Consider a dataset $D = (x_i, y_i)^N$ of N data points, where $x_i = (x_{i1}, x_{i2}, \cdots, x_{iM})$ is a feature vector with M features, and y_i is the target, i.e., the response, variable. Let x_j denote the jth variable in feature space. A **typical linear regression model** can then be expressed mathematically as:

$$\underbrace{y}_{\text{target}} = \overbrace{\beta_0}^{\text{slope}} + \beta_1 \overbrace{x_1}^{\text{feature vector}} + \beta_2 x_2 + \cdots + \beta_M x_M$$

This model assumes that the relationships between the target variable y_i and features x_j are linear and can be captured in slope terms $\beta_1, \beta_2, \ldots, \beta_M$.