

In the process of minimizing the Residual Sum of Squares (RSS) in multiple linear regression, we differentiate the RSS expression with respect to the coefficient vector β . Here's the detailed step-by-step explanation of how we perform the differentiation:

Step 1: Start with the RSS Expression

We have the Residual Sum of Squares (RSS) as:

$$RSS = \|\mathbf{y} - \mathbf{X}\beta\|^2$$

Expanding this expression in terms of the matrix operations:

$$RSS = (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta)$$

Step 2: Expanding the RSS Expression

Now, we expand this expression. Using the distributive property of matrix multiplication:

$$RSS = \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\beta + \beta^T \mathbf{X}^T \mathbf{X}\beta$$

- $\mathbf{y}^T \mathbf{y}$ is a scalar (the dot product of \mathbf{y} with itself, i.e., the squared norm of the response vector \mathbf{y}).
- $-2\mathbf{y}^T \mathbf{X}\beta$ is a scalar as well.
- $\beta^T \mathbf{X}^T \mathbf{X}\beta$ is a quadratic form and is also a scalar.