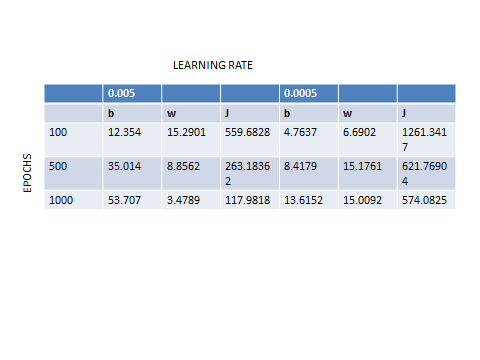
ASSESSMENT 1 : MACHINE LEARNING LAB

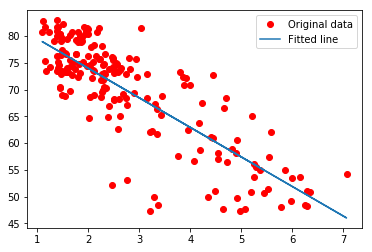
PART-1

c.



d. At epoch, 6100 for learning rate 0.005 the cost remains 15.113 for the consecutive epochs.

PART-2



1. In the above graph, x axis=epochs

Y axis=Cost function.

This particular graph is for learning rate 0.005. We observe that the cost function decreases as the epochs increase until a particular threshold value is reached. Later the value remains constant after certain epochs.

1. The time taken by Batch gradient decent is less as compared to Stochastic as compared to Gradient decent.

The time complexity for Stochastic Gradient decent is C/e where C is time complexity of one sample. Whereas the time complexity of Gradient Decent is n\*C\*log(1/e).

Therefore, the order is Gradient Decent > Stochastic Gradient Decent > Batch Gradient Decent.

1. From the table in Part-1 c. We observe that the min value of cost is for values theta0=53.707 and theta1=3.4789 with the initial values zero.

PART-3

1. The optimization can be increased by changing the learning rate and the epochs. By using different types of optimizers like Stochastic Gradient Decent if the training data set is very humongous. Gradient descent may take too long because in every iteration when we are updating the values of the parameters, we are running through the entire data set. Whereas in Stochastic Gradient Descent we use only a subset of the training sample and it starts improving itself right away from the first sample.
2. The cost function used in this Gradient Decent is "Least squares". This means that the overall solution minimizes the sum of the squares of the difference between an observed value and the actual value.