Assignment 3

Nabin Chapagain
1001551151

October 1, 2020

1 TASK 1

Program is attached along with this file in the same folder.

for degree =1, λ =0:

- w0 = -6.3872
- w1 = 0.0276
- w2 = 0.0432
- w3 = 0.0126
- w4 = 0.0176
- w5 = 0.0080
- w6 = -0.0058
- w7 = -0.0081
- w8 = 0.0714
- w9 = -0.0153
- w10 = -0.0190
- w11 = 0.0117
- w12 = 0.0222
- w13 = -0.0018
- w14 = -0.0013
- w15 = 0.0091
- w16 = 0.0382

Last line being:
ID= 3498, output= 3.8514, target value = 4.0000, squared error = 0.0221

for degree = 1, λ = 1:

- w0 = -6.2611
- w1 = 0.0275
- w2 = 0.0428
- w3 = 0.0126
- w4 = 0.0172

- w5 = 0.0078
- w6 = -0.0059
- w7 = -0.0081
- w8 = 0.0713
- w9 = -0.0154
- w10 = -0.0191
- w11 = 0.0116
- w12 = 0.0221
- w13 = -0.0018
- w14 = -0.0017
- w15 = 0.0090
- w16 = 0.0383

last line:
ID= 3498, output= 3.8528, target value = 4.0000, squared error = 0.0217

for degree = 2 λ **= 0**:

- w0 = -7.5608
- w1 = 0.0223
- w2 = 0.0001
- w3 = 0.0352
- w4 = 0.0000
- w5 = 0.0049
- w6 = -0.0000
- w7 = -0.0299
- w8 = 0.0002
- w9 = 0.0327
- w10 = -0.0001
- w11 = 0.0694
- w12 = -0.0004
- w13 = 0.0079
- w14 = -0.0002
- w15 = 0.0596
- w16 = -0.0003
- w17 = -0.0184
- w18 = -0.0000
- w19 = 0.0093
- w20 = 0.0002
- w21 = 0.0162
- w22 = -0.0000
- w23 = 0.0398
- w24 = -0.0002
- w25 = -0.0041
- w26 = 0.0001
- w27 = 0.0538
- w28 = -0.0007

- w29 = -0.0149
- w30 = 0.0002
- w31 = 0.1215
- w32 = -0.0007

last line:
ID= 3498, output= 3.6074, target value = 4.0000, squared error = 0.1542

for degree = 2 λ = 1:

- w0 = -7.0384
- w1 = 0.0219
- w2 = 0.0001
- w3 = 0.0310
- w4 = 0.0001
- w5 = 0.0043
- w6 = -0.0000
- w7 = -0.0345
- w8 = 0.0002
- w9 = 0.0315
- w10 = -0.0001
- w11 = 0.0678
- w12 = -0.0004
- w13 = 0.0077
- w14 = -0.0002
- w15 = 0.0574
- w16 = -0.0003
- w17 = -0.0192
- w18 = -0.0000
- w19 = 0.0091
- w20 = 0.0002
- w21 = 0.0156
- w22 = -0.0000
- w23 = 0.0401
- w24 = -0.0002
- w25 = -0.0050
- w26 = 0.0001
- w27 = 0.0536
- w28 = -0.0007
- w29 = -0.0155
- w30 = 0.0002
- w31 = 0.1208
- w32 = -0.0007

last line:
ID= 3498, output= 3.6001, target value = 4.0000, squared error = 0.1599

2 TASK 2

Qsn: We are given these training examples for a linear regression problem:
$x_1 = 5.3$, $t_1 = 9.6$ $x_2 = 7.1$, $t_2 = 4.2$ $x_3 = 6.4$, $t_3 = 2.2$
We just want to fit a line to this data, and we want to find the 2-dimensional vector w that minimizes $E_D(w)$ as defined in slide 60 of the linear regression slides. What is the value of w in the limit where λ approaches positive infinity? Justify your answer. Correct answers with insufficient justification will not receive credit.

Ans: We know, as it is given in side 61 that the value of w that minimizes $E_D(w)$ is:

$$w = \left(\lambda I + \phi^T \phi\right)^{-1} \phi^T t \tag{2.1}$$

for $\lambda = 0$, it will just be like a linear regression without any regularization and it would go like:

$$w = \left(\phi^T \phi\right)^{-1} \phi^T t$$

for $\lambda = 1, 2, 3$, the same equation will be:

$$w = \left(I + \phi^T \phi\right)^{-1} \phi^T t$$

$$w = \left(2I + \phi^T \phi\right)^{-1} \phi^T t$$

$$w = \left(3I + \phi^T \phi\right)^{-1} \phi^T t$$

And, as $\lim_{\lambda \to \infty}$,

$$w \to \left(\infty I + \phi^T \phi\right)^{-1} \phi^T t$$
$$w \to (\infty)^{-1} \phi^T t$$

A little sidenote here:
for $A = \begin{pmatrix} j & 0 \\ 0 & j \end{pmatrix}$, $A^{-1} = \begin{pmatrix} \frac{1}{j} & 0 \\ 0 & \frac{1}{j} \end{pmatrix}$ and as j approaches ∞, $A^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ So, w will be a zero matrix of M× 1 dimensions where, M is the dimension of ϕ. It will be a column matrix at the end, thanks to 't'. It can be seen as:

$$lim_{\lambda \to \infty} w = \emptyset_{M \times M} \phi^T t$$

$$lim_{\lambda \to \infty} w = \emptyset_{M \times M} t_{M \times 1}$$

$$lim_{\lambda \to \infty} w = \emptyset_{M \times 1}$$

So, to sum everything up, as $\lim_{\lambda \to \infty}$, $w \to \emptyset_{M \times 1}$ where, $\emptyset_{M \times 1}$ is a zero matrix.

3 Task 3

Qsn: We are given these training examples for a linear regression problem:
$x_1 = 5.3$, $t_1 = 9.6$
$x_2 = 7.1$, $t_2 = 4.2$
$x_3 = 6.4$, $t_3 = 2.2$

We are also given these two lines as possible solutions:

$$f(x) = 3.1x + 4.2 \tag{3.1}$$

$$f(x) = 2.4x - 1.5 \tag{3.2}$$

Which of these lines is a better solution according to the sum-of-squares criterion? This criterion is defined as function $E_D(w)$ in slide 25 of the linear regression slides. Justify your answer. Correct answers with insufficient justification will not receive credit.

Ans: We will be using formula for standard sum of squares error for both equations 3.1 and 3.2.
We have the formula for $E_D(w)$ as:

$$E_D(w) = \frac{1}{2}\sum_{n=1}^{N}\left[\left(t_n - w^T\phi(x_n)\right)^2\right] \tag{3.3}$$

So, we start with equation 3.1 where $w = \begin{pmatrix} 3.1 \\ 4.2 \end{pmatrix}$. Now, we use each t_n and x_n and plug them into the formula.

Let's calculate $t_n - w^T\phi(x_n)^2$ first for each data set given.

$$a)\left(9.6 - \begin{pmatrix} 3.1 & 4.2 \end{pmatrix}\begin{pmatrix} 1 \\ 5.3 \end{pmatrix}\right)^2$$

$$= (9.6 - (3.1 + 4.2 \times 5.3))^2$$

$$= (-15.76)^2$$

$$= 248.3776$$

Similarly,

$$b)\left(4.2 - \begin{pmatrix} 3.1 & 4.2 \end{pmatrix}\begin{pmatrix} 1 \\ 7.1 \end{pmatrix}\right)^2$$

$$= (4.2 - (3.1 + 4.2 \times 7.1))^2$$

$$= (-28.72)^2$$

$$= 824.8384$$

$$c)\left(2.2-\begin{pmatrix}3.1 & 4.2\end{pmatrix}\begin{pmatrix}1\\6.4\end{pmatrix}\right)^2$$

$$= (2.2-(3.1+4.2\times 6.4))^2$$

$$= (-27.78)^2$$

$$= 771.728$$

$$E_D(w) = \frac{1}{2}(248.3776+824.8384+771.728)$$

$$\boxed{E_D(w)_{3.1} = 922.4717}$$

The approach is the same for equation 3.2 as well so we can go ahead and calculate $E_D(w)$ for this equation as well.

$$d)\left(9.6-\begin{pmatrix}2.4 & -1.5\end{pmatrix}\begin{pmatrix}1\\5.3\end{pmatrix}\right)^2$$

$$= (9.6-2.4+7.95)^2$$

$$= (15.15)^2$$

$$= 229.5225$$

$$e)\left(4.2-\begin{pmatrix}2.4 & -1.5\end{pmatrix}\begin{pmatrix}1\\7.1\end{pmatrix}\right)^2$$

$$= (4.2-2.4+10.65)^2$$

$$= (12.15)^2$$

$$= 155.0025$$

$$f)\left(2.2-\begin{pmatrix}2.4 & -1.5\end{pmatrix}\begin{pmatrix}1\\6.4\end{pmatrix}\right)^2$$

$$= (2.2-2.4+9.6)^2$$

$$= (9.4)^2$$

$$= 88.36$$

And, for the error calculation, we proceed as:

$$E_D(w) = \frac{1}{2}(229.5225+155.0025+88.36)$$

$$\boxed{E_D(w)_{3.2} = 236.425}$$

which is the sum of squares error for equation 3.2.

Upon comparison, we see equation 3.2 has much less standard sum of squares error which means **equation 3.2** is much better solution than equation 3.1.