Part I

We calculate the two conditional pdf using Figure 1, the result as follows:

$$P(x|\omega_1) = \begin{cases} \frac{x}{2} - \frac{1}{2}, 1 \leq x \leq 2 \\ -\frac{x}{6} + \frac{5}{6}, 2 \leq x \leq 5 \end{cases}, \qquad P(x|\omega_2) = \begin{cases} \frac{x}{12} - \frac{1}{4}, 3 \leq x \leq 6 \\ \frac{5}{16}, 6 \leq x \leq 8 \end{cases}$$

By the condition that both classes are equally likely, so $P(\omega_1) = P(\omega_2) = \frac{1}{2}$ Therefore,

$$\omega^* = arg \min_{\omega_j} P(x|\omega_j)P(\omega_j) \triangleq arg \min_{\omega_j} P(x|\omega_j)$$

(a) Because $x = 4 < x* = 5.1$, then $x \in \omega_1$.
(b) Beh's classifier will always correctly classify ω_2.
(c)The class with none-zero error is ω_1.
Because $P(\text{error}|\omega_1) = \text{P(x} \in \omega_2\text{|x is assigned to } \omega_1)$, using Figure 1 we have

$$P(\text{error}|\omega_1) = \int_3^{5.1} \frac{\text{x}}{12} - \frac{1}{4}\text{dx} = 0.1838$$

(d)Solve the equation systems below, we can obtain the intersection of p(x|ω_1) and p(x|ω_2):

$$\begin{cases} p(x|\omega_1) = -\frac{x}{6} + \frac{5}{6}, 2 \leq x \leq 5 \\ p(x|\omega_2) = \frac{x}{12} - \frac{1}{4}, 3 \leq x \leq 6 \end{cases} \rightarrow (\frac{13}{3}, \frac{1}{9})$$

In order to choose the decision rule for Bayes' classifer, we minimize the probability of error that Baye's classifier will make for their two classes separately.
Target function is :

$$\begin{aligned} \min P(error) &= \min\{P(\text{error}|\omega_1)\text{P}(\omega_1) + \text{P(error}|\omega_2)\text{P}(\omega_2)\} \\ &\triangleq min\{P(\text{error}|\omega_1) + \text{P(error}|\omega_2)\} \end{aligned} \tag{1}$$

The $P(\text{error}|\omega_1)$ is the figure acreage enclosed by pdf of ω_2 and the decision boundary under rectangular coordinate system; the $P(\text{error}|\omega_2)$ is the figure acreage enclosed by pdf of ω_1 and the decision boundary under rectangular coordinate system.

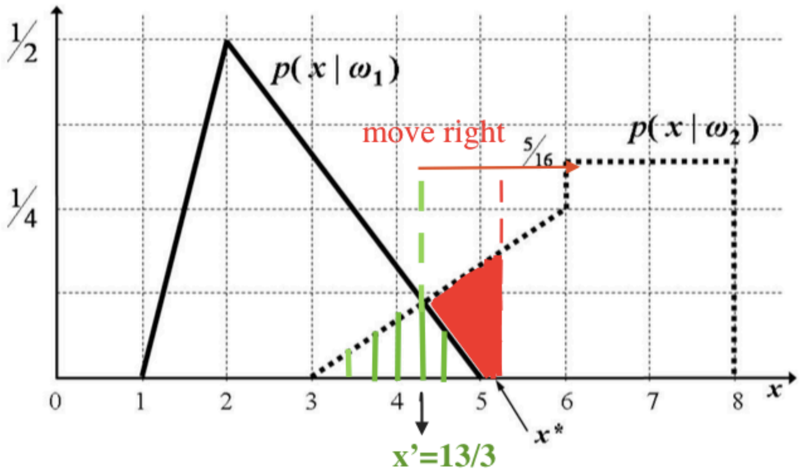


Figure 1: Class-conditional pdfs for a single feature x.

Using Figure1, we see that the intersection is $x' = \frac{13}{3}$, assume $x' = \frac{13}{3}$ is the decision boundary. The region of integration at x' is the green area enclosed by pdf of ω_1, pdf of ω_2, and the decision boundary under rectangular coordinate system.
if $x' = \frac{13}{3}$ moves to the right, the region of integration is the sum of green and red area, it becomes large, so the probability of error become large. Consider the reverse situation, when the $x' = \frac{13}{3}$ moves to the right, the region of integration will become large, too.

Therefore, the decision boundary at $x' = \frac{13}{3}$ can minimizing the area of integration i.e. the probability of error. We sets a threshold $x' = \frac{13}{3}$, and the decision rule as following:

$$\text{If } x \geq x' \text{ then decide } \omega_2\text{, else decide } \omega_1.$$

Part II

Q1.Confusion Matrix

cofusion matrix of PCA

7.	2.	0.	0.	0.	3.	0.	0.	0.	0.
0.	12.	0.	0.	0.	0.	0.	0.	0.	0.
0.	2.	8.	0.	2.	0.	0.	0.	0.	0.
0.	2.	0.	8.	2.	0.	0.	0.	0.	0.
0.	2.	0.	0.	10.	0.	0.	0.	0.	0.
0.	2.	0.	0.	0.	7.	0.	0.	3.	0.
0.	2.	0.	0.	0.	2.	8.	0.	0.	0.
0.	2.	0.	0.	0.	1.	0.	9.	0.	0.
0.	2.	0.	0.	0.	0.	0.	0.	10.	0.
0.	2.	0.	0.	0.	0.	0.	0.	0.	10.

cofusion matrix of PCA

12.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	12.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	12.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	12.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	12.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	12.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	12.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	12.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	12.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	12.

cofusion matrix of fusion(alpha=0.5)

9.	2.	0.	0.	0.	1.	0.	0.	0.	0.
0.	12.	0.	0.	0.	0.	0.	0.	0.	0.
0.	2.	8.	0.	2.	0.	0.	0.	0.	0.
0.	2.	0.	8.	2.	0.	0.	0.	0.	0.
0.	2.	0.	0.	10.	0.	0.	0.	0.	0.
0.	2.	0.	0.	0.	7.	0.	0.	3.	0.
0.	2.	0.	0.	0.	1.	9.	0.	0.	0.
0.	2.	0.	0.	0.	1.	0.	9.	0.	0.
0.	2.	0.	0.	0.	0.	0.	0.	10.	0.
0.	2.	0.	0.	0.	0.	0.	0.	0.	10.

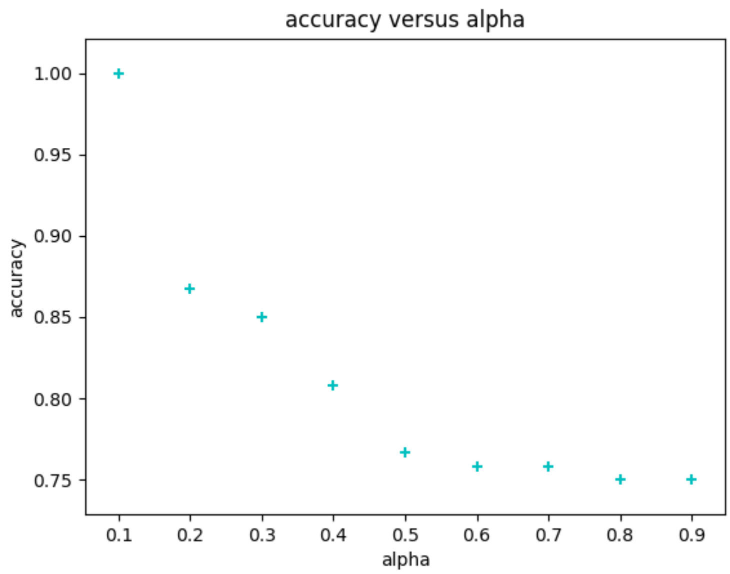
The accuracy of PCA is: 0.742.

The accuracy of LDA is: 1.0.

The accuracy of fusion is: 0.767

Q2. In this experiment, LDA outperforms PCA. Because LDA deals directly with class discrimination and PCA is less sensitive to different training data classes.

Q3.



As the α goes up, the accuracy is reduced.

Q4.The fused feature doesn't always outperform both PCA and LDA feature. In this experience, it's better than PCA. In Q3, we obtain that LDA is better than PCA. The result of fusion is becoming worse because the proportion of LDA is reduced and that of the PCA is increasing, especially when $\alpha > 0.5$, i.e. $\alpha = 0.6, 0.7, 0.8, 0.9$.