BHRAMARI SARKAR
Roll No-2105364
Section-ML_CS3
ML ASSIGNMENT 1

Question 1: Use linear regression to fit a straight line to the given database.Set your learning rate to 0.5. What are the cost function value and learning parameters values after convergence? Also,mention the convergence criteria you need.

Solution:

Cost function value=0.5561400558913063
Slope(m)=0.6619554709942285
Intercept(c)= -2.4064644777449575e-15

The convergence criteria are based on the number of epochs:

```
Slope (m): 0.6619554709942285
Intercept (c): -2.4064644777449575e-15

Epoch 1000, Cost: 0.5561400558913063
```

Question 2: The cost function we are using in this assignment is different than the one we used in class.Can you think of the advantage of averaging the cost?

Solution:

Yes, there are several advantages in averaging the cost in cost function.
The advantages are:

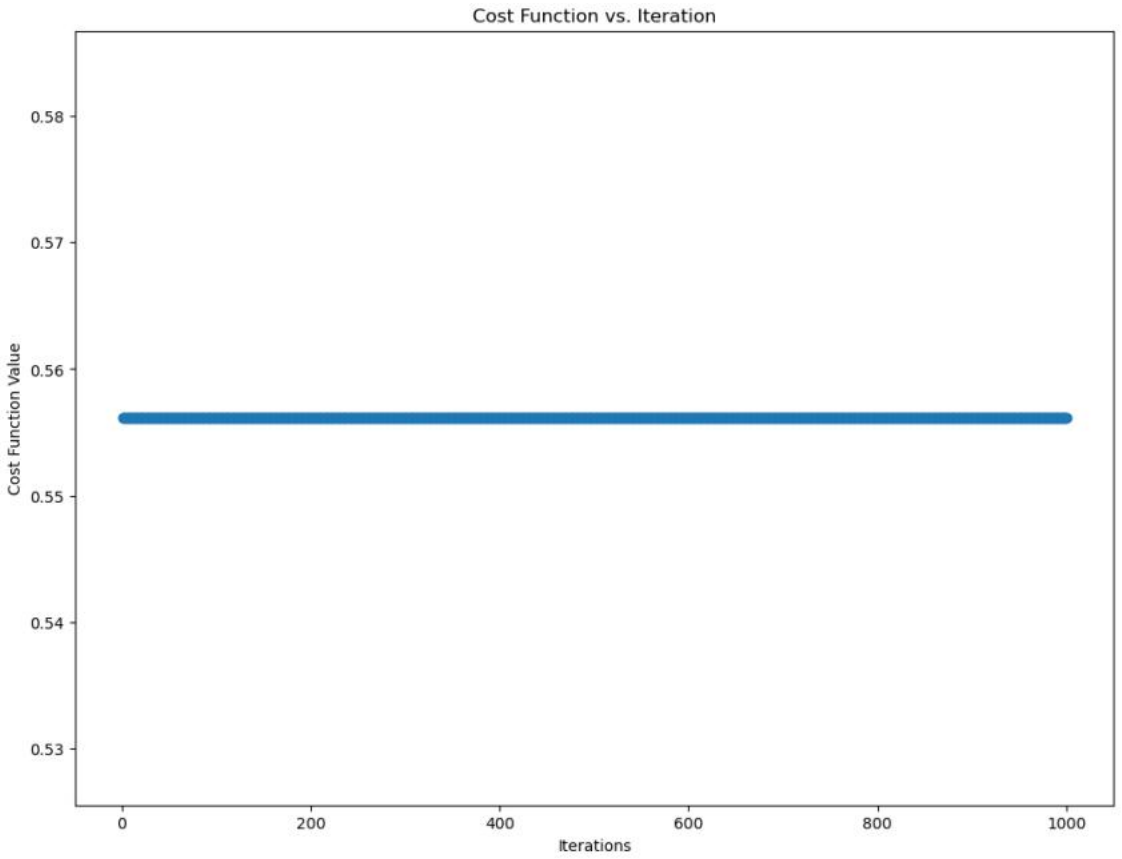
i. **Normalization:**It helps in normalizing the cost function thus making it easier for comparing it among different datasets with varying number of samples.It provides a clear view of how well the model is performing.

ii. **Mean squared Error:**The known formula is for calculating the mean squared error function, which is used in regression problems. After calculating the average

in the given context , it is helpful as it gives an idea of how off the predictions are across all the data points.

iii. **Stabilization:**By averaging the cost ,it aids in stabilizing and speeding up the gradient descent optimization process, as smaller cost values can prevent the learning algorithm from oscillating excessively.

<u>Question 3:</u> Plot cost function v/s iteration graph for the model in question 1.
<u>Question 4 :</u> Plot the given dataset on a graph and also print the straight line you obtained in question1 to show how it fits the data.
<u>Solution:</u>



<u>Question 5:</u> Test your regression model with the learning rates lr=0.005,lr=0.5,lr=5 for each learning rate,plot a graph showing how a cost function changes for each iteration and write your observation.
<u>Solution:</u>

For L=0.005

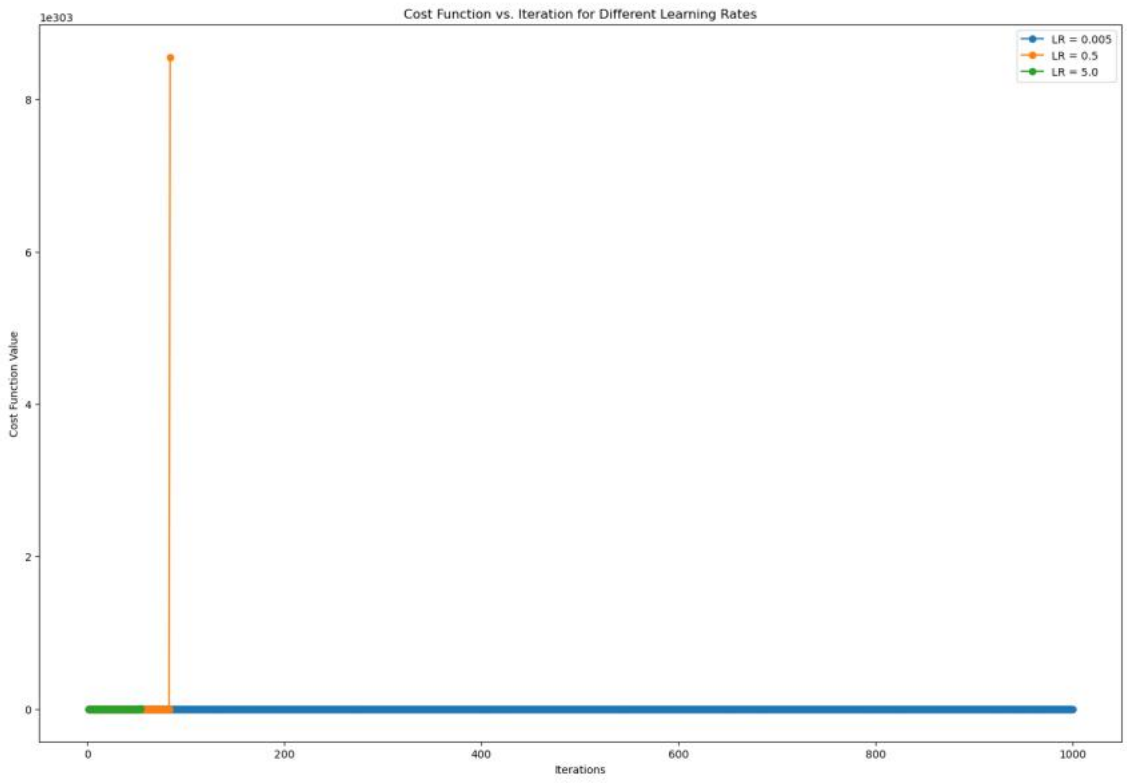
Cost function value=0.5561400569026452
Slope(m)=0.6619238240082247
Intercept(c)= -2.4148058692955056e-15

For L=0.5

Cost function value=0.5561400558913063
Slope(m)=0.6619554709942285
Intercept(c)=-2.4064644777449575e-15

For L=5
Cost function value=nan
Slope(m)=nan
Intercept(c)= nan

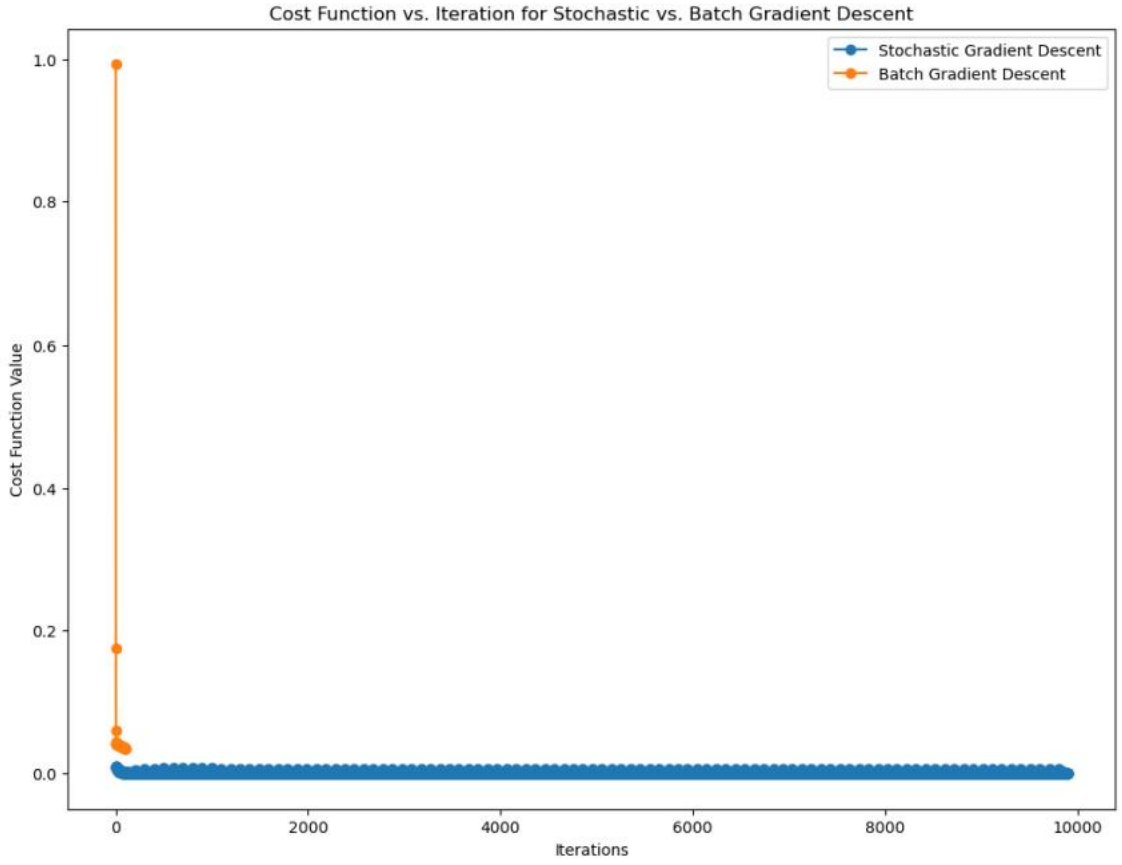
My observation:
A learning rate of 5 is relatively high and may lead to overshooting the minimum point of the cost function. In gradient descent, the learning rate determines the step size taken during each iteration to update the model parameters. A very high learning rate can cause the algorithm to oscillate or diverge instead of converging to the minimum.



<u>Question 6:</u> Choose a suitable learning rate then implement stochastic and min-batch gradient descent,plot the cost function against iteration, and observe how your cost function changes compared to batch gradient descent.
<u>Solution:</u>

Learning rate=0.01
Epochs=100

The cost function is high for Batch Gradient Descent and low and normal for Stochastic Gradient Descent.



GITHUB LINK: https://github.com/Bhrammm/2105364_MLAssignment1