HW 3 - Linear Regression

Martin Kraus

February 21, 2020

1 Theory

1.1 Compute the coefficients for the linear regression using least squares estimate (LSE)?

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -5 \\ 1 & -3 \\ 1 & 0 \\ 1 & -8 \\ 1 & -2 \\ 1 & 1 \\ 1 & 5 \\ 1 & -1 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 10 & -9 \\ -9 & 169 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 10 & -9 \\ -9 & 169 \end{bmatrix}^{-1} = \frac{1}{1690 - 81} \begin{bmatrix} 169 & 9 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 0.1050 & 0.0056 \\ 0.0056 & 0.0062 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -2 & -5 & -3 & 0 & -8 & -2 & 1 & 5 & -1 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ -4 \\ 1 \\ 3 \\ 11 \\ 5 \\ 0 \\ -1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 14 \\ -79 \end{bmatrix}$$

$$\theta = (X'X)^{-1}X'Y = \begin{bmatrix} 0.1050 & 0.0056 \\ 0.0056 & 0.0062 \end{bmatrix} \begin{bmatrix} 14 \\ -79 \end{bmatrix} = \begin{bmatrix} 1.0286 \\ -0.4127 \end{bmatrix}$$

1.2 For the function $J = (x_1 + x_2 - 2)^2$

1.2.1 What are the partial gradients, $\frac{\partial J}{\partial x_1}$ and $\frac{\partial J}{\partial x_2}$?

$$\begin{aligned}\frac{\partial J}{\partial x_1} &= \frac{\partial}{\partial x_1}(x_1 + x_2 - 2)^2 \\ &= 2(x_1 + x_2 - 2)(1 + 0 + 0) \\ &= 2(x_1 + x_2 - 2)\end{aligned}$$

$$\begin{aligned}\frac{\partial J}{\partial x_2} &= \frac{\partial}{\partial x_2}(x_1 + x_2 - 2)^2 \\ &= 2(x_1 + x_2 - 2)(0 + 1 + 0) \\ &= 2(x_1 + x_2 - 2)\end{aligned}$$

1.2.2 Create a 3D plot of x_1 vs x_2, vs J

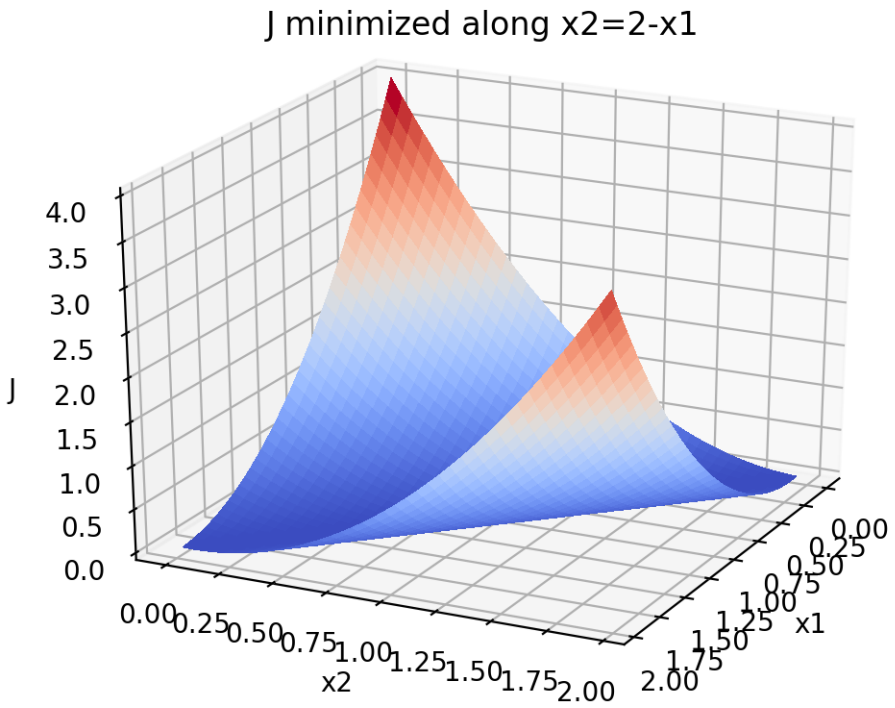

Figure 1: 3d plot of x_1, x_2, and J. J is minimized $\forall x_1, x_2 : x_2 = 2 - x_1$

1.2.3 Based on your plot, what are the values of x_1 and x_2 that minimize J

According to figure 1 and the equation $J = (x_1 + x_2 - 2)^2$, J is minimized when $x_2 = 2 - x_1$.

2 Gradient Descent

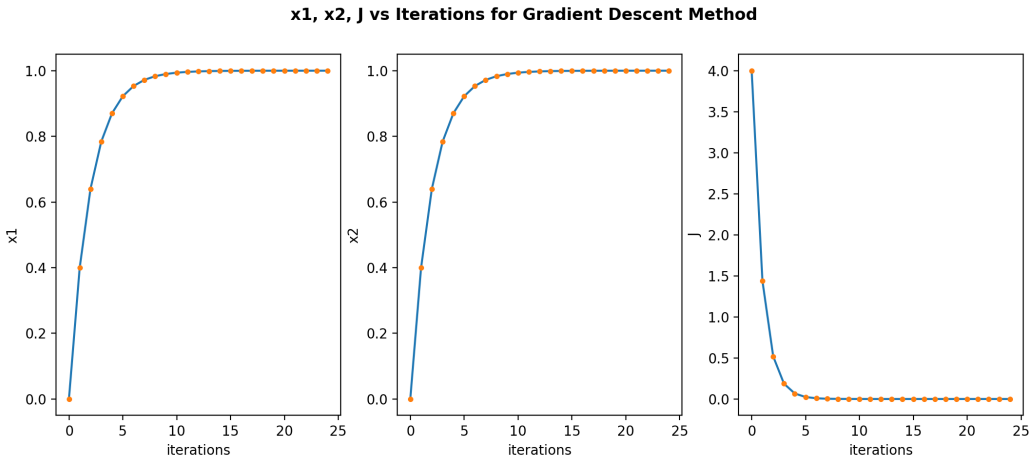

Figure 2: Graphs of x_1, x_2, J vs iterations. x_1 and x_2 approach 1 as J minimizes at 24 iterations

3 Closed Form Linear Regression

According the results of the code for the closed form linear regression

$y = -131.0496 + 4.1599x_1 + .0382x_2$

RMSE: 20.068

4 S-Folds Cross-Validation

4.1 The average and standard deviation of the root mean squared error for $S = 2$ over the 20 different seed value

$Avg = 22.4175$
$Std = 1.0246$

4.2 The average and standard deviation of the root mean squared error for $S = 4$ over the 20 different seed value

$Avg = 22.0495$
$Std = 0.5596$

4.3 The average and standard deviation of the root mean squared error for $S = 22$ over the 20 different seed value

$Avg = 21.9020$
$Std = 0.1263$

4.4 The average and standard deviation of the root mean squared error for $S = 44$ over the 20 different seed value

$Avg = 21.8804$
$Std = 0.00$