Q1) e)

$$g_1(x) = \ln(P(x|\omega_1)) + \ln P(\omega_1)$$

ω_1, ω_2 : class 1, class 2 resp.

$$g_2(x) = \ln(P(x|\omega_2)) + \ln P(\omega_2)$$

As number of training samples for both samples is equal to $P(\omega_1) = P(\omega_2) = \frac{1}{2}$ hence the term $\ln(P(\omega_1))$ and $\ln(P(\omega_2))$ can be ignored.

As there is independence b/w the 2 Bernoulli RVs :

$$P(x|\omega) = P(x_1|\omega) \cdot P(x_2|\omega)$$

$$= \theta_1^{x_1}(1-\theta_1)^{(1-x_1)} \cdot \theta_2^{x_2}(1-\theta_2)^{(1-x_2)}$$

θ_i : parameter of Bernoulli distribution of i^{th} dimension

$$\ln(P(x|\omega)) = x_1 \ln\theta_1 + (1-x_1)\ln(1-\theta_1) + x_2 \ln\theta_2 + (1-x_2)\ln(1-\theta_2) = g(x)$$

We will compare the values of the discriminant corresponding to both classes and whichever is larger, it will classified to that class

For Bernoulli RV, the parameter θ ~~can~~ is same as the mean and hence θ in the expression can be replaced with μ or the μ found from MLE.

$$\therefore \quad g_i(x) = x_{1i} \ln(MLE_{1i}) + (1 - x_{1i}) \ln(1 - MLE_{1i}) + x_{2i} \ln(MLE_{2i}) + (1 - x_{2i}) \ln(1 - MLE_{2i})$$

(i is for i^{th} class)

Q2)

Prior for θ $= \theta_1 \cdot \theta_2 \dots \theta_d \cdot e^{-(\theta_1+\theta_2 \dots \theta_d)}$

$= \theta_1 e^{-\theta_1} \cdot \theta_2 e^{-\theta_2} \dots \theta_d e^{-\theta_d}$

For MAP we need to maximise the product of likelihood and prior

$$\underbrace{P(D|\theta)}_{\text{likelihood}} \cdot \underbrace{P(\theta)}_{\text{prior}}$$

likelihood $\left(\prod_{i=1}^{N} P(x_i|\theta)\right)$ $= \prod_{i=1}^{N} \prod_{j=1}^{d} (\theta_j)^{x_{ij}} (1-\theta_j)^{(1-x_{ij})}$

x_{ij} = i^{th} sample's j^{th} dimension

For max a priori :

We need to maximise

$$F(\theta) \quad \ln\left(\left(\prod_{i=1}^{N}\prod_{j=1}^{d} \left(\theta_j^{x_{ij}} (1-\theta_j)^{(1-x_{ij})}\right)\right)(\theta_1\theta_2 \dots \theta_d)\, e^{-(\theta_1+\theta_2 - \theta_d)}\right)$$

$$\frac{\partial}{\partial \theta_j} F(\theta) = \sum_{i=1}^{N} \left(\frac{x_{ij}}{\theta_j} - \frac{1-x_{ij}}{1-\theta_j}\right) + \frac{\partial}{\partial \theta_j}\left(\ln(\theta_1 e^{-\theta_1}) + \ln(\theta_2 e^{-\theta_2}) - \ln(\theta_d e^{-\theta_d})\right) = 0$$

$$\Rightarrow \sum_{i=1}^{N} \left(\frac{x_{ij}}{\theta_j} - \frac{(1-x_{ij})}{1-\theta_j}\right) + \frac{(1-\theta_j)\,\cancel{e^{-\theta_j}}}{\theta_j\, \cancel{e^{-\theta_j}}} = 0$$

$$\Rightarrow \frac{\sum_i x_{ij}}{\theta_j(1-\theta_j)} = \frac{N\theta_j}{\theta_j(1-\theta_j)} + \left(\frac{\theta_j - 1}{\theta_j}\right)$$

$$\frac{\sum x_{ij}}{\cancel{\theta_j (1-\theta_j)}} = \frac{N\theta_j - \theta_j^2 + 2\theta_j - 1}{\cancel{\theta_j (1-\theta_j)}}$$

$$\Rightarrow \quad \theta_j^2 - (N+2)\theta_j + \left(1 - \sum x_{ij}\right) = 0$$

$$\Rightarrow \quad \theta_j = \frac{(N+2) \pm \sqrt{(N+2)^2 - 4 + 4\sum x_{ij}}}{2}$$

b) Using the above expression

$$X = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

here $N = 4$

$$\sum_{i=1}^{N} x_{i1} = 3, \quad \sum_{i=1}^{N} x_{i2} = 1$$

$$\theta_{1_{MAP}} = \frac{6 \pm \sqrt{44}}{2} = 3 \pm \sqrt{11}$$

$$\theta_{2_{MAP}} = \frac{6 \pm \sqrt{36}}{2} = 6 \text{ or } 0$$

Q3 a) Arbitrary 2x2 matrix

$$X = \begin{bmatrix} 4 & 7 \\ 2 & 3 \end{bmatrix}$$

$$\mu_x = \begin{bmatrix} \frac{4+7}{2} \\ \frac{2+3}{2} \end{bmatrix} = \begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$$

$$X_c = X - \mu_x = \begin{bmatrix} -1.5 & 1.5 \\ -0.5 & 0.5 \end{bmatrix}$$

$$\text{cov matrix}(X_c) = S_{X_c} = \begin{bmatrix} Var(X_{c_1}) & cov(X_{c_2}, X_{c_1}) \\ cov(X_{c_1}, X_{c_2}) & Var(X_{c_2}) \end{bmatrix}$$

$$Var(X_{c_1}) = \left(\left(-1.5 - \left(\frac{1.5-1.5}{2}\right)\right)^2 + \left(1.5 - \left(\frac{1.5-1.5}{2}\right)\right)^2\right) / 2$$

$$= \frac{2.25 \times 2}{2} = \frac{4.5}{2} = 2.25$$

$$Var(X_{c_2}) = \left(\left(-0.5 - \left(\frac{0.5-0.5}{2}\right)\right)^2 + \left(0.5 - \left(\frac{0.5-0.5}{2}\right)\right)^2\right) / 2$$

$$= \frac{0.25 \times 2}{2} = 0.5/2 = 0.25$$

$$cov(X_{c_1}, X_{c_2}) = cov(X_{c_2}, X_{c_1})$$

$$= \left(\left(4 - \left(\frac{4+7}{2}\right)\right)\left(2 - \left(\frac{2+3}{2}\right)\right) + \left(7 - \left(\frac{4+7}{2}\right)\right)\left(3 - \left(\frac{2+3}{2}\right)\right)\right) / 2$$

$$= \cancel{1.25}\ 0.75$$

$$S_{X_c} = \begin{bmatrix} 2.25 & 0.75 \\ 0.75 & 0.25 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 4.5 & 1.5 \\ 1.5 & 0.5 \end{bmatrix}$$

For finding eigen values, $\det(S_{x_c} - \lambda I) = 0$

$$\det\begin{pmatrix} 4.5-\lambda & 1.5 \\ 1.5 & 0.5-\lambda \end{pmatrix} = 0$$

$\Rightarrow$ $(4.5-\lambda)(0.5-\lambda) - 2.25 = 0$

$\Rightarrow$ $\lambda^2 - 5\lambda = 0$

$\lambda_1 = 5 \quad \lambda_2 = 0$

for $\lambda = 5$

$$\begin{bmatrix} 4.5 & 1.5 \\ 1.5 & 0.5 \end{bmatrix}\begin{bmatrix} V_{11} \\ V_{12} \end{bmatrix} = 5\begin{bmatrix} \cancel{4.5} & \cancel{1.5} \\ 1.5 & \cancel{0.5} \end{bmatrix} = \begin{bmatrix} \cancel{22.5} & \cancel{7.5} \\ \cancel{7.5} & 2.5 \end{bmatrix} \quad 5\begin{bmatrix} V_{11} \\ V_{12} \end{bmatrix}$$

$4.5 V_{11} + 1.5 V_{12} = 5 V_{11} \quad \Rightarrow \quad 3V_{12} = V_{11}$

$1.5 V_{11} + 0.5 V_{12} = 5 V_{12}$

- eigen vector corresponding to $\lambda = 5$

$= \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

for $\lambda = 0$

$$\begin{bmatrix} 4.5 & 1.5 \\ 1.5 & 0.5 \end{bmatrix}\begin{bmatrix} V_{21} \\ V_{22} \end{bmatrix} = 0$$

$4.5V_{21} = -1.5V_{22}$

$1.5V_{21} = -0.5V_{22}$

$\Rightarrow 3V_{21} = -V_{22}$

corresponding eigen vector $= \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

$U = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}$ (with determinant 1)

$Y = U' x_c$

$$= \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix} \begin{bmatrix} -1.5 & 1.5 \\ -0.5 & 0.5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-5}{\sqrt{10}} & \frac{5}{\sqrt{10}} \\ 0 & 0 \end{bmatrix}$$

b) $UY + \text{mean}(X) = UY + \mu_X$

$$= \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} \frac{-5}{\sqrt{10}} & \frac{5}{\sqrt{10}} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-15}{10} & \frac{15}{10} \\ \frac{-5}{10} & \frac{5}{10} \end{bmatrix} + \begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 7 \\ 2 & 3 \end{bmatrix}$$

$$X = \begin{bmatrix} 4 & 7 \\ 2 & 3 \end{bmatrix}$$

MSE between $UY + \mu_X$ and X $= \frac{(4-4)^2 + (7-7)^2 + (2-2)^2 + (3-3)^2}{4}$

$= 0$

The calculations match with the code's output