PRML Homework 1

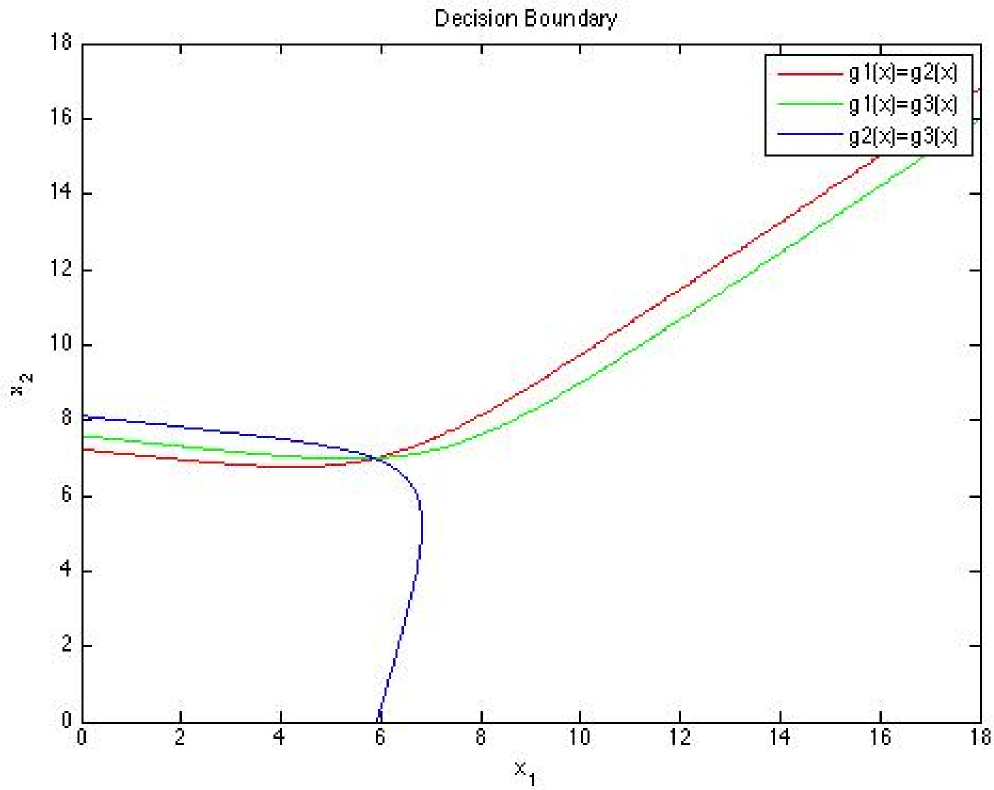
Hao Chen (904547539)

October 16, 2014

1 Problem 1

1. Solution: $g_i(x) = \sum_{y=j} \lambda(\alpha(x) = i|y = j)p(x|y = j)p(y = j)$.
   As three class models $p(x|y = 1), p(x|y = 2), p(x|y = 3)$ are supposed to be 2D Gaussian distributions with the same covariance matrix $\Sigma = 9I$.
   Thus, $\begin{cases} p(x|y=1) & = \frac{1}{18\pi}\exp\{-\frac{(x_1-4)^2+(x_2-12)^2}{18}\} \\ p(x|y=2) & = \frac{1}{18\pi}\exp\{-\frac{(x_1-12)^2+(x_2-3)^2}{18}\} \\ p(x|y=3) & = \frac{1}{18\pi}\exp\{-\frac{(x_1-3)^2+(x_2-5)^2}{18}\} \end{cases}$
   Then, $\begin{cases} g_1(x) & = \frac{1}{15\pi}\exp\{-\frac{(x_1-12)^2+(x_2-3)^2}{18}\} + \frac{1}{45\pi}\exp\{-\frac{(x_1-3)^2+(x_2-5)^2}{18}\} \\ g_2(x) & = \frac{2}{45\pi}\exp\{-\frac{(x_1-4)^2+(x_2-12)^2}{18}\} + \frac{1}{90\pi}\exp\{-\frac{(x_1-3)^2+(x_2-5)^2}{18}\} \\ g_3(x) & = \frac{1}{15\pi}\exp\{-\frac{(x_1-4)^2+(x_2-12)^2}{18}\} + \frac{1}{45\pi}\exp\{-\frac{(x_1-12)^2+(x_2-3)^2}{18}\} \end{cases}$

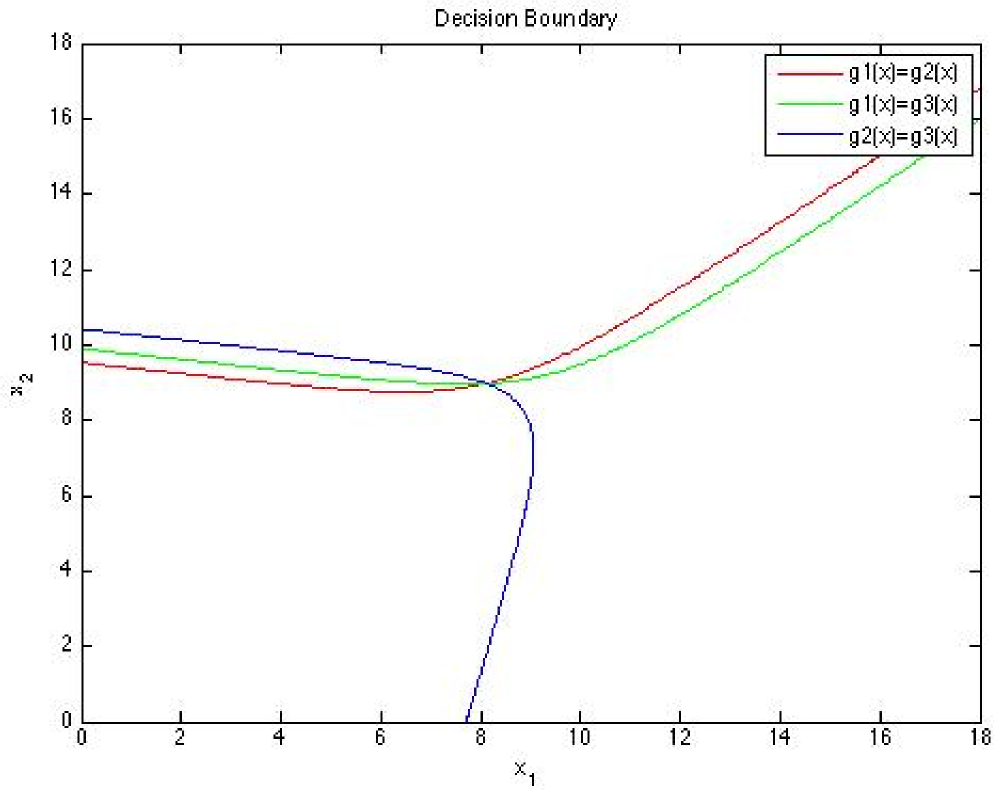
2. Solve the $g_i(x) = g_j(x)$ by MATLAB and draw the boundary as following figure.



3. When the prior distributions change, the discriminant functions change to:

$$\begin{cases} g_1(x) &= \frac{1}{30\pi}\exp\{-\frac{(x_1-12)^2+(x_2-3)^2}{18}\} + \frac{1}{15\pi}\exp\{-\frac{(x_1-3)^2+(x_2-5)^2}{18}\} \\ g_2(x) &= \frac{1}{45\pi}\exp\{-\frac{(x_1-4)^2+(x_2-12)^2}{18}\} + \frac{1}{30\pi}\exp\{-\frac{(x_1-3)^2+(x_2-5)^2}{18}\} \\ g_3(x) &= \frac{1}{30\pi}\exp\{-\frac{(x_1-4)^2+(x_2-12)^2}{18}\} + \frac{1}{90\pi}\exp\{-\frac{(x_1-12)^2+(x_2-3)^2}{18}\} \end{cases}$$

We re-draw the decision boundary:



2 Problem 2

1. Solution

$$\begin{aligned} R_{ran} &= \int_{\Omega^d} R(\alpha|x)p(x)\mathrm{d}x. \\ &= \int_{\Omega^d} \sum_{j\in\Sigma^c} \lambda(\alpha|y=j)p(y=j|x)p(x)\mathrm{d}x. \end{aligned}$$

Due to it is 0-1 loss function, and the decision is made by a randomized decision rule, which decides x to class i following $\alpha(x) = y \sim p(y|x)$, so

$$\begin{aligned} \lambda(\alpha|y=j) &= \sum_{k\in\Omega^c, k\neq j} p(\alpha = k) \\ &= \sum_{k\in\Omega^c, k\neq j} p(y=k|x) \\ &= 1 - p(y=j|x) \end{aligned}$$

which leads to

$$
\begin{aligned}
R_{ran} &= \int_{\Omega^d} \sum_{j\in\Omega^c} (1 - p(y=j|x))p(y=j|x)p(x)\mathrm{d}x. \\
&= \int_{\Omega^d} (1 - \sum_{j\in\Omega^c} p^2(y=j|x))p(x)\mathrm{d}x.
\end{aligned}
$$

2. Proof.

$$
\begin{aligned}
R_{bayes} &= \int_{\Omega^d} R(\alpha|x)p(x)\mathrm{d}x. \\
&= \int_{\Omega^d} (1 - p(y=t|x))p(x)\mathrm{d}x. \\
where\ t &= argmax_{t\in\Omega^c}\ \ p(y|x).
\end{aligned}
$$

Compare integral parts of R_{ran}, R_{hayes}. Now, we consider each possible $x \in \Omega^d$, we want to prove $(1 - p(y = t|x))p(x) \leq (1 - \sum_{j\in\Omega^c} p^2(y = j|x))p(x)$ $(t = argmax_{t\in\Omega^c}\ \ p(y|x))$, which can lead to R_{ran} is always larger than or equal to R_{bayes}.

$$
\begin{aligned}
& (1 - \sum_{j\in\Omega^c} p^2(y=j|x)) - (1 - p(y=t|x)) \\
= \ & p(y=t|x) - \sum_{j\in\Omega^c} p^2(y=j|x) \\
= \ & p(y=t|x)(1 - p(y=t|x)) - \sum_{j\in\Omega^c, j\neq t} p^2(y=j|x) \\
= \ & p(y=t|x) \sum_{j\in\Omega^c, j\neq t} p(y=j|x) - \sum_{j\in\Omega^c, j\neq t} p^2(y=j|x) \\
= \ & \sum_{j\in\Omega^c, j\neq t} p(y=j|x) \sum_{j\in\Omega^c, j\neq t} (p(y=t|x) - p(y=j|x))
\end{aligned}
$$

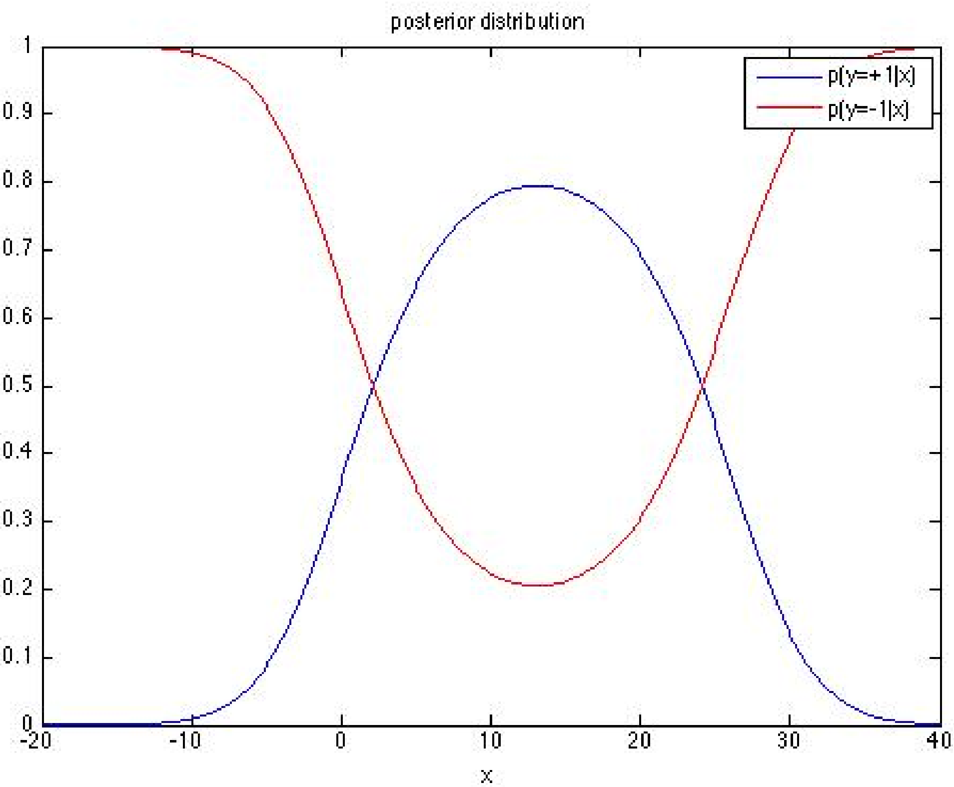
Due to $t = argmax_{t\in\Omega^c} p(y|x)$, $(p(y = t|x) - p(y = j|x)) > 0$, which leads to $\sum_{j\in\Omega^c, j\neq t} p(y = j|x) \sum_{j\in\Omega^c, j\neq t}(p(y = t|x) - p(y = j|x)) \geq 0$. Therefore R_{ran} is always larger than or equal to R_{hayes}.

3. As we can see the equation in (2), when $p(y = t|x) = p(y = j|x), \forall j \in \Omega^c$, the $R_{ran} = R_{hayes}$, which means that all $p(y = j|x)$ are the same value $= \frac{1}{|\Omega^c|}$.

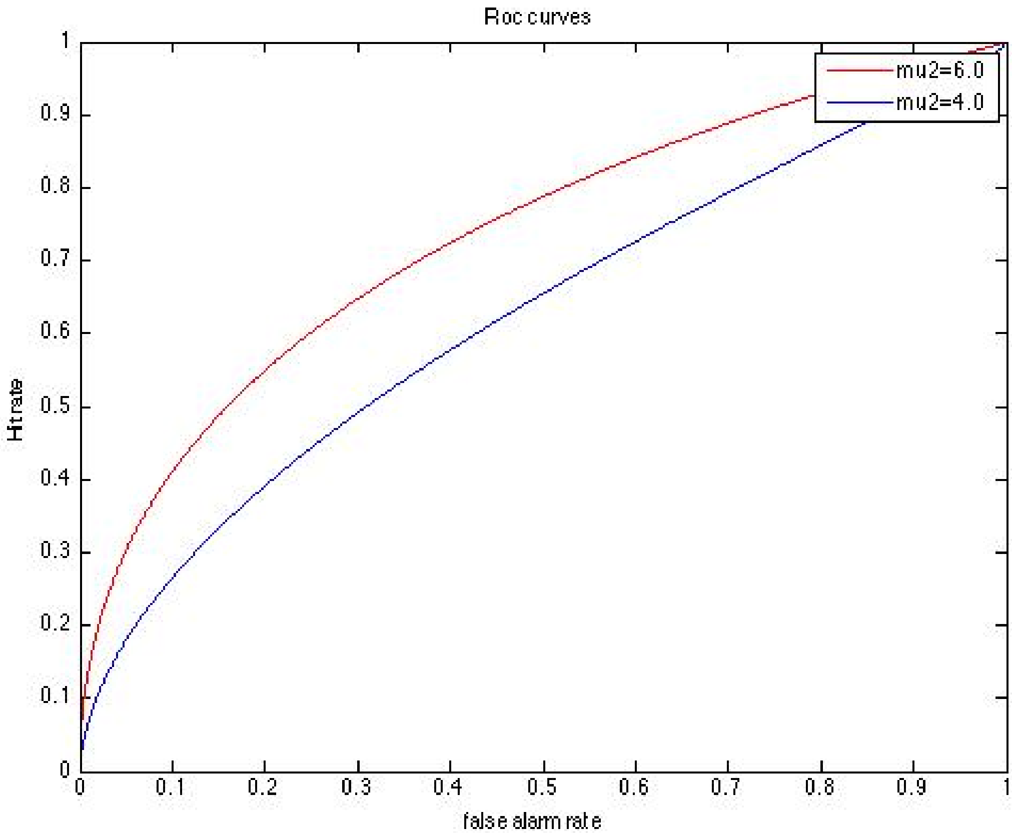
3 Problem 3

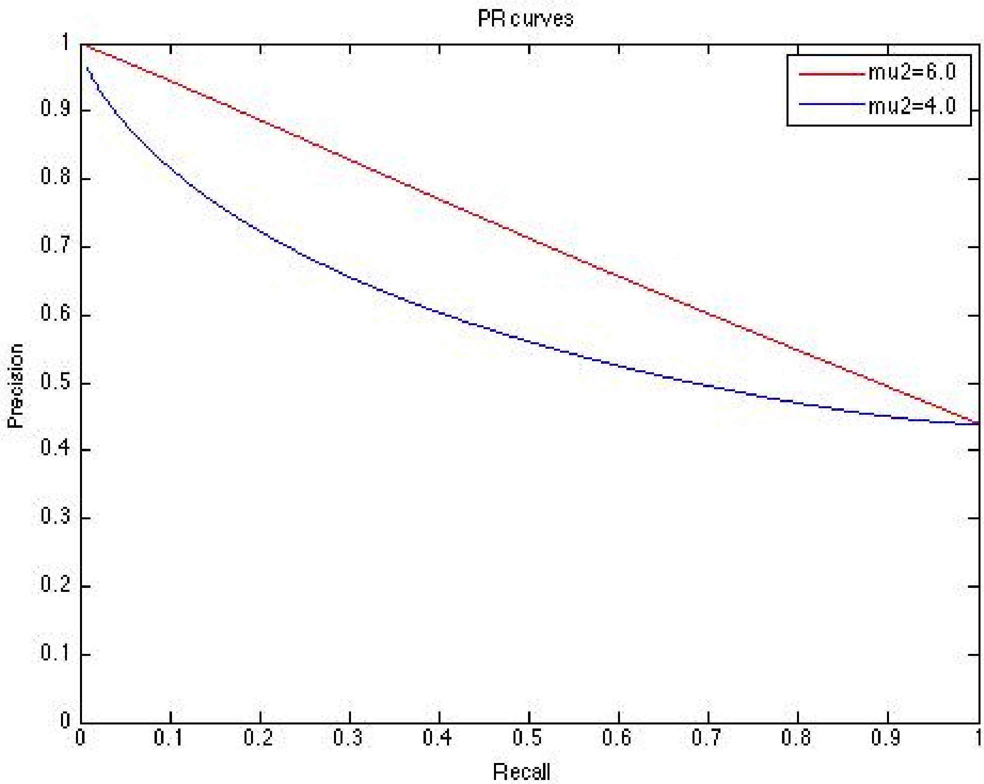
$$
\begin{aligned}
p(y=+1|x) &= \frac{p(x|y=+1)p(y=+1)}{f(x)} \\
&= \frac{p(x|y=+1)p(y=+1)}{p(x|y=+1)p(y=+1)+p(x|y=-1)p(y=-1)} \\
p(y=-1|x) &= \frac{p(x|y=-1)p(y=-1)}{p(x|y=+1)p(y=+1)+p(x|y=-1)p(y=-1)}
\end{aligned}
$$

plot in MATLAB:



The ROC and PR curves are showing as:


Roc curves
mu2=6.0
mu2=4.0
Hit rate
false alarm rate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


PR curves
mu2=6.0
mu2=4.0
Precision
Recall
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1