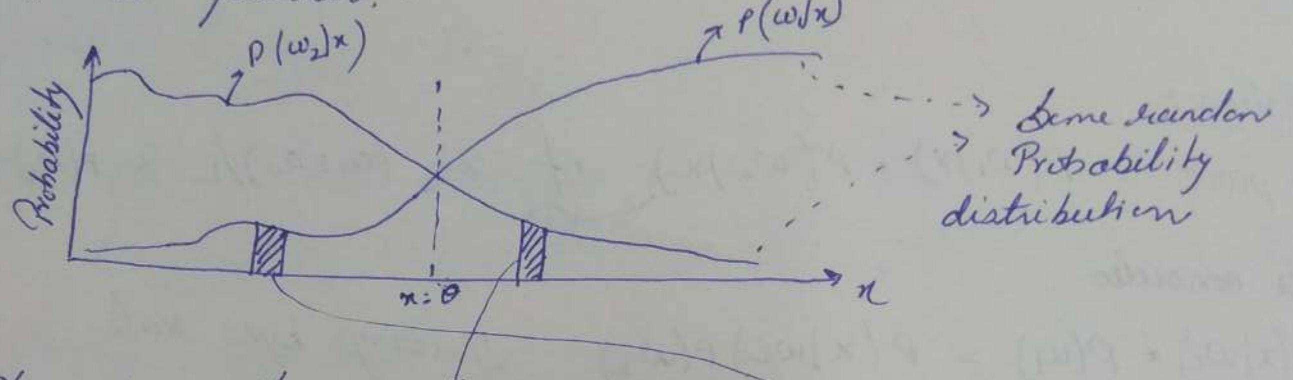
question 1

(a) here $x = \theta$ is the decision boundary.
Since there is no explicit information about probability distribution function, therefore I am not explicitly assuming to be gaussian.



$$P(error) = \int_{\theta}^{\infty} P(\omega_2|x)P(x)\,dx + \int_{-\infty}^{\theta} P(x)\,P(\omega_1|x)\,dx$$

$$\Rightarrow P(error) = \int_{\theta}^{\infty} \frac{P(x|\omega_2)P(\omega_2)}{\cancel{P(x)}} \cancel{P(x)}\,dx + \int_{-\infty}^{\theta} \frac{P(x|\omega_1)P(\omega_1)}{\cancel{P(x)}} \times \cancel{P(x)}\,dx$$

// using byes therem

$$\Rightarrow P(error) = P(\omega_1)\int_{-\infty}^{\theta} P(x|\omega_1)\,dx + \int_{\theta}^{\infty} P(x|\omega_2)\;P(\omega_2)\,dx \quad \rightarrow (1)$$

(b) In order to minimize P(error) we need to differentiate eq (1) using the Leibniz's rule.

Leibniz's Rule : $$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)\,dt\right) = f(x,b(x))\frac{d}{dx}b(x) - f(x,a(x))\frac{d}{dx}a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\,dt$$

$$P'(error) = 0$$

$$\Rightarrow P(\omega_1)\frac{d}{d\theta}\int_{-\infty}^{\theta} P(x|\omega_1)\,dx + P(\omega_2)\frac{d}{d\theta}\int_{\theta}^{\infty} P(x|\omega_2)\,dx = 0$$

$$\Rightarrow P(\omega_1)\left[P(\theta|\omega_1)\frac{d\theta}{d\theta} - P(-\infty|\omega_1)\frac{d\infty}{d\theta}\right] + P(\omega_2)\left[P(\infty|\omega_2)\frac{d\infty}{d\theta} - P(\theta|\omega_2)\frac{d\theta}{d\theta}\right] = 0$$

$\Rightarrow P(\omega_1)P(\theta|\omega_1) - P(\omega_2)P(\theta|\omega_2) = 0 \quad \because P(-\alpha|\omega_1) = 0, P(\alpha|\omega_2) = 0$

$\Rightarrow$

$P(\theta|\omega_1)P(\omega_1) = P(\omega_2)P(\theta|\omega_2)$

Hence Proved

Question 2

To prove :- $P(\omega_1|x) = P(\omega_2|x)$ if $x = (a_1+a_2)/2$ & $P(\omega_1) = P(\omega_2)$ ①

Let's consider

$\Rightarrow P(x|\omega_1) * P(\omega_1) = P(x|\omega_2)P(\omega_2)$ // using byes rule

$\Rightarrow \frac{1}{\pi b}\left[\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}\right]P(\omega_1) = P(\omega_2)\frac{1}{\pi b}\left[\frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}\right] \quad \because P(x|\omega_i) = \frac{1}{\pi b}\frac{1}{1+\left(\frac{x-a_i}{b}\right)^2}$ $i=1,2$

$\Rightarrow$ rearranging the terms and put $P(\omega_2) = P(\omega_1)$ [given]

$\Rightarrow \frac{1}{\pi b}P(\omega_1)\left[\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}\right] = \frac{1}{\pi b}P(\omega_1)\left[\frac{1}{1+\left(\frac{x-a_2}{b}\right)^2}\right]$

= put $x = \left(\frac{a_1+a_2}{2}\right)$

$\Rightarrow \frac{1}{\pi b}P(\omega_1)\left[\frac{1}{1+\left[\frac{\frac{a_1+a_2}{2}-a_1}{b}\right]^2}\right] = \frac{1}{\pi b}P(\omega_1)\left[\frac{1}{1+\left(\frac{\frac{a_1+a_2}{2}-a_2}{b}\right)^2}\right]$

$\Rightarrow \frac{1}{\pi b}P(\omega_1)\left[\frac{1}{1+\left(\frac{a_2-a_1}{b}\right)^2}\right] = \frac{1}{\pi b}P(\omega_1)\left(\frac{1}{1+\left(\frac{a_1-a_2}{b}\right)^2}\right)$

since $\left(\frac{a_2-a_1}{b}\right)^2 = \left(\frac{a_1-a_2}{b}\right)^2 \quad \therefore LHS = RHS$

hence the min error decision boundary is a point midway between the peaks of the two distributions regardless of b's value.

Question :3

I am making use of discriminant funct on posterior probabilities in order to classify $x=\begin{bmatrix}0.3\\0.3\end{bmatrix}$ in one among three classes

$$g_i(x) = \ln \frac{1}{(2\pi)^{d/2}} + \ln \frac{1}{|\Sigma_i|^{1/2}} - \frac{1}{2}(x-\mu_i)^T \Sigma^{-1}(x-\mu_i)$$

here $\Sigma = I$ for all three classes and classes are equiprobable

$\therefore$ $\ln \frac{1}{(2\pi)^{d/2}}$ & $\ln \frac{1}{|\Sigma|^{1/2}}$ terms can be ignored to get the following "standard expression"

$$g_i(x) = \frac{\mu_i^T x}{\sigma^2} - \frac{\mu_i^T \mu}{2\sigma^2} \longrightarrow 1$$

(i) for $N(0, I)$

$$g_i(x) = [0 \quad 0]\begin{bmatrix}0.3\\0.3\end{bmatrix} - \frac{[0 \quad 0]}{2}\begin{bmatrix}0\\0\end{bmatrix} = 0 \longrightarrow a$$

(ii) for $N\left(\begin{bmatrix}1\\1\end{bmatrix}, I\right)$

$$g_j(x) = [1 \quad 1]\begin{bmatrix}0.3\\0.3\end{bmatrix} - \frac{1}{2}[1 \quad 1]\begin{bmatrix}1\\1\end{bmatrix} = -0.4 \longrightarrow b$$

(iii) for $0.5N\left(\begin{bmatrix}0.5\\0.5\end{bmatrix}, I\right) + 0.5N\left(\begin{bmatrix}-0.5\\0.5\end{bmatrix}, I\right)$

$$g_k(x) = 0.5\left([0.5 \quad 0.5]\begin{bmatrix}0.3\\0.3\end{bmatrix} - \frac{1}{2}[0.5 \quad 0.5]\begin{bmatrix}0.5\\0.5\end{bmatrix}\right) + 0.5\left([0.5 \quad 0.5]\begin{bmatrix}0.3\\0.3\end{bmatrix} - \frac{[0.5 \quad 0.5]}{2} \times \begin{bmatrix}0.5\\0.5\end{bmatrix}\right)$$

$$g_k(x) = -0.1 \longrightarrow c$$

since out of $g_i(x)$, $g_j(x)$ & $g_k(x)$ from eq^n a, b, c $g_i(x)$ is greatest $\therefore$ pt $x=\begin{bmatrix}0.3\\0.3\end{bmatrix}$ will be classified under $p(x|\omega_1) \sim N(0, I)$

Ques 4 (c) I have made use of script to classify $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in class 1 or class 2.

The mathematical derivation is as follows

$$g_i(x) = \ln(P(\omega_i)) + \frac{\mu_i^T x}{\sigma^2} - \frac{\mu_i^T \mu}{2\sigma^2}$$

(i) for class 1 $N\left(\mu_1 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \Sigma_1 = \sigma^2 I\right)$

$$g_i(x) = \ln(1/3) + \frac{[0.5 \ 0.5]}{1}\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{[0.5 \ 0.5]}{2}\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

$$g_i(x) = \ln(1/3) + 1 - \frac{1}{2} \times 0.50 = -0.3496$$

(ii) for class 2 $N\left(\mu_2 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}, \Sigma_2 = \sigma^2 I\right)$

$$g_j(x) = \ln(2/3) + [-0.5 \ 0.5]\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{[-0.5 \ 0.5]}{2}\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

$$= \ln(2/3) + 0 - \frac{1}{2} \times 0.50 = -0.6549$$

since $g_i(x) > g_j(x)$

$\therefore x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ belongs to $N\left(\mu = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \Sigma = \sigma^2 I\right)$ class I.

// Same result is inferred from the python script of ques 4-b as well.