Robert Roche

HW1-PDF Writeup

Part 1:

1.)

X= $\begin{pmatrix} -2 \\ -5 \\ -3 \\ 0 \\ -8 \\ -2 \\ 1 \\ 5 \\ -1 \\ 6 \end{pmatrix}$ Y= $\begin{pmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{pmatrix}$ add bias to X X= $\begin{pmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{pmatrix}$ Y= $\begin{pmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{pmatrix}$

The formula to compute thetas is, $\theta = (X^TX)^{-1}(X^TY)$

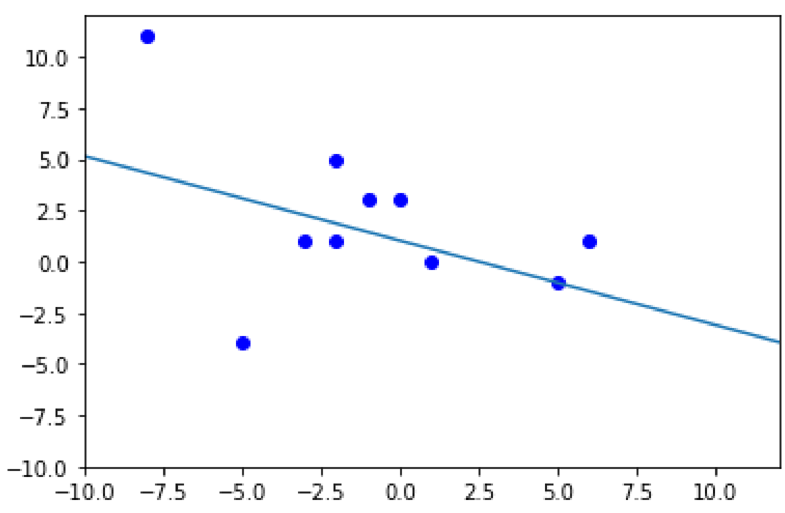
Using this with our X and Y matrices gives us the following θ matrix

$$\begin{pmatrix} 1.028589 \\ -0.4126787 \end{pmatrix}$$

$\theta_0 = 1.028589$, $\theta_1 = -0.4126787$

The equation for our line of best fit is: y = (-0.4126787)x + 1.028589

The line with the dataset points are plotted below



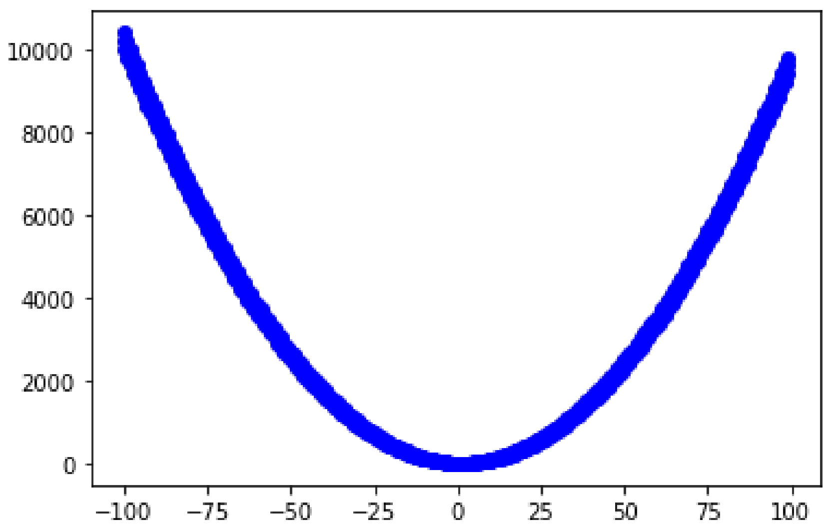
2.)

a. Given $J = (x_1 + x_2 - 2)^2$

$\frac{\partial J}{\partial x1}$ = $2x_1 + 2x_2 - 4$ $\frac{\partial J}{\partial x2}$ = $2x_1 + 2x_2 - 4$

For both x_1 and x_2: $\frac{\partial J}{\partial x1,2} (x_1 + x_2 - 2)^2 = 2(x_1 + x_2 - 2)(1) = 2x_1 + 2x_2 - 4$

b.



c. $X_1,X_2=1$ gives us the minimum of J which is 0

Part 2:

Final Model:

θ_0=2986.896551724138
θ_1=1088.4862496885312
θ_2=-341.67869902828295

$y=2986.89655+(1088.4862497)x_1+(-341.6786990)x_2$

RMSE=709.897098

Part 3:

RMSE=351.6779056674261

Part 4:

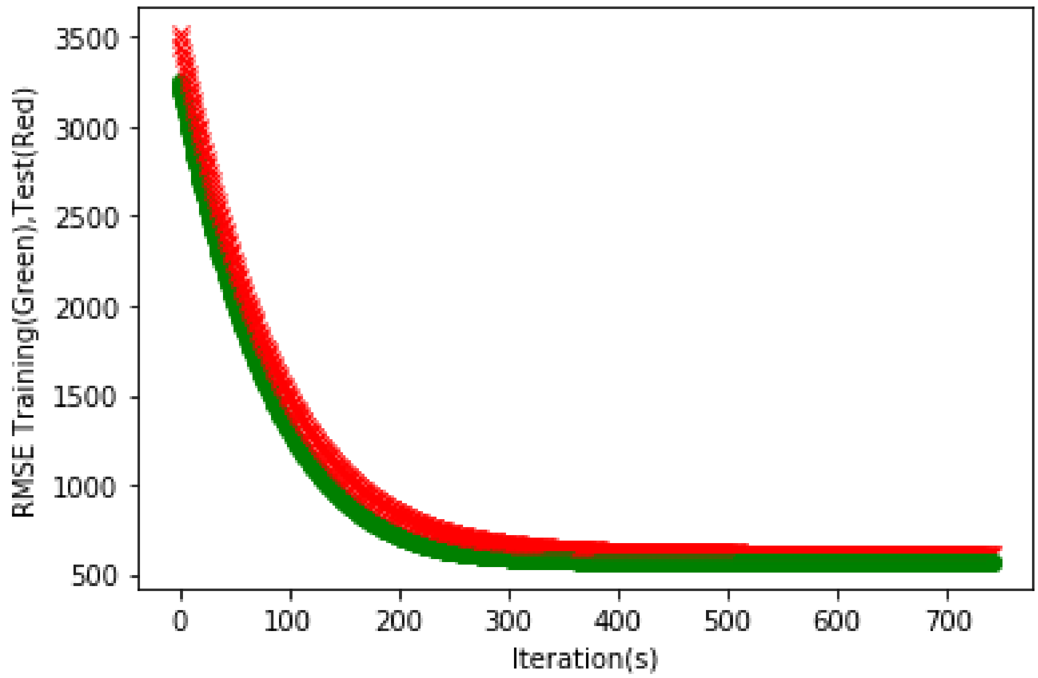
Final Model:

θ_0= 2964.90782605
θ_1= 1278.04815434
θ_2=- 200.89824176

$y=2964.90782605 +(1278.04815434)x_1+(-200.89824176)x_2$

RMSE=570.170565711668

From the plot we can see the training RMSE is reaching the convergent point faster than the testing RMSE.



3500
3000
2500
2000
1500
1000
500
RMSE Training(Green),Test(Red)
0
100
200
300
400
500
600
700
Iteration(s)