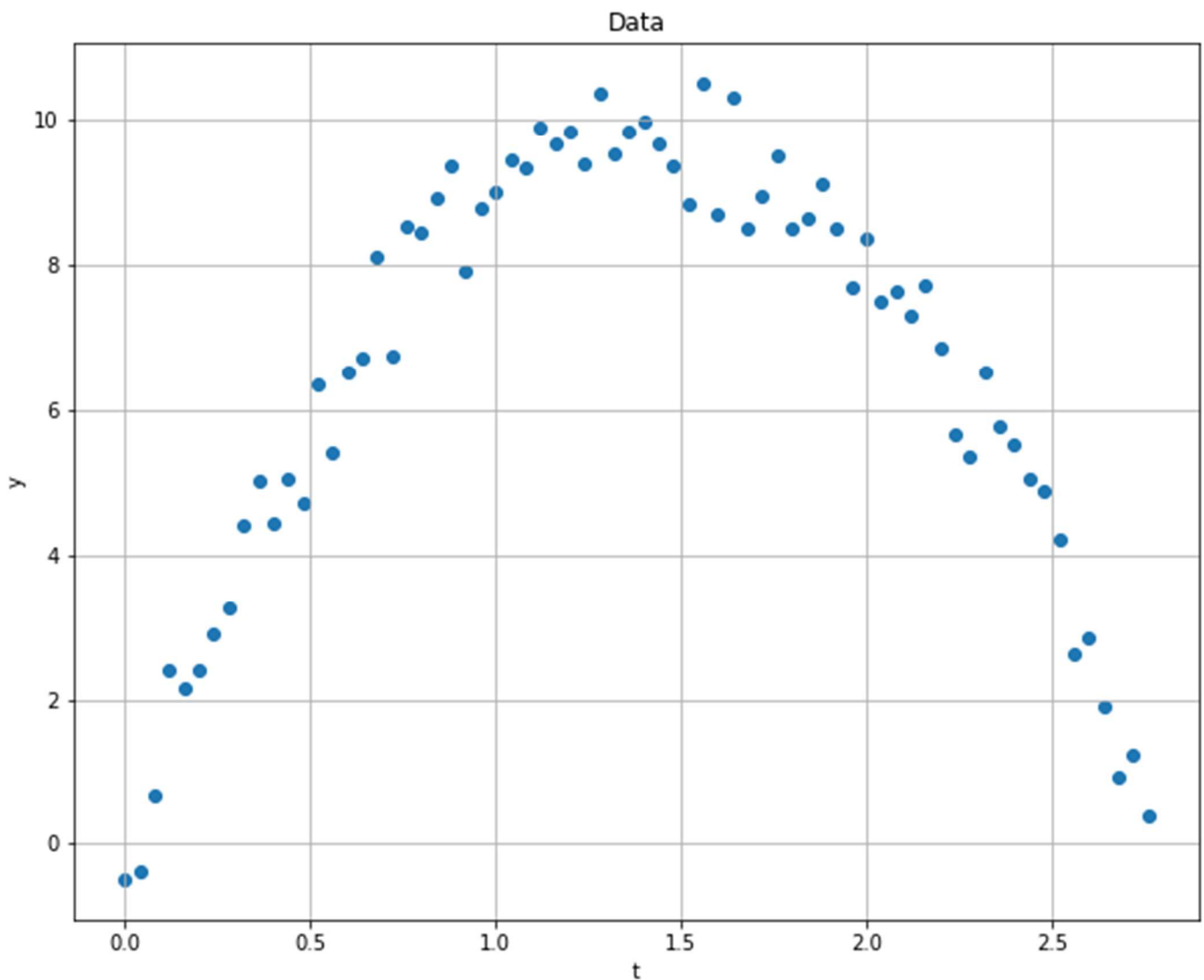


Name – Mohit Garg

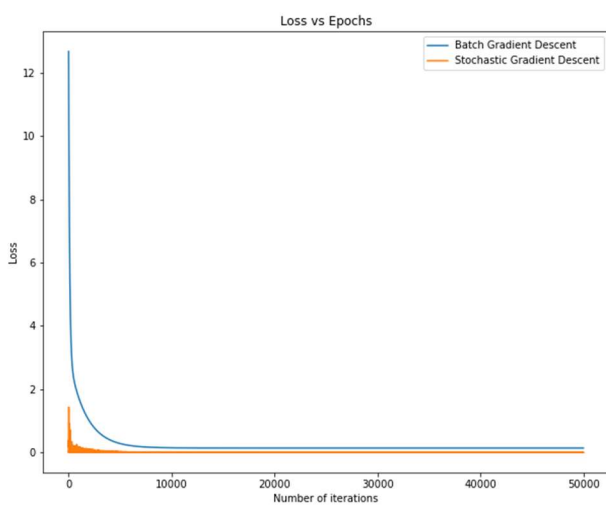
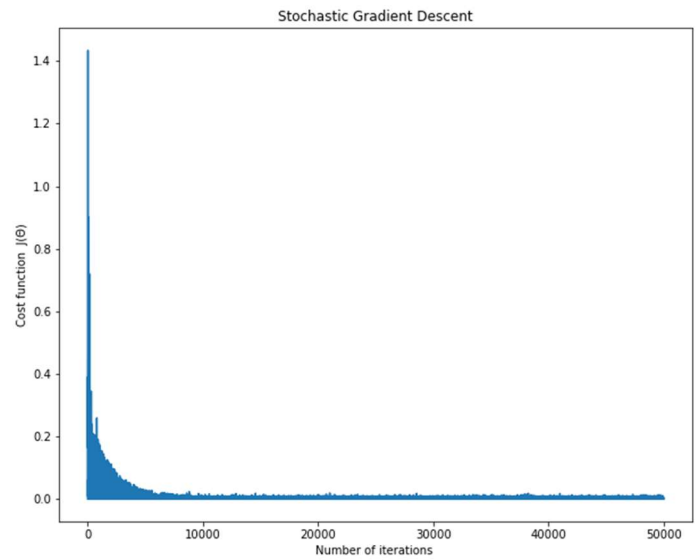
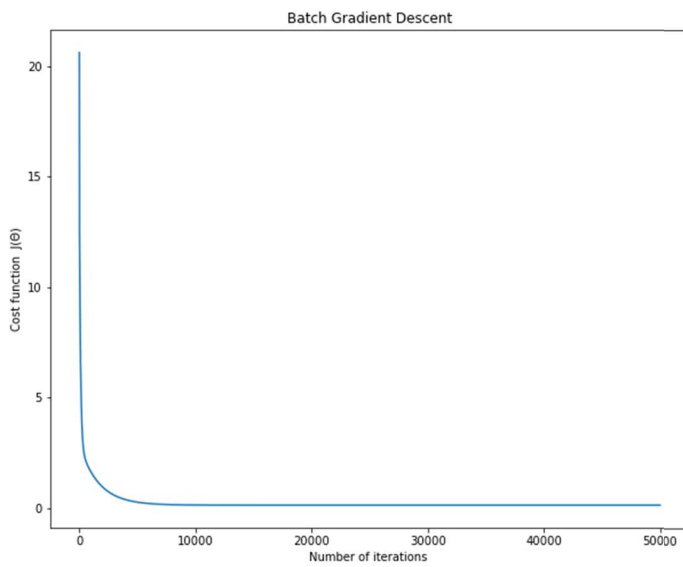
Entry Number – 2020AM10657

Assignment 2

Problem 1



First plot is y v/s t plot for data given. As one can see, parabolic nature is fairly obvious.



Here, we have plotted the loss with number of iterations.

Here are some of the observations found and difference between the SGD and BGD model:

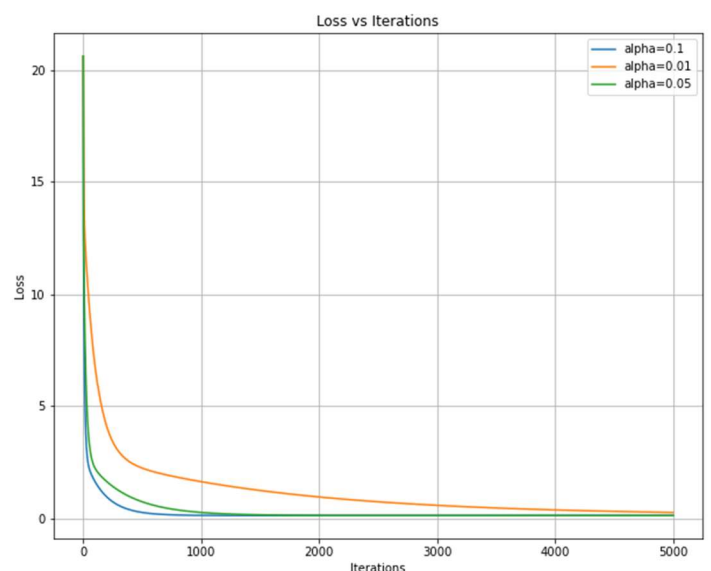
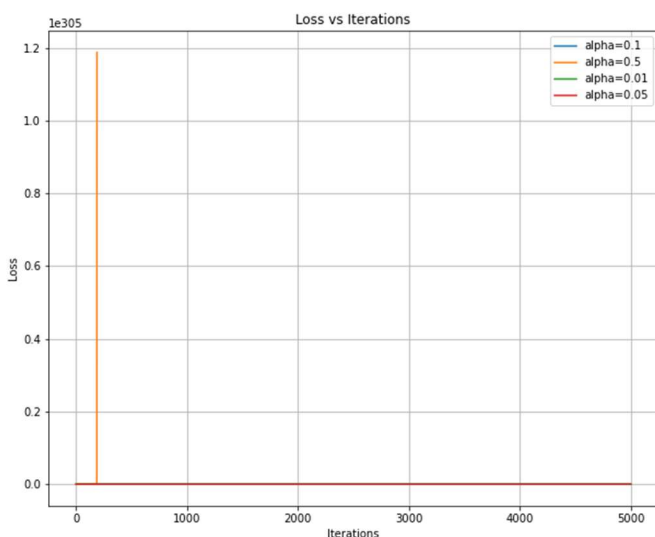
- SGD losses have fluctuations (although small)
- SGD loss function is lower as compared to BGD and convergence is also relatively early as compared to BGD.
- Both SGD and BGD almost converged near 6000

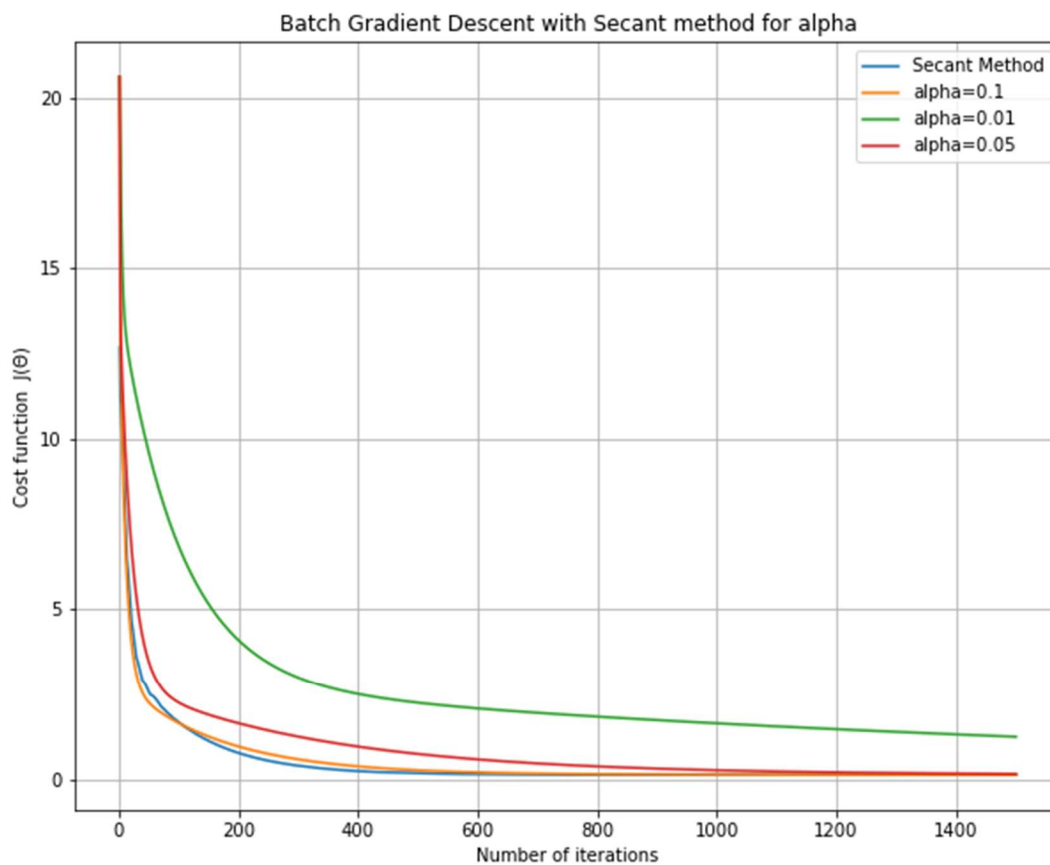
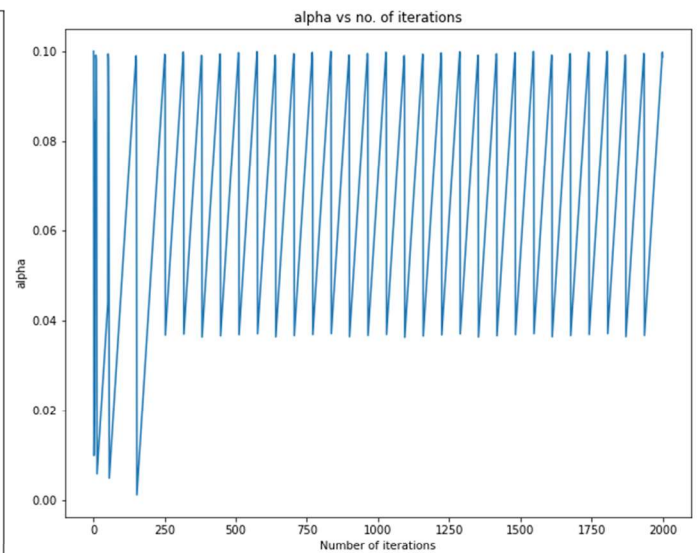
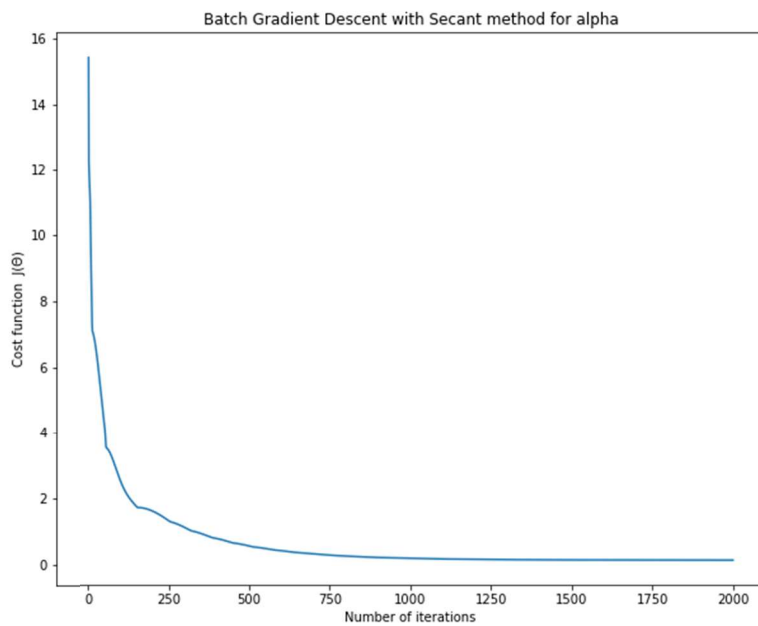
The plot of the loss J with number of iterations is shown here for α in $\{0.1, 0.5, 0.01, 0.05\}$.

We got overflow for $\alpha = 0.5$ so plot excluding 0.5 is shown.

We find that plot of 0.1 is lowest followed by 0.05 while 0.01 is the topmost plot.

From here, we can get idea that $\alpha = 0.1$ is fast as well as near maximum α value for convergence as $\alpha = 0.5$ diverged as shown before.

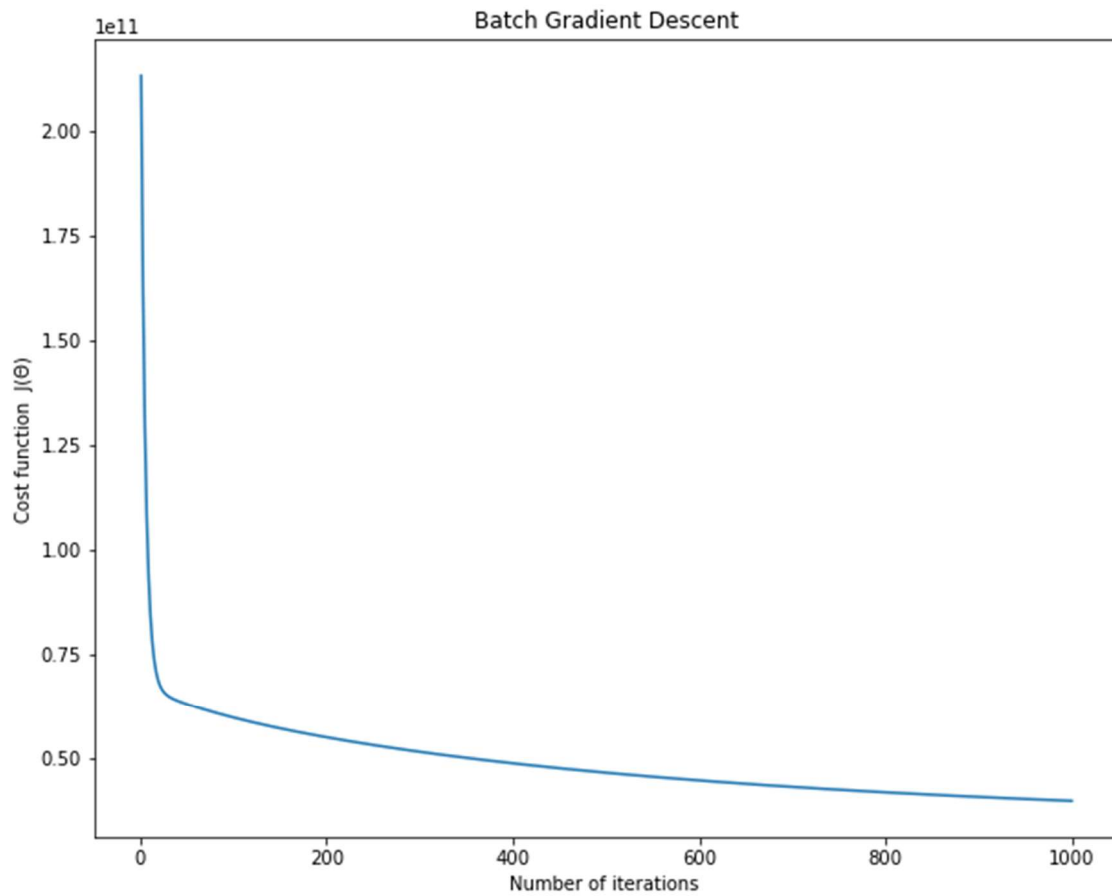




The plot of the loss J with number of iterations is shown here where α is calculated based on secant search method using batch gradient descent method.

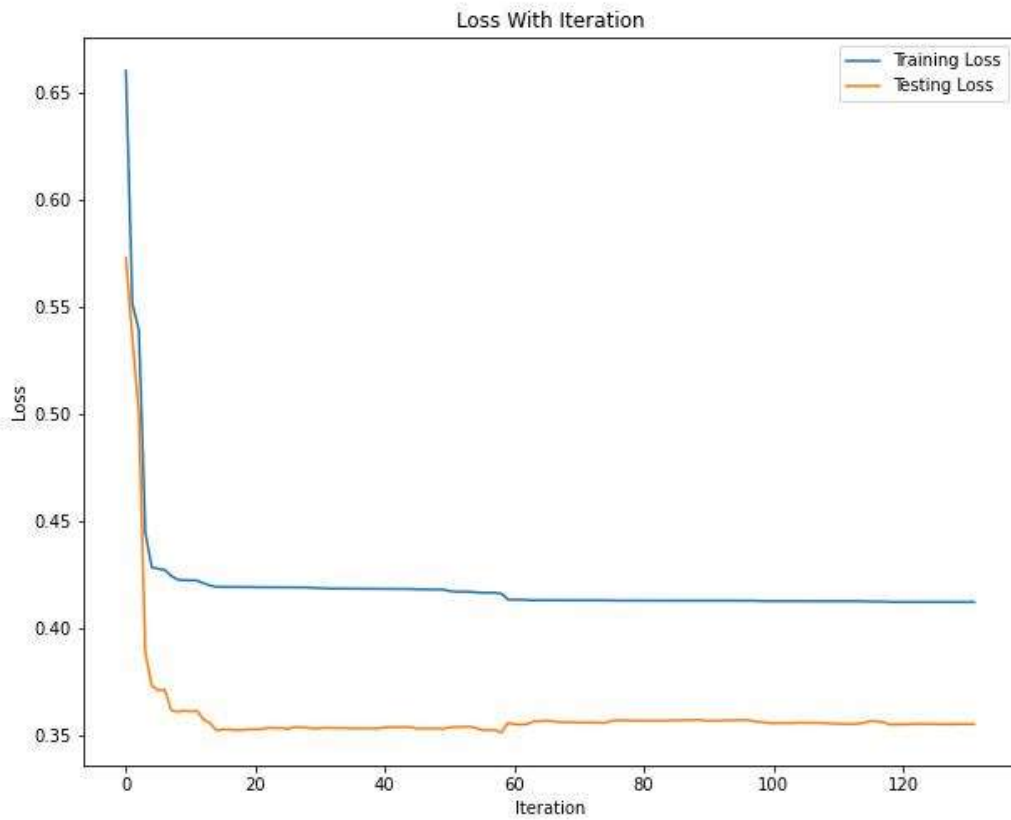
As one can see that it converges relatively earlier and performance is optimized based on comparison with previous data.

Problem 2



Plot representing how the loss is changing with number of iterations. We can see that the loss is decreasing very steeply.

Problem 3



Plot of the cost history (J) vs. the number of iterations for training and test data. Since loss function for test data is lower than training data this means there is no overfitting. Also one can notice that both plots converge at ~ 15 iterations and subsequent change is almost negligible.