Homework 1 Solutions

1. Problem 1

Google "Monty Hall Problem" for solutions.

2. Problem 2

We have:

$$R(\alpha_1|\mathbf{x}) = \lambda_{11}P(\omega_1|\mathbf{x}) + \lambda_{12}P(\omega_2|\mathbf{x})$$
$$R(\alpha_2|\mathbf{x}) = \lambda_{21}P(\omega_1|\mathbf{x}) + \lambda_{22}P(\omega_2|\mathbf{x})$$

Thus we decide ω_1 if

$$(\lambda_{21} - \lambda_{11})P(\mathbf{x}|\omega_1)P(\omega_1) > (\lambda_{12} - \lambda_{22})P(\mathbf{x}|\omega_2)P(\omega_2)$$

Otherwise, decide ω_2.
Because $\lambda_{12} = 2$, $\lambda_{21} = 1$, and $\lambda_{11} = \lambda_{22} = 0$,
Then if we decide ω_1, We get:

$P(\mathbf{x}|\omega_1)P(\omega_1) > 2P(\mathbf{x}|\omega_2)P(\omega_2)$

$$\begin{aligned}
&\implies ln(P(\mathbf{x}|\omega_1)) + ln(P(\omega_1)) > ln2 + ln(P(\mathbf{x})|\omega_2) + ln(P(\omega_2))\\
&\implies -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) - \frac{d}{2}ln(2\pi) - \frac{1}{2}ln|\boldsymbol{\Sigma}| + ln(P(\omega_1))\\
&\quad > ln2 - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) - \frac{d}{2}ln(2\pi) - \frac{1}{2}ln|\boldsymbol{\Sigma}| + ln(P(\omega_2))\\
&\implies -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + ln(P(\omega_1)) > ln2 - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) + ln(P(\omega_2))
\end{aligned}$$

otherwise Decide ω_2.

Let $g_1(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + ln(P(\omega_1))$,
Let $g_2(\mathbf{x}) = ln2 - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) + ln(P(\omega_2))$.

the decision boundary is

$$
\begin{aligned}
& g_1(\mathbf{x}) = g_2(\mathbf{x}) \\
& \implies g_1(\mathbf{x}) - g_2(\mathbf{x}) = 0 \\
& \implies -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^t \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) + lnP(\omega_1) + \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^t \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2) + lnP(\omega_2) = ln2 \\
& \implies -\frac{1}{2}(\mathbf{x}^t\boldsymbol{\Sigma}^{-1}\mathbf{x} - \boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x}^t\boldsymbol{\Sigma}^{-1}\mathbf{x} - \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2 + \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2) \\
& \quad + ln\frac{P(\omega_1)}{P(\omega_2)} = ln2 \\
& \implies \frac{1}{2}(2\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1}\mathbf{x} - \boldsymbol{\mu}_1\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1) + \frac{1}{2}(-2\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}_1\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1) + ln\frac{P(\omega_1)}{P(\omega_2)} = ln2 \\
& \implies (\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1} - \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1})\mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_1\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2) + ln\frac{P(\omega_1)}{P(\omega_2)} = ln2 \qquad (1)
\end{aligned}
$$

Let $\mathbf{w} = \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$, and $\mathbf{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) - \frac{ln[P(\omega_1)/P(\omega_2)]}{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)} \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$.
Then,

$$
\begin{aligned}
\mathbf{w}^t\mathbf{x} &= (\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1} - \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1})\mathbf{x} \\
\mathbf{w}^t\mathbf{x}_0 &= (\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1} - \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1}) \cdot \Big[\frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) - \frac{ln[P(\omega_1)/P(\omega_2)]}{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)} \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\Big] \\
&= \frac{1}{2}\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 - \frac{1}{2}\boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2 - ln\frac{P(\omega_1)}{P(\omega_2)}
\end{aligned}
$$

Thus,

$$
\mathbf{w}^t\mathbf{x} - \mathbf{w}^t\mathbf{x}_0 = (\boldsymbol{\mu}_1^t\boldsymbol{\Sigma}^{-1} - \boldsymbol{\mu}_2^t\boldsymbol{\Sigma}^{-1})\mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_1\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2) + ln\frac{P(\omega_1)}{P(\omega_2)} = (1) = ln2
$$

Therefore the decision boundary is:

$$
\boxed{\mathbf{w}^t(\mathbf{x} - \mathbf{x}_0) = ln2}
$$

With:

$$
\begin{aligned}
\mathbf{w} &= \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2); \\
\mathbf{x}_0 &= \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) - \frac{ln[P(\omega_1)/P(\omega_2)]}{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^t\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)} \cdot (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)
\end{aligned}
$$

3. Problem 3

a)

$$
\begin{aligned}
g_i(x) &= P(x|\omega_i)P(\omega_i) \\
g_1(x) &= \begin{cases} \frac{-x+1}{2} \cdot 0.5 = \frac{-x+4}{4},\ x \in [-1,1] \\ 0,\ otherwise \end{cases} \\
g_2(x) &= \begin{cases} \frac{x+1}{2} \cdot 0.5 = \frac{x+4}{4},\ x \in [-1,1] \\ 0,\ otherwise \end{cases}
\end{aligned}
$$

The decision boundary is $g_1(x) = g_2(x)$

$\implies \frac{-x+1}{4} = \frac{x+1}{4},\ x \in [-1, 1]$
$\implies -x = x$
$\implies \boxed{\text{decision boundary is } x = 0}$

Since $g_1(x) > g_2(x)$ when $x \in [-1, 0)$,
decide ω_1 for $x \in [-1, 0)$, and ω_2 otherwise.

$$\begin{aligned}
P(error) &= \int\limits_{R_2} P(x|\omega_1)P(\omega_1)dx + \int\limits_{R_1} P(x|\omega_2)P(\omega_2)dx \\
&= \int_0^1 \frac{-x+1}{4} dx + \int_{-1}^0 \frac{x+1}{4} dx \\
&= \frac{1}{4}(-\frac{x^2}{2} + x)\Big|_0^1 + \frac{1}{4}(-\frac{x^2}{2} + x)\Big|_{-1}^0 \\
&= \frac{1}{4}(-\frac{1}{2} + 1) + \frac{1}{4}(-\frac{1}{2} + 1) \\
&= \frac{1}{4}
\end{aligned}$$

Therefore, the $\boxed{\text{Bayes Error} = \frac{1}{4}}$

b)

$$\begin{aligned}
g_i(x) &= P(x|\omega_i)P(\omega_i) \\
g_1(x) &= \begin{cases} \frac{-x+1}{2} \cdot 0.7 = \frac{-7x+7}{20},\ x \in [-1, 1] \\ 0,\ otherwise \end{cases} \\
g_2(x) &= \begin{cases} \frac{x+1}{2} \cdot 0.3 = \frac{3x+3}{20},\ x \in [-1, 1] \\ 0,\ otherwise \end{cases}
\end{aligned}$$

The decision boundary is $g_1(x) = g_2(x)$
$\implies \frac{-7x+7}{20} = \frac{3x+3}{20},\ x \in [-1, 1]$
$\implies 10x = 4$
$\implies \boxed{\text{decision boundary is } x = \frac{2}{5}}$

Since $g_1(x) > g_2(x)$ when $x \in [-1, \frac{2}{5})$,
decide ω_1 for $x \in [-1, \frac{2}{5})$, and ω_2 otherwise.

$$\begin{aligned}
P(error) &= \int\limits_{R_2} P(x|\omega_1)P(\omega_1)dx + \int\limits_{R_1} P(x|\omega_2)P(\omega_2)dx \\
&= \int_{0.4}^1 \frac{-7x+7}{20} dx + \int_{-1}^{0.4} \frac{3x+3}{20} dx \\
&= \frac{7}{20}(-\frac{x^2}{2} + x)\Big|_{0.4}^1 + \frac{3}{20}(-\frac{x^2}{2} + x)\Big|_{-1}^{0.4} \\
&= 0.21
\end{aligned}$$

Therefore, the $\boxed{\text{Bayes Error} = 0.21}$

c)
Decide ω_1 if

$$(\lambda_{21} - \lambda_{11})P(x|\omega_1)P(\omega_1) > (\lambda_{12} - \lambda_{22})P(x|\omega_2)P(\omega_2)$$

Otherwise, decide ω_2.

Thus,

$$\begin{aligned}
g_1(x) &= (\lambda_{21} - \lambda_{11})P(x|\omega_1)P(\omega_1) \\
&= P(x|\omega_1)P(\omega_1) \\
&= \begin{cases} \frac{-x+1}{2} \cdot 0.5 = \frac{-x+1}{4}, \ x \in [-1, 1] \\ 0, \ otherwise \end{cases} \\
g_2(x) &= (\lambda_{12} - \lambda_{22})P(x|\omega_2)P(\omega_2) \\
&= 2P(x|\omega_2)P(\omega_2) \\
&= \begin{cases} \frac{x+1}{2} \ x \in [-1, 1] \\ 0, \ otherwise \end{cases}
\end{aligned}$$

Then
$g_1(x) = g_2(x) \implies \frac{-x+1}{4} = \frac{x+1}{2}$
$\implies -x + 1 = 2x + 2 \implies 3x = -1$
$\implies \boxed{\text{The decision boundary is } x = -\dfrac{1}{3}}$
Since $g_1(x) > g_2(x)$ when $x \in [-1, -\frac{1}{3})$,
decide ω_1 for $x \in [-1, -\frac{1}{3})$, and ω_2 otherwise.

Then,

$$\begin{aligned}
R &= \int R(\alpha(x)|x)P(x)dx \\
&= \int_{R_1} R(\alpha_1|x)P(x)dx + \int_{R_2} R(\alpha_2|x)P(x)dx \\
&= \int_{R_1} \lambda_{12}P(\omega_2|x)P(x)dx + \int_{R_2} \lambda_{21}P(\omega_1|x)P(x)dx \\
&= \int_{R_1} \lambda_{12}P(x|\omega_2)P(\omega_2)dx + \int_{R_2} \lambda_{21}P(x|\omega_1)P(\omega_1)dx \\
&= \int_{-1}^{-\frac{1}{3}} 2 \cdot (\frac{x+1}{4})dx + \int_{-\frac{1}{3}}^{1} (\frac{x+1}{4})dx \\
&= \frac{1}{2}(\frac{x^2}{2} + x)\Big|_{-1}^{-\frac{1}{3}} + \frac{1}{4}(-\frac{x^2}{2} + x)\Big|_{-\frac{1}{3}}^{1} \\
&= \frac{1}{3}
\end{aligned}$$

Thus the $\boxed{\text{Bayes Risk} = 0.33}$.

4. Problem 4

a)

$$
\begin{aligned}
P(w_0|x_1) &= P(w_0, x_1)/P(x_1) \\
&= \sum_{Y,Z} P(x_1, Y, Z, w_0)/P(x_1) \\
&= \sum_{Y,Z} P(x_1)P(Y|x_1)P(Z|Y)P(w_0|Z)/P(x_1) \\
&= P(y_0|x_1)P(z_0|y_0)P(w_0|z_0) \\
&\quad + P(y_1|x_1)P(z_0|y_1)P(w_0|z_0) \\
&\quad + P(y_0|x_1)P(z_1|y_0)P(w_0|z_1) \\
&\quad + P(y_1|x_1)P(z_1|y_1)P(w_0|z_1) \\
&= (1-0.4)(1-0.6)(1-0.3) \\
&\quad + (0.4)(1-0.25)(1-0.3) \\
&\quad + (1-0.4)(0.6)(1-0.45) \\
&\quad + (0.4)(0.25)(1-0.45) \\
&= \boxed{0.631}
\end{aligned}
$$

b)

$$
\begin{aligned}
P(x_0|w_1) &= P(x_0, w_1)/P(w_1) \\
&= \frac{\sum_{Y,Z} P(x_0)P(Y|x_0)P(Z|Y)P(w_1|Z)}{\sum_{X,Y,Z} P(X)P(Y|X)P(Z|Y)P(w_1|Z)} \\
&= \frac{\sum_{Y,Z} P(x_0)P(Y|x_0)P(Z|Y)P(w_1|Z)}{\sum_{Y,Z} P(x_0)P(Y|x_0)P(Z|Y)P(w_1|Z) + \sum_{Y,Z} P(x_1)P(Y|x_1)P(Z|Y)P(w_1|Z)} \\
&= \boxed{0.403}
\end{aligned}
$$

Q5. True-or-False: For a two-class classification problem using the minimum-error-rate rule, in general the decision boundary can take any form. However, if the underlying class-conditionals are Gaussian densities, then the decision boundary is linear (hyperplanes).

[ ] True [X] False

Brief explanation of your answer: whether the decision boundary is linear depends on the covariance matrices of the two classes. For example, in the three cases discussed in the lecture notes, only the first two cases lead to linear decision boundaries.