

Multivariate Linear Regression

Linear Regression is a statistical learning algorithm used within a supervised setting which can be used for predicting a continuous numerical value. Linear regression can be split into two different types: **univariate** or simple regression and **multivariate** regression. What differentiates these two variations of the algorithm are whether it's fit to multiple independent variables for prediction. **Univariate** models use only one independent variable as a predictor, while **multivariate** models use multiple to estimate a prediction. This algorithm predates its modern machine learning application significantly, as it is considered one of statistics most fundamental mechanisms.

In a nutshell, linear regression fits a line or hyperplane to the data. This regression line represents the best fit for measuring a constant rate of change between a dependent variable or **response variable** and one or multiple independent variables or **features**. The slope of the regression line is dependent on coefficients or weights which are independently assigned to the features within the model. For example, if our objective was to predict the temperature on a given day, one feature which may intuitively seem like a candidate for a high weight is the time of the year (season). As we can assume that the season most likely has a high correlation with the response variable or in other words is an important factor in deriving its value mathematically.

Equations

The equation of simple regression is certainly one most people have encountered in their grade school math curriculum, as it is simply the equation for slope, or change of a dependent variable as a ratio of change in an independent variable. The equation is as follows:

$$y = \beta_0 + \beta_1 x$$

Linear Regression's Objective Function

Pros and Cons of Linear Regression

Applying Linear Regression