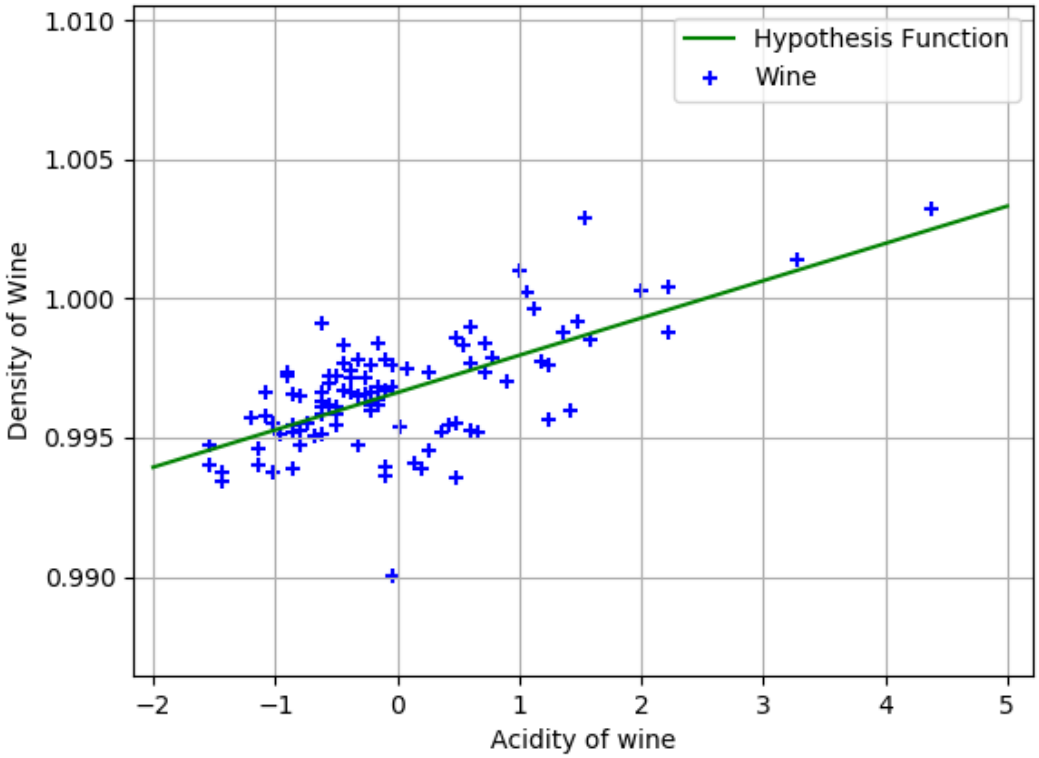
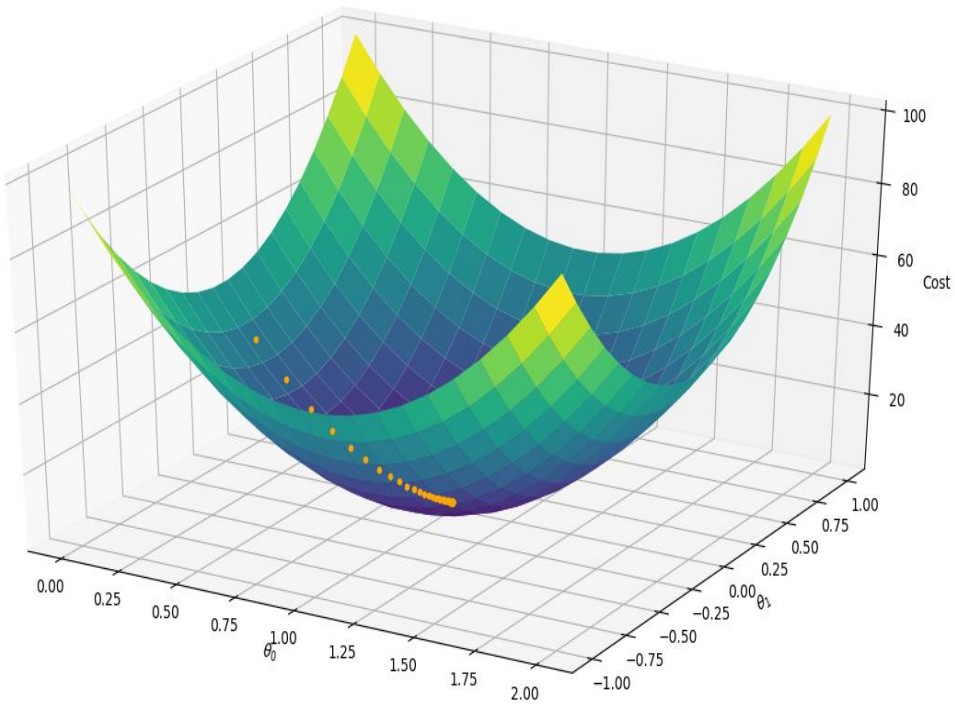
COL 774: Machine Learning. Assignment 1

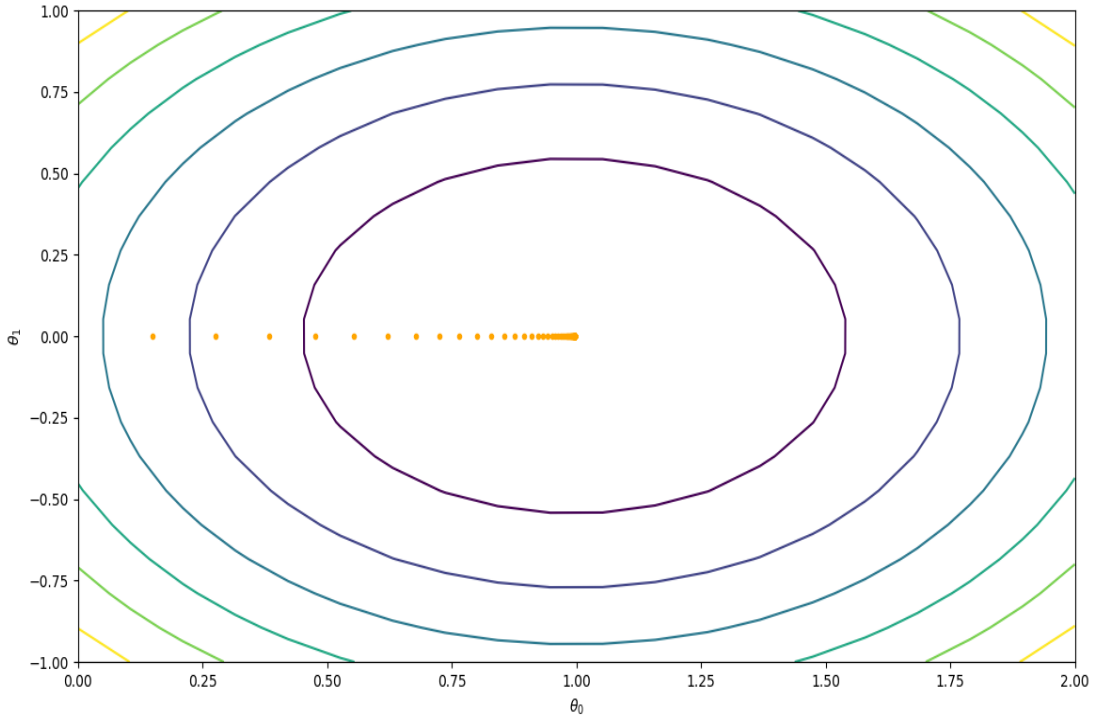
1. A) Learning rate = 0.0015,
   theta = $[0.9966201\ 0.0013402]^T$
   Convergence Condition: error < 0.001 or # of iterations > 1000
   B)



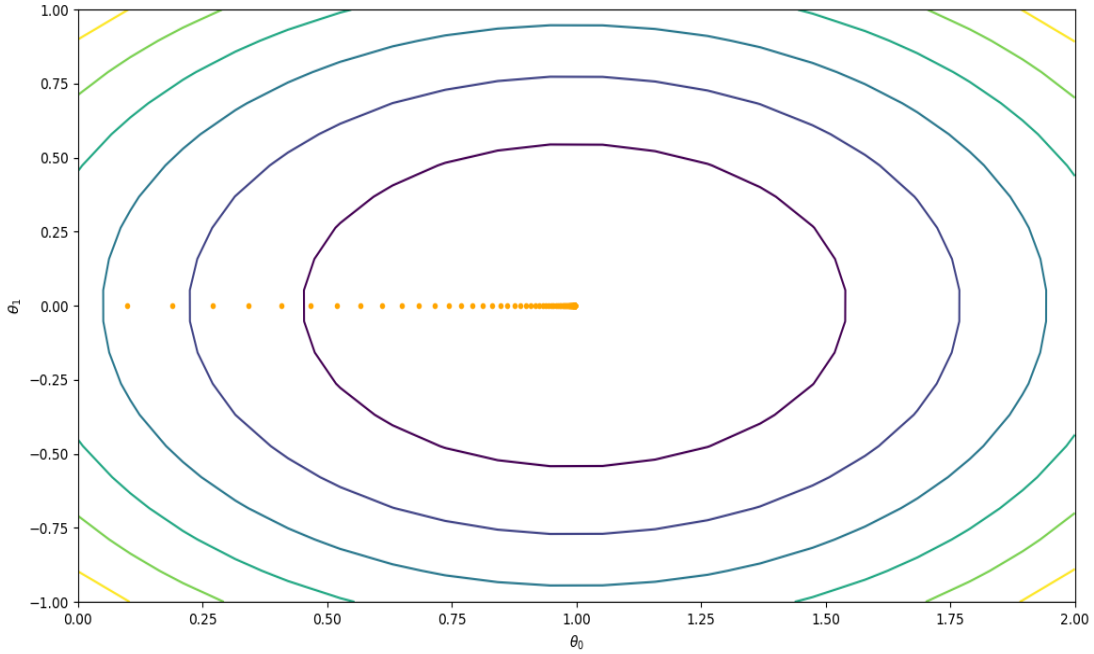
C)



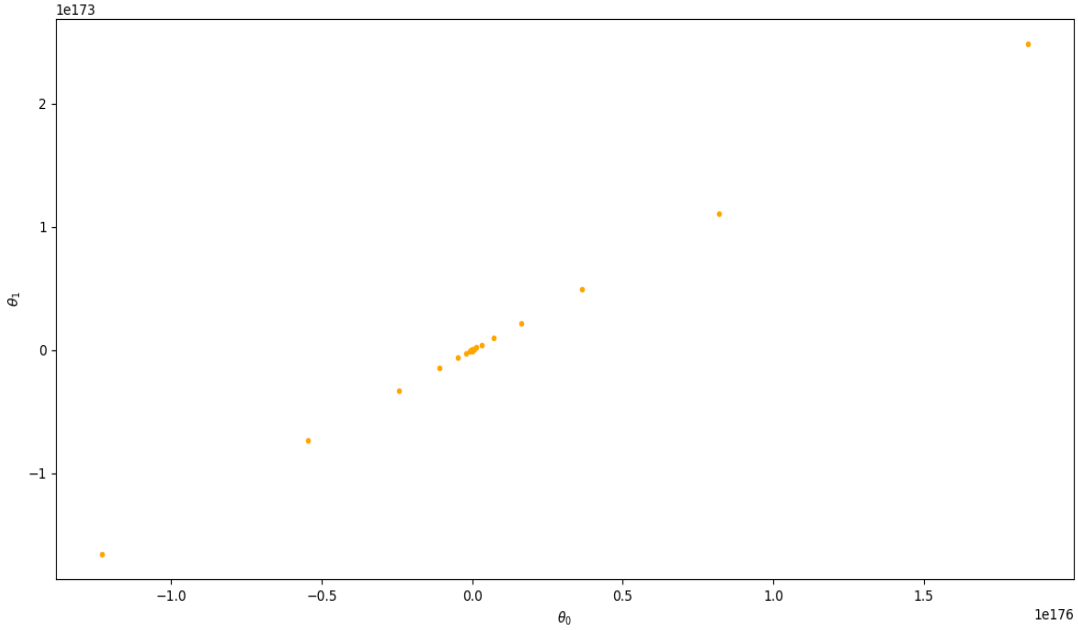
D) $\eta = 0.0015$



E) $\eta = 0.001$

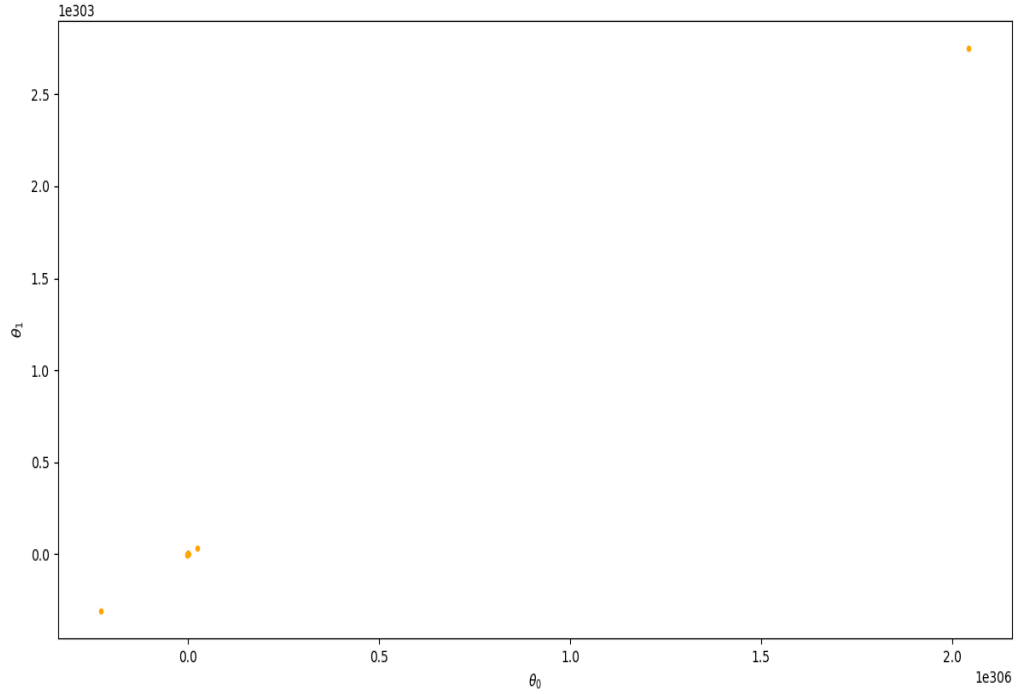


η = 0.025



η = 0.1

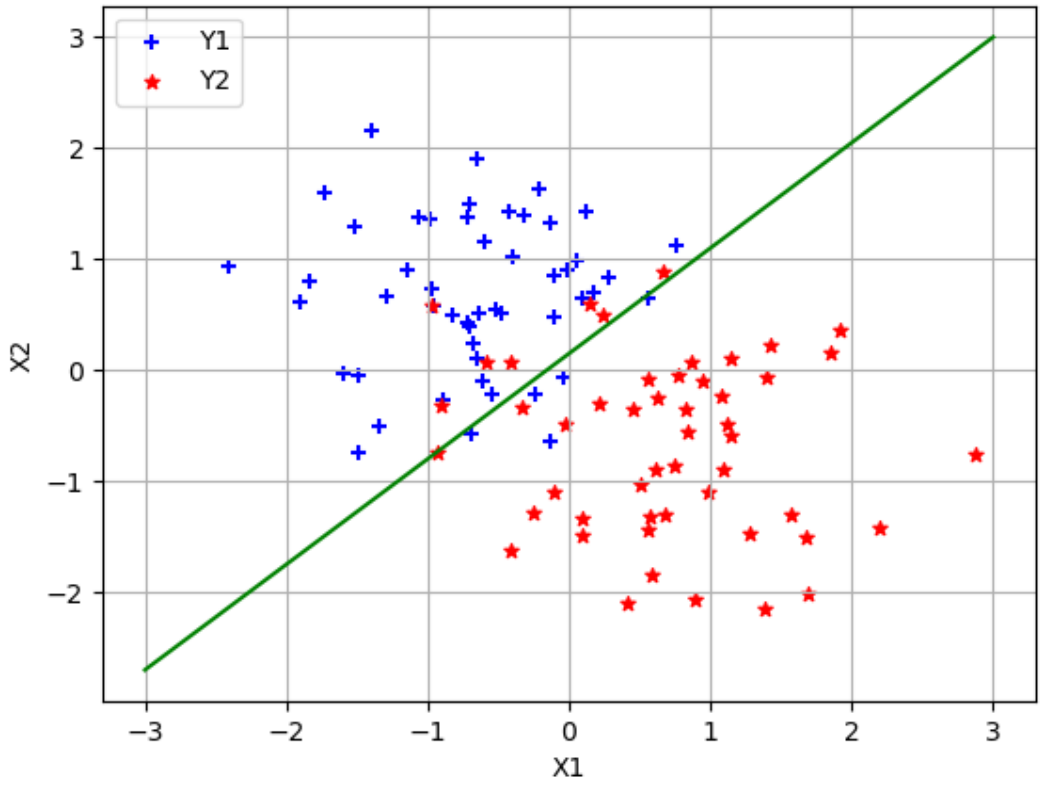
We observe that as the learning rate increases beyond a point the cost function starts diverging. Like for η = 0.001 its converging but diverging for η = 0.025 , 0.1 . Initially there are small oscillations and the function still converges . As η increases there are more oscillations near the minimum, each time overshooting even more .



3.)

a) theta = $[\,0.40125316 \;\; 2.5885477 \;\; -2.72558849]^{\mathrm{T}}$

b)



The green line is the decision Boundary

4.)

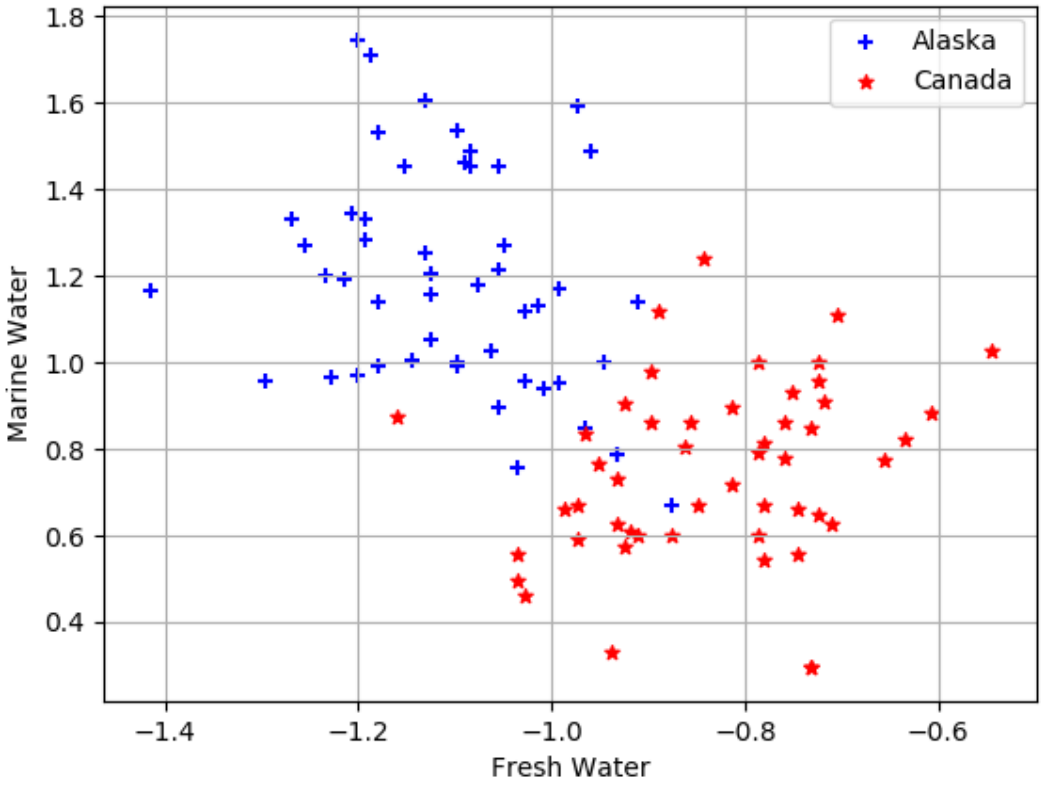
a) $\mu_0 = [-1.10106488 \;\; 1.18368785]^T$

$\mu_1 = [-0.8315402 \;\; 0.74891723]^T$

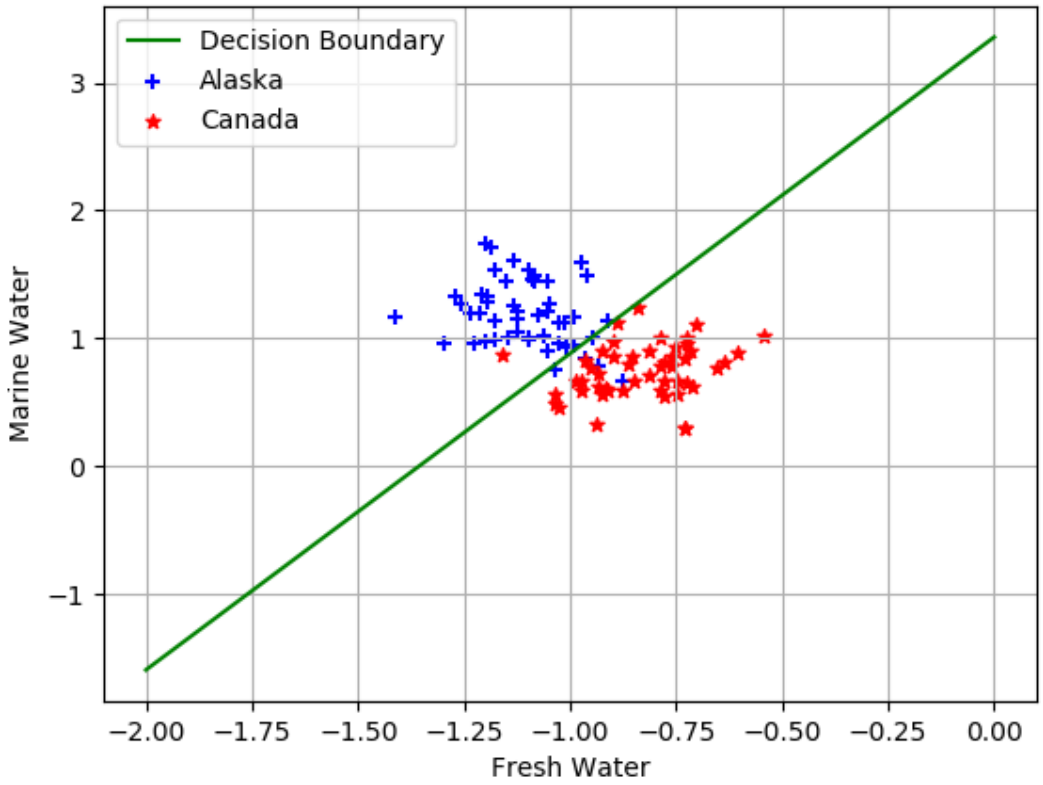
$\Sigma = [[\, 0.0136741 \;\; -0.00127227] \;\; [-0.00127227 \;\; 0.05342744]]$

y[i] = 0 corresponds to Alaska

b)


Alaska
Canada
Marine Water
Fresh Water
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
−1.4
−1.2
−1.0
−0.8
−0.6

c)



d)

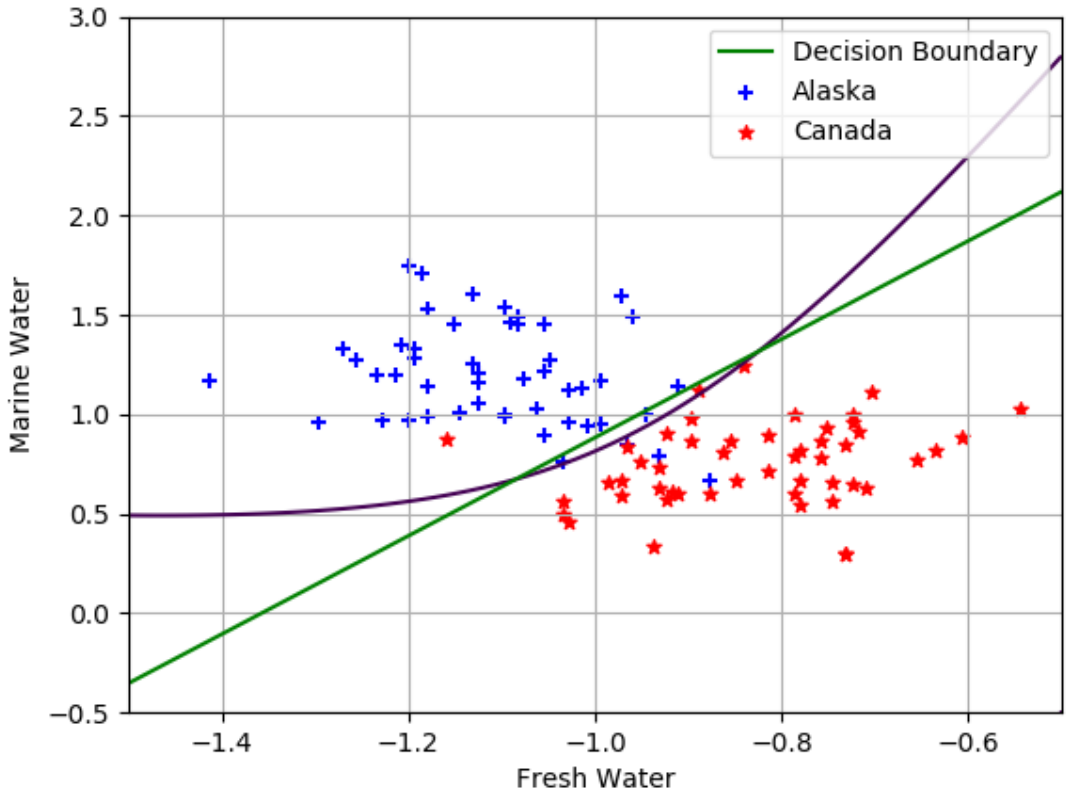
$\Sigma_0 = [[\ 0.0121479 \quad -0.0087677\ ]\ [-0.0087677 \quad 0.06521665]]$

$\Sigma_1 = [[0.01520029\ 0.00622316] \quad [0.00622316\ 0.04163824]]$

$\mu_0 = [-1.10106488 \quad 1.18368785]^T$

$\mu_1 = [-0.8315402 \quad 0.74891723]^T$

e)



f) Quadratic decision Boundary does a better job than the linear decision boundary (Logistic Regression). The assumption that the underlying data belong to gaussian distribution is valid to a great extent.

2)

b) [3.00371 0.99475 1.9981] 65183 1

[2.999 0.99 1.99] 28455 100

[2.922 1.01 1.99] 13100 10000

[2.983 1.1081 1.99547] 6674 1000000

Theta (# of iteration) r

c)Number of iterations inversely proportional to r

The value of theta very close to the given theta

Speed of convergence is inversely proportional to r

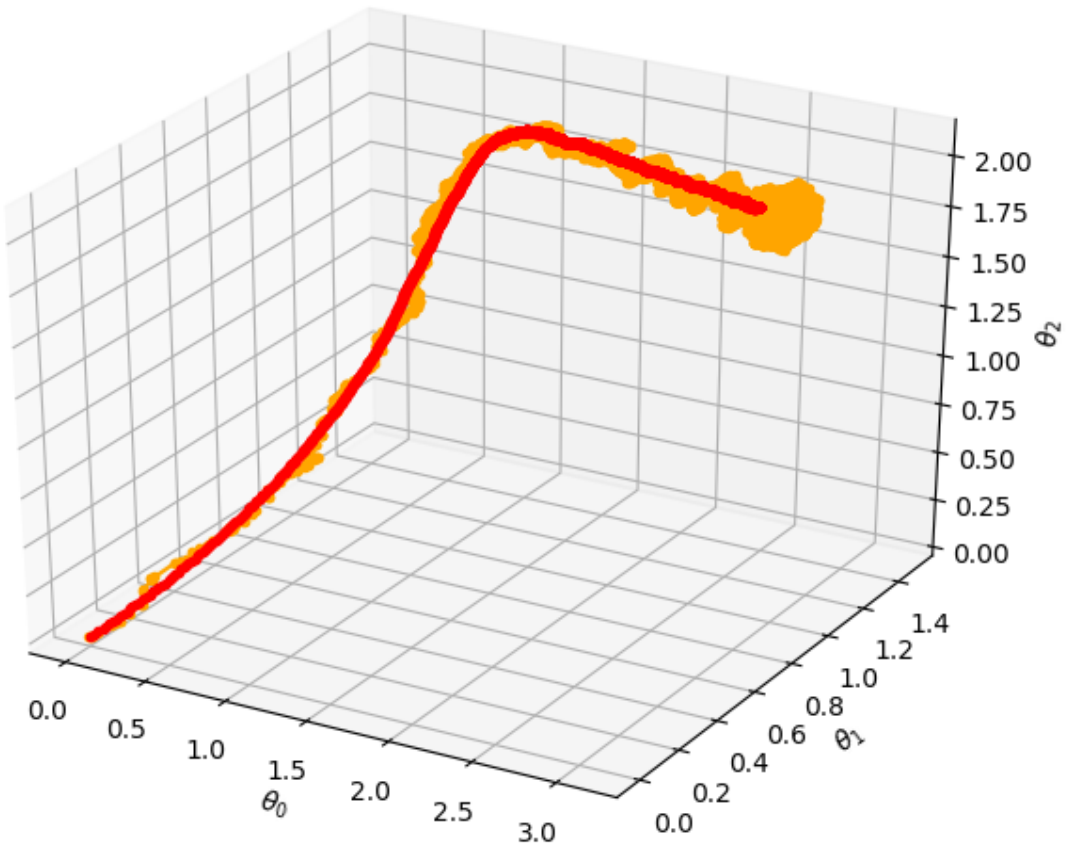
0.9850678759592312 – error for r = 1

0.9958009697879703 – error for r = 100

0.98 – error for original Hypothesis

SGD Converges to the original Hypothesis

d)



red corresponds to r = 100

orange corresponds to r = 1

It is noticed that SGD gets theta close to minimum but the parameters keep oscillating near the minimum value.

Report By:

Shivam Jadhav

2017CS10378