Jocelyn Parra

425 HW 2 Write- Up

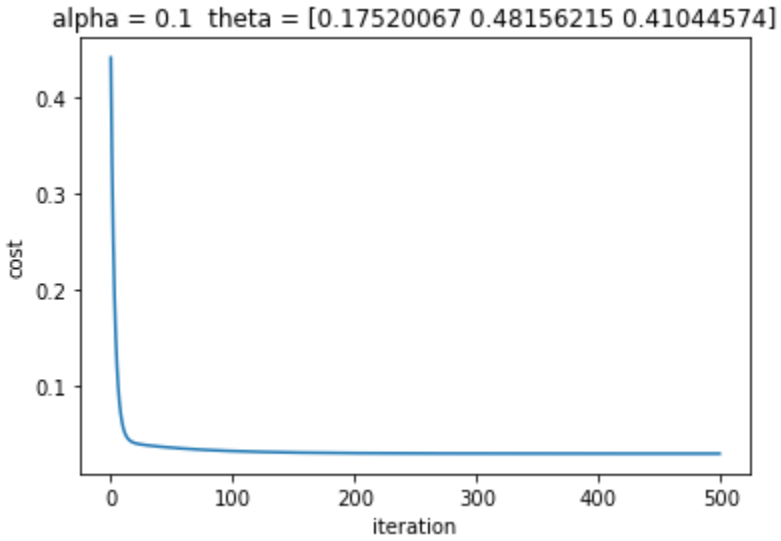
In order to implement a Linear Regression algorithm to predict a student's university GPA, we need to construct the gradient decent program into python. In class we have concluded that the GD is an optimization algorithm used to minimize the cost function in a linear regression model.

The cost function in linear regression, on the other hand, measures the difference between the predicted values and the actual values, i.e., a student's GPA in university using math SAT and Verb SAT scores. It's used to optimize the model's parameters so that the predicted values are as close as possible to the actual values.

The gradient tells us about the direction in which the cost function increases the most, and the algorithm moves in the opposite direction to minimize the cost. This process continues until the cost function converges to a minimum.

All in all, of the cost function in linear regression measures the error between the predicted and actual values, and the gradient descent algorithm uses this information to optimize the model parameters and reduce the error.

I have used a Min-Max scaling normalization technique between [0,1] that takes minimum and maximum values in the data frame while subtracting the minimum value from each value in the data frame and divides the result by the difference between the maximum and minimum values. In doing this, it results in a new data frame where the minimum value is 0 and the maximum value is 1.



In the example, the first figure shows a testing rate of 0.1 with 500 iterations and an alpha of 0.1. The cost function in the figure converges very abruptly, which suggests that the learning rate of 0.1 is appropriate for this problem. On the other hand, the second figure shows a testing rate of 0.1 with 500 iterations and an alpha of 0.001. The cost function converges slower in this case, which is expected with a smaller learning rate.

In conclusion, both the alpha and testing rate play an important role in determining the performance of the gradient descent algorithm. The alpha affects the step size and convergence rate, while the testing rate affects the size of the test set and the model's performance evaluation. The values for these parameters can be adjusted and tuned to find the optimal solution for a specific problem.