - b_1\bar{X}$. It requires X′X to be **nonsingular** (invertible); X′X is always square and symmetric. Once b is obtained, the fitted values are $\hat{Y}$ = Xb and the residuals are $Y - \hat{Y}$.

4. Key Points

- Each coefficient measures the effect of its predictor **after adjusting for all the others** — unlike simple regression, where only one predictor is present.
- Degrees of freedom for the error are **n − (k + 1)**, since k + 1 parameters (the intercept and k slopes) are estimated from the data.
- Individual coefficients are tested with **t-tests**; the overall usefulness of the model is tested with an **F-test**.
- Strong correlation among predictors (**multicollinearity**) makes X′X nearly singular and the estimates unstable.

In one line: multiple linear regression fits Y = Xβ + ε by least squares, giving $b = (X'X)^{-1}X'Y$ — the same principle as the straight line, expressed in matrix form for many predictors.