1. $P(x|C_1) = \frac{1}{2(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{8}(x+1)^2\right)$

$P(x|C_2) = \frac{1}{2(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{8}(x-2)^2\right)$

(1) $P(x|C_1)P(C_1) = P(x|C_2)P(C_2)$

$(x+1)^2 = (x-2)^2$

$x = \frac{1}{2}$

$\therefore$ $x < \frac{1}{2}$时，$P(x, C_1) > P(x, C_2)$，判决为C_1

$x > \frac{1}{2}$时，$P(x, C_1) < P(x, C_2)$，判决为C_2

(2) $e = \int_{-\infty}^{\frac{1}{2}} P(x, C_2)\,dx + \int_{\frac{1}{2}}^{\infty} P(x, C_1)\,dx$

$= \frac{1}{4(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\frac{1}{2}} \exp\left(-\frac{1}{8}(x-2)^2\right) dx \cdot 2$

$= \frac{1}{2(2\pi)^{\frac{1}{2}}} \int_{\frac{3}{2}}^{\infty} \exp\left(-\frac{1}{8}x^2\right) dx$

2. $P(x|C_i) = \frac{1}{(2\pi)^{\frac{M}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)\right)$

$g_i(x) = \ln[P(x|C_i)P(C_i)]$

$= -\frac{M}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i) + \ln P(C_i)$

去掉与i无关的项，展开：

$g_i(x) = -\frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}x^T\Sigma_i^{-1}x + \frac{1}{2}x^T\Sigma_i^{-1}\mu_i + \frac{1}{2}\mu_i^T\Sigma_i^{-1}x - \frac{1}{2}\mu_i^T\Sigma_i^{-1}\mu_i + \ln P(C_i)$

$= -\frac{1}{2}x^T\Sigma_i^{-1}x + \Sigma_i^{-1}\mu_i x$

$= x^T W_i x + w_i x + w_{i0}$

其中 $W_i = -\frac{1}{2}\Sigma_i^{-1}$

$w_i = \Sigma_i^{-1}\mu_i$

$w_{i0} = -\frac{1}{2}\mu_i^T\Sigma_i^{-1}\mu_i + \ln P(C_i)$ ， $i = 1, 2, \cdots, K$

由贝叶斯公式

$P(C_i|x) = \frac{P(x|C_i)P(C_i)}{P(x)}$

$= \frac{P(x|C_i)P(C_i)}{\sum_{i=1}^{K} P(x|C_i)P(C_i)}$

$$= \frac{\exp[\ln(x|C_i)P(C_i)]}{\sum_{i=1}^{k}\exp[\ln(x|C_i)P(C_i)]}$$

$$= \frac{\exp(g_i(x))}{\sum_{i=1}^{k}\exp(g_i(x))} = \frac{\exp(a_i)}{\sum_{i=1}^{k}\exp(a_i)}$$

其中 $a_i = g_i(x) = x^T W_i x + w_i x + w_{i0}$

3.4(1) $P(C_i|x) = \frac{1}{(2\pi)^{\frac{1}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu_i)^T\Sigma^{-1}(x-\mu_i)\right)$

由样本集

$P(C_1) = P(C_0) = \frac{1}{2}$

$\mu_1 = \frac{1}{4}\left[\begin{pmatrix}0.5\\1.5\end{pmatrix} + \begin{pmatrix}1.5\\1.5\end{pmatrix} + \begin{pmatrix}0.5\\0.5\end{pmatrix} + \begin{pmatrix}1.5\\0.5\end{pmatrix}\right] = \begin{pmatrix}1\\1\end{pmatrix}$

$\mu_0 = \frac{1}{4}\left[\begin{pmatrix}1.5\\0.5\end{pmatrix} + \begin{pmatrix}2.5\\0.5\end{pmatrix} + \begin{pmatrix}1.5\\-0.5\end{pmatrix} + \begin{pmatrix}2.5\\-0.5\end{pmatrix}\right] = \begin{pmatrix}2\\0\end{pmatrix}$

$$\Sigma = \frac{1}{8}\sum_{i=1}^{N}(x_i-\mu_i)(x_i-\mu_i)^T$$

$$= \frac{1}{8}\left[\begin{pmatrix}-0.5\\0.5\end{pmatrix}(-0.5, 0.5) + \begin{pmatrix}0.5\\0.5\end{pmatrix}(0.5, 0.5) + \cdots + \begin{pmatrix}-0.5\\-0.5\end{pmatrix}(-0.5, -0.5)\right]$$

$$= \begin{pmatrix}\frac{1}{4} & 0\\ 0 & \frac{1}{4}\end{pmatrix}$$

$\therefore \Sigma^{-1} = \begin{pmatrix}4 & 0\\ 0 & 4\end{pmatrix}$, $|\Sigma| = \frac{1}{16}$

$$\therefore P(x|C_1) = \frac{1}{(2\pi)\ |\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)$$

$$= \frac{2}{\pi}\exp\left(-\frac{1}{2}(x_1-1, x_2-1)\begin{pmatrix}4 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}\right)$$

$$P(x|C_0) = \frac{1}{(2\pi)\ |\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)$$

$$= \frac{2}{\pi}\exp\left(-\frac{1}{2}(x_1-2, x_2)\begin{pmatrix}4 & 0\\ 0 & 4\end{pmatrix}\begin{pmatrix}x_1-2\\x_2\end{pmatrix}\right)$$

$$\therefore g(x_1) = \ln[P(x|C_1)P(C_1)]$$

$$= \ln\frac{2}{\pi} - 2(x_1-2, x_2)\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}$$

$$= -2(x_1-1, x_2-1)\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}$$

$$g(x_2) = -2(x_1-2, x_2)\begin{pmatrix}x_1-2\\x_2\end{pmatrix}$$

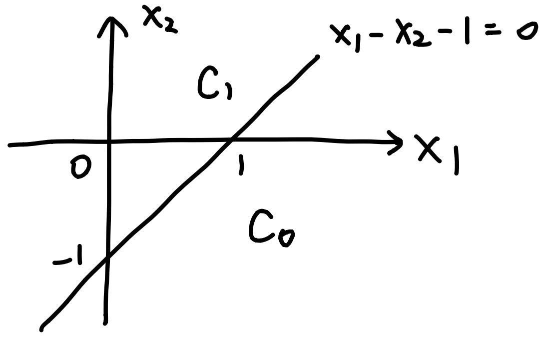
令 $g(x_1) > g(x_2)$

$$-2(x_1-1)^2-2(x_2-1)^2 > -2(x_1-2)^2-2x_2^2$$

$$x_1^2-2x_1+1+x_2^2-2x_2+1 < x_1^2-4x_1+4+x_2^2$$

$$2x_1-2x_2-2<0$$

$$x_1-x_2-1<0$$



(2) $P(x)=\sum P(x|C_i)P(C_i)$

$$=P(x|C_1)P(C_1)+P(x|C_0)P(C_0)$$

$$=-\frac{1}{\pi}\exp\left(-2(x_1-1,x_2-1)\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}\right)-\frac{1}{\pi}\exp\left(-2(x_1-2,x_2)\begin{pmatrix}x_1-2\\x_2\end{pmatrix}\right)$$

$$P(C_1|x)=\frac{P(x|C_1)P(C_1)}{P(x)}$$

$$=\frac{\exp\left(-2(x_1-1,x_2-1)\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}\right)}{\exp\left(-2(x_1-1,x_2-1)\begin{pmatrix}x_1-1\\x_2-1\end{pmatrix}\right)+\exp\left(-2(x_1-2,x_2)\begin{pmatrix}x_1-2\\x_2\end{pmatrix}\right)}$$

(3) $x=\begin{pmatrix}2.0\\0.3\end{pmatrix}$

$$P(C_1|x)=\frac{\exp(-2(1+0.49))}{\exp(-2(1+0.49))+\exp(-2(0+0.49))}$$

$$=\frac{\exp(-2.98)}{\exp(-2.98)+\exp(-0.98)}$$

$$=0.119$$

$$P(C_0|x)=\frac{\exp(-0.98)}{\exp(-2.98)+\exp(-0.98)}$$

$$=0.881$$

$\therefore P(C_0|x)>P(C_1|x)$，可分为 C_0

4.1 $\hat{y}=w_0+w_1x=(w_0,w_1)\begin{pmatrix}1\\x\end{pmatrix}=w^Tx$

$$w=(X^TX)^{-1}X^Ty$$

$$= \left(\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1.5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1.5 \\ 1 & 2 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1.5 & 2 \end{pmatrix} \begin{pmatrix} 0.8 \\ 0.9 \\ 1.2 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 4.5 \\ 4.5 & 7.25 \end{pmatrix}^{-1} \begin{pmatrix} 2.9 \\ 4.55 \end{pmatrix}$$

$$= \begin{pmatrix} 4.83 & -3 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2.9 \\ 4.55 \end{pmatrix}$$

$$= \begin{pmatrix} 0.367 \\ 0.4 \end{pmatrix}$$

$\therefore w_0 = 0.367,\ w_1 = 0.4$

$\hat{y} = 0.367 + 0.4x$

4.3 $$P(w|\alpha, \sigma^2) = \prod_{k=0}^{K} \frac{1}{2\beta\Gamma(\frac{1}{\alpha})} \exp\left(-\frac{|w_k|^\alpha}{\beta^\alpha}\right)$$

$$\beta = \sigma\sqrt{\frac{\Gamma(\frac{1}{\alpha})}{\Gamma(\frac{3}{\alpha})}},\quad \Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt,\ \alpha > 0$$

$$P(x_i|w) = \frac{1}{(2\pi)^{\frac{m}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma_i^{-1}(x-\mu)\right)$$

$$P(w|x_i) = \frac{P(x_i|w)P(w)}{P(x_i)}$$

$$\propto \frac{1}{(2\pi)^{\frac{m}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma_i^{-1}(x-\mu)\right) \prod_{k=0}^{K} \frac{1}{2\beta\Gamma(\frac{1}{\alpha})} \exp\left(-\frac{|w_k|^\alpha}{\beta^\alpha}\right)$$

$$\propto \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma_i^{-1}(x-\mu)\right) \prod_{k=0}^{K} \exp\left(-\frac{|w_k|^\alpha}{\beta^\alpha}\right)$$

$$J = -\ln P(w|x_i)$$

$$= \frac{1}{2}(x-\mu)^T(x-\mu) + \lambda\sum_{k=0}^{K}|w_k|^2$$

5.5 $$\sigma(a) = \frac{1}{1+e^{-a}}$$

$$\sigma(-a) = \frac{1}{1+e^{a}} = \frac{e^{-a}}{e^{-a}+1} = 1 - \frac{1}{1+e^{-a}} = 1 - \sigma(a)$$

$$\frac{d}{da}\sigma(a) = \frac{d}{da}\left(\frac{1}{1+e^{-a}}\right)$$

$$= -\frac{1}{(1+e^{-a})^2} \cdot e^{-a} \cdot (-1)$$

$$= \frac{1}{1+e^{-a}} \frac{e^{-a}}{1+e^{-a}}$$

$$= \sigma(a)(1-\sigma(a))$$

$$\frac{\exp(w^T\varphi_n(x))}{\sum_{i=1}^{N}\exp(w^T\varphi_i(x))}$$

$$\frac{d}{da_n}\hat{y}_n = \sigma(a_n)(1-\sigma$$

5.6 $J(w) = -\sum_{n=1}^{N} y_n \ln \hat{y}_n + (1-y_n)\ln(1-\hat{y}_n)$

$$\hat{y}_n = \sigma(w^T \varphi_n) = \sigma(a_n)$$

$$\frac{\partial}{\partial w}\hat{y}_n = \frac{\partial}{\partial a_n}\hat{y}_n \frac{\partial}{\partial w} a_n$$

$$= \sigma(a_n)(1-\sigma(a_n))\varphi_n$$

$$= \hat{y}_n(1-\hat{y}_n)\varphi_n$$

$$\nabla J(w) = \frac{\partial J(w)}{\partial w}$$

$$= -\sum_{n=1}^{N} y_n \frac{\partial \ln \hat{y}_n}{\partial w} + (1-y_n)\frac{\partial \ln(1-\hat{y}_n)}{\partial w}$$

$$= -\sum_{n=1}^{N} \frac{y_n}{\hat{y}_n}\frac{\partial \hat{y}_n}{\partial w} - \frac{1-y_n}{1-\hat{y}_n}\frac{\partial y_n}{\partial w}$$

$$= -\sum_{n=1}^{N} y_n(1-\hat{y}_n)\varphi_n - \frac{1-y_n}{1-\hat{y}_n}\hat{y}_n(1-\hat{y}_n)\varphi_n$$

$$= \sum_{n=1}^{N}(\hat{y}_n - y_n)\varphi_n$$

$$= \phi^T(\hat{y} - y)$$

其中 $\phi = [\phi(x_1), \phi(x_2), \cdots, \phi(x_N)]^T$

$$\nabla^2 J(w) = \frac{\partial}{\partial w}\sum_{n=1}^{N}(\hat{y}_n - y_n)\varphi_n$$

$$= \sum_{n=1}^{N}\frac{\partial}{\partial w}\hat{y}_n \cdot \varphi_n$$

$$= \sum_{n=1}^{N}\hat{y}_n(1-\hat{y}_n)\varphi_n\varphi_n^T$$

$$= \phi^T R \phi$$