Avik Bag
Professor Matt Burlick
CS 383 – Assignment 3
9th February 2017

Question 1

Dataset

$$\begin{bmatrix} -2 & 1 \\ -5 & -4 \\ -3 & 1 \\ 0 & 3 \\ -8 & 11 \\ -2 & 5 \\ 1 & 0 \\ 5 & -1 \\ -1 & -3 \\ 6 & 1 \end{bmatrix}$$

$$\mu = -0.9$$
$$\sigma = 4.22$$

$$X(After\ standardizing\ and\ bias) = \begin{bmatrix} 1 & -0.26 \\ 1 & -0.97 \\ 1 & -0.497 \\ 1 & 0.213 \\ 1 & -1.679 \\ 1 & -0.26 \\ 1 & 0.449 \\ 1 & 1.395 \\ 1 & -0.024 \\ 1 & 1.632 \end{bmatrix}$$

$$Y = \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

Therefore, calculating the coefficients for linear regression is done by using the following equation.

$$\theta = (X^T X)^{-1} X^T Y$$

$$\theta = [1.399, -1.744]^T$$

Thus, the equation is

$$y = 1.399 - 1.744x$$

Question 2

RMSE calculated from closed form linear regression

$$\theta = [3275.674, 1097.618, -259.336]$$

$$y = 3275.674 + 1097.618x_1 - 259.336x_2$$

RMSE = 601.94

Question 3

RMSE calculated from S-Fold Cross Validation

RMSE = 714.10