**Aim**

To practice the Linear Regression Algorithm on the existing data and analyse the algorithm.

**Tool used:** MATLAB

**Linear Regression**

It refers to the linear approach of defining the relationship between a dependent or output variable (Y) and independent variable(s): (X). The objective is to predict the output while constantly attempting to reduce the error.

Gradient Descent is used to find the minimum point of model’s cost function (defines the error term) by iteratively getting better approximation.

**Why this method?**

It is simple, easy to implement method which approximately predicts the output as a linear function of input features. Also, by plotting learning curve, we could understand how the algorithm is underfitting or overfitting the data and correct the algorithm accordingly.

**Code:**

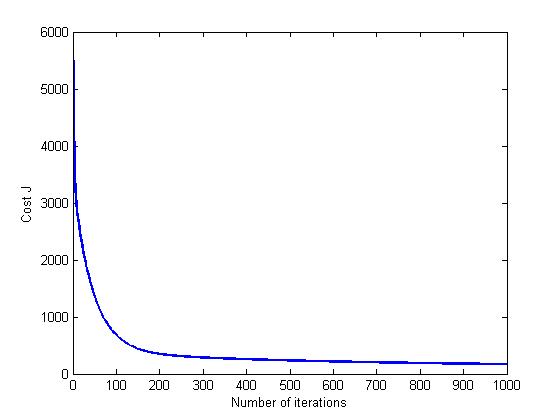
run.m : This is the main file which calls the other functions. Here one can choose to process the output equation for Diastolic BP or Systolic BP.

Extract features.m : It is used to extract the features from the given BP matrix which contains the data and information of 94 samples.

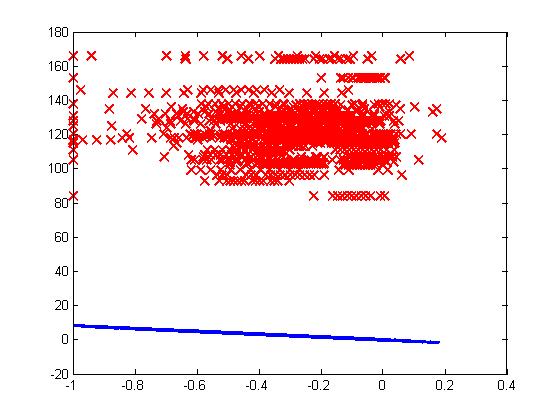
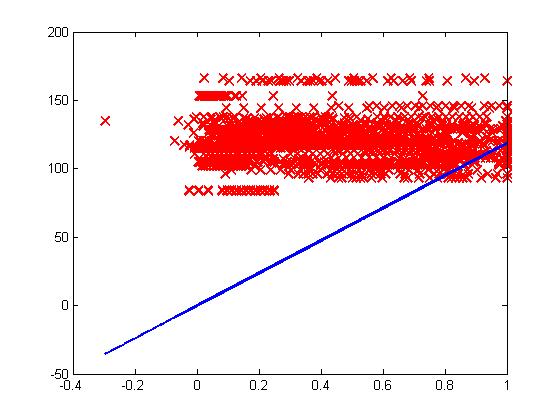
gradientDescentMulti.m: Calculates the cost function and iteratively tries to decrease it as much as possible.

**Observations:**

* The following graph depicts the performance error rate of the algorithm. As seen, the cost function decreases with increase in number of iterations.

****

* The next two images show the data fitting curve of the linear function and the data. The red ‘x’ marks here denote the data and the blue line is the prediction line. It is clearly seen from the graph that prediction line is either hardly fitting the data or not able to fit properly.

****

**Result:**  Although algorithm is able to predict the required BP, the prediction function fails to completely fit on the feature variable. Thus it was needed to find feature importance for the proper prediction of Blood Pressure.