Michael Kwan
Intro to Machine Learning
Homework 1

1a. Regression, Inference, N = 500, d = 4
b. Classification, Prediction, N = 20, d = 14

2a. - Classification would be useful in self-driving cars to differentiate obstacles as people, signs, roads, and various other thing that could be in a street. Some of the features could be size and speed of the object. This would be a prediction - based on the features, the car should identify an object -
- Another example would be testing the quality of fruits using features such as shape and color of the fruits being processed. This would be an inference to find patterns based on the features to determine which fruits with certain features tend to be ripe and ready for consumption.
- A third example is identifying faces on a photo using features such as the eyes, nose, and mouth. This can be a prediction becase we want it so that a consumer can take a photo and the phone can identify faces which can be useful in applications such as tagging people.

2b. • Regression will be useful in going through data for studies like determining salary for people using features such as age, profession, and region. This would be an infernce because it will be used to find trends of which features impact salary and which don't.
• Another example is when a stadium wants to find out how many people attend their game using features such as the sport, the teams playing, and the competition. This can help them do an inferance to see what ~~the~~ features influence the attendance of a game and what they might have to do to increase attendance.

3a. target variable: college GPA
b. continuous
c. Predictor variable = high school GPA
d. A linear model would be reasonable. I would expect a positive slope.

4a. $\bar{x} = 2 \quad \bar{y} = 6$
b. $S_x^2 = 2 \quad S_y^2 = 37.2 \quad S_{xy} = 8$
c. $w_1 = \frac{S_{xy}}{S_x^2} = \frac{8}{2} = 4$
$w_0 = \bar{y} - w_1\bar{x} = 6 - 4(2) = -2$
d. $y = w_0 + w_1(x) = -2 + 4(2.5) = 8$
e. $E_{in} = MSS = \frac{1}{N}\sum(y-\hat{y})^2 = \frac{1}{5}\left((0--2)^2 + (2-2)^2 + (3-6)^2 + (8-10)^2 + (17-14)^2\right) = 5.2$
g. In the first change the result would decrease $S_{xy}$, which consequently decreases $w_1$ and increases $w_0$ to be $w_1 = 3.8$ and $w_0 = -1.6$.
In the second change the $S_{xy}$ will further decrease and cause $w_1$ to decrease and $w_0$ to increase. A big jump in the change in data will translate to a drastic change in the parameters.

5. $w_0 = 0,\ w_1 = 0,\ a = .1$

$w_0 = w_0 - a\left(\frac{1}{N}\right)\Sigma(w_0 + w_1 x - y)$

$w_1 = w_1 - a\left(\frac{1}{N}\right)\Sigma(w_0 + w_1 x - y)(x)$

Step 1

$w_0 = 0 - (.1)\left(\frac{1}{5}\right)(0 + (-2) + (-3) + (-8) + (-17))$

$w_0 = 0 - \frac{1}{50}(-30)$

$\boxed{w_0 = \frac{3}{5}}$

$w_1 = 0 - (.1)\left(\frac{1}{5}\right)(0 + (-2) + (-6) + (-24) + (-68))$

$w_1 = 0 - \frac{1}{50}(-100)$

$\boxed{w_1 = 2}$

Step 2

$w_0 = \frac{3}{5} - (.1)\left(\frac{1}{5}\right)\left(\frac{3}{5} + \left(\frac{3}{5} + 2 - 2\right) + \left(\frac{3}{5} + 4 - 3\right) + \left(\frac{3}{5} + 6 - 8\right) + \left(\frac{3}{5} + 8 - 17\right)\right)$

$w_0 = \frac{3}{5} - \frac{1}{50}\left(\frac{3}{5} + \frac{3}{5} + \frac{8}{5} + \left(-\frac{7}{5}\right) + \left(-\frac{42}{5}\right)\right)$

$w_0 = \frac{3}{5} - \frac{1}{50}\left(-\frac{35}{5}\right)$

$w_0 = \frac{3}{5} + \frac{7}{50}$

$\boxed{w_0 = \frac{37}{50}}$

$w_1 = 2 - (.1)\left(\frac{1}{5}\right)\left(\frac{3}{5}(0) + \frac{3}{5}(1) + \frac{8}{5}(2) + \left(-\frac{7}{5}\right)(3) + \left(-\frac{42}{5}\right)(4)\right)$

$w_1 = 2 - \frac{1}{50}\left(0 + \frac{3}{5} + \frac{16}{5} + \left(-\frac{21}{5}\right) + \left(-\frac{168}{5}\right)\right)$

$w_1 = 2 - \frac{1}{50}\left(-\frac{170}{5}\right)$

$w_1 = 2 + \frac{1}{50}(34)$

$w_1 = 2\frac{34}{50}$

$\boxed{w_1 = 2\frac{17}{25}}$

Scanned with CamScanner

6. $w_0 = -1 \quad w_1 = 4$

$MSS = \frac{1}{N}\left((0-(-1))^2 + (2-3)^2 + (3-7)^2 + (8-11)^2 + (17-15)^2\right)$

$MSS = \frac{1}{N}\left(1^2 + (-1)^2 + (-4)^2 + (-3)^2 + (2)^2\right)$

$MSS = \frac{1}{5}(1 + 1 + 16 + 9 + 4)$

$MSS = \frac{1}{5}(31)$

$MSS = 6.2$

$w_0 = -2, \ w_1 = 4$

$MSS = \frac{1}{N}\left((0-(-2))^2 + (2-2)^2 + (3-6)^2 + (8-10)^2 + (17-14)^2\right)$

$MSS = \frac{1}{N}\left(2^2 + 0^2 + (-3)^2 + (-2)^2 + 3^2\right)$

$MSS = \frac{1}{5}(4 + 0 + 9 + 4 + 9)$

$MSS = \frac{1}{5}(26)$

$MSS = 5.2$

$w = -2, \ w_1 = 3$

$MSS = \frac{1}{N}\left((0-(-2))^2 + (2-1)^2 + (3-4)^2 + (8-7)^2 + (17-10)^2\right.$

$MSS = \frac{1}{5}\left(2^2 + 1^2 + (-1)^2 + 1^2 + 7^2\right.$

$MSS = \frac{1}{5}(4 + 1 + 1 + 1 + 49)$

$MSS = \frac{1}{5}(56)$

$MSS = 11.2$

$w_0 = -2$ and $w_1 = 4$ is the most likely choice of parameters where you can assume the noise is IID and Gaussian of 0 and variance $\sigma^2 = 5.2$ because the MSS of those parameters is 5.2.

7a. $z(t) = z_0 e^{-at}$

$\log(z(t)) = \log(z_0 e^{-at})$

$\log(z(t)) = \log(z_0) + \log(e^{-at})$

$\log(z(t)) = \log(z_0) + (-at)$

$\log(z(t)) = \log(z_0) - at$ ← linear form

b. $w_{lin} = (t^T t)^{-1} t^T z(t)$

$w_0 = \log(z_0)$

$e^{w_0} = z_0$

$z_0 = e^{w_0}$

$\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = (t^T t)^{-1} t^T z$

$w_1 = -a$

$-a = w_1$

$a = -w_1$

$\begin{bmatrix} \log(z_0) \\ -a \end{bmatrix} = (t^T t)^{-1} t^T z$

c.

```
import math
import numpy as np
a=np.ones
t=np.hstack((a,t))
q=np.linalg.inv(np.dot(t.T,t))
w=np.dot(t.T,z)
w_vec=np.dot(q,w)
w0=w_vec[0][0]
w1=w_vec[1][0]
z0=math.exp(w0)
a=-1*w1
```

8a. $w_0 = 0$

$$RSS = \sum_{n=1}^{N} (\hat{y} - y)^2$$

$$RSS = \sum_{i=1}^{N} (w_1 x_i - y_i)^2$$

b. $$\frac{\partial RSS}{\partial w} = 2 \sum_{i=1}^{N} (w x_i - y_i) x_i = 0$$

$$= \sum_{i=1}^{N} (w x_i x_i - x_i y_i) = 0$$

$$\sum_{i=1}^{N} (w x_i x_i) - \sum_{i=1}^{N} (x_i y_i) = 0$$

$$w \sum_{i=1}^{N} (x_i x_i) - \sum_{i=1}^{N} (x_i y_i) = 0$$

$$w \sum_{i=1}^{N} (x_i x_i) = \sum_{i=1}^{N} (x_i y_i)$$

$$\boxed{w = \frac{\sum_{i=1}^{N} (x_i y_i)}{\sum_{i=1}^{N} (x_i x_i)}}$$