0.0.1 Theory

The basic idea of linear regression is to find the set of coefficients of that satisfy

$$y = X\beta$$

, where X is the data matrix. It's unlikely that for the given values of X, we will find a set of coefficients that exactly satisfy the equation; an error term gets added if there is an inexact specification or measurement error. Therefore, the equation becomes $y = X\beta + \varepsilon$, where ε is assumed to be normally distributed and independent of the X values. Geometrically, this means that the error terms are perpendicular to X.

In order to find the set of betas that map the X values to y, we minimize the error term. This is done by minimizing the residual sum of squares. This problem can be solved analytically, with the solution being .