Assignment 1 Report

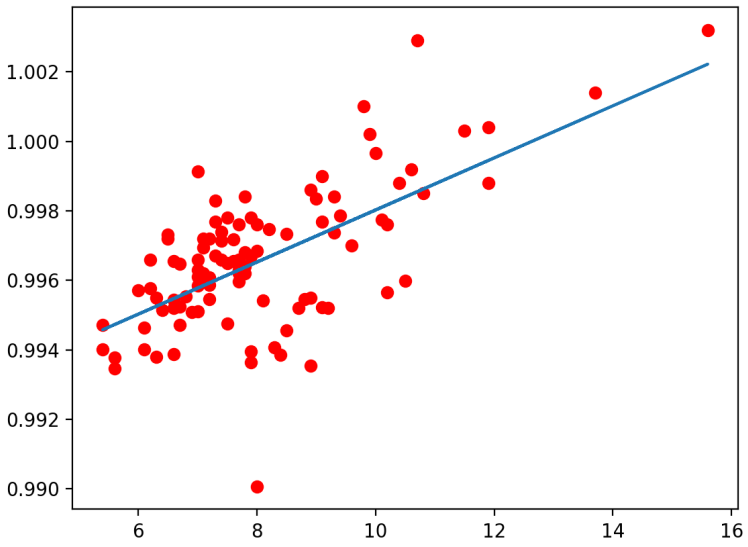
Pradyumna Meena - 2016CS10375

Q1

Observations for various learning rates where test-set consists the last 20 data-points from the given data-set

Learning Rate	Iterations	Cost (training-set) (10^{-6})	Cost (test-set) (10^{-6})
0.001	9220	1.4100187465131329	0.9759962195011227
0.003	3513	1.3305684468289872	0.8737000319310789
0.01	1206	1.3008637736256783	0.8186746547546557
0.03	448	1.2921961609153327	0.7928671381016129
0.1	148	1.2891887819930612	0.7777339925937648
0.3	52	1.288329165073939	0.7698203011594116
0.5	31	1.288173482202881	0.7674149549060678
0.9	16	1.2880859256030308	0.7654913305532566

- Stopping criteria: Change in cost < 10^{-9}
- Results
  - Learning Rate = 0.03 (To avoid having large steps in other cases)
  - Parameters = [*0.99657156 , 0.00222505*]



Q2

The equation can be written as $J(\theta) = (X\theta-Y)^T W(X\theta-Y)$

$= (\theta^T X^T - Y)W(X\theta-Y)$

$= \theta^T X^T W X\theta - \theta^T X^T WY - YWX\theta + YWY$

Put $\nabla_\theta J(\theta) = 0$

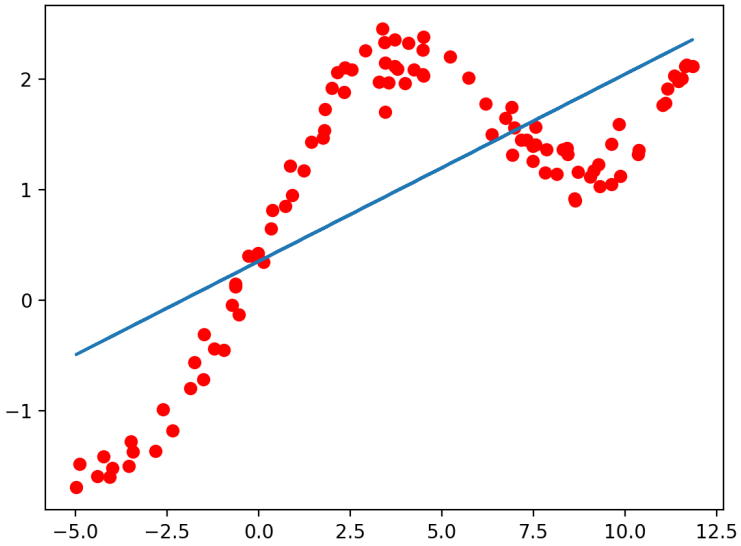
$\Rightarrow 2X^T WX\theta - X^T WY = 0$

$\Rightarrow 2\theta = (X^T W)^{-1}(X^T WY)$

Hence $\theta = ((X^T W)^{-1}(X^T WY))/2$

Part A:

- Convergence criteria: change in cost < 10^{-6}.

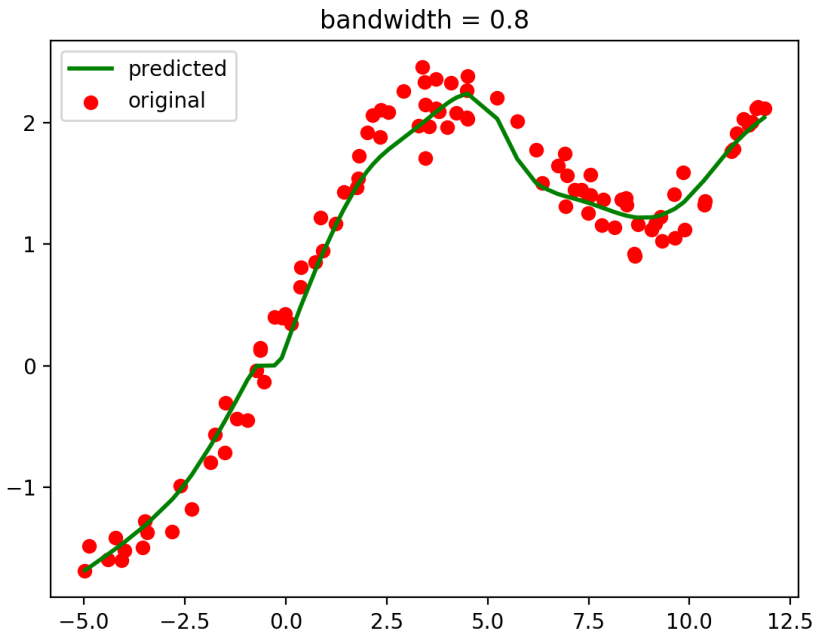


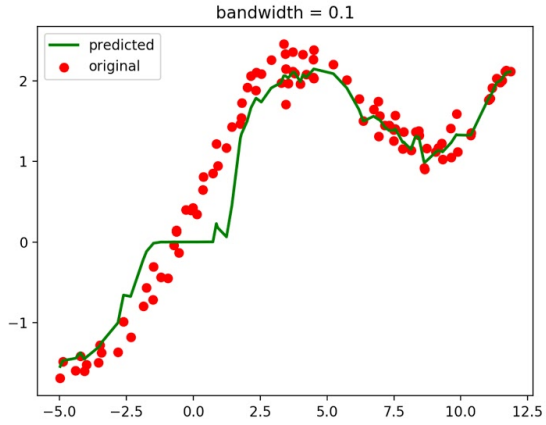
Learning Rate	Iterations	Cost (test-set) (10^{-6})	Cost (training-set) (10^{-6})
0.001	39777	300126.28362072624	351988.90767426966
0.003	17258	311764.03390232527	344707.5024437655
0.03	2563	325761.0696255268	341430.4292296926
0.3	338	330830.6240484761	341103.2128473918

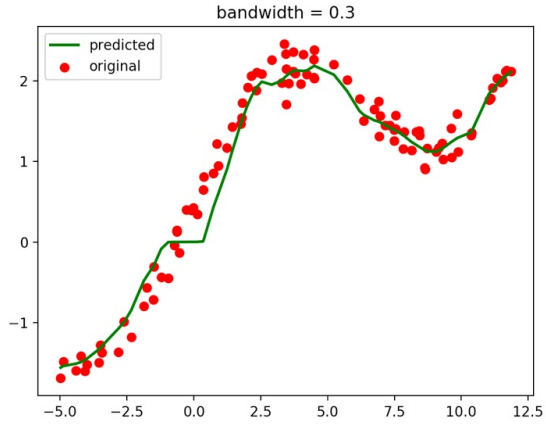
Since its a non-linear curve, using linear hypothesis is giving huge errors in prediction. Above plot is for ***α = 0.03*** and parameters obtained are [*1.03128116 , 3.84809705*].

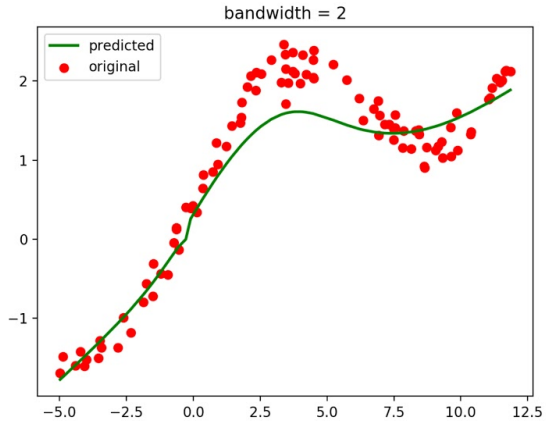
Part B:

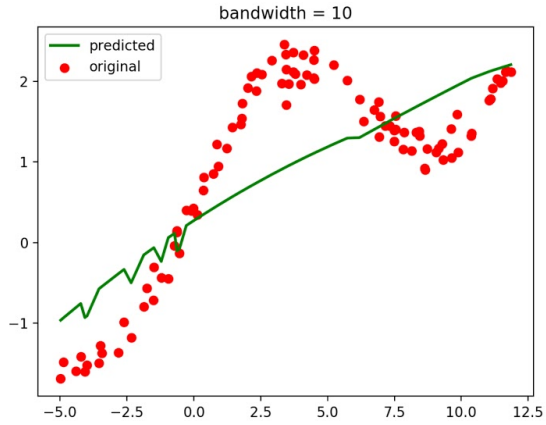
- Convergence criteria for underlying gradient descent: Change in predicted value<10^{-3}
- Learning rate for gradient descent = 0.03





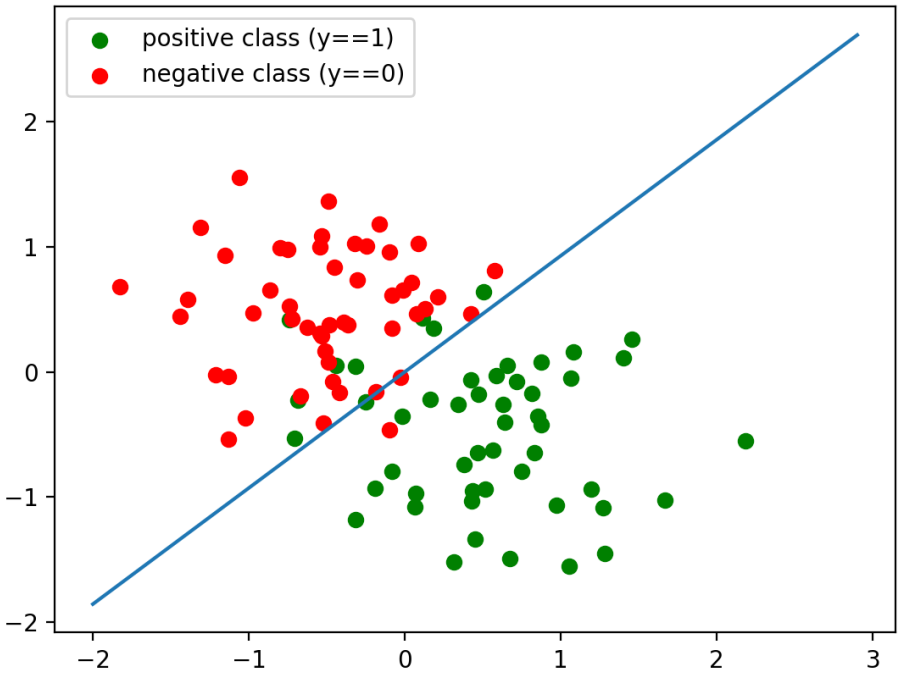






As the plot suggests choosing choosing 0.8 as the bandwidth parameter gives the best results.

Q3



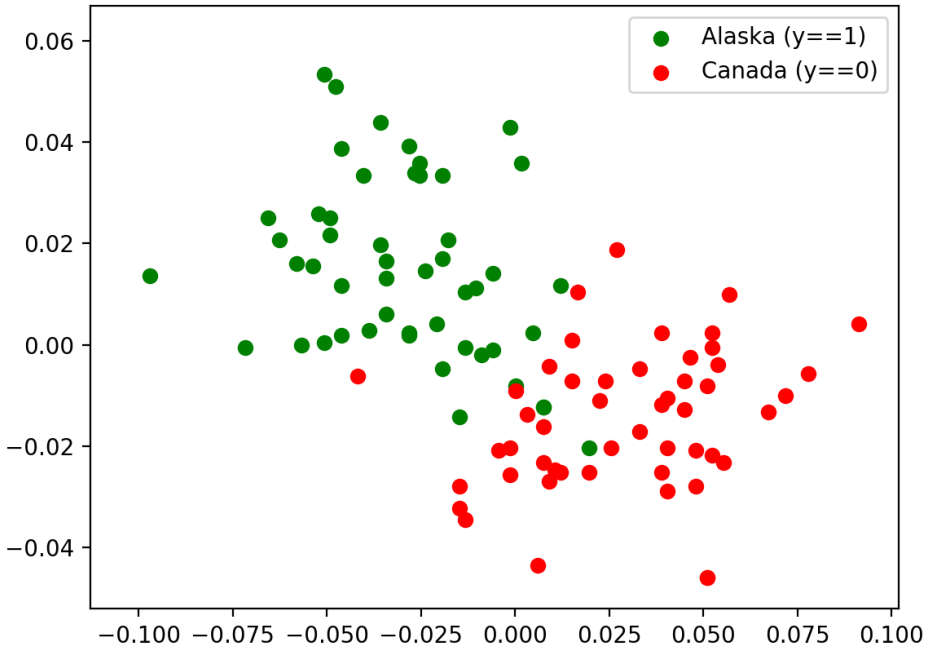
The parameters obtained are [-1.67966358e-16 , 1.18352831 , -1.27634840]

Q4

μ_0 = [1 , 0.02919496 , -0.01489068] μ_1 = [1 , -0.02919496 , 0.01489068]

$\sum_0$ = [[0. 0. 0.]
[0. 0.0007134 0.]
[0. 0. 0.00019537]]

$\sum_1$ = [[0. 0. 0.]
[0. 0.00057014 0.]
[0. 0. 0.000306]]


Alaska (y==1)
Canada (y==0)
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.100
−0.075
−0.050
−0.025
0.000
0.025
0.050
0.075
0.100