機器學習 Hw 1

陳栢翔
4109061012

Problem 01 :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} \quad , -\infty < x < \infty$$

(1) $Y = X^2 \Rightarrow \{Y \le y\}$ event occurs when $\{X^2 \le y\} = \{-\sqrt{y} \le X \le \sqrt{y}\}$ for $y \ge 0$

$$\Rightarrow f_Y(y) = \begin{cases} 0 & , y < 0 \\ f_X(\sqrt{y}) - f_X(-\sqrt{y}) & , y > 0 \end{cases}$$

$$= \frac{f_X(\sqrt{y})}{2\sqrt{y}} - \frac{f_X(-\sqrt{y})}{-2\sqrt{y}} \quad , y > 0$$

$$= \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{y^2}{2\sigma_x^2}} \quad , y > 0$$

(2)

$$E[Y] = E[X^2] = VAR[X] + E[X]^2$$

$$= \sigma_x^2 + 0^2 = \underline{\underline{\sigma_x^2}}$$

$$E[Y^2] =$$

Problem 02 :

(1)

$$Y = 2X_1 + 3X_2 \Rightarrow \text{CDF of } Y : F_Y(y) \qquad \begin{cases} 0 < X_1 < 1 \\ 0 < X_2 < 1 \end{cases} \Rightarrow 0 < Y < 1$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{y-2x_1}{3}} f_{X_1X_2}(x_1, x_2)\, dx_2\, dx_1$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{y-2x_1}{3}} f_{X_1}(x_1) f_{X_2}(x_2)\, dx_2\, dx_1 \quad \text{since } X_1, X_2 \text{ are independent}$$

$$= \int_0^1 \int_0^{\frac{y-2x_1}{3}} 1\, dx_2\, dx_1 = \int_0^1 \frac{y-2x_1}{3} dx_1 = \frac{1}{3}(yx_1 - x_1^2)\Big|_0^1 = \frac{1}{3}y - \frac{1}{3} \quad , 0 < y < 1$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(\frac{1}{3}y - \frac{1}{3}\right) = \underline{\underline{\frac{1}{3} \quad , 0 < y < 1}}$$

(2)

$$E[Y] = \int_0^1 y \cdot \frac{1}{3} dy = \underline{\underline{\frac{1}{3}}}$$

$$E[Y^2] = \int_0^1 y^2 \cdot \frac{1}{3} dy = \frac{1}{3} \cdot \frac{1}{3} y^3 \Big|_0^1 = \underline{\underline{\frac{1}{9}}}$$

Problem 03 :

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \frac{\partial y_3}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_3}{\partial x_2} \end{bmatrix} = \underline{\underline{\begin{bmatrix} 2x_1 & -6x_1+2x_2 & 2x_1 \\ 4x_2 & 2x_1+2x_2 & 2x_2 \end{bmatrix}}}$$

Problem 04 :

(1)
$$y = Ax = \begin{bmatrix} -1 & 2 & 3 \\ 2 & 0 & 1 \\ 3 & -1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_1+2x_2+3x_3 \\ 2x_1+0+x_3 \\ 3x_1-x_2+5x_3 \end{bmatrix}$$

$$\Rightarrow \frac{\partial y}{\partial x} = \underline{\underline{\begin{bmatrix} -1 & 2 & 3 \\ 2 & 0 & -1 \\ 3 & 1 & 5 \end{bmatrix}}}$$

(2)
$$y = x^T A x = [x_1\ x_2\ x_3]\begin{bmatrix} -x_1+2x_2+3x_3 \\ 2x_1+0+x_3 \\ 3x_1-x_2+5x_3 \end{bmatrix} = -x_1^2+2x_1x_2+3x_1x_3+2x_1x_2+x_2x_3+3x_1x_3-x_2x_3+5x_3^2$$

$$= -x_1^2+4x_1x_2+6x_1x_3+5x_3^2$$

$$\Rightarrow \frac{\partial y}{\partial x} = \underline{\underline{\begin{bmatrix} -2x_1+4x_2+6x_3 \\ 4x_1 \\ 6x_1+10x_3 \end{bmatrix}}}$$

Problem 05 :

$X_{new} = (2, 2)$

Distance :

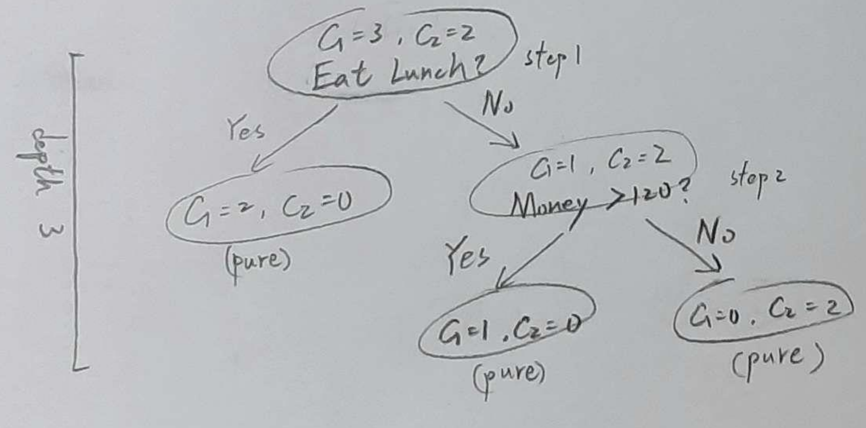
x	class	$p=1$	$p=2$	$p=\infty$
(1, 2)	1	1	1	1
(3, 2)	2	1	1	1
(0, 3)	1	3	$\sqrt{5}$	2
(4, 1)	2	3	$\sqrt{5}$	2
(2, 0)	1	2	2	2
(2, 3)	2	1	1	1

K = 3

$\Rightarrow$ 1、 class 2 2、 class 2 3、 class 2

Problem 06 :

令 C_1 = Have tea
C_2 = Not have tea



1.

Information Gain (step 1)

$$= \left[-\left(\tfrac{3}{5}\right)\log\left(\tfrac{3}{5}\right) - \left(\tfrac{2}{5}\right)\log\left(\tfrac{2}{5}\right)\right]$$
$$- \left\{ \tfrac{2}{5}\cancel{[-0-0]} + \tfrac{3}{5}\left[-\left(\tfrac{1}{3}\right)\log\left(\tfrac{1}{3}\right) - \left(\tfrac{2}{3}\right)\log\left(\tfrac{2}{3}\right)\right] \right\}$$
$$= 0.97095 - 0.55098 = 0.41997$$

Information Gain (step 2)

$$= \left[-\left(\tfrac{1}{3}\right)\log\left(\tfrac{1}{3}\right) - \left(\tfrac{2}{3}\right)\log\left(\tfrac{2}{3}\right)\right] - 0$$
$$= 0.55098$$

2. Gini Index after step 1 splitting :

$$= \frac{2}{5}\left(1-(1+0)\right) + \frac{3}{5}\left(1-\left(\left(\tfrac{1}{3}\right)^2+\left(\tfrac{2}{3}\right)^2\right)\right) = \frac{3}{5}\times\left(1-\frac{5}{9}\right) = 0.2667$$

Gini Index after step 2 splitting :

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

3.

Instrinsic Information (step 1) $= -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log\frac{3}{5} = 0.97095$

(step 2) $= -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0.91830$

$\Rightarrow$ Information Gain Ratio (step 1) $= 0.43254$

(step 2) $= 0.600$

Problem 07 :

(x, y) data : $(1,2)\ (3,5)\ (5,4)\ (7,4)\ (9,8)$, model $\hat{y} = w_0 + w_1 x$

$$Aw = b \Rightarrow \begin{array}{cc} 1 & x \end{array} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \\ 1 & 9 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{array}{c} y \end{array} \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 & 9 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \\ 1 & 9 \end{bmatrix} = \begin{bmatrix} 5 & 25 \\ 25 & 165 \end{bmatrix}, \quad A^Tb = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 23 \\ 137 \end{bmatrix}$$

$$w = (A^TA)^{-1}A^Tb = \begin{bmatrix} \frac{33}{40} & -\frac{5}{40} \\ -\frac{1}{8} & \frac{1}{40} \end{bmatrix} \begin{bmatrix} 23 \\ 137 \end{bmatrix} = \begin{bmatrix} 1.85 \\ 0.55 \end{bmatrix} \Rightarrow \underline{\underline{\hat{y} = 1.85 + 0.55x}}$$

Problem 08 :

$$Aw = b \Rightarrow \begin{array}{ccc} 1 & x & x^2 \end{array} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 7 & 49 \\ 1 & 9 & 81 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{array}{c} y \end{array} \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 & 9 \\ 1 & 9 & 25 & 49 & 81 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \\ 1 & 7 & 49 \\ 1 & 9 & 81 \end{bmatrix} = \begin{bmatrix} 5 & 25 & 165 \\ 25 & 165 & 1225 \\ 165 & 1225 & 9669 \end{bmatrix}, \quad A^Tb = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 & 9 \\ 1 & 9 & 25 & 49 & 81 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 23 \\ 137 \\ 991 \end{bmatrix}$$

$$w = (A^TA)^{-1}A^Tb = \begin{bmatrix} \frac{773}{280} \\ \frac{1}{70} \\ \frac{3}{56} \end{bmatrix} \Rightarrow \underline{\underline{\hat{y} = 2.761 + 0.014x + 0.054x^2}}$$

Problem 09 :

$$Aw = b \Rightarrow \begin{array}{ccc} 1 & x_1 & x_2 \\ \end{array} \qquad \begin{array}{c} y \end{array}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & 5 & 2 \\ 1 & 7 & 4 \\ 1 & 9 & 2 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix} , \quad A^T A = \begin{bmatrix} 5 & 23 & 12 \\ 23 & 157 & 60 \\ 12 & 60 & 34 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 5 & 7 & 9 \\ 1 & 3 & 2 & 4 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 4 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 23 \\ 127 \\ 57 \end{bmatrix}$$

$$w = (A^T A)^{-1} A^T b = \begin{bmatrix} 869/608 & -31/608 & -63/152 \\ -31/608 & 13/608 & -3/152 \\ -63/152 & -3/152 & 4/19 \end{bmatrix} \begin{bmatrix} 23 \\ 127 \\ 57 \end{bmatrix} = \begin{bmatrix} \frac{843}{304} \\ \frac{127}{304} \\ -\frac{3}{76} \end{bmatrix} = \begin{bmatrix} 2.773 \\ 0.418 \\ -0.0395 \end{bmatrix}$$

$$\hat{y} = 2.773 + 0.418 x_1 - 0.0395 x_2$$

```python
# Problem 10    4109061012 陳柏翔

import numpy as np
from numpy.linalg import inv
import matplotlib.pyplot as plt


SIGMA_LIST = [0.01, 0.1, 1, 2]
SAMPLES = 40


if __name__ == "__main__":
    # Problem.1 ~ Problem.4
    for sigma_idx in range(len(SIGMA_LIST)):

        # Generate Dataset
        dataset = np.empty((SAMPLES, 2))
        for i in range(SAMPLES):
            x = np.random.uniform(0, 8, size=None)
            y = 3 + 2*x + 0.2*(x**2) + np.random.normal(0, SIGMA_LIST[sigma_idx], size=None)
            dataset[i] = x, y

        # Linear Regression
        mat_X = np.ones((SAMPLES, 3))
        mat_X[:, 1] = dataset[:, 0]
        mat_X[:, 2] = dataset[:, 0] * dataset[:, 0]
        mat_Y = np.empty((SAMPLES, 1))
        mat_Y[:, 0] = dataset[:, 1]

        mat_W = inv(mat_X.T @ mat_X) @ mat_X.T @ mat_Y

        # Plot
        x = np.linspace(0, 8, num=200)
        y = mat_W[2] * x**2 + mat_W[1] * x + mat_W[0]
        plt.subplot(2, 2, sigma_idx+1)
        plt.title(f"Regression {sigma_idx+1} (sigma = {SIGMA_LIST[sigma_idx]})")
        plt.plot(dataset[:, 0], dataset[:, 1], "ro", markersize=2, label="dataset")
        plt.plot(x, y, label="model")
        plt.legend(loc="upper left")
```