COL774 : Machine Learning Assignment 1

Q1) Implemented Gradient Descent with function name as **gradientDescent()**

'eta' is learning rate and 'I' is the number of iterations.

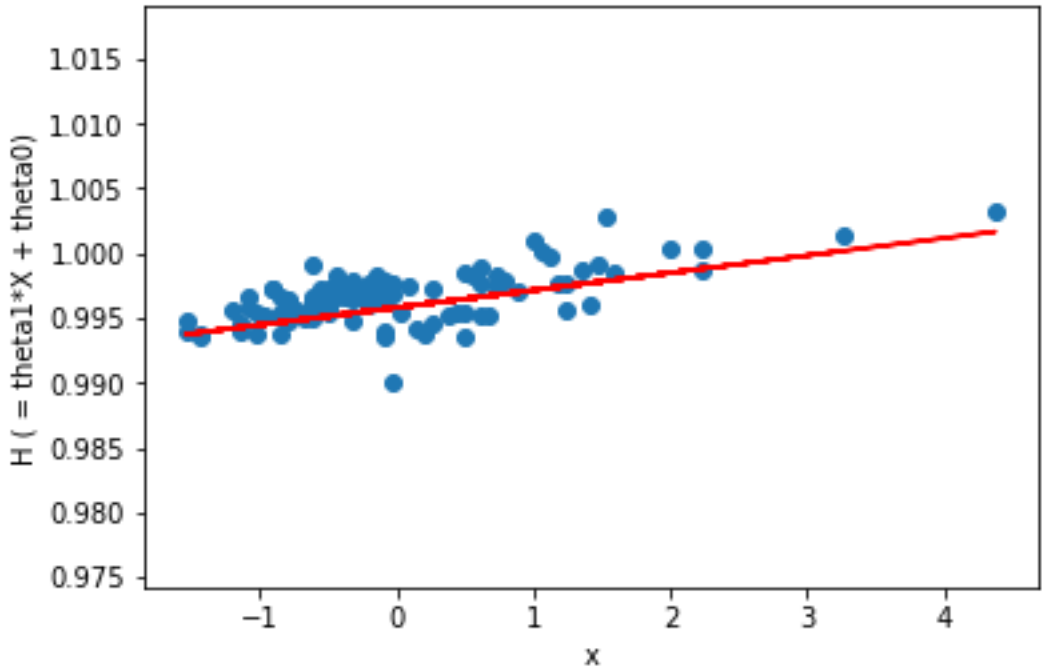
'tracktheta0',' tracktheta1',' tracktheta2' records the respective values of theta0,theta1 and theta2 respectively throughout the number of iterations.

a) Stoppage Criteria: if the absolute difference between current cost (i.e. the loss) and the previous loss is less than '0.00000001' or if the current cost is greater than the immediate previous cost, then we will say that our model has converged.

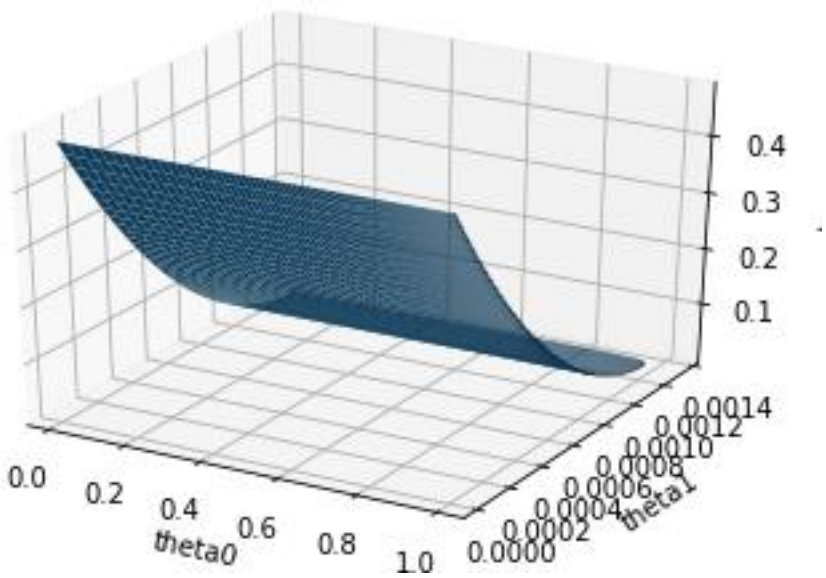
Learning rate: 0.015

Final set of parameters: **theta0=0.99581288 ; theta1=0.00133911**

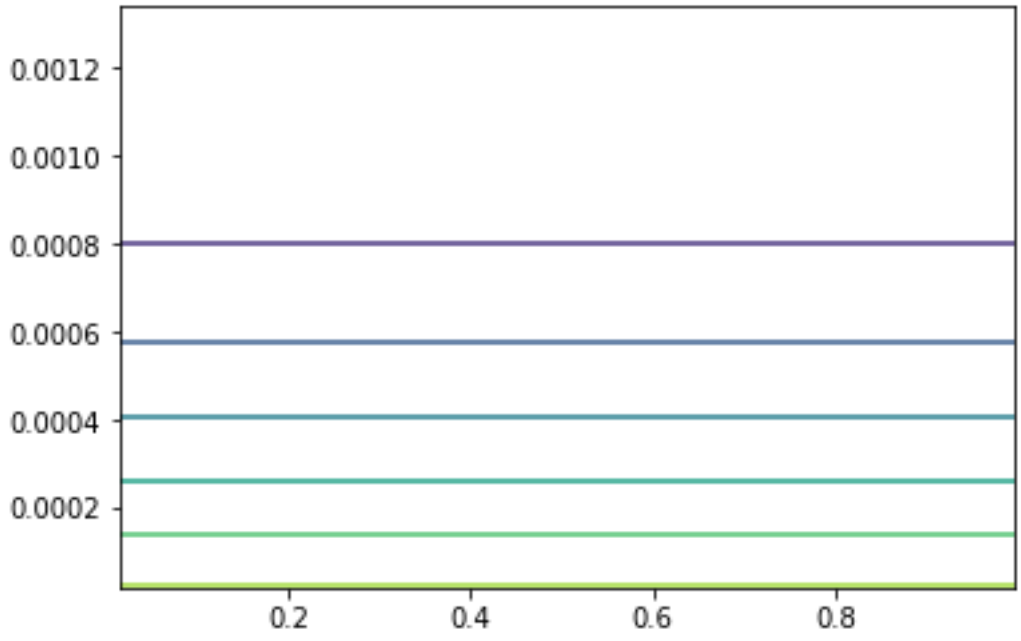
b)



c) *could not draw the animation*

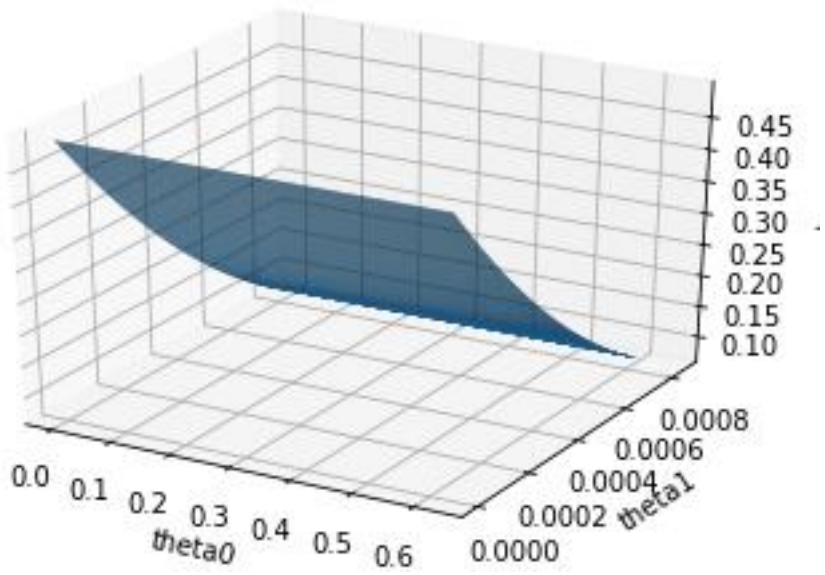


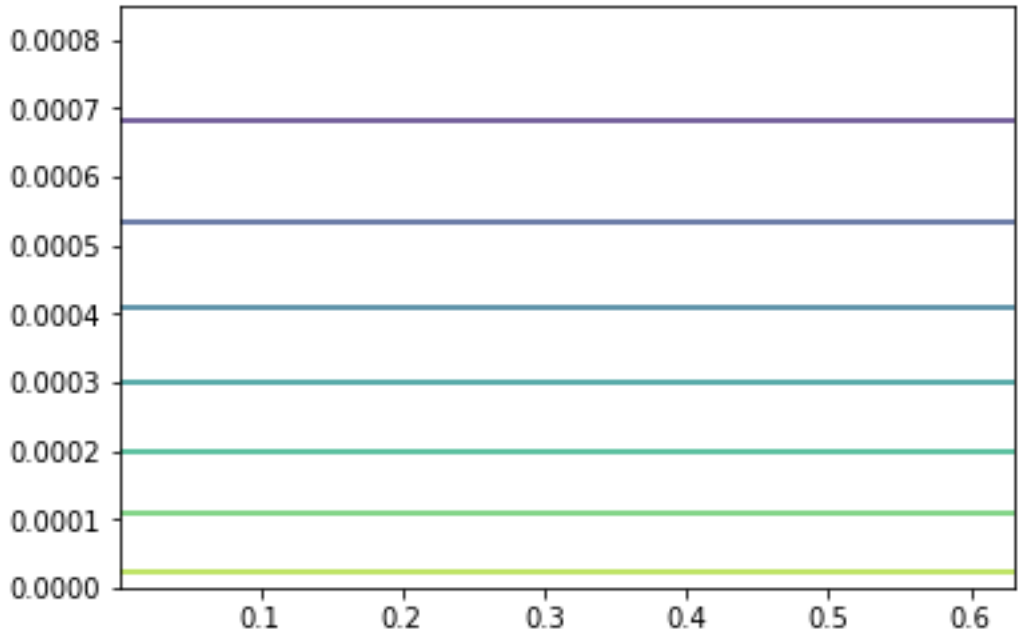
d) *could not draw the animation*



e)

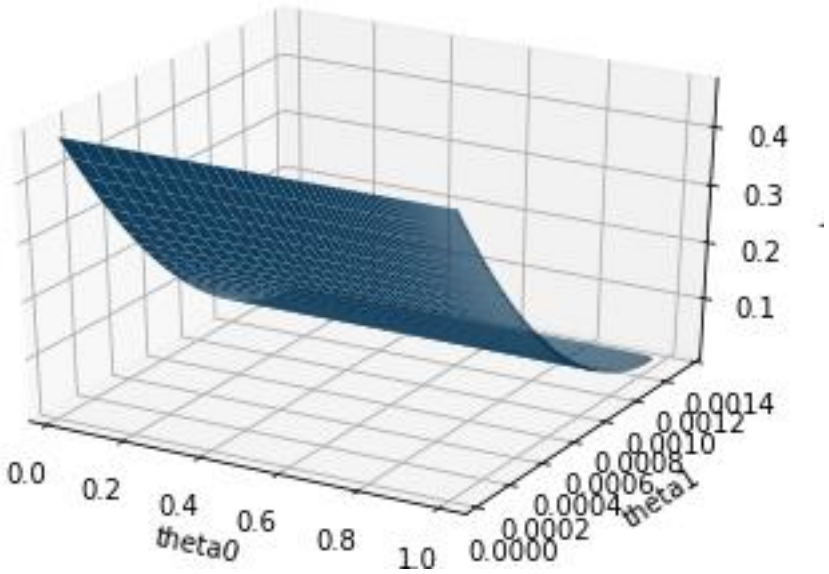
for eta=0.001:

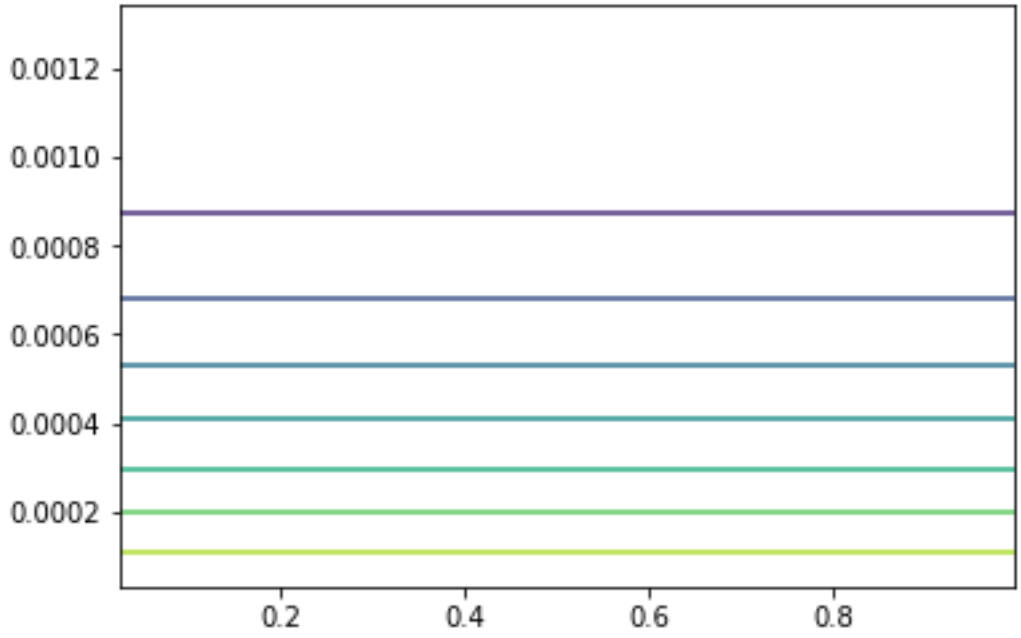




In this case the loss function curve is more slant rather than the dip in the middle.

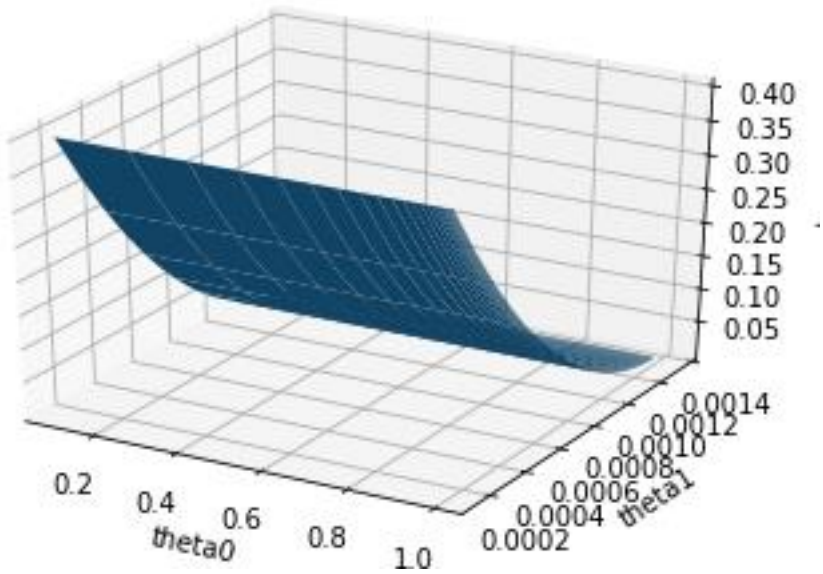
for eta=0.025:

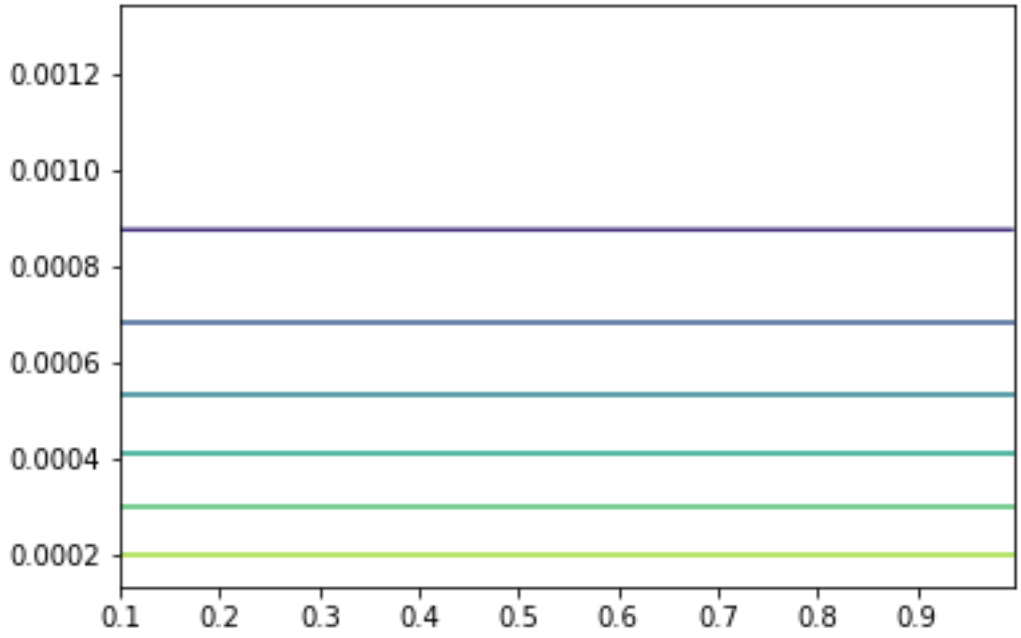




The curve of loss function is more curvy than the base curve.

For eta=0.1:





Observation:

These graphs are plotted after normalization of the input variables.

Before, normalization the graphs were more evident and dramatic to se the curves.

Q2)

a) Data is sampled as Xnos for input data

b) theta calculated in each case:

i. 3.01816

0.968684

2.06034

ii. 3.0427

0.954625

2.0554

iii. 1.99376

1.21999

1.9276

iv. 1.86543

1.2483

1.91702

Here I have kept the number of iterations as the (1000000/number of batches) for each batch

c) for case i and case ii, the theta parameters are almost equal to the original assigned theta

while for case iii and iv, they are not ran for long iterations.

By using the above computed theta from theta, we get cost as follows by running the test data provided:

1.2133679733288216

1.2391441344317442

3.901478667445829

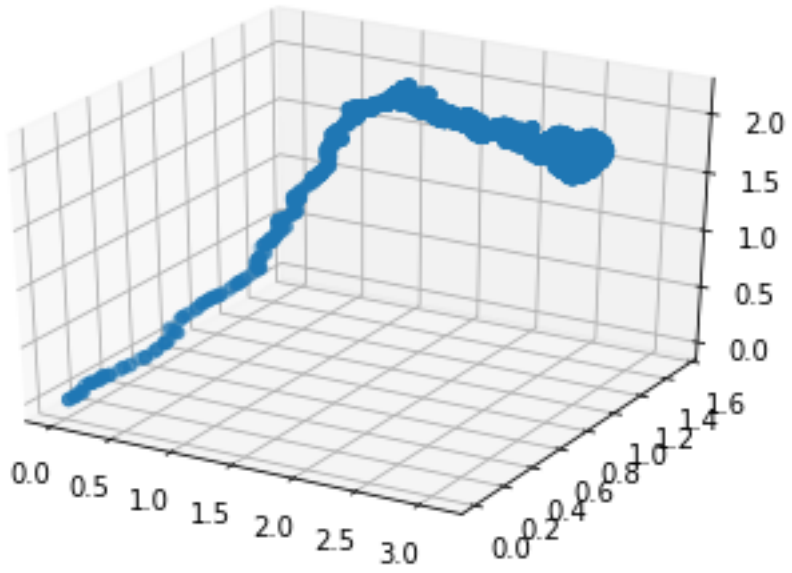
4.70960874794496

Error gets increased as going from case i to iv

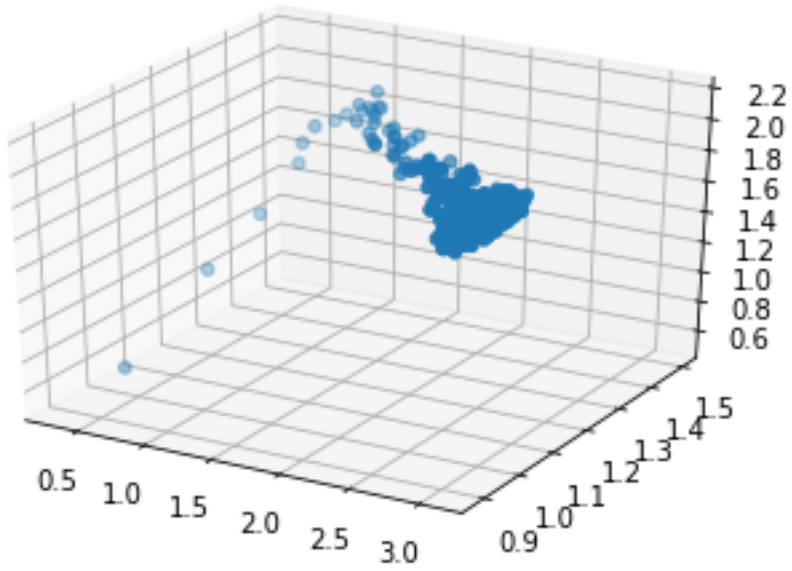
d)

3d plots for,

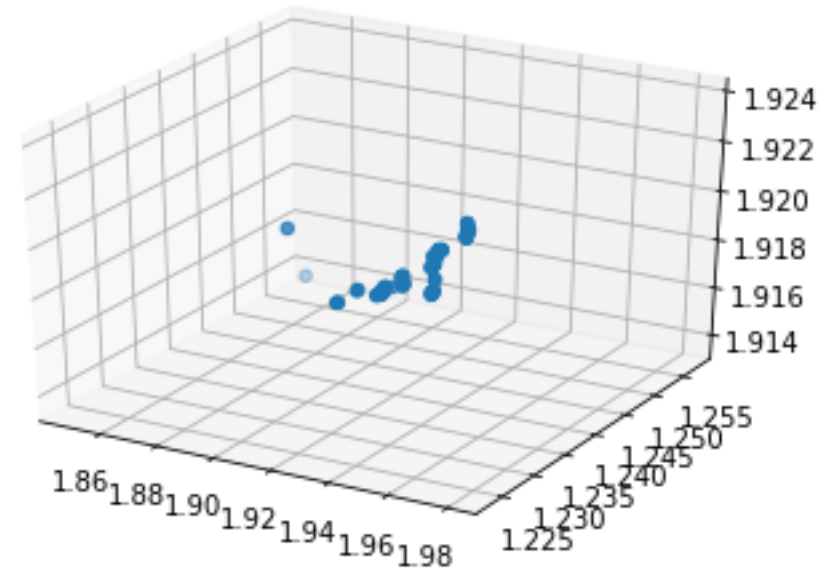
Case i:



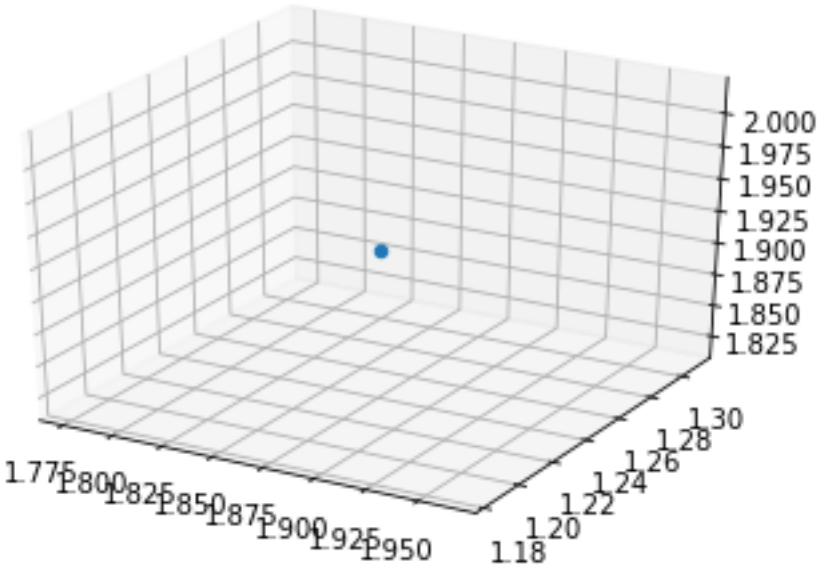
Case ii:



Case iii:



Case iv:



-has single iteration

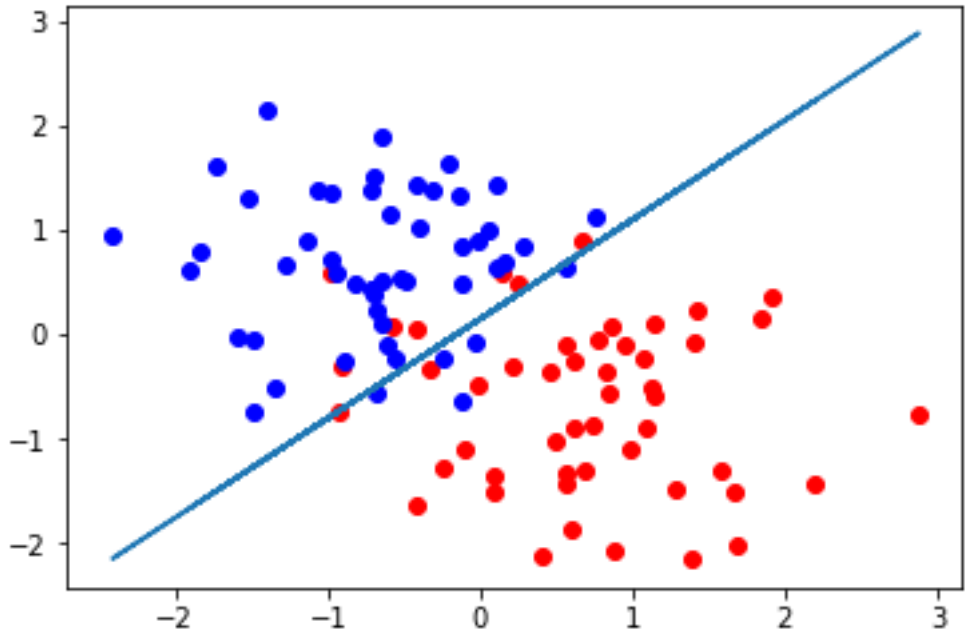
Q3) Logistic regression

Resulting thetas: **[[-0.]**

[-0.33839046]

[0.34126055]]

Figure for decision boundary:



Q4) Gaussian Discriminant Analysis

a)

phi= 0.5

mu0= -0.755294

0.685094

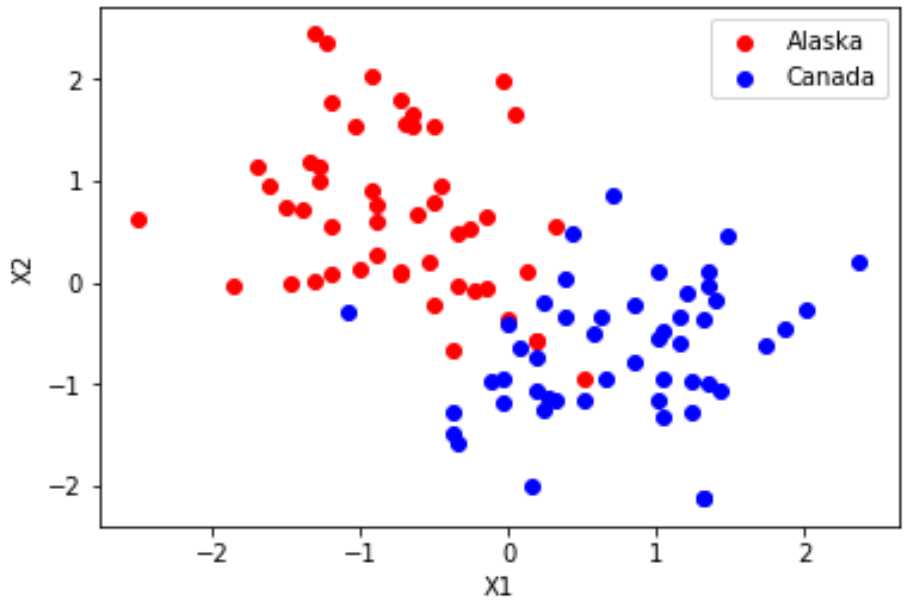
mu1= 0.755294

-0.685094

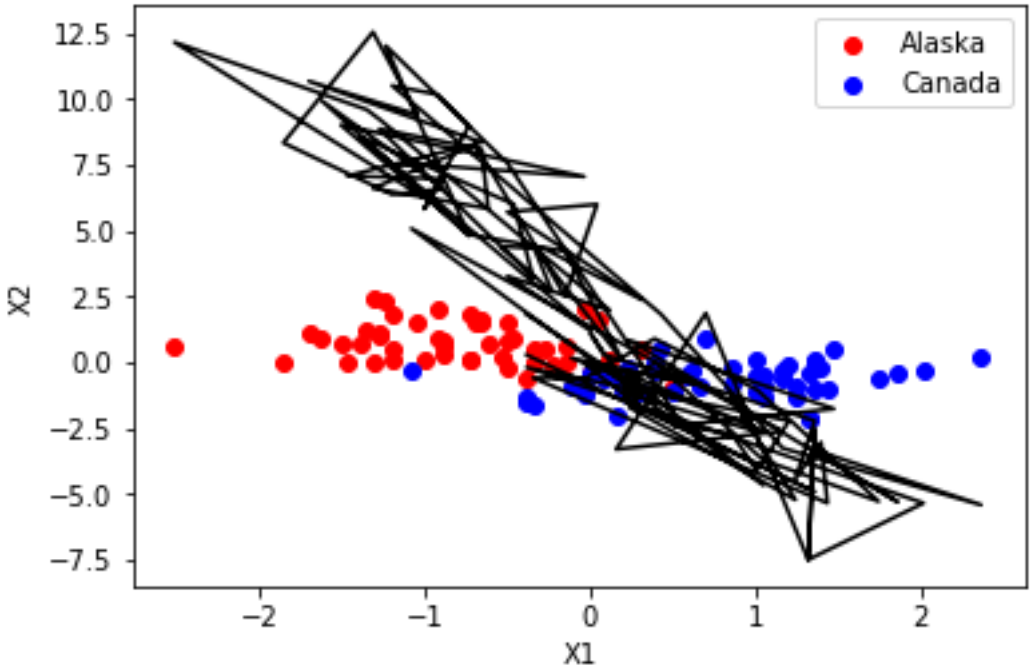
Covariance matrix= 0.42953 -0.0224723

-0.0224723 0.530646

b)



c)



d)

cov1= 2.90222 0.693882

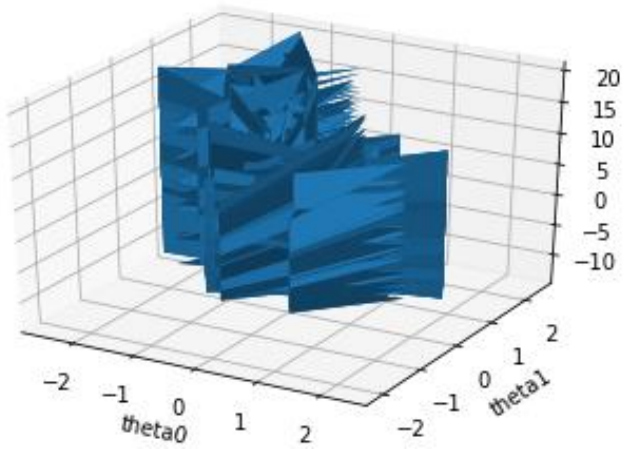
0.693882 1.70973

cov2= 2.23087 -0.592955

-0.592955 2.57567

mu0 and mu1 same as above

e)



- mesh for quadratic separator

f) as there was no proper separator for my implementation can't see on the plots.

But linear separator will be same as logistics separator