

ML ASSIGNMENT1

Utkal Sahoo (2105846)

Question 1

Cost function value = 0.5561400558913063

Slope (m): 0.6619554709942285

Intercept (c): -2.4064644777449575e-15

the convergence criteria are based on the number of epochs (epochs).

```
Epoch 1000, Cost: 0.5561400558913063  
Slope (m): 0.6619554709942285  
Intercept (c): -2.4064644777449575e-15
```

Question 2

Yes, averaging the cost in a cost function, as shown in the image, can have several advantages:

Normalization: It helps in normalizing the cost function, making it easier to compare across different datasets with varying numbers of samples. This normalization provides a clearer view of how well the model is performing.

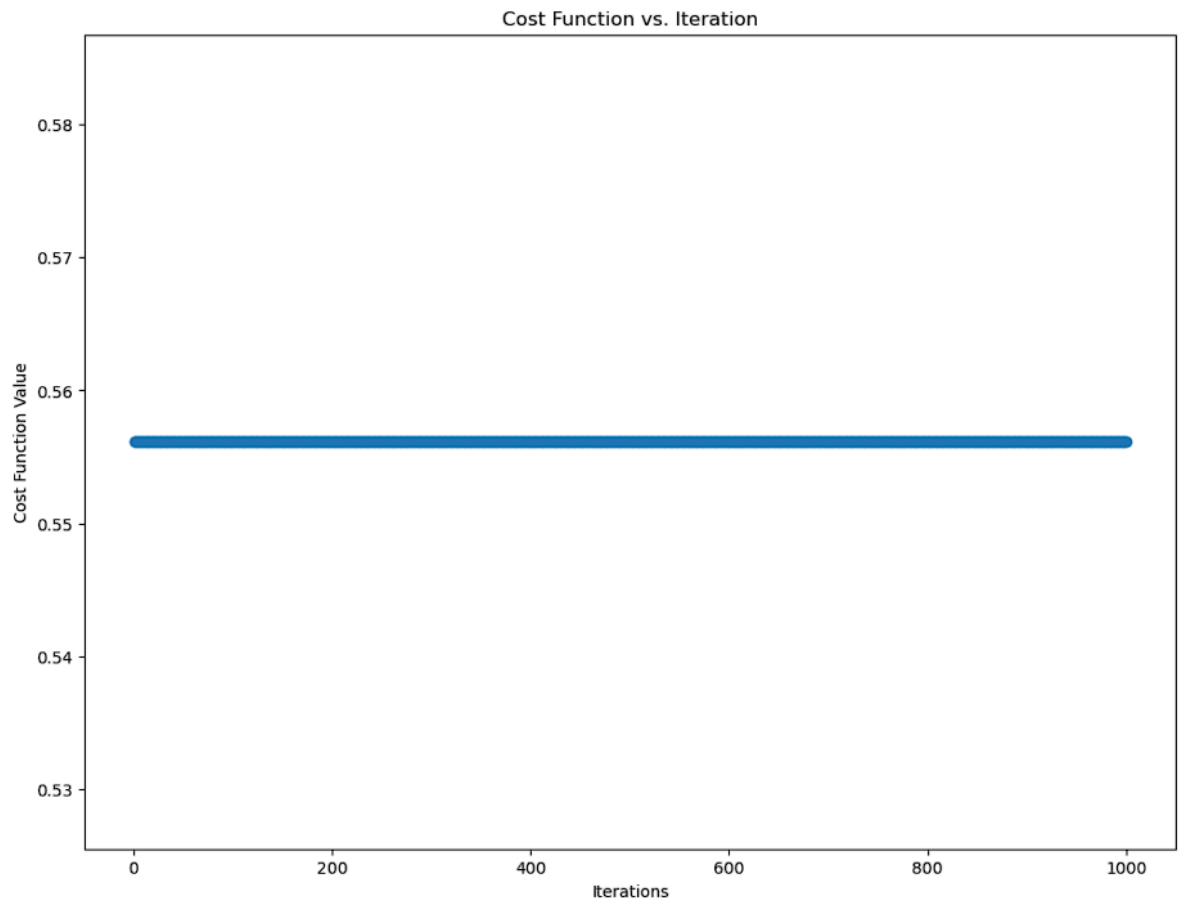
Stabilization: Averaging the cost aids in stabilizing and speeding up the gradient descent optimization process. This is because smaller cost values (resulting from averaging) can prevent the learning algorithm from oscillating excessively.

Mean Squared Error: The formula you provided is for calculating the mean squared error cost function, often used in regression problems. Averaging in this context is beneficial as it gives an idea of how off the predictions are, on average, across all data points.

ML ASSIGNMENT1

Utkal Sahoo (2105846)

Question 3 & 4.



Question 5.

For $\eta = 0.005$

Cost function value : 0.5561400569026452

Slope (m): 0.6619238240082247

Intercept (c): -2.4148058692955056e-15

ML ASSIGNMENT1

Utkal Sahoo (2105846)

For $\eta = 0.5$

Cost function value = 0.5561400558913063

Slope (m): 0.6619554709942285

Intercept (c): -2.4064644777449575e-15

For $\eta = 5$

Cost function value: nan

Slope (m): nan

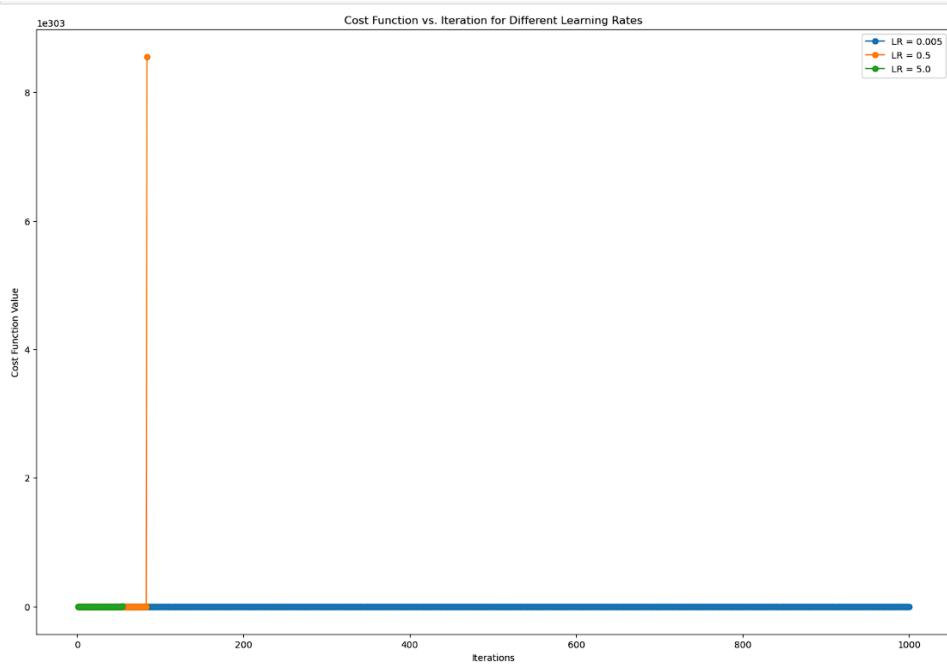
Intercept (c): nan

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.

ML ASSIGNMENT 1

Utkal Sahoo (2105846)

Comparisons :-



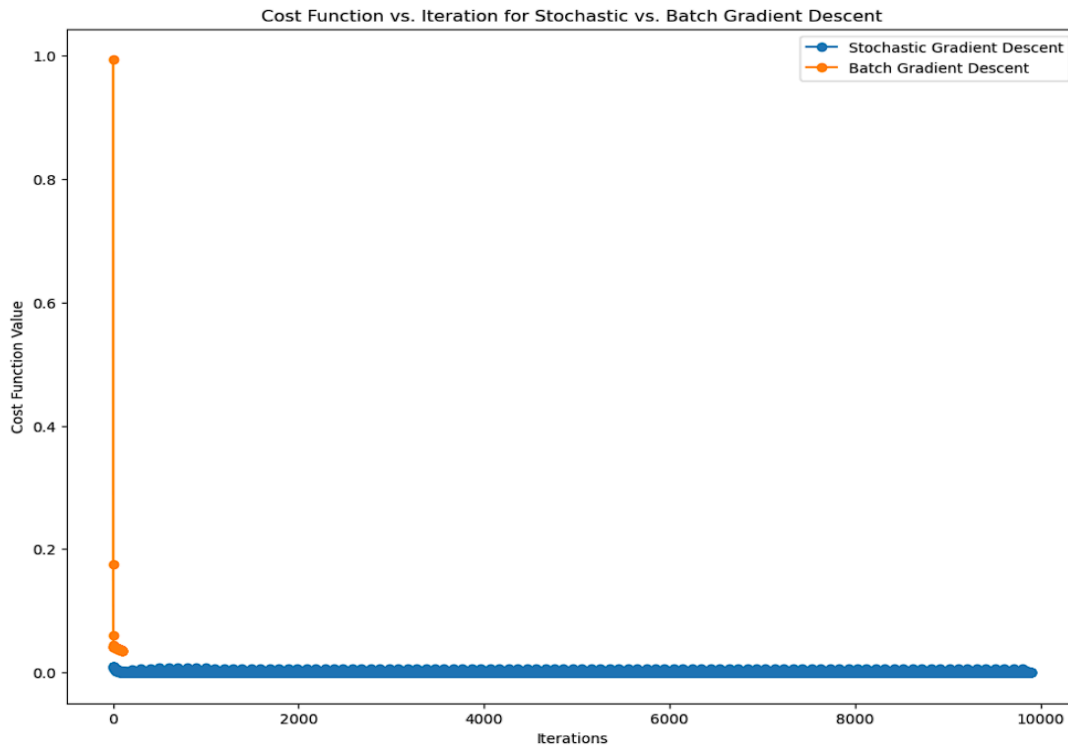
Question 6.

Learning Rate = 0.01

Epochs = 100

ML ASSIGNMENT 1

Utkal Sahoo (2105846)



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

GITHUB LINK - <https://github.com/utkal2048/ML-ASSIGNMENT-1>