Regression

Simple Linear Regression, Multiple Linear Regression, Least Square Gradient Method

by Dr. Tumpa Banerjee

Introduction

- The objective of regression is to determine the value of one or more continuous target variables Y based on the d-dimensional input vector X. Given a set of N observations $\{X_n\}$, together with the target values $\{Y_n\}$.
- The objective is to predict the value of y for an input x.
- A linear regression model assumes that the regression function $E(Y|X)$ is linear in the inputs $X_1, X_2, \dots, X_p$.

Introduction

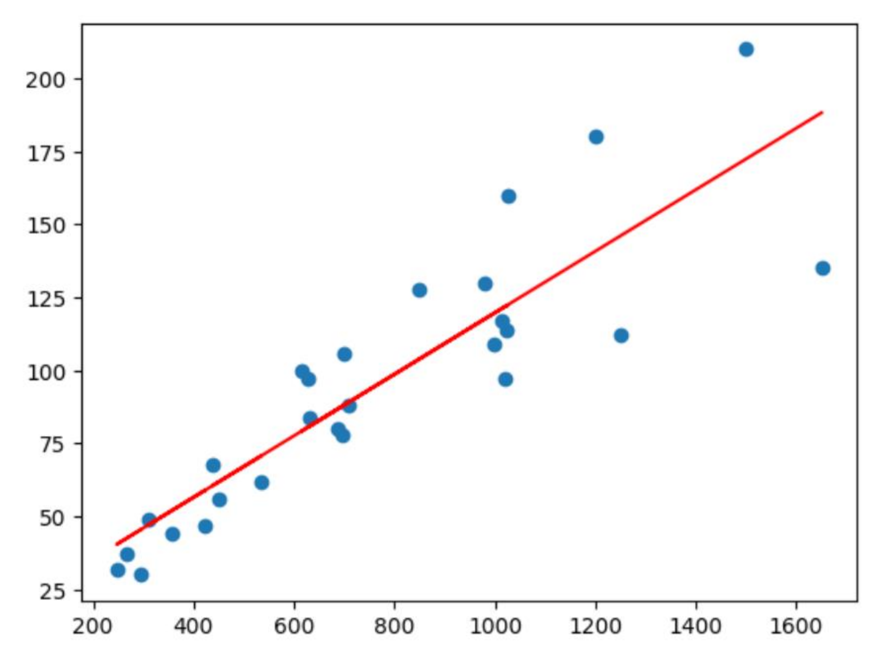
- This scenario can be represented as $Y = f(X) + \epsilon$, i.e. the target variable is a function of input variables X.
- Uncertainty always exists in the process of determining the value of the target variable and this motivates to represent this as a probability distribution $P(Y|X)$.

Linear Regression Model

- The linear regression model has the form $Y = \beta_0 + \sum_{j=1}^{d} \beta_j X_j$
- Here β_j are unknown parameter and X_j can come come from different sources:
- quantitative inputs
- transformation of quantitative inputs, such as log, square root or square
- basic expansion
- numeric or dummy coding of the qualitative inputs

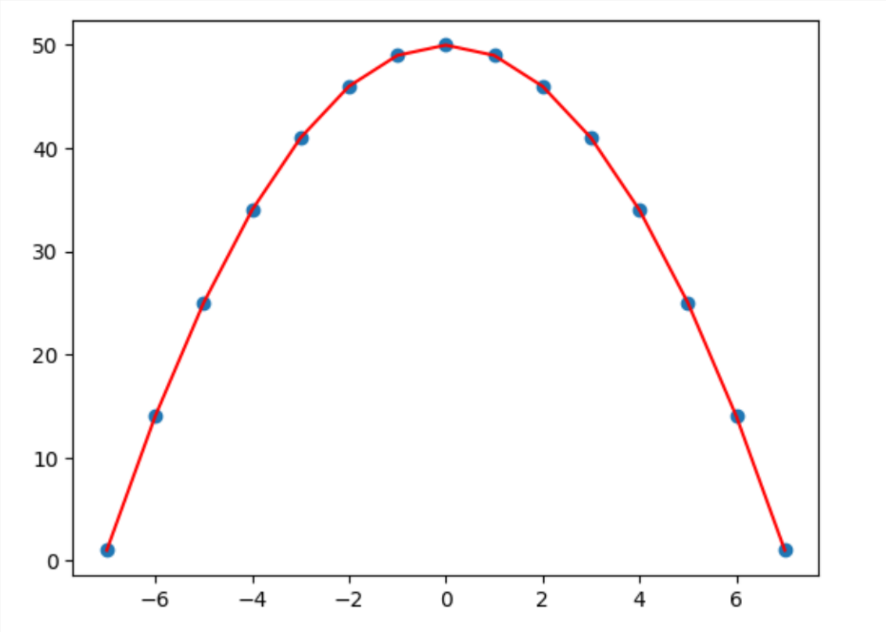
Least Square Method

- The most popular estimation method is least squares, in which we pick the coefficients $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ to minimize the residual sum of squares.
- $RSS = \sum_{i=1}^{N} \left(y_j - f(x_j)\right)^2$
- $RSS = \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{d} \beta_j x_{ij}\right)^2$



Least Square Method

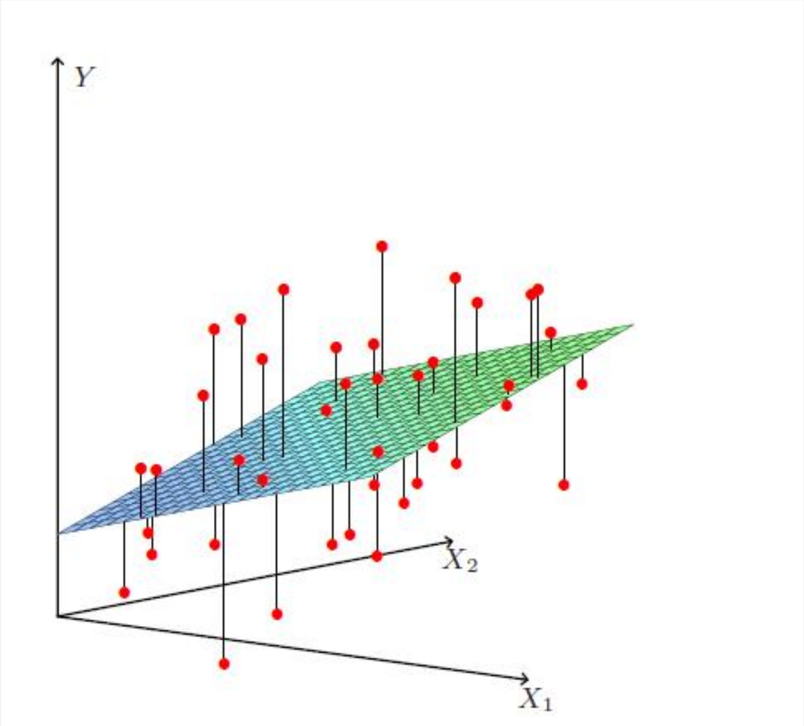
- $\mathcal{L} = (Y - X\beta)^T (Y - X\beta)$
- $\frac{\partial \mathcal{L}}{\partial \beta} = -2X^T (Y - X\beta)$
- $\frac{\partial^2 \mathcal{L}}{\partial \beta^2} = 2X^T X$
- Set the first order derivative to zero $\frac{\partial \mathcal{L}}{\partial \beta} = 0$
- $X^T (Y - X\beta) = 0 \Rightarrow \hat{\beta} = (X^T X) X^T Y$



Multiple Regression

- The linear model with (d>1) inputs is called the multiple linear regression model. The least squares estimate for linear regression are best understood in terms of univariate model.
- Suppose we have single variable input with no intercept, that is
- $Y = X\beta + \epsilon$
- The least square estimate and residuals are
- $\hat{\beta} = \frac{\sum x_i y_i}{x_i^2}$
- $r = y_i - x_i\hat{\beta}$

- In vector notation $< x, y > = \sum x_i y_i = x^T y$
- Then the parameter and residual can be written as
- $\hat{\beta} = \frac{<x,y>}{<x,x>}$
- $r = y_i - x\hat{\beta}$



Linear Classification