ECS 171: Machine Learning

Summer 2023

Edwin Solares
easolares@ucdavis.edu

Optimizing Linear Regression

Minimize on RSS

Find an approximate ***y***
We will call ***ŷ***.
Model maps ***f(X)* = *ŷ*** → ***y***
For all seen ***X*** and unseen ***X***

Residual Sum of Squares

$$RSS = \epsilon^T \epsilon = \sum_{i=1}^{m} (\epsilon_i)^2 = \sum_{i=1}^{m} (y_i - w^T x_i)^2$$

Observations - Predictions

$$RSS_t := min(RSS_{t-1}), \, where \, w \, is \, changed \, at \, each \, time \, step \, t$$

How to minimize the RSS?

1. **Ordinary Least Squares (OLS) : Method 1 – Analytical approach**
2. Gradient Descent (GD) : Method 2 – Numerical approach

Ordinary Least Squares (OLS)

$$RSS_t := \ min(RSS_{t-1}), \ where\ w\ is\ changed\ at\ each\ time\ step\ t$$

$$RSS \ = \ \sum_{i=1}^{m} (\ y_i \ - \ w^T x_i)^2$$

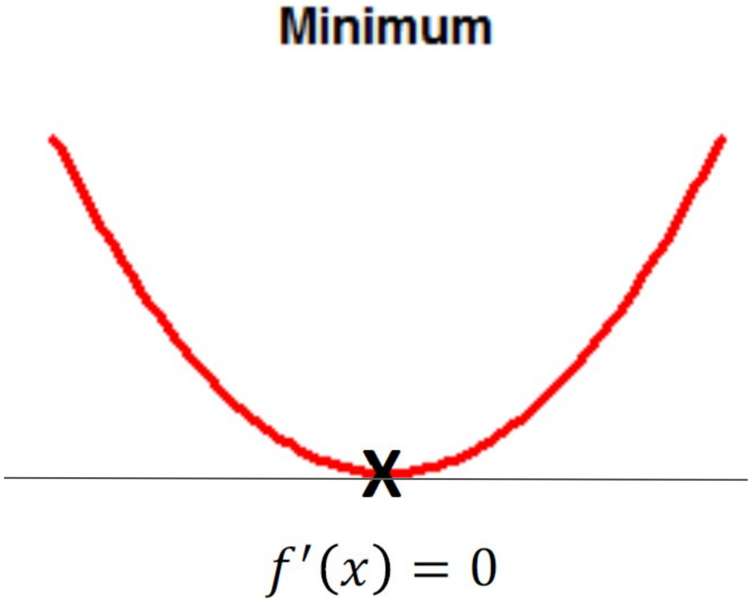
Observations - Predictions

We want to find some change in ***w*** where (Observations - Predictions)^2 = 0
Rate of change = 0

$$find\ \frac{\delta}{\delta w} RSS \ = \ 0$$

OLS Method

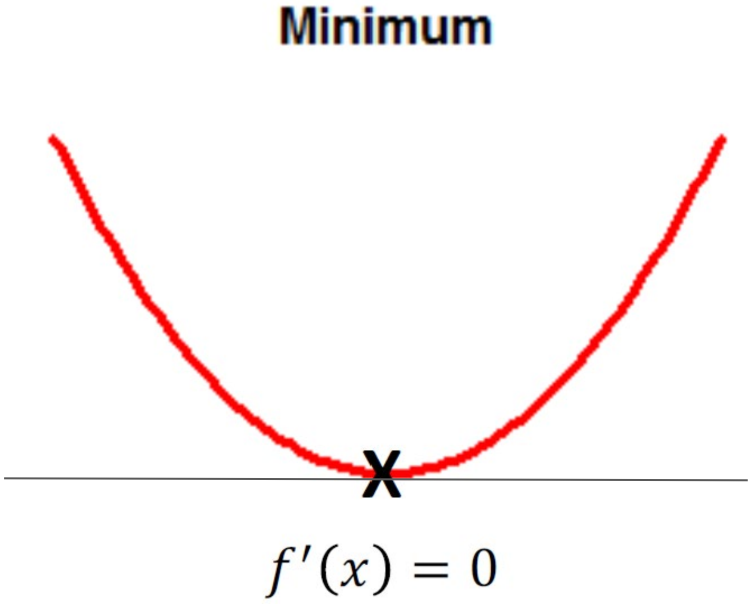
Recall for some function ***f(x)*** we can find **rate of change** by taking its **derivative** at some point x



OLS Method

Recall for some function ***f(x)*** we can find **rate of change** by taking its **derivative** at some point x

$$\hat{w} = \frac{\sum_{i=1}^{m} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{m} (x_i - \bar{x})^2}$$



OLS Method

$$\hat{w} = \frac{\sum_{i=1}^{m} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{m} (x_i - \bar{x})^2}$$

$$\sigma_{xy} = \sum_{i=1}^{m} (y_i - \bar{y})(x_i - \bar{x})$$

$$\sigma_x = \sum_{i=1}^{m} (x_i - \bar{x})$$

$$\hat{w} = \frac{\sigma_{xy}}{\sigma_x^2}$$