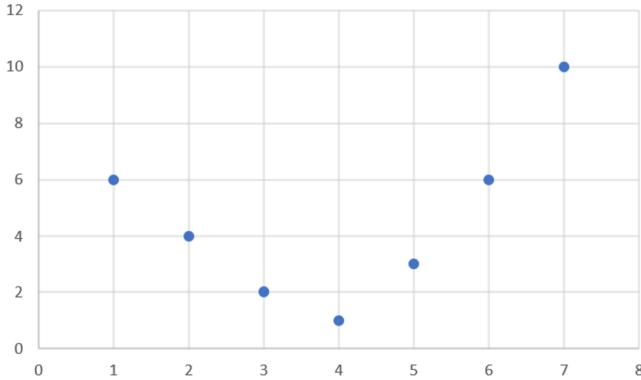


# Assignment 1

Sunday, January 12, 2020 10:46 PM

## Problem 1

- See the picture



- $\hat{Y}_i = wX_i + b$

$$\begin{aligned}
 \text{MSE} &= \frac{1}{2n} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2 = \frac{1}{2n} \sum_{i=1}^N [Y_i - (wX_i + b)]^2 \\
 &= \frac{1}{2n} \sum_{i=1}^N [Y_i^2 - 2Y_i(wX_i + b) + (wX_i + b)^2] \\
 &= \frac{1}{2n} \sum_{i=1}^N [Y_i^2 - 2wX_i Y_i - 2bY_i + w^2 X_i^2 + 2wbX_i + b^2] \\
 &= \frac{1}{2n} \sum_{i=1}^N [X_i^2 w^2 + b^2 + 2X_i wb - 2X_i Y_i w - 2Y_i b + Y_i^2] \\
 &\quad \parallel \\
 &= \frac{1}{2n} \sum_{i=1}^N [Aw^2 + Bb^2 + Cwb + Dw + Eb + F]
 \end{aligned}$$

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

- $\nabla_{\text{MSE}} = 0 \Rightarrow \begin{cases} \nabla_w \text{MSE} = 0 & \textcircled{1} \\ \nabla_b \text{MSE} = 0 & \textcircled{2} \end{cases}$

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

$$= \frac{1}{2n} \sum_{i=1}^n [2x_i^2 w + 2x_i b - 2x_i y_i] = 0 \quad \textcircled{3}$$

$$\nabla_b \text{MSE} = \frac{1}{2n} \sum_{i=1}^n [2b + 2x_i w - 2y_i] = 0 \quad \textcircled{4}$$

$$\begin{cases} w \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i - \sum_{i=1}^n x_i y_i = 0 & \textcircled{3} \\ N b + w \sum_{i=1}^n x_i - \sum_{i=1}^n y_i = 0 & \textcircled{4} \end{cases}$$

$$\textcircled{5}: \quad b = \frac{\sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \quad \textcircled{7}$$

$\textcircled{5} \Rightarrow \textcircled{6}$ :

$$N \cdot \frac{\sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} + w \sum_{i=1}^n x_i - \sum_{i=1}^n y_i = 0$$

$$\Rightarrow N \left( \sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2 \right) + w \left( \sum_{i=1}^n x_i \right)^2 - \sum_{i=1}^n x_i \sum_{i=1}^n y_i = 0$$

$$\Rightarrow N \sum_{i=1}^n x_i y_i - N w \sum_{i=1}^n x_i^2 + w \left( \sum_{i=1}^n x_i \right)^2 - \sum_{i=1}^n \sum_{j=1}^n x_i y_j = 0$$

$$\Rightarrow N \sum_{i=1}^n x_i y_i + \left[ \left( \sum_{i=1}^n x_i \right)^2 - N \sum_{i=1}^n x_i^2 \right] w - \sum_{i=1}^n \sum_{j=1}^n x_i y_j = 0$$

$$\Rightarrow w = \frac{\sum_{i=1}^n \sum_{j=1}^n x_i y_j - N \sum_{i=1}^n x_i y_i}{\left( \sum_{i=1}^n x_i \right)^2 - N \sum_{i=1}^n x_i^2} \quad \textcircled{8}$$

$$\textcircled{8} \rightarrow \textcircled{7}: \quad b = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n \sum_{j=1}^n x_i y_j - N \sum_{i=1}^n x_i y_i}{\left( \sum_{i=1}^n x_i \right)^2 - N \sum_{i=1}^n x_i^2} \cdot \sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i}$$

$$\textcircled{8} \rightarrow \textcircled{7} : b = \frac{\sum_{i=1}^N x_i y_i - \frac{\sum_{i=1}^N \sum_{j=1}^N x_i y_j - N \sum_{i=1}^N x_i \sum_{j=1}^N y_j}{\left(\sum_{i=1}^N x_i\right)^2 - N \sum_{i=1}^N x_i^2} \cdot \sum_{i=1}^N x_i^2}{\sum_{i=1}^N x_i}$$

$$= \frac{\left[\left(\sum_{i=1}^N x_i\right)^2 - N \sum_{i=1}^N x_i^2\right] \sum_{i=1}^N x_i y_i - \left(\sum_{i=1}^N \sum_{j=1}^N x_i y_j - N \sum_{i=1}^N x_i y_i\right) \cdot \sum_{i=1}^N x_i^2}{\left(\sum_{i=1}^N x_i\right)^3 - N \sum_{i=1}^N x_i^2 \cdot \sum_{i=1}^N x_i} \quad \textcircled{9}$$

\textcircled{9}'s numerator:

$$\begin{aligned} & \left(\sum_{i=1}^N x_i\right)^2 \sum_{i=1}^N x_i y_i - N \sum_{i=1}^N x_i^2 \sum_{i=1}^N x_i y_i - \sum_{i=1}^N \sum_{j=1}^N x_i y_j \sum_{i=1}^N x_i^2 + N \sum_{i=1}^N x_i y_i \sum_{i=1}^N x_i^2 \\ &= \left(\sum_{i=1}^N x_i\right)^2 \sum_{i=1}^N x_i y_i - \sum_{i=1}^N \sum_{j=1}^N x_i y_j \sum_{i=1}^N x_i^2 \\ \therefore b &= \frac{\left(\sum_{i=1}^N x_i\right)^2 \sum_{i=1}^N x_i y_i - \sum_{i=1}^N \sum_{j=1}^N x_i y_j \sum_{i=1}^N x_i^2}{\left(\sum_{i=1}^N x_i\right)^3 - N \sum_{i=1}^N x_i^2 \cdot \sum_{i=1}^N x_i} \end{aligned}$$

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)
```

```

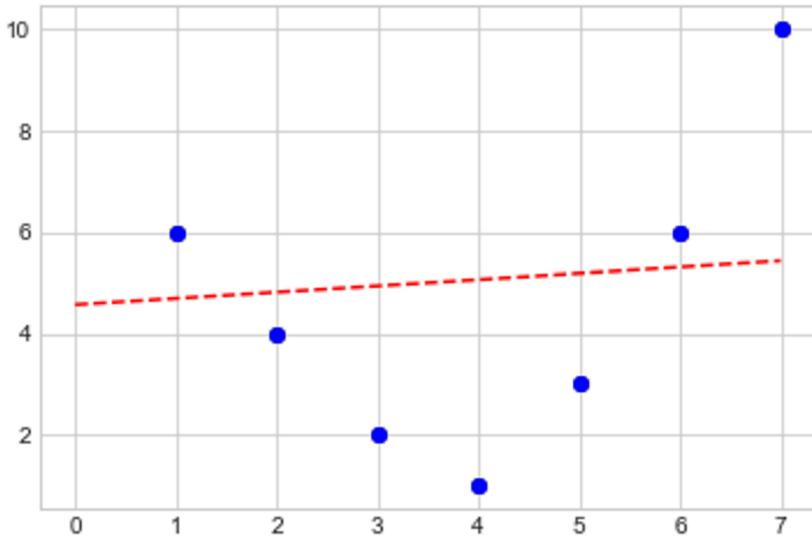
# draw the X Y scatter plot
plt.plot(X, Y, 'o', color='black')

# 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.

$$1. \quad \mathbf{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{id} \end{bmatrix}, \quad \mathbf{x}_i \mathbf{x}_i^T = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{id} \end{bmatrix} \begin{bmatrix} x_{i1} & \dots & x_{id} \end{bmatrix}$$

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

$$\mathbf{N} = \mathbf{n} \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & \dots & x_{i1}x_{id} \end{bmatrix}$$

$$\therefore A = \sum_{i=1}^N X_i X_i^T = \sum_{i=1}^N \begin{bmatrix} X_{i1}^2 & X_{i1}X_{i2} \dots X_{i1}X_{id} \\ X_{i2}X_{i1} & X_{i2}^2 \dots X_{i2}X_{id} \\ \vdots & \ddots \vdots \\ X_{id}X_{i1} & \dots X_{id}^2 \end{bmatrix}$$

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

2.

$$\begin{aligned} \nabla L(\bar{w}, D) &= \nabla_w \left[ \frac{1}{2N} \sum_{i=1}^N (Y_i - \bar{w} X_i)^2 + \frac{\lambda}{2} \|\bar{w}\|_2^2 \right] \\ &= \nabla_w \left[ \frac{1}{2N} \sum_{i=1}^N (\bar{w} X_i - Y_i)^2 + \frac{\lambda}{2} \|\bar{w}\|_2^2 \right] \\ &= \frac{1}{2N} \sum_{i=1}^N \left[ 2(\bar{w} X_i - Y_i) \cdot X_i \right] + \lambda \bar{w} \\ &= \frac{1}{N} \sum_{i=1}^N (\bar{w} X_i X_i^T - Y_i X_i) + \lambda \bar{w} \\ &= \frac{1}{N} \left[ \bar{w} \sum_{i=1}^N X_i X_i^T - \sum_{i=1}^N Y_i X_i \right] + \lambda \bar{w} \\ &= \frac{1}{N} (A \bar{w} - B) + \lambda \bar{w} \end{aligned}$$

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

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

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

$(A + N\lambda I)$  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  $a = x_i x_i^T$

$$\begin{aligned} \text{let: } x \in R^n, \quad x^T a x &= x^T x_i x_i^T x \\ &= (x^T x_i) (x^T x_i)^T \\ &= \|x_i x_i^T\|^2 \geq 0 \end{aligned}$$

therefore,  $a$  is a positive-semidefinite

therefore,  $A$  is a positive-semidefinite

therefore, all  $A$ 's eigenvalues are non-negative

5- let:  $a = x_i x_i^T$ ,  $x \in R^n$

$$\begin{aligned} x^T [a + \lambda N \text{Id}] x &= x^T a x + x^T (\lambda N \text{Id}) x \\ &= \|x_i x_i^T\|^2 + \lambda N \|x\|^2 \geq 0 \end{aligned}$$

therefore,  $A + \lambda N \text{Id}$  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

$$b. \text{ Let: } \frac{1}{N} (A\bar{\omega} - B) + \lambda \bar{\omega} = 0$$

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

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

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