***Question* 1**

解:

$$R(\alpha_i|x)=\sum_{j=1}^{c}\lambda\left(\alpha_i\middle|w_j\right)P(w_j|x)$$

当$i=j$时，$R(\alpha_i|x)=0$

当$i=1$，$2,\dots,c$且$i\neq j$时，$R(\alpha_i|x)=\lambda_s\sum_{j=1}^{c}P\left(w_j\middle|x\right)=\lambda_s(1-P(w_i|x))$

当$i=c+1$时，$R(\alpha_i|x)=\lambda_r$

$$R(\alpha_i|x)=\begin{cases}\lambda_s\big(1-P(w_i|x)\big) & i=1,\ 2,\dots,c\\ \lambda_r & reject\end{cases}$$

如果判决为w_i，则最小化风险

$$R(\alpha_i|x)=\lambda_s\big(1-P(w_i|x)\big)>\lambda_r$$

$$P(w_i|x)\geq 1-\frac{\lambda_r}{\lambda_s}$$

否则拒绝。

如果$\lambda_r=0$，则做出任何决策的风险都大于拒绝决策，则系统对所有输入都做出拒绝决策。
如果$\lambda_r>\lambda_s$，则做出任何类别的决策风险都小于拒绝决策，则系统对所有输入都作出判决。

***Question* 2**

解:

(a)假设$u_1>u_2$，决策面为 $x=\theta$，，

$$P_e=\int_{R_2}p(x|w_1)P(w_1)dx+\int_{R_1}p(x|w_2)P(w_2)dx$$
$$=\frac{1}{2\sqrt{2\pi}\,\sigma}\left(\int_{-\infty}^{\theta}\exp\left(-\frac{(x-u_1)^2}{2\sigma^2}\right)dx+\int_{\theta}^{\infty}\exp\left(-\frac{(x-u_2)^2}{2\sigma^2}\right)\right)dx$$

则当$\theta=\frac{u_1+u_2}{2}$时，P_e取得最小值，即

$$P_e=\frac{1}{\sqrt{2\pi}\sigma}\int_{\frac{u_1+u_2}{2}}^{\infty}\exp\left(-\frac{(x-u_2)^2}{2\sigma^2}\right))dx$$

令$u=\frac{x-u_2}{\sigma}$ 则$dx=\sigma du$，代入得到

$$P_e=\frac{1}{\sqrt{2\pi}}\int_{\frac{u_1-u_2}{2\sigma}}^{\infty}\exp\left(-\frac{u^2}{2}\right)du$$

假设$u_2>u_1$，

$$P_e=\frac{1}{\sqrt{2\pi}}\int_{\frac{u_2-u_1}{2\sigma}}^{\infty}\exp\left(-\frac{u^2}{2}\right)du$$

综上

$$P_e = \frac{1}{\sqrt{2\pi}}\int_a^{\infty} \mathrm{e}^{-\frac{u^2}{2}} du$$

(b)

由于$\lim_{a\to\infty}\frac{1}{\sqrt{2\pi}a}\mathrm{e}^{-\frac{a^2}{2}} = 0$,且

$$0 < P_e \le \frac{1}{\sqrt{2\pi}a}\mathrm{e}^{-\frac{a^2}{2}}$$

由夹逼定理可得

$$\lim_{a\to\infty}\frac{1}{\sqrt{2\pi}}\int_a^{\infty}\mathrm{e}^{-\frac{u^2}{2}}du = 0$$

故当$\frac{|u_2-u_1|}{\sigma}$趋近于无穷时，P_e趋近于 0。

Question3

解：

(a)条件概率密度:

$$P(x|w_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}}\exp\left[-\frac{1}{2}(\mathrm{x}-u_i)^t\Sigma_i^{-1}(\mathrm{x}-u_i)\right]$$

(b)判别函数:

$$g_i(x) = -\frac{1}{2}(\mathrm{x}-u_i)^t\Sigma_i^{-1}(\mathrm{x}-u_i) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(w_i)$$

在相关矩阵相等时:

$$\begin{aligned} g_i(x) &= w_i^t x + w_{i0} \\ &= \Sigma^{-1}u_i x - \frac{1}{2}{u_i}^t\Sigma_i^{-1}u_i + \ln P(w_i) \end{aligned}$$

在相关矩阵不相等时:

$$\begin{aligned} g_i(x) &= x^t W_i x + w_i^t x + w_{i0} \\ &= -\frac{1}{2}x^t\Sigma_i^{-1}x + \Sigma_i^{-1}u_i x - \frac{1}{2}{u_i}^t\Sigma_i^{-1}u_i - \frac{1}{2}\ln|\Sigma_i| + \ln P(w_i) \end{aligned}$$

(c) 运用 PCA 对数据进行特征提取，筛去本征值为 0 得特征，从而得到可以求逆的协方差矩阵。

Question4

解：

由第三题可知，决策函数为

$$\begin{aligned} g_i(x) &= w_i^t x + w_{i0} \\ &= \Sigma^{-1}{u_i}^t x - \frac{1}{2}{u_i}^t\Sigma_i^{-1}u_i + \ln P(w_i) \end{aligned}$$

则，决策面为

$$g_1(x) = g_2(x)$$
$$w^t(x - x_0) = 0$$

$$x_0 = \frac{1}{2}(u_1 + u_2) - \frac{\ln\left[\frac{P(w_1)}{P(w_2)}\right]}{(u_1 - u_2)^t\Sigma^{-1}(u_1 - u_2)}(u_1 - u_2)$$

假设$u_1 > u_2$，要满足决策面不在两个均值之间，则应满足

$$① \qquad w^t(u_1 - x_0) > 0 \text{ 且} w^t(u_1 - x_0) > 0$$

或者

$$② \qquad w^t(u_1 - x_0) < 0 \text{ 且} w^t(u_1 - x_0) < 0$$

恒成立

$$\begin{aligned} & w^t(u_1 - x_0) \\ & = (u_1 - u_2)^t \Sigma^{-1} \left(\frac{1}{2}(u_1 - u_2) + \frac{\ln\left[\frac{P(w_1)}{P(w_2)}\right]}{(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2)} (u_1 - u_2) \right) \end{aligned}$$

$$= \frac{1}{2}(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) + \ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

$$w^t(u_2 - x_0) = \frac{1}{2}(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) + \ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

在情况①时，需满足

$$\frac{1}{2}(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) + \ln\left[\frac{P(w_1)}{P(w_2)}\right] > 0$$

$$-\frac{1}{2}(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) + \ln\left[\frac{P(w_1)}{P(w_2)}\right] > 0$$

化简得

$$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) > -2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

且

$$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) < 2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

同理，在情况②时，需满足

$$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) < -2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

且

$$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) > 2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$$

易知,

当$P(w_1) > P(w_2)$时，在$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) < 2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$时，满足情况①.

当$P(w_1) < P(w_2)$时，在$(u_1 - u_2)^t \Sigma^{-1} (u_1 - u_2) > 2\ln\left[\frac{P(w_1)}{P(w_2)}\right]$时，满足情况②.

此时，决策面不通过两个均值之间。

$\boldsymbol{Question5}$

解:

由连续、独立可知:

$$P[z_{ik}=1|p(w_i)]=p(w_i)$$
$$P[z_{ik}=0|p(w_i)]=1-p(w_i)$$

两个可以写为通式:

$$P\left(z_{ik}\middle|P(w_i)\right)=[P(w_i)]^{z_{ik}}\cdot[1-P(w_i)]^{(1-z_{ik})}$$

根据其独立性，可以得到

$$P\left(z_{i1},z_{i2},\dots,z_{in}\,\middle|P(w_i)\right)=\prod_{k=1}^{n}[P(w_i)]^{z_{ik}}\cdot[1-P(w_i)]^{(1-z_{ik})}$$

$P(w_i)$的最大似然函数为

$$l(P(w_i))=lnP\left(z_{i1},z_{i2},\dots,z_{in}\,\middle|P(w_i)\right)$$
$$=\ln\,[\prod_{i=1}^{n}[P(w_i)]^{z_{ik}}\cdot[1-P(w_i)]^{(1-z_{ik})}$$
$$=\sum_{k=1}^{n}(z_{ik}\ln\,[\,P(w_i)]+(1-z_{ik})\ln\,[1-P(w_i)])$$

最大似然函数应该满足

$$\nabla_{P(w_i)}l\left(P(w_i)\right)=\frac{1}{P(w_i)}\sum_{k=1}^{n}z_{ik}-\frac{1}{1-P(w_i)}\sum_{k=1}^{n}(1-z_{ik})=0$$

$$\frac{1}{P(w_i)}\sum_{k=1}^{n}z_{ik}=\frac{1}{1-P(w_i)}\sum_{k=1}^{n}(1-z_{ik})$$

$$\hat{P}(w_i)=\frac{1}{n}\sum_{k=1}^{n}z_{ik}$$

Question6

解:

(a)

$$\hat{u}_{n+1}=\frac{1}{n+1}\sum_{k=1}^{n+1}x_k=\frac{n\hat{u}_n+x_{n+1}}{n+1}=\hat{u}_n+\frac{1}{n+1}(x_{n+1}-\hat{u}_n)$$

$$C_{n+1}=\frac{1}{n}\sum_{k=1}^{n+1}(x_k-\hat{u}_{n+1})(x_k-\hat{u}_{n+1})^t$$

$$=\frac{1}{n}(\sum_{k=1}^{n}(x_k-\hat{u}_{n+1})(x_k-\hat{u}_{n+1})^t+(x_{n+1}-\hat{u}_{n+1})(x_{n+1}-\hat{u}_{n+1})^t$$

$$= \frac{1}{n}\left(\sum_{k=1}^{n}\left(x_k - \hat{u}_n - \frac{1}{n+1}(x_{n+1} - \hat{u}_n)\right)\left(x_k - \hat{u}_n - \frac{1}{n+1}(x_{n+1} - \hat{u}_n)\right)^t\right.$$

$$\left. + \left(x_{n+1} - \hat{u}_n - \frac{1}{n+1}(x_{n+1} - \hat{u}_n)\right)\left(x_{n+1} - \hat{u}_n - \frac{1}{n+1}(x_{n+1} - \hat{u}_n)\right)^t\right)$$

$$= \frac{1}{n}\left(\sum_{k=1}^{n}(x_k - \hat{u}_n)(x_k - \hat{u}_n)^t + \sum_{k=1}^{n}\frac{1}{(n+1)^2}(x_{n+1} - \hat{u}_n)(x_{n+1} - \hat{u}_n)^t + \frac{n^2}{(n+1)^2}(x_{n+1}\right.$$

$$\left. - \hat{u}_n)(x_{n+1} - \hat{u}_n)^t)\right)$$

$$= \frac{1}{n}[(n-1)C_n + \frac{n+n^2}{(n+1)^2}(x_{n+1} - \hat{u}_n)(x_{n+1} - \hat{u}_n)^t]$$

$$= \frac{n-1}{n}C_n + \frac{1}{n+1}(x_{n+1} - \hat{u}_n)(x_{n+1} - \hat{u}_n)^t$$

(b)均值计算复杂度为$O(dn)$;

协方差计算复杂度为$O(dn^2)$.