Pattern Recognition Homework2

楊憲閔 613410047

1.

(a)

$$\begin{aligned}
p(D|\theta) &= \prod_{i=1}^{n} p(x_i|\theta) \\
&= \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} \\
l(\theta) &= ln[p(D|\theta)] \\
&= \sum_{i=1}^{n} ln[\theta^{x_i}(1-\theta)^{1-x_i}] \\
&= \sum_{i=1}^{n} [x_i ln(\theta) + (1-x_i) ln(1-\theta)] \\
\nabla_\theta ln[l(\theta)] &= \sum_{i=1}^{n} \nabla_\theta [x_i ln(\theta) + (1-x_i) ln(1-\theta)] \\
&= \sum_{i=1}^{n} (\frac{x_i}{\theta} - \frac{1-x_i}{1-\theta}) \\
&= \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{n - \sum_{i=1}^{n} x_i}{1-\theta} \\
&= 0
\end{aligned}$$

$$\Rightarrow \frac{\sum_{i=1}^{n} x_i}{\theta} = \frac{n - \sum_{i=1}^{n} x_i}{1-\theta}$$
$$\Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n}$$

(b) $E[\hat{\theta}] = E[\frac{\sum_{i=1}^{n} x_i}{n}]$

Since $\mathrm{x_1, x_2, ..., x_n}$ are independent and identically distributed (i.i.d.).

Bernoulli random variables with $E[\hat{\theta}] = \theta,$ the expectation of their sum is :

$E[\sum_{i=1}^{n} x_i] = \sum_{i=1}^{n} E[x_i] = n\theta.$

$\Rightarrow E[\hat{\theta}] = \frac{E[\sum_{i=1}^{n} x_i]}{n} = \frac{n\theta}{n} = \theta$

Since $E[\hat{\theta}] = \theta \Rightarrow$ unbiased

2.

$p(x|D) = \int p(x|\theta)p(\theta|D)d\theta$

$p(D|\theta) = \prod_{i=1}^{n} p(x_i|\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$ for $x_i \geq 0$

$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} = \alpha_1 p(D|\theta)p(\theta) = \alpha_1 \prod_{k=1}^{n} p(x_k|\theta)p(\theta) = \alpha_2 \theta^n e^{-\theta \sum_{i=1}^{n} x_i} \cdot \lambda e^{-\lambda\theta}$

$p(\theta|D) = \alpha_3 \theta^n e^{-\theta(\sum_{i=1}^{n} x_i + \lambda)}$

$logp(\theta|D) = nlog\theta - \theta(\sum_{i=1}^{n} x_i + \lambda) + constant$

$\nabla_\theta[logp(\theta|D)] = \nabla_\theta[nlog\theta - \theta(\sum_{i=1}^{n} x_i + \lambda) + constant] = 0$

$\Rightarrow \frac{n}{\theta} - (\sum_{i=1}^{n} x_i + \lambda) = 0$

$\Rightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i + \lambda}$

$\Rightarrow$ The maximum a posterior estimate of θ is $\frac{n}{\sum_{i=1}^{n} x_i + \lambda}$

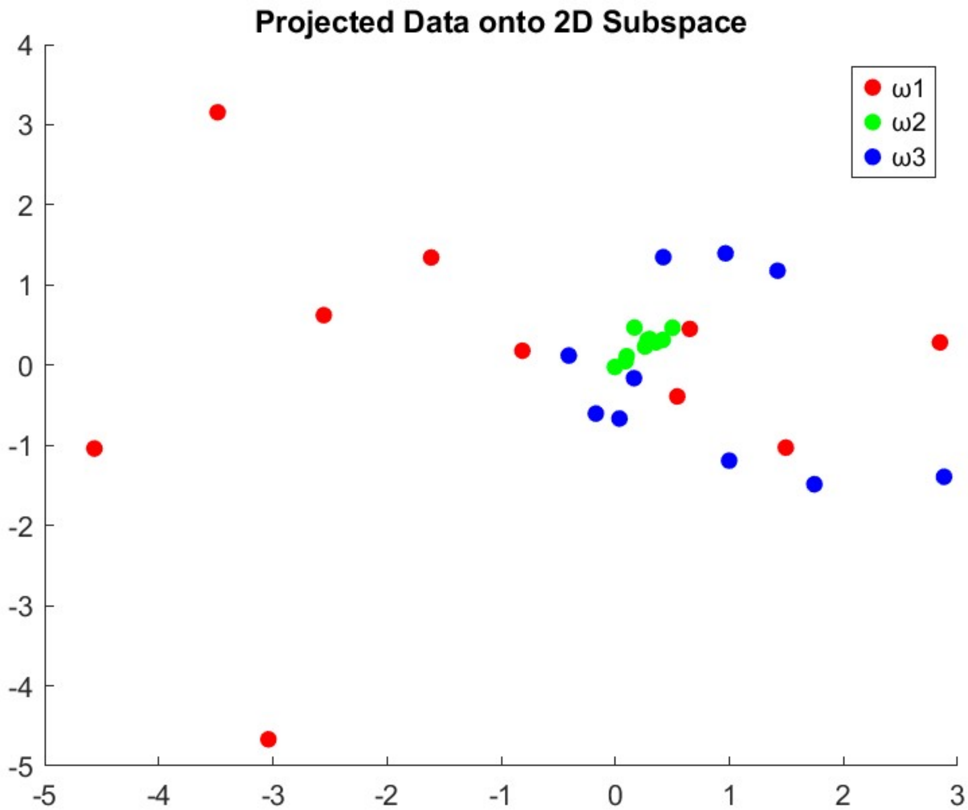
3.

(a) scatter matrix $= \begin{bmatrix} 13.5236 & 10.3097 & 4.6601 \\ 10.3097 & 55.2695 & 13.0659 \\ 4.6601 & 13.0659 & 70.8436 \end{bmatrix}$

(b) two largest eigenvalues of the scatter matrix = 79.5607 and 49.0038

corresponding eigenvectors $= \begin{bmatrix} 0.1400 \\ 0.5145 \\ 0.8460 \end{bmatrix}$ and $\begin{bmatrix} 0.1710 \\ 0.8290 \\ -0.5324 \end{bmatrix}$

(c)



(d) ω_1 error rate = 50%

ω_2 error rate = 0%

ω_3 error rate = 0%

4.

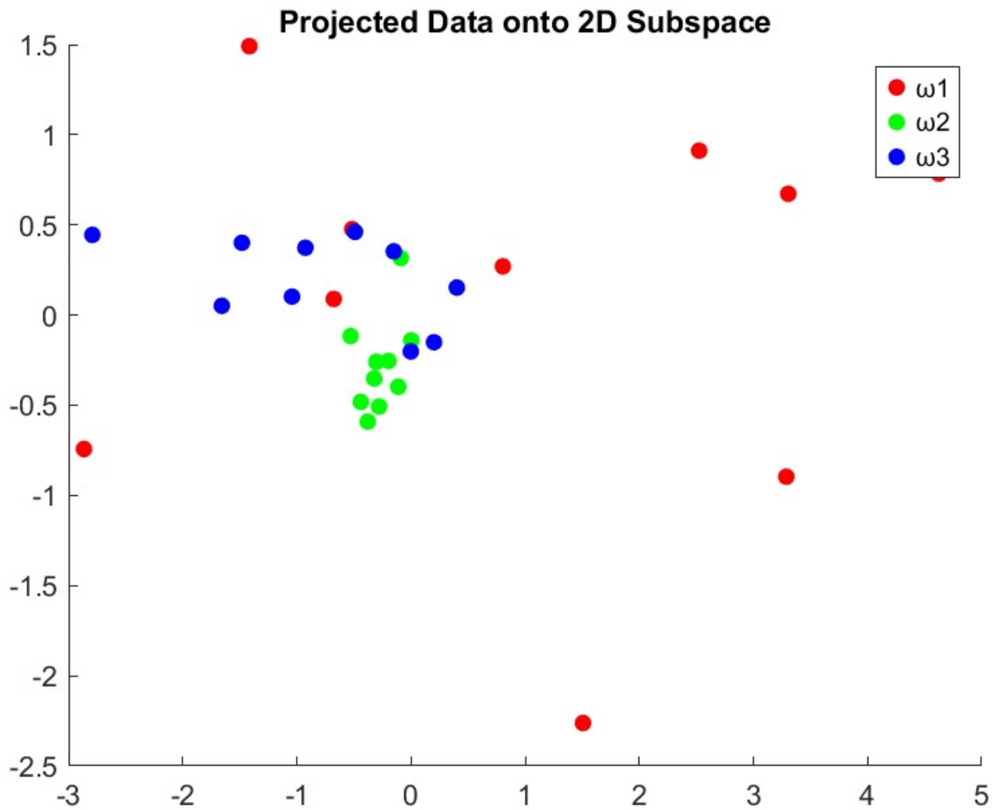
(a) $$\text{within} - \text{class scatter matrix} = \begin{bmatrix} 12.6196 & 9.5787 & 2.0848 \\ 9.5787 & 48.9165 & 5.4091 \\ 2.0848 & 5.4091 & 58.1136 \end{bmatrix}$$

$$\text{between} - \text{class scatter matrix} = \begin{bmatrix} 0.9039 & 0.7309 & 2.5753 \\ 0.7309 & 6.3530 & 7.6568 \\ 2.5753 & 7.6568 & 12.7300 \end{bmatrix}$$

(b) two largest eigenvalues of the scatter matrix = 0.2972 and 0.1028

$$\text{corresponding eigenvectors} = \begin{bmatrix} -0.1386 \\ -0.5653 \\ -0.8132 \end{bmatrix} \text{ and } \begin{bmatrix} 0.9081 \\ -0.4046 \\ 0.1079 \end{bmatrix}$$

(c)



(d) ω_1 error rate $= 30\%$

ω_2 error rate $= 10\%$

ω_3 error rate $= 10\%$