Homework 2

2. We are given that the density is of the form $p(x|\omega_i) = ke^{-|x-a_i|/b_i}$.

(a) We seek k so that the function is normalized, as required by a true density. We integrate this function, set it to 1.0,

$$k\left[\int_{-\infty}^{a_i} \exp[(x-a_i)/b_i]dx + \int_{a_i}^{\infty} \exp[-(x-a_i)/b_i]dx\right] = 1,$$

which yields $2b_i k = 1$ or $k = 1/(2b_i)$. Note that the normalization is independent of a_i, which corresponds to a shift along the axis and is hence indeed irrelevant to normalization. The distribution is therefore written

$$p(x|\omega_i) = \frac{1}{2b_i}e^{-|x-a_i|/b_i}.$$

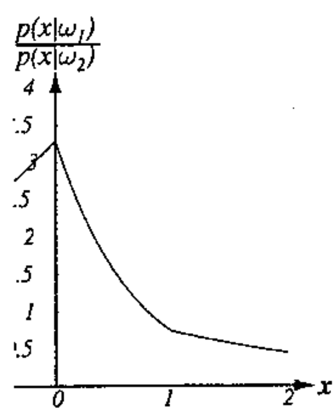
(b) The likelihood ratio can be written directly:

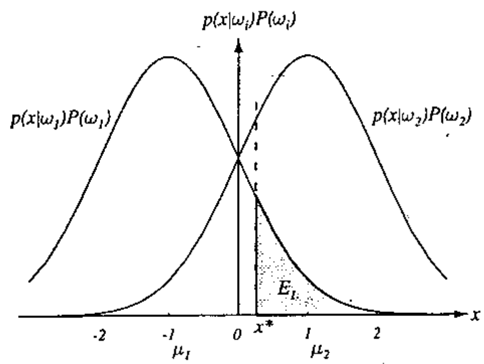
$$\frac{p(x|\omega_1)}{p(x|\omega_2)} = \frac{b_2}{b_1}\exp\left[-\frac{|x-a_1|}{b_1} + \frac{|x-a_2|}{b_2}\right].$$

(c) For the case $a_1 = 0$, $a_2 = 1$, $b_1 = 1$ and $b_2 = 2$, we have the likelihood ratio is

$$\frac{p(x|\omega_2)}{p(x|\omega_1)} = \begin{cases} 2e^{(x+1)/2} & x \le 0 \\ 2e^{(1-3x)/2} & 0 < x \le 1 \\ 2e^{(-x-1)/2} & x > 1, \end{cases}$$

as shown in the figure.





6. We let x^* denote our decision boundary and $\mu_2 > \mu_1$, as shown in the figure.

(a) The error for classifying a pattern that is actually in ω_1 as if it were in ω_2 is:

$$\int_{\mathcal{R}_2} p(x|\omega_1)P(\omega_1)\, dx = \frac{1}{2}\int_{x^*}^{\infty} N(\mu_1, \sigma_1^2)\, dx \leq E_1.$$

Our problem demands that this error be less than or equal to E_1. Thus the bound on x^* is a function of E_1, and could be obtained by tables of cumulative normal distributions, or simple numerical integration.

(b) Likewise, the error for categorizing a pattern that is in ω_2 as if it were in ω_1 is:

$$E_2 = \int_{\mathcal{R}_1} p(x|\omega_2)P(\omega_2)\, dx = \frac{1}{2}\int_{-\infty}^{x^*} N(\mu_2, \sigma_2^2)\, dx.$$

(c) The total error is simply the sum of these two contributions:

$$\begin{aligned} E &= E_1 + E_2 \\ &= \frac{1}{2}\int_{x^*}^{\infty} N(\mu_1, \sigma_1^2)\, dx + \frac{1}{2}\int_{-\infty}^{x^*} N(\mu_2, \sigma_2^2)\, dx. \end{aligned}$$

(d) For $p(x|\omega_1) \sim N(-1/2, 1)$ and $p(x|\omega_2) \sim N(1/2, 1)$ and $E_1 = 0.05$, we have (by simple numerical integration) $x^* = 0.2815$, and thus

$$\begin{aligned} E &= 0.05 + \frac{1}{2}\int_{-\infty}^{0.2815} N(\mu_2, \sigma_2^2)\, dx \\ &= 0.05 + \frac{1}{2}\int_{-\infty}^{0.2815} \frac{1}{\sqrt{2\pi}0.05}\exp\left[-\frac{(x-0.5)^2}{2(0.5)^2}\right] dx \\ &= 0.168. \end{aligned}$$

(e) The decision boundary for the (minimum error) Bayes case is clearly at $x^* = 0$. The Bayes error for this problem is:

$$E_B = 2\int_{0}^{\infty} \frac{1}{2}N(\mu_1, \sigma_1^2)\, dx$$

$$= \int_{0}^{\infty} N(1,1)\, dx = \text{erf}[1] = 0.159,$$

which of course is lower than the error for the Neyman-Pearson criterion case. Note that if the Bayes error were lower than $2 \times 0.05 = 0.1$ in this problem, we would use the Bayes decision point for the Neyman-Pearson case, since it too would ensure that the Neyman-Pearson criteria were obeyed *and* would give the lowest total error.

37. We are given that $P(\omega_1) = P(\omega_2) = 0.5$ and

$$\begin{aligned} p(\mathbf{x}|\omega_1) &\sim N(\mathbf{0}, \mathbf{I}) \\ p(\mathbf{x}|\omega_2) &\sim N(\mathbf{1}, \mathbf{I}) \end{aligned}$$

where $\mathbf{1}$ is a two-component vector of 1s.

(a) The inverse matrices are simple in this case:

$$\boldsymbol{\Sigma}_1^{-1} = \boldsymbol{\Sigma}_2^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We substitute these into Eqs. 53–55 in the text and find

$$\begin{aligned} g_1(\mathbf{x}) &= \mathbf{w}_1^t\mathbf{x} + w_{10} \\ &= \mathbf{0}^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 0 + \ln(1/2) \\ &= \ln(1/2) \end{aligned}$$

and

$$\begin{aligned} g_2(\mathbf{x}) &= \mathbf{w}_2^t\mathbf{x} + w_{20} \\ &= (1,1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \frac{1}{2}(1,1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \ln(1/2) \\ &= x_1 + x_2 - 1 + \ln(1/2). \end{aligned}$$

We set $g_1(\mathbf{x}) = g_2(\mathbf{x})$ and find the decision boundardy is $x_1 + x_2 = 1$, which passes through the midpoint of the two means, that is, at

$$(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)/2 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

This result makes sense because these two categories have the same prior and conditional distributions except for their means.

(b) We use Eqs. 76 in the text and substitute the values given to find

$$\begin{aligned} k(1/2) &= \frac{1}{8}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)^t \left[\frac{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}{2}\right]^{-1} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) + \frac{1}{2}\ln\frac{\left|\frac{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}{2}\right|}{\sqrt{\left|\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right|\left|\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right|}} \\ &= \frac{1}{8}\begin{pmatrix} 1 \\ 1 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}\ln\frac{1}{1} \\ &= 1/4. \end{aligned}$$

Equation 77 in the text gives the Bhatacharyya bound as

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}e^{-k(1/2)} = \sqrt{0.5 \cdot 0.5}e^{-1/4} = 0.3894.$$

(c) Here we have $P(\omega_1) = P(\omega_2) = 0.5$ and

$$\mu_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad , \quad \Sigma_1 = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 2 \end{pmatrix}$$
$$\mu_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad \Sigma_2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

The inverse matrices are

$$\Sigma_1^{-1} = \begin{pmatrix} 8/5 & -2/15 \\ -2/15 & 8/15 \end{pmatrix}$$
$$\Sigma_2^{-1} = \begin{pmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{pmatrix}.$$

We use Eqs. 66–69 and find

$$\begin{aligned} g_1(\mathbf{x}) &= -\frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 8/5 & -2/15 \\ -2/15 & 8/15 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \left(\begin{pmatrix} 8/5 & -2/15 \\ -2/15 & 8/15 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &\quad -\frac{1}{2}\begin{pmatrix} 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} 8/5 & -2/15 \\ -2/15 & 8/15 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{2}\ln\left|\begin{pmatrix} 2 & 0.5 \\ 0.5 & 2 \end{pmatrix}\right| + \ln\frac{1}{2} \\ &= -\frac{4}{15}x_1^2 + \frac{2}{15}x_1x_2 - \frac{4}{15}x_2^2 - 0.66 + \ln\frac{1}{2}, \end{aligned}$$

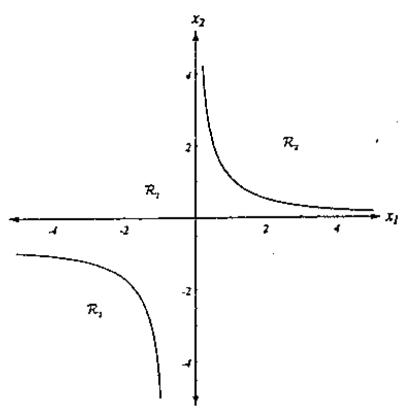
and

$$\begin{aligned} g_2(\mathbf{x}) &= -\frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \left(\begin{pmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &\quad -\frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}^t \begin{pmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2}\ln\left|\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}\right| + \ln\frac{1}{2} \\ &= -\frac{5}{18}x_1^2 + \frac{8}{18}x_1x_2 - \frac{5}{18}x_2^2 + \frac{1}{9}x_1 + \frac{1}{9}x_2 - \frac{1}{9} - 1.1 + \ln\frac{1}{2}. \end{aligned}$$

The Bayes decision boundary is the solution to $g_1(\mathbf{x}) = g_2(\mathbf{x})$ or

$$x_1^2 + x_2^2 - 28x_1x_2 - 10x_1 - 10x_2 + 50 = 0,$$

which consists of two hyperbolas, as shown in the figure.



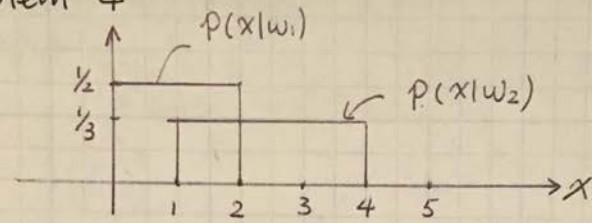
We use Eqs. 76 and 77 in the text and find

$$
\begin{aligned}
k(1/2) &= \frac{1}{8}\left(\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right)^t \left[\frac{\begin{pmatrix}2 & 0.5\\0.5 & 2\end{pmatrix}+\begin{pmatrix}5 & 4\\4 & 5\end{pmatrix}}{2}\right]^{-1}\left(\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\right)+\ln\frac{\left|\frac{\begin{pmatrix}2 & 0.5\\0.5 & 2\end{pmatrix}+\begin{pmatrix}5 & 4\\4 & 5\end{pmatrix}}{2}\right|}{\sqrt{\left|\begin{pmatrix}2 & 0.5\\0.5 & 2\end{pmatrix}\right|\left|\begin{pmatrix}5 & 4\\4 & 5\end{pmatrix}\right|}} \\
&= \frac{1}{8}\begin{pmatrix}1\\1\end{pmatrix}^t\begin{pmatrix}3.5 & 2.25\\2.25 & 3.5\end{pmatrix}^{-1}\begin{pmatrix}1\\1\end{pmatrix}+\frac{1}{2}\ln\frac{7.1875}{5.8095} \\
&= 0.1499.
\end{aligned}
$$

Equation 77 in the text gives the Bhatacharyya bound as

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}e^{-k(1/2)} = \sqrt{0.5\cdot 0.5}e^{-1.5439} = 0.4304.$$

Problem 4



$P(\omega_1) = P(\omega_2) = 0.5$

1. Decide ω_1 if $\dfrac{p(x|\omega_1)}{p(x|\omega_2)} > \dfrac{P(\omega_2)}{P(\omega_1)} = 1$

It can be seen from the graph above that this is true when $x < 2$.

Hence, the decision boundary $x^* = 2$.

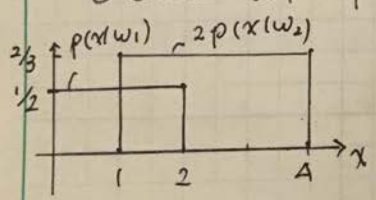
<u>Decide ω_1 if $x<2$, otherwise decide ω_2.</u>

<u>P(error)</u> $= p(x \in R_2|\omega_1)P(\omega_1) + p(x \in R_1|\omega_2)P(\omega_2)$

$= 0 + \frac{1}{2}\int_1^2 \frac{1}{3}\, dx = \frac{1}{2} \times \frac{1}{3} = \underline{\frac{1}{6}}$

2. To find the decision boundary that minimizes the overall risk with $\lambda_{12} = 2\lambda_{21}$

Decide ω_1 if $\dfrac{p(x|\omega_1)}{p(x|\omega_2)} > \dfrac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}} \cdot \dfrac{P(\omega_2)}{P(\omega_1)} = 2$

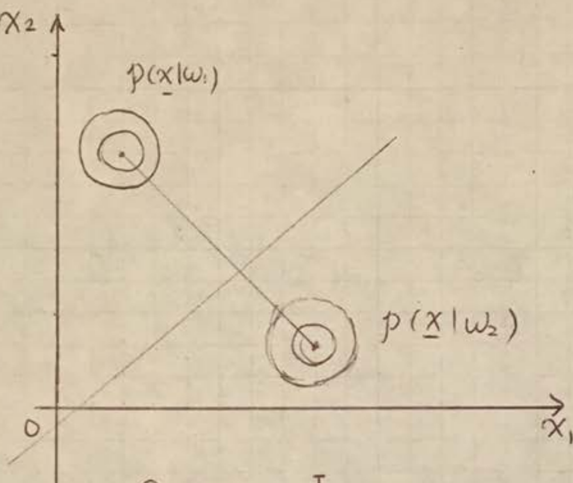


$p(x|\omega_1) > 2p(x|\omega_2)$

It can be seen $x^* = 1$

<u>Decide ω_1 if $x<1$, otherwise decide ω_1.</u>

Problem 5



$P(\omega=1)=0.6, \quad P(\omega=2)=0.4$

$\underline{\mu}_1=\begin{pmatrix}4\\16\end{pmatrix}, \quad \underline{\mu}_2=\begin{pmatrix}16\\4\end{pmatrix}$

$\Sigma_1=\Sigma_2=\begin{pmatrix}4&0\\0&4\end{pmatrix}$

This is case I.

$g_i(\underline{x}) = W_i^T \underline{x} + \omega_{i0}$

where $W_i = \frac{1}{\sigma^2}\mu_i, \quad \omega_{i0} = -\frac{1}{2\sigma^2}\mu_i^T\mu_i + \ln P(\omega_i)$

$W_1 = \frac{1}{4}\mu_1 = \begin{pmatrix}1\\4\end{pmatrix}; \quad \omega_{10} = -\frac{1}{8}(4\ 16)\begin{pmatrix}4\\16\end{pmatrix} + \ln(0.6) = -\frac{1}{8}\times(16+256) + \ln 0.6$

$= -34 + \ln(0.6)$

$\underline{g_1(\underline{x}) = (1\ 4)\begin{pmatrix}x_1\\x_2\end{pmatrix} - 34 + \ln(0.6) = x_1 + 4x_2 - 34 + \ln(0.6)}$

$W_2 = \frac{1}{4}\mu_2 = \begin{pmatrix}4\\1\end{pmatrix}; \quad \omega_{20} = -\frac{1}{8}(16\ 4)\begin{pmatrix}16\\4\end{pmatrix} + \ln(0.4) = -\frac{1}{8}(256+16) + \ln 0.4$

$= -34 + \ln(0.4)$

$\underline{g_2(\underline{x}) = (4\ 1)\begin{pmatrix}x_1\\x_2\end{pmatrix} - 34 + \ln(0.4) = 4x_1 + x_2 - 34 + \ln(0.4)}$

$g_1(\underline{x}) = g_2(x)$

$x_1 + 4x_2 - 34 + \ln(0.6) = 4x_1 + x_2 - 34 + \ln(0.4)$

$3x_1 - 3x_2 + \ln\frac{4}{6} = 0$

$\underline{x_1 - x_2 = 0.135}$