CS 383 HW 3

Rakeen Rouf

Part 1: Theory

Question 1

X = $\begin{bmatrix} -2 \\ -5 \\ -3 \\ 0 \\ -8 \\ -2 \\ 1 \\ 5 \\ -1 \\ 6 \end{bmatrix}$, Y= $\begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$

Since we are only dealing with one feature, we do not need to standardize X.

X = [1 X]

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

Calculating theta:

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

$X^T X$ = $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix}$ = $\begin{bmatrix} 10 & -9 \\ -9 & 168 \end{bmatrix}$

$(X^T X)^{-1}$ = $\frac{1}{1599} \begin{bmatrix} 168 & 9 \\ 9 & 10 \end{bmatrix}$ = $\begin{bmatrix} 0.105 & 0.0056 \\ 0.0056 & 0.0063 \end{bmatrix}$

$(X^T X)^{-1} X^T$ = $\begin{bmatrix} 0.105 & 0.0056 \\ 0.0056 & 0.0063 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix}$

$$= \begin{bmatrix} 0.0938 & 0.077 & 0.0882 & 0.105 & 0.0602 & 0.0938 & 0.1106 & 0.133 & 0.0994 & 0.1386 \\ -0.007 & -0.0259 & -0.0133 & 0.0056 & -0.0448 & -0.007 & 0.0119 & 0.0371 & -0.0007 & 0.0434 \end{bmatrix}$$

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

=

$$\begin{bmatrix} 0.0938 & 0.077 & 0.0882 & 0.105 & 0.0602 & 0.0938 & 0.1106 & 0.133 & 0.0994 & 0.1386 \\ -0.007 & -0.0259 & -0.0133 & 0.0056 & -0.0448 & -0.007 & 0.0119 & 0.0371 & -0.0007 & 0.0434 \end{bmatrix} \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1.0276 \\ -0.4193 \end{bmatrix}$$

Final Model:

Y = 1.0279 – 0.4193X

Question 2

a) $g(x,y) = (x + y - 2)^2$
Using chain rule
$$\frac{\partial g}{\partial x} = 2(x + y - 2)\left(\frac{\partial(x + y - 2)}{\partial x}\right)$$
$$= 2(x + y - 2)(1)$$
$$= 2(x + y - 2)$$

$g(x,y) = (x + y - 2)^2$
Using chain rule
$$\frac{\partial g}{\partial y} = 2(x + y - 2)\left(\frac{\partial(x + y - 2)}{\partial x}\right)$$
$$= 2(x + y - 2)(1)$$
$$= 2(x + y - 2)$$

b)

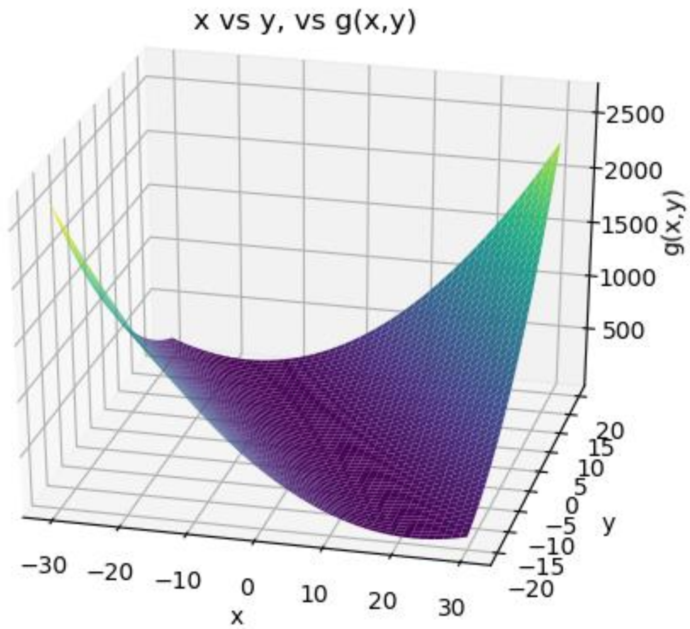

Figure 1: 3D plot of x vs y, vs g(x,y)

c) X =- -16.24, y=18.24

Part 2: Gradient Descent

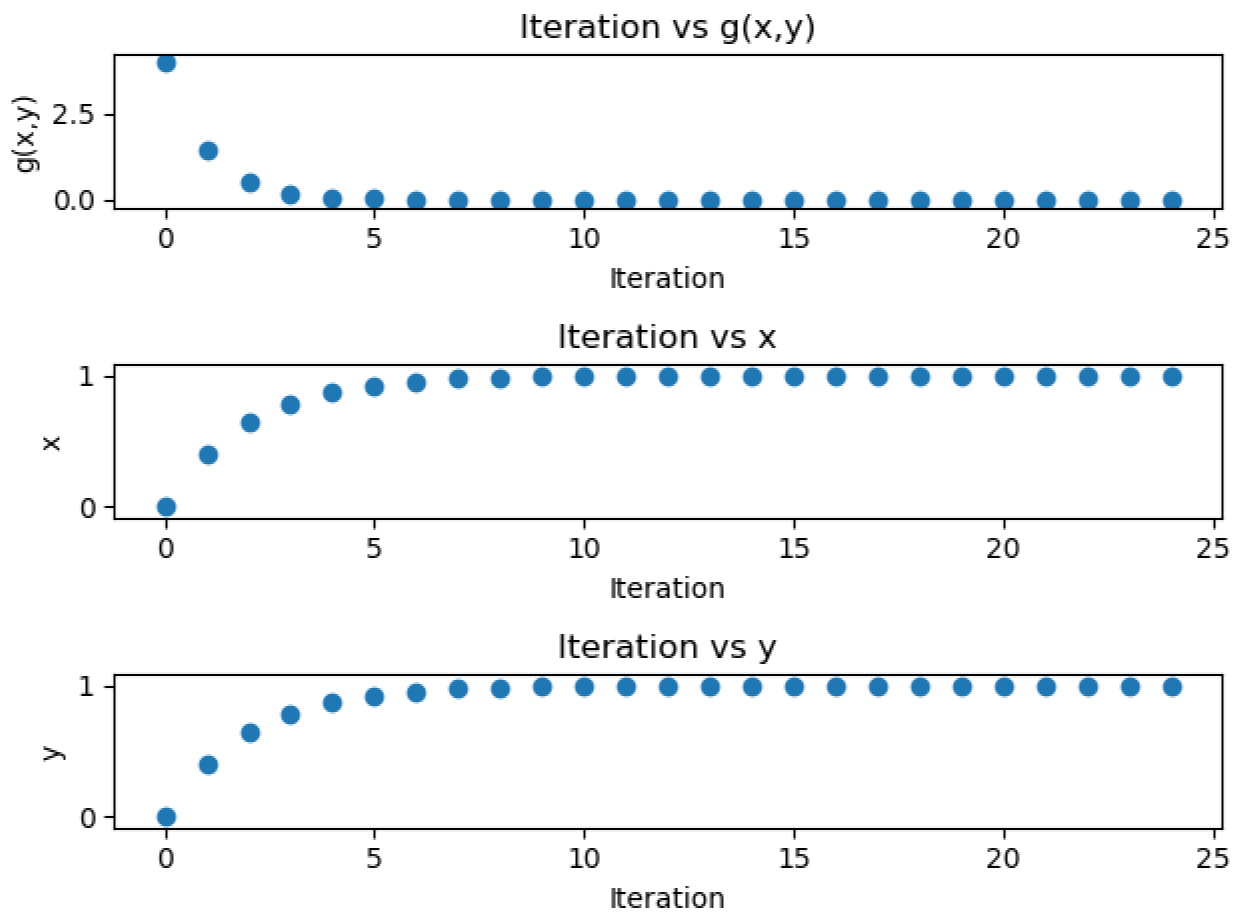

Figure 2: Part 2 results

Part 3: Closed Form Linear Regression

```
Final Model: Y = 81.80297668522786 + 8.57422120068162*X1 + 40.00036910523971*X2
RMSE: 20.15734398732259
```

Part 4: S-Folds Cross Validation

```
Number of fold: 3 ; Average of RMSE: 21.34318141314621; STD of RMSE: 1.2103632892385927
Number of fold: 5 ; Average of RMSE: 20.55948287382471; STD of RMSE: 0.39185041847749363
Number of fold: 20 ; Average of RMSE: 18.516630476065806; STD of RMSE: 0.6319608707633639
Number of fold: 44 ; Average of RMSE: 16.3148765559188; STD of RMSE: 4.203627514058621e-15
```