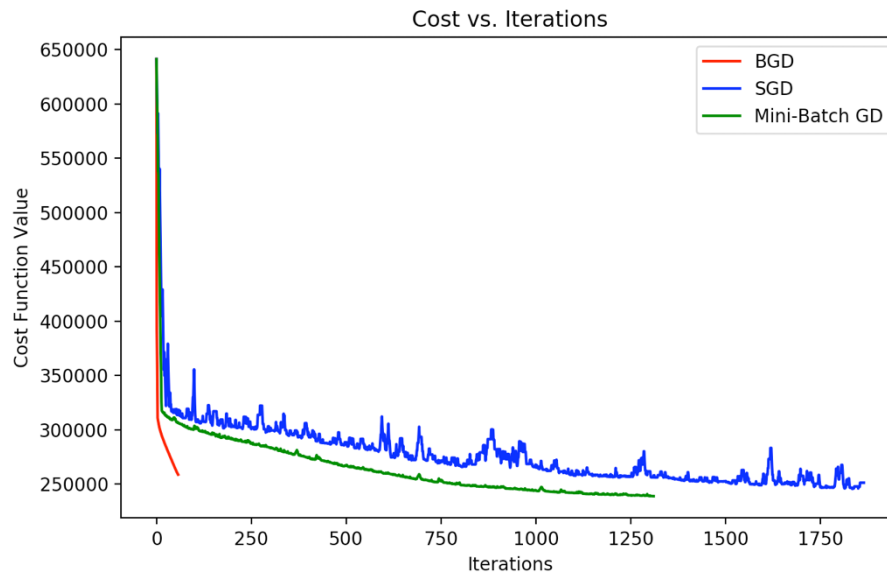


Problem 1 Part b



Time for BGD: 9.216s

Number of iterations for BGD: 57

Time for SGD: 148.120s

Number of iterations for SGD: 1864

Time for Mini-Batch GD: 103.825s

Number of iterations for Mini-Batch GD: 1310

From the graph and the results, batch gradient descent takes the least number of iterations and stochastic gradient descent takes the largest number of iterations. Mini batch is in between. The cost function value of stochastic gradient descent has lots of oscillations because this technique updates the parameters based on one sample. The line of the cost function value of BGD is most smooth. Mini batch is also in between. For each iteration, BGD takes the longest time and SGD and Mini batch takes shorter time for each iteration.