Experiment No.:

Title: Study and implementation of simple linear regression

Objectives: To learn simple linear regression

Theory:

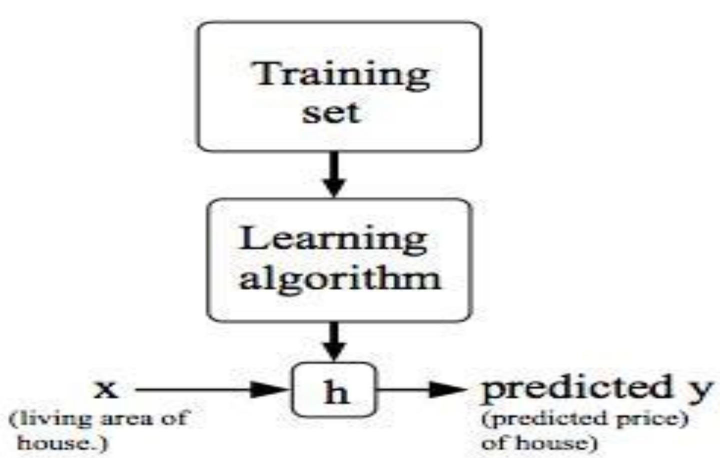
Regression is a method of modeling a target value based on independent predictors. This method is mostly used for forecasting and finding out cause and effect relationship between variables. There are two main types:

1. **Simple Regression**
2. **Multivariable Regression**

- Simple linear regression uses traditional slope-intercept form where m and c are the variables our algorithm will try to "learn" to produce the most accurate predictions. x represents our input data and y represents our prediction. y=mx+c.
- Multivariable regression having more complex form.

Hypothesis Function

- Given a training set, try to learn mapping function
- h : X → Y so that h(x) is a "good" predictor for the corresponding value of y.
- For historical reasons, this function h is called a hypothesis. The job of the hypothesis is a function that takes as input for eg. the size of a house like maybe the size of the new house your friend's trying to sell so it takes in the value of x and it tries to output the estimated value of y for the corresponding house. So h is a function that maps from x's to y's. When the target variable that we're trying to predict is continuous, such as in our housing example, we call the learning problem a **regression problem**.



- When designing a learning algorithm, the next thing is we need to decide is how do we represent this hypothesis h. Normally initial h is considered as
- $h\ (x) = \ _0 + \ _1 x$
-
-
- **0=intercept and 1=the slope**
- The intercept is level of **y** when **x is 0** and slope is the slope is the rate of predicted **increase or decrease in *y*** for each unit **increase in *X*.** Sometimes you can use **h(x)** instead of **h (x)**. Meaning of above equation is we are going to predict that y is a linear function of x. Data set and what this function is doing, is predicting that y is some straight line function of x.

Cost Function

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(\hat{y}_i - y_i\right)^2 = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_i) - y_i\right)^2$$

To break it apart, it is $\frac{1}{2}\ \bar{x}$ where $\bar{x}$ is the mean of the squares of $h_\theta(x_i) - y_i$, or the difference between the predicted value and the actual value.

This function is otherwise called the "Squared error function", or "Mean squared error". The mean is halved $\left(\frac{1}{2}\right)$ as a convenience for the computation of the gradient descent, as the derivative term of the square function will cancel out the $\frac{1}{2}$ term. The

Gradient Descent

Gradient descent algorithm

repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1) \quad (\text{for } j = 0 \text{ and } j = 1)$$

}

$$\theta_0 := \theta_0 - \alpha \frac{d}{d\theta_0} J(\theta_0, \theta_1)$$

$$\theta_1 := \theta_1 - \alpha \frac{d}{d\theta_1} J(\theta_0, \theta_1)$$

Derivatives:

$$\frac{d}{d\theta_0} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)$$

$$\frac{d}{d\theta_1} J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) \cdot x^{(i)}$$

Repeat until convergence

{

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}$$

}

Case Study:

Suppose you are the CEO of a restaurant franchise and are considering different cities for opening a new outlet. The chain already has trucks in various cities and you have data for profits and populations from the cities. You would like to use this data to help you select which city to expand to next. The file ex1data1.txt contains the dataset for linear regression problem. The first column is the population of a city and the second column is the profit of a food truck in that city. A negative value for profit indicates a loss. Make the predictions on profits in areas of 35000 and 70,000 people.

Algorithm:

1. Plot the data form ex1data1.txt

   1.a keep markersize 10, xlable is Population of City in 10,000s and ylabel is Profit in $10,000s

   1.b set axis as a [4,24,-5,25]

2. Calculate Gradient Descent

3. Compute Cost Function
4. Predict profit for 35000 and 70,000 population

Keywords: linear regression, cost function, gradient descent.