

MACHINE LEARNING

ASSIGNMENT 1

Abhishek Agarwal (2014MCS2114)

Question 1

(a)

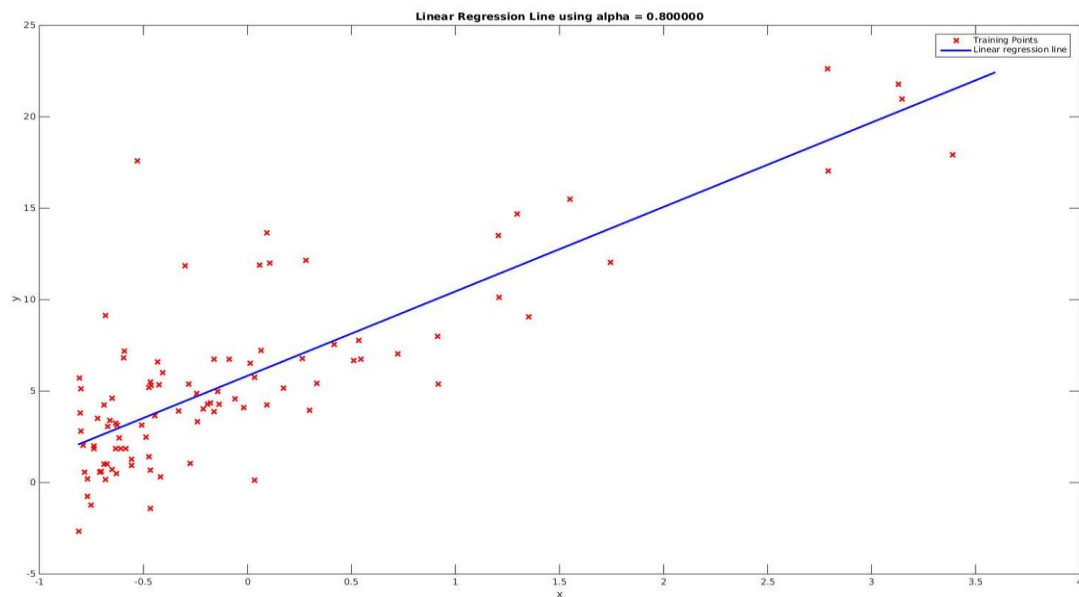
For Implementing Batch Gradient Descent Method

Learning Rate (eta) is 0.8

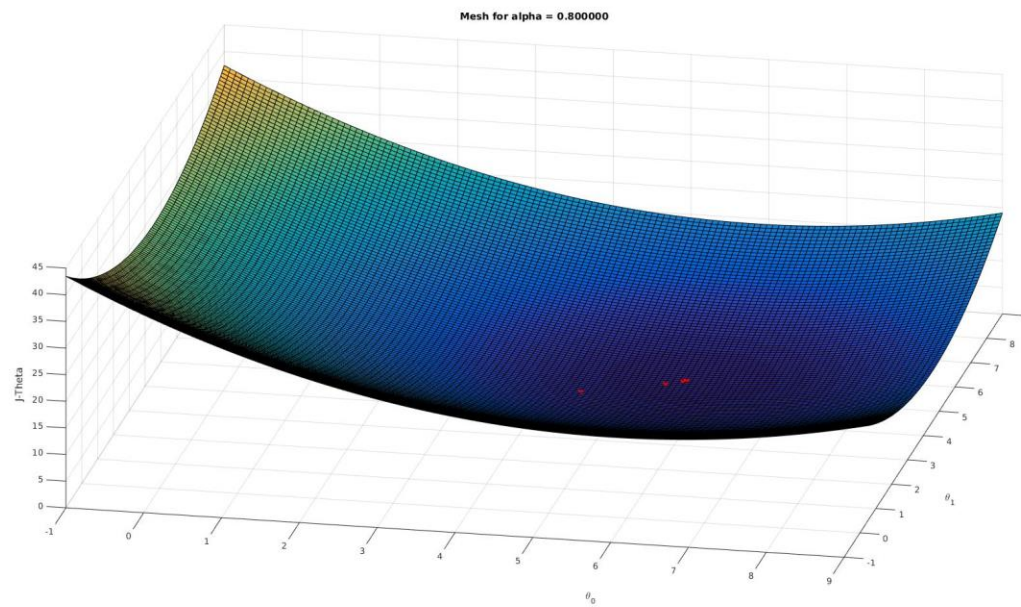
Stopping Criteria is chosen as Difference in error function is less than a small value ϵ .

Final Value of θ obtained – [5.8391, 4.6169].

(b) Plot of Hypothesis Function.



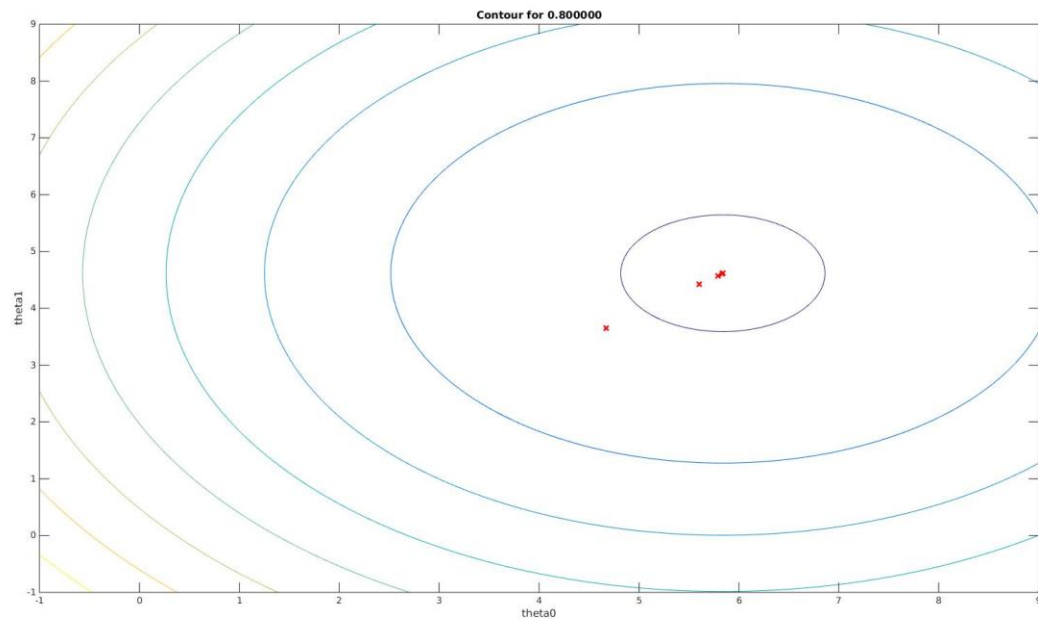
(c)



Three dimensional mesh showing $j(\theta)$ on z axis and parameters theta0 and theta1 and x and y axis respectively.

(d)

Shown below is the contour plot of $j(\theta)$ shown at each iteration of gradient descent.



(e)

For very small values of eta say 0.1, the convergence takes place slowly but accurate values are obtained. For moderate values of eta 0.9 or 1.3, the convergence takes place relatively fast and good approximation of parameters is obtained. For large values such as 2.1 or 2.5 the cross marker keeps on oscillating and convergence takes place extremely slowly or may not take place depending upon how large eta is chosen.