Regression

Regression

What is Regression ?

It is supervised machine learning technique used to estimate the relationship between dependent variable and one or more independent variable.

It is used when output/outcome variable is continuous in nature.

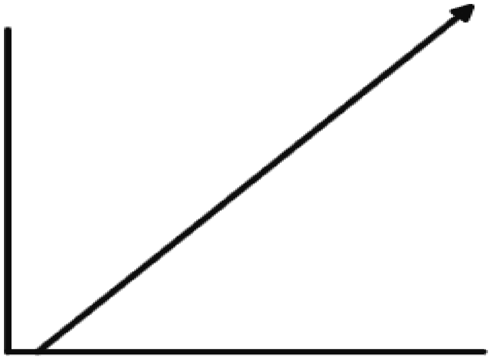
Types of Regression

SLR – Simple Linear Regression
MLR – Multiple Linear Regression
SLR with Gradient Descent
MLR with Gradient Descent
Polynomial Regression
Ridge Regression
Lasso Regression
Elastic Net Regression

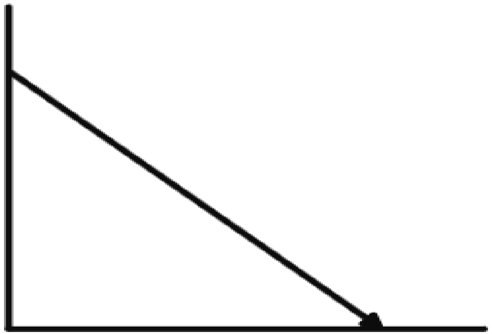
Simple Linear Regression

It is statistical technique which models the relation between one dependent and one independent variable.

In linear relationship when independent variable increases(or decreases), dependent variable also increases. (or decreases)



Positive Linear Relationship



Negative Linear Relationship

Correlation

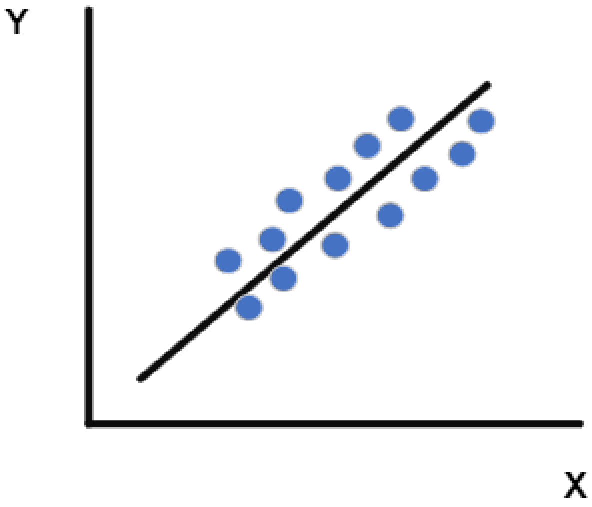
Correlation is defined as statistical measure which estimate the strength of relation between the quantitative variables.

The range of correlation coefficient is between -1 to 1 .

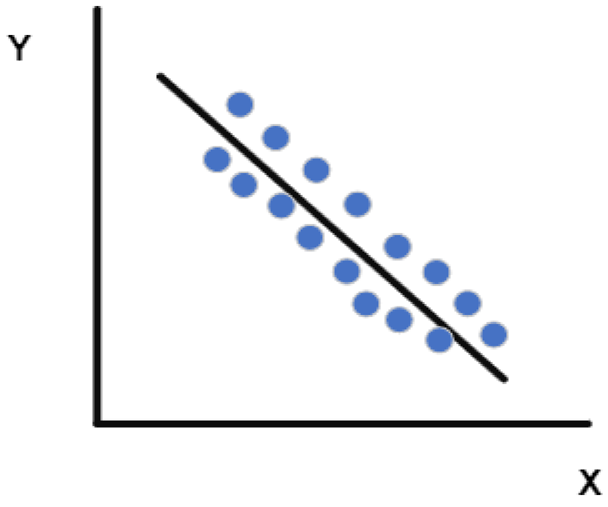
Positive correlation :- When the value of an independent variable increases, the value of dependent variable also increases.

For Ex. When the size of the house increases the price of the house also increases.

Correlation



Positive Correlation



Negative Correlation

Positive Correlation

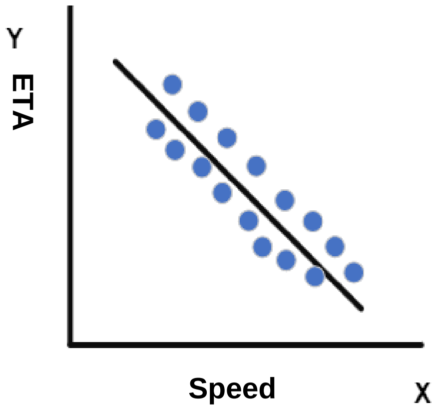
When an independent variable increases, the dependent variable also increases.

For Ex.   When the size of the house increases, the price of house also increases.

Negative Correlation

When an independent variable increases, the dependent variable deceases.

For Ex. When the speed of the vehicle increases, the estimated time of arrival (ETA) decreases.



Pearson Correlation Co-efficient

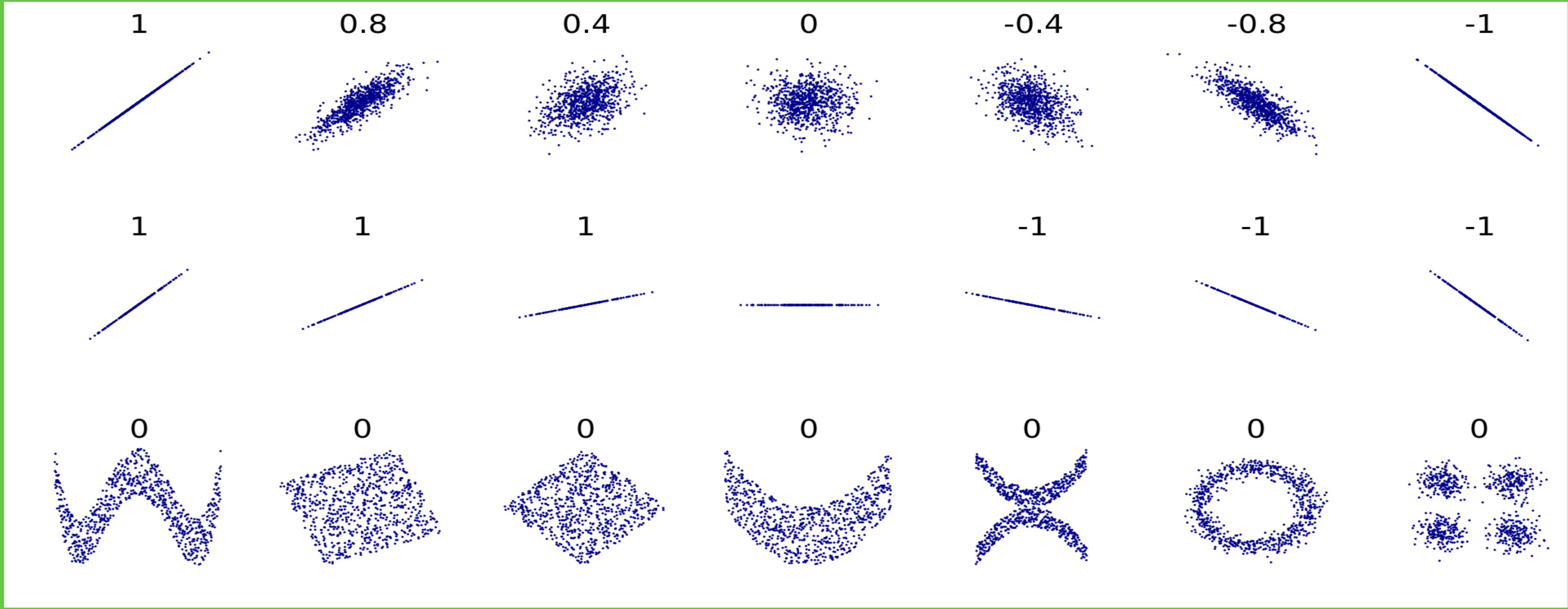


Image Source :   Correlation - Wikipedia

Multicollinearity

When there is a high correlation between independent variables, it is referred as multi-collinearity.

Simple Linear Regression

Simple Linear Regression is statistical technique which models the relationship between dependent variable and independent variable. The prediction is performed on based on single independent variable.

$$Y = \theta 0 + \theta 1.X$$

$\theta 0$ = Intercept at Y.

$\theta 1$ = slope.

- $\theta 0$ and $\theta 1$ are called as coefficients or parameters

Fitting Simple Linear Regression

- In Simple Linear Regression model the fitting operation is performed with training data and the co-efficients $\theta 0$ and $\theta 1$ are estimated.

- The predicted value can be calculated for the specific value of x with the equation

  Y-hat $\hat{y} = \theta 0 + \theta 1.X$

- *$\hat{y}$ is known as predicted y and calculated based on the value of x*
- *We have to find the model co-efficient an intercept $\theta 0$ and $\theta 1$ a slope such that the resulted line will be closer to all the data points.*

Fitting Simple Linear Regression

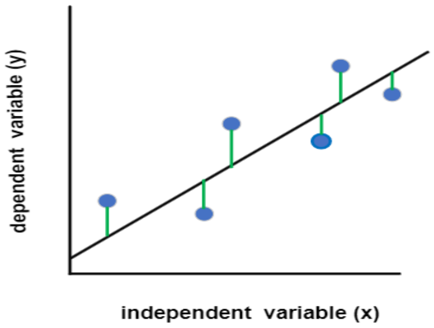
- It is used to estimate the parameters by creation of model which minimizes the sum of the squared errors between the actual and predicted value.

- ith residual is the difference between ith observed and predicted response. $e_i = y_i - \hat{y}_i$

  $= y_i - (\theta_0 + \theta_1 . X)$

- Residual Sum of Squares is defined as **RSS** = $e_1^2 + e_2^2 + \ldots + e_n^2$

- The least squares approach choose those values θ_0 and θ_1 RSS is minimum.



Simple Linear Regression

Hypothesis = $h_\Theta(x) = \Theta_0 + \Theta_1 x_1$

Co-efficient/parameters = Θ_0, Θ_1

Cost Function = $J(\Theta_0, \Theta_1) = 1/2m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2$

Equation Source : Coursera

Cost Function

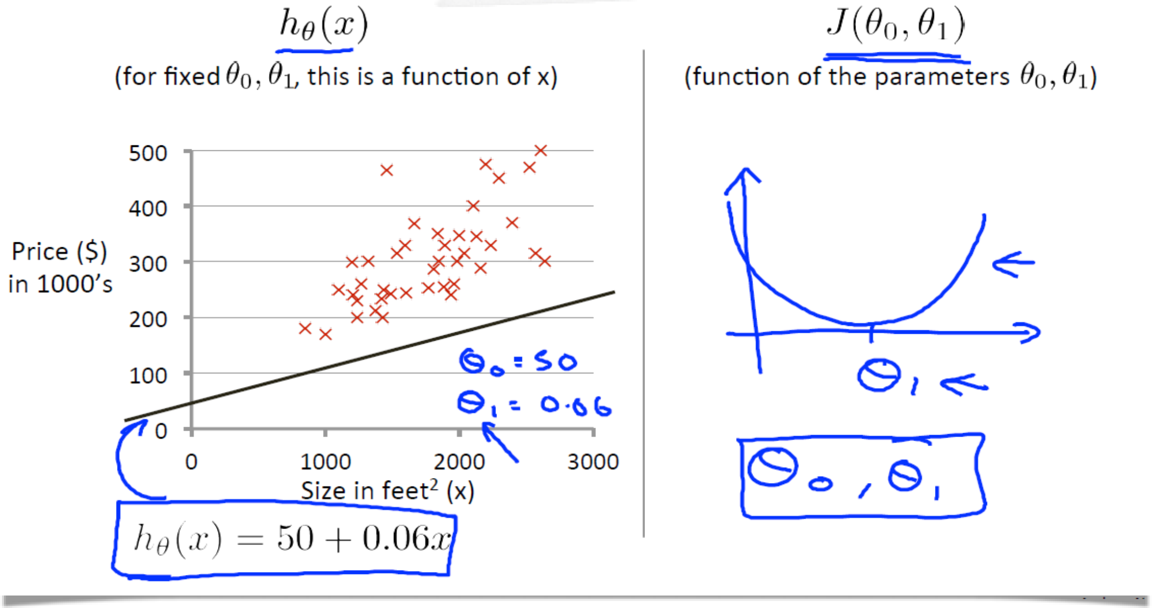


Image Source : Coursera

Cost Function

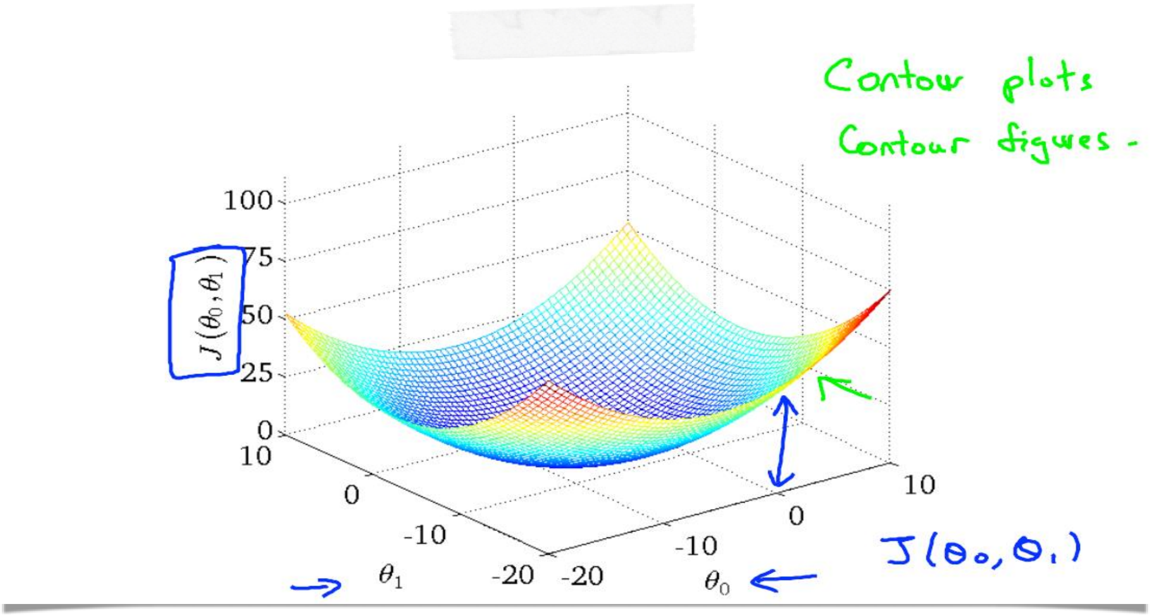
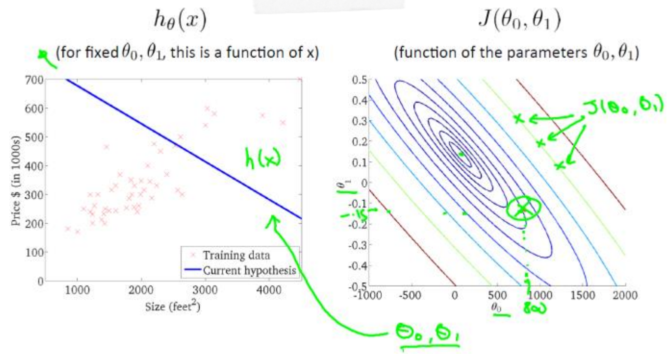
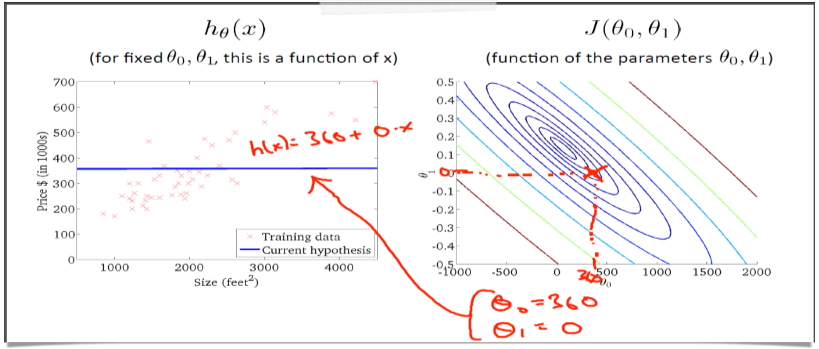
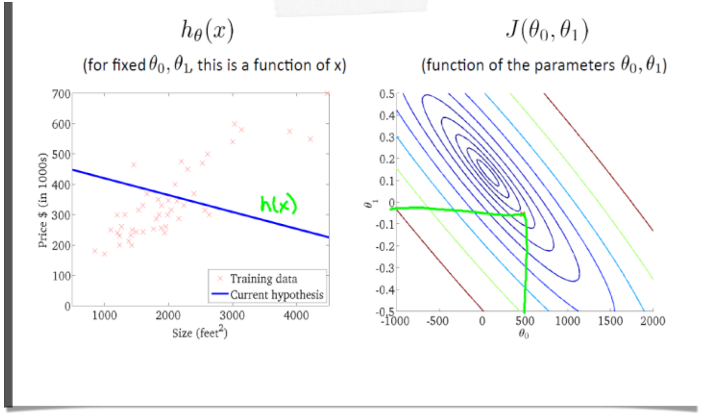


Image Source : Coursera

Hypothesis and Cost Function







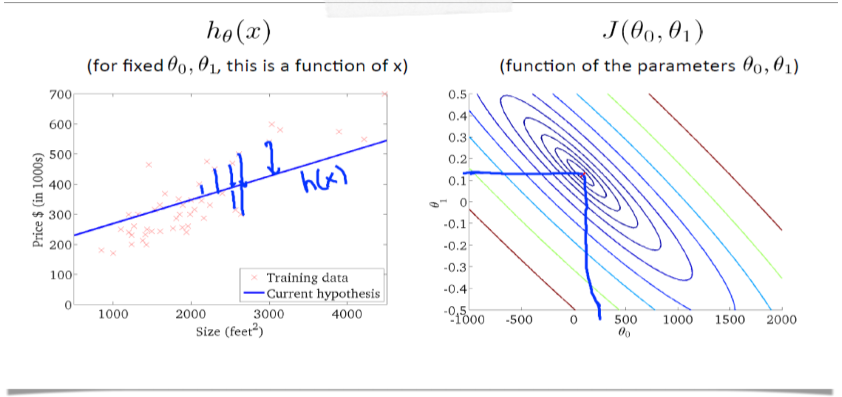


Image Source : Coursera

Gradient Descent Algorithm

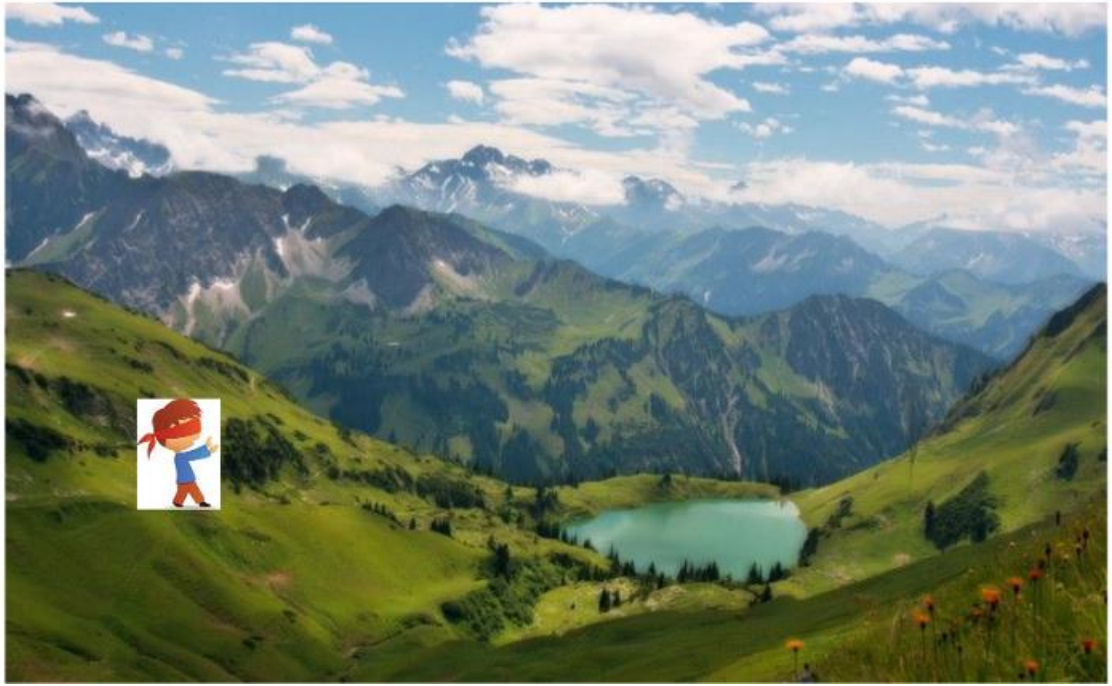
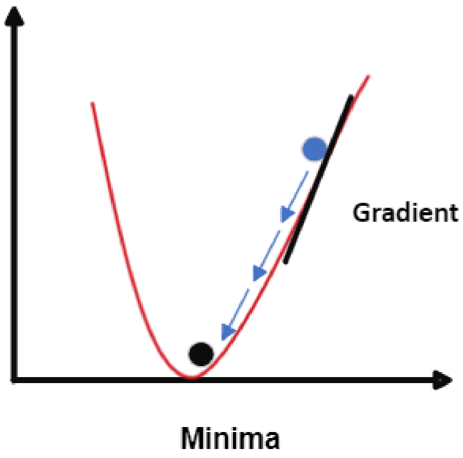


Image source : https://medium.com/@sunil.jangir07/the-outline-of-gradient-descent-da7763a0d66c

Gradient Descent Algorithm

Gradient descent is an iterative algorithm used for optimization of the cost. The goal here is to find out the value of the parameters where the cost is minimum.



Algorithm : initialization of parameter by some value

For every iteration, the parameters are calculated to minimize the cost until we reach to global minimum.

$$\Theta_j = \Theta_j - \alpha \frac{\partial}{\partial x} J(\Theta_0, \Theta_1)$$

j = 0, 1

Equation Source : Coursera

Simple Linear Regression with Gradient Descent

Gradient Descent Algorithm

Algorithm : initialization of parameter by some value

For every iteration, the parameters are calculated to minimize the cost until we reach to global minimum.

$$\Theta_j = \Theta_j - \alpha \frac{\partial}{\partial x} J(\Theta_0, \Theta_1)$$

j = 0, 1

Simple Linear Regression

Hypothesis = $h_\Theta(x) = \Theta_0 + \Theta_1 x_1$

Co-efficient/parameters = Θ_0, Θ_1

Cost Function = $J(\Theta_0, \Theta_1) = 1/2m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2$

Equation Source : Coursera

Simple Linear Regression with Gradient Descent

Keep on changing values of Θ_0, Θ_1 until we reach to minima .

$$\Theta_0 = \Theta_0 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2$$

$$\Theta_1 = \Theta_1 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2 x^i$$

Equation Source : Coursera

Multiple Linear Regression

Multiple Linear Regression is statistical technique which models the relation between one dependent variable and multiple independent variables.

Hypothesis

$$h_\Theta(x) = \Theta_0 + \Theta_1 x_1 + \Theta_2 x_2 + \Theta_n x_n$$

Parameters

$$\Theta_0, \Theta_1, \Theta_2, \Theta_n$$

Cost Function

$$J(\Theta_0, \Theta_1, \Theta_n) = 1/2m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2$$

Equation Source : Coursera

Multiple Linear Regression with Gradient Descent

Simple Linear Regression with Gradient Descent

Keep on changing values of Θ_0, Θ_1 until we reach to minima .

$$\Theta_0 = \Theta_0 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2$$

$$\Theta_1 = \Theta_1 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2 x^i$$

Equation Source : Coursera

Multiple Linear Regression with Gradient Descent

Keep on changing values of Θ_0, Θ_1 until we reach to minima .

$$\Theta_j = \Theta_j - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2 x_j^i$$

j = 0, 1, ..n

$$\Theta_0 = \Theta_0 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2 x_0^i$$

$$\Theta_1 = \Theta_1 - \alpha 1/m \sum_{i=1}^{m} (h_\Theta(x^i) - y^i)^2 x_1^i$$

Equation Source : Coursera

Evaluation Metrics

$$MAE = 1/N \sum_{i=1}^{N} |y_i - \hat{y}_i|$$

$$MSE = 1/n \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Thank You !!!!!