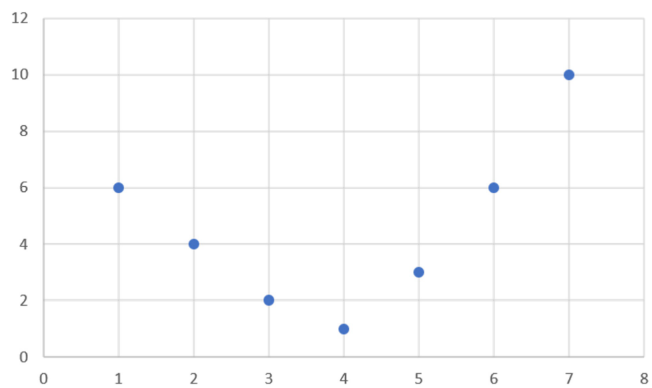
Assignment 1

Sunday, January 12, 2020 10:46 PM

Problem 1

1. See the picture



2. $\hat{Y}_i = w X_i + b$

ti

$$MSE = \frac{1}{2n} \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2 = \frac{1}{2n} \sum_{i=1}^{N} [Y_i - (wX_i + b)]^2$$

$$= \frac{1}{2n} \sum_{i=1}^{N} [Y_i^2 - 2Y_i(wX_i + b) + (wX_i + b)^2]$$

$$= \frac{1}{2n} \sum_{i=1}^{N} [Y_i^2 - 2wX_iY_i - 2bY_i + w^2X_i^2 + \underline{2wbX_i} + b^2]$$

$$= \frac{1}{2n} \sum_{i=1}^{N} [X_i^2 w^2 + b^2 + 2X_i wb - 2X_iY_i w - 2Y_i b + Y_i^2]$$

$$\|$$

$$\frac{1}{2n} \sum_{i=1}^{N} [Aw^2 + Bb^2 + Cwb + Dw + Eb + F]$$

$$\therefore \boxed{\begin{cases} A = X_i^2 \\ B = 1 \\ C = 2X_i \\ D = -2X_iY_i \\ E = -2Y_i \\ F = Y_i^2 \end{cases}}$$

3. $\nabla MSE = 0 \Rightarrow \begin{cases} \nabla_w MSE = 0 & ① \\ \nabla_b MSE = 0 & ② \end{cases}$

① $\nabla_w MSE = \nabla_w \left\{ \frac{1}{2n} \sum_{i=1}^{N} [X_i^2 w^2 + b^2 + 2X_i wb - 2X_iY_i w - 2Y_i b + Y_i^2] \right\}$

$$= \frac{1}{2n} \sum_{i=1}^{N} \left[2x_i^2 w + 2X_i b - 2X_i Y_i \right] = 0 \quad ③$$

$$\nabla_b MSE = \frac{1}{2n} \sum_{i=1}^{N} \left[2b + 2X_i w - 2Y_i \right] = 0 \quad ④$$

$$③④: \begin{cases} w \sum_{i=1}^{N} x_i^2 + b \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} x_i Y_i = 0 & ⑤ \\ Nb + w \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} Y_i = 0 & ⑥ \end{cases}$$

$$⑤: \quad b = \frac{\sum_{i=1}^{N} x_i Y_i - w \sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} X_i} \quad ⑦$$

$$⑦ \Rightarrow ⑥: \quad N \cdot \frac{\sum_{i=1}^{N} x_i Y_i - w \sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} X_i} + w \sum_{i=1}^{N} X_i - \sum_{i=1}^{N} Y_i = 0$$

$$\Rightarrow N \left(\sum_{i=1}^{N} X_i Y_i - w \sum_{i=1}^{N} X_i^2 \right) + w \left(\sum_{i=1}^{N} X_i \right)^2 - \boxed{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i} = 0$$

$$\Rightarrow N \sum_{i=1}^{N} X_i Y_i - N w \sum_{i=1}^{N} X_i^2 + w \left(\sum_{i=1}^{N} X_i \right)^2 - \boxed{\sum_{i=1}^{N} \sum_{j=1}^{N} X_i Y_j} = 0$$

$$\Rightarrow N \sum_{i=1}^{N} X_i Y_i + \left[\left(\sum_{i=1}^{N} X_i \right)^2 - N \sum_{i=1}^{N} X_i^2 \right] w - \sum_{i=1}^{N} \sum_{j=1}^{N} X_i Y_j = 0$$

$$\Rightarrow \boxed{w = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} X_i Y_j - N \sum_{i=1}^{N} X_i Y_i}{\left(\sum_{i=1}^{N} X_i \right)^2 - N \sum_{i=1}^{N} X_i^2}} \quad ⑧$$

$$⑧ \rightarrow ⑦: \quad b = \sum_{i=1}^{N} X_i Y_i - \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} X_i Y_j - N \sum_{i=1}^{N} X_i Y_i}{\left(\sum_{i=1}^{N} X_i \right)^2 - N \sum_{i=1}^{N} X_i^2} \cdot \sum_{i=1}^{N} X_i^2$$

⑧ → ⑦ :

$$b = \frac{\sum_{i=1}^{N} X_i Y_i - \dfrac{\sum_{i=1}^{N}\sum_{j=1}^{N} X_i Y_j - N\sum_{i=1}^{N} X_i Y_i}{\left(\sum_{i=1}^{N} X_i\right)^2 - N\sum_{i=1}^{N} X_i^2} \cdot \sum_{i=1}^{N} X_i^2}{\sum_{i=1}^{N} X_i}$$

$$= \frac{\left[\left(\sum_{i=1}^{N} X_i\right)^2 - N\sum_{i=1}^{N} X_i^2\right]\sum_{i=1}^{N} X_i Y_i - \left(\sum_{i=1}^{N}\sum_{j=1}^{N} X_i Y_j - N\sum_{i=1}^{N} X_i Y_i\right)\cdot \sum_{i=1}^{N} X_i^2}{\left(\sum_{i=1}^{N} X_i\right)^3 - N\sum_{i=1}^{N} X_i^2 \cdot \sum_{i=1}^{N} X_i} \quad ⑨$$

⑨'s numerator:

$$\left(\sum_{i=1}^{N} X_i\right)^2 \sum_{i=1}^{N} X_i Y_i - \underline{N\sum_{i=1}^{N} X_i^2 \sum_{i=1}^{N} X_i Y_i} - \sum_{i=1}^{N}\sum_{j=1}^{N} X_i Y_j \sum_{i=1}^{N} X_i^2 + \underline{N\sum_{i=1}^{N} X_i Y_i \sum_{i=1}^{N} X_i^2}$$

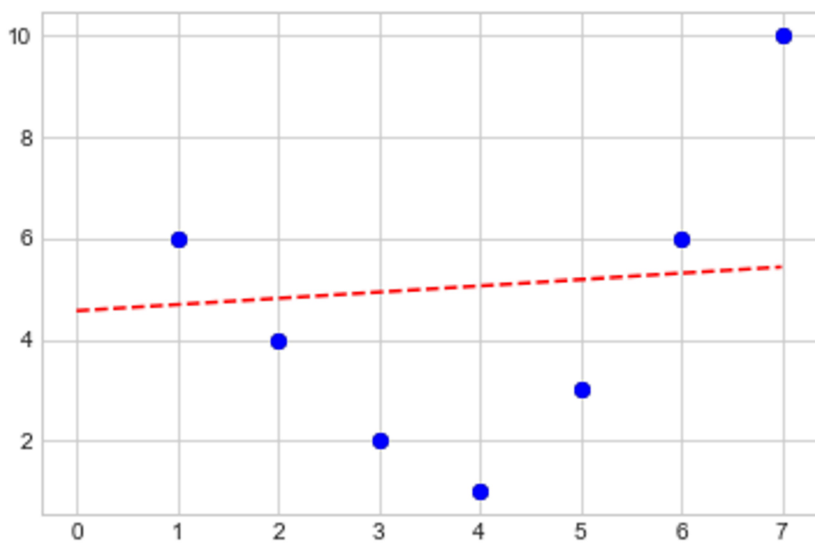
$$= \left(\sum_{i=1}^{N} X_i\right)^2 \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N}\sum_{j=1}^{N} X_i Y_i \sum_{i=1}^{N} X_i^2$$

$$\therefore \boxed{b = \frac{\left(\sum_{i=1}^{N} X_i\right)^2 \sum_{i=1}^{N} X_i Y_i - \sum_{i=1}^{N}\sum_{j=1}^{N} X_i Y_j \sum_{i=1}^{N} X_i^2}{\left(\sum_{i=1}^{N} X_i\right)^3 - N\sum_{i=1}^{N} X_i^2 \cdot \sum_{i=1}^{N} X_i}}$$

4. 5. Find attached doc.

```
"""
Created on Mon Jan 13 19:50:23 2020

This is the script for assignment 1

@author: Zonghao Li
"""
import numpy as np
import matplotlib.pyplot as plt
plt.style.use('seaborn-whitegrid')

# input data
N = 7
X = np.array([1,2,3,4,5,6,7])
Y = np.array([6,4,2,1,3,6,10])
# shape the vector
Y.shape = (1,7)
X.shape = (1,7)
```

```
# draw the X Y scatter plot
plt.plot(X, Y, 'o', color='black')

# w's numerator
w_num = np.sum(np.matmul(np.transpose(X),Y)) - N * np.sum(X*Y)
# w's denominator
w_den = (np.sum(X))^2 - N * np.sum(np.square(X))
w = float(w_num) / float(w_den)                        # w = 0.122

# b's numerator
b_num = (np.sum(X))^2 * np.sum(X*Y) - np.sum(np.matmul(np.transpose(X),Y)) * np.sum(np.square(X))
# b's denominator
b_den = (np.sum(X))^3 - N * np.sum(np.square(X)) * (np.sum(X))
b = float(b_num) / float(b_den)                        # b = 4.56

# draw the picture of bot X Y scatter plot and linear approx.
x = np.arange(8)
y = np.arange(10)
Y_hat = w * x + b
plt.plot(X, Y, 'bo', x, Y_hat, 'r--')
```



Problem 2.

1. $x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{id} \end{bmatrix}$, $x_i x_i^T = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{id} \end{bmatrix} [x_{i1} \cdots x_{id}]$

$$= \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & \cdots & x_{i1}x_{id} \\ x_{i2}x_{i1} & x_{i2}^2 & \cdots & x_{i2}x_{id} \\ \vdots & & \ddots & \vdots \\ x_{id}x_{i1} & \cdots & & x_{id}^2 \end{bmatrix}$$

$$N \qquad - \qquad N \; [\; x_{i1}^2 \quad x_{i1}x_{i2} \cdots x_{i1}x_{id} \;]$$

$$-x_{id}x_{i1} \quad - - \cdot \quad x_{id} \quad -$$

$$\therefore A = \sum_{i=1}^{N} x_i x_i^T = \sum_{i=1}^{N} \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & \cdots & x_{i1}x_{id} \\ x_{i2}x_{i1} & x_{i2}^2 & \cdots & x_{i2}x_{id} \\ \vdots & & \ddots & \vdots \\ x_{id}x_{i1} & \cdots & & x_{id}^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{N} x_{i1}^2 & \sum_{i=1}^{N} x_{i1}x_{i2} & \cdots & \sum_{i=1}^{N} x_{i1}x_{id} \\ \sum_{i=1}^{N} x_{i2}x_{i1} & \sum_{i=1}^{N} x_{i2}^2 & \cdots & \sum_{i=1}^{N} x_{i2}x_{id} \\ \vdots & & \ddots & \\ \sum_{i=1}^{N} x_{id}x_{i1} & \cdots & & \sum_{i=1}^{N} x_{id}^2 \end{bmatrix}$$

Ok, correct

2.

$$\nabla L(\bar{w}, D) = \nabla_w \left[\frac{1}{2N} \sum_{i=1}^{N} (Y_i - \bar{w}^T x_i)^2 + \frac{\lambda}{2} \|\bar{w}\|_2^2 \right]$$

$$= \nabla_w \left[\frac{1}{2N} \sum_{i=1}^{N} (\bar{w}^T x_i - Y_i)^2 + \frac{\lambda}{2} \|\bar{w}\|_2^2 \right]$$

$$= \frac{1}{2N} \sum_{i=1}^{N} \left[2 (\overset{x_i^T w}{\bar{w} x_i} - \overset{\text{scalar}}{Y_i}) \cdot x_i \right] + \lambda \bar{w}$$

→ property : for col. vector w & x

$$\underbrace{(w^T x) x}_{} = x x^T w$$

$$\hookrightarrow = x \overbrace{(w^T x)}^{\text{scalar}}$$

$$= x x^T w$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left(\boxed{\overset{x_i^T w x_i}{\bar{w}^T x_i x_i}} - Y_i x_i \right) + \lambda \bar{w}$$

$$= \frac{1}{N} \left[\boxed{\bar{w} \left(\sum_{i=1}^{N} x_i x_i^T \right) \bar{w}} - \sum_{i=1}^{N} Y_i x_i \right] + \lambda \bar{w}$$

This is fine, full marks

$$= \frac{1}{N} (A\bar{w} - B) + \lambda \bar{w}$$

3. let: $\nabla L(\bar{w}, D) = \frac{1}{N} (A\bar{w}^* - B) + \lambda \bar{w}^* = 0$

$$\Rightarrow (A\bar{w}^* - B) + N\lambda \bar{w}^* = 0 \Rightarrow (A + N\lambda I)\bar{w}^* - B = 0$$

$$\therefore \bar{w}^* = (A + N\lambda I)^{-1} B$$

$(A + N\lambda I)$ is invertible as it is a square matrix whose det. is non-zero, which is proven from Q4 & Q5)

4. since $A = \sum_{i=1}^{N} X_i X_i^T$, let $a = X_i X_i^T$

let : $x \in R^n$, $x^T a x = x^T X_i X_i^T x$

$$= (x^T X_i) (x^T x_i)^T$$

$$= \|x X_i\|^2 \geq 0$$

therefore, a is a positive-semidefinite

therefore, A is a positive-semidefinite

therefore, all A's eigenvalues are non-negative

5. let : $a = X_i X_i^T$, $x \in R^n$

$$x^T [a + \lambda N I_d] x = x^T a x + x^T (\lambda N I_d) x$$

$$= \|x X_i\|^2 + \lambda N \|x\|^2 > 0$$

therefore, $A + \lambda N I_d$ is a positive definite matrix with all eigenvalues larger than zero ; in addition, it is a square matrix with dimension d ; therefore it is invertible

6. let: $\frac{1}{N}(A\bar{w} - B) + \lambda \bar{w} = 0$

$$\Rightarrow A\bar{w} - B + \lambda N \bar{w} = 0$$

$$\Rightarrow (A + \lambda N I)\bar{w} = B$$

$$\Rightarrow \bar{w}^* = (A + \lambda N I)^{-1} B$$