模式识别

(Pattern Recognition)

武汉大学计算机学院

Email: 1896211797@189.cn

第3章 Bayes决策理论(Cont.)

3.1 Bayes决策的引入
3.2 分类器的描述方法
3.3 最小错误率Bayes决策
3.4 最小风险Bayes决策
3.5 Neyman-Pearson决策
3.6 最小最大风险决策
3.7 正态分布决策
3.8 Bayes分类器算法和例题
3.9 序惯分类（*:了解）

3.6 最小最大风险决策 (Minimax Criterion)

最小错误概率Bayes决策 (最大后验概率判决准)是使分类的平均错误概率最小，最小风险Bayes决策是使分类的平均风险最小，Neyman-Pearson决策是假设在一类错误率固定的条件下使另一类的错误率最小。

在实际应用中经常遇到的是各类先验概率不能精确知道或者在分析过程中发生变动。 这就使得判决结果不能达到最佳，实际分类器的平均损失要变大，甚至变得很大。在这种情况下要采用最小最大风险(Minimize the maximum possible overall risk)判决准则，它的基本思想是在最差的(即最大的可能风险)情况下争取最好的结果。

条件风险(Conditional risk)：

$$R(\alpha_i \mid x) = E\left[\lambda\left(\alpha_i, \omega_j\right)\right] = \sum_{j=1}^{M} \lambda\left(\alpha_i, \omega_j\right) P\left(\omega_j \mid x\right), i = 1, 2, \ldots, c. (c \le M)$$

前面的三种判决准则讨论的都是假定先验概率不变的情况，现在讨论在$P(\omega_i)$变化时如何使最大可能风险最小，先验概率$P(\omega_1)$与**期望风险(总风险：overall risk)**R间的变化关系如下：

$$R = \int_{R^d} R\left(\alpha(x) \mid x\right) P(x)\, dx = \sum_{j=1}^{M} \int_{R_j} R(\alpha_j \mid x) p(x)\, dx$$

$$= \int_{R_1} R\left(\alpha_1(x) \mid x\right) P(x)\, dx + \int_{R_2} R\left(\alpha_2(x) \mid x\right) P(x)\, dx$$

$$= \int_{R_1} \{\lambda(\alpha_1, \omega_1) P(\omega_1 \mid \mathbf{x}) p(\mathbf{x}) + \lambda(\alpha_1, \omega_2) P(\omega_2 \mid \mathbf{x}) p(\mathbf{x})\} d\mathbf{x}$$

$$+ \int_{R_2} \{\lambda(\alpha_2, \omega_1) P(\omega_1 \mid \mathbf{x}) p(\mathbf{x}) + \lambda(\alpha_2, \omega_2) P(\omega_2 \mid \mathbf{x}) p(\mathbf{x})\} d\mathbf{x}$$

期望风险R反映对整个特征空间上所有本样$\boldsymbol{x}$采取相应决策所带来的平均风险。

对二类情况有：

$$P(\omega_1)+P(\omega_2)=1, \quad \int_{R_2} P(x|\omega_1)dx = 1-\int_{R_1} P(x|\omega_1)dx$$

$$\therefore R=\lambda_{22}+(\lambda_{12}-\lambda_{22})\int_{R_2} P(x|\omega_2)dx+$$

$$P(\omega_1)\left[(\lambda_{11}-\lambda_{22})+(\lambda_{21}-\lambda_{11})\int_{R_2} P(x|\omega_1)dx-(\lambda_{12}-\lambda_{22})\int_{R_1} P(x|\omega_2)dx\right]$$

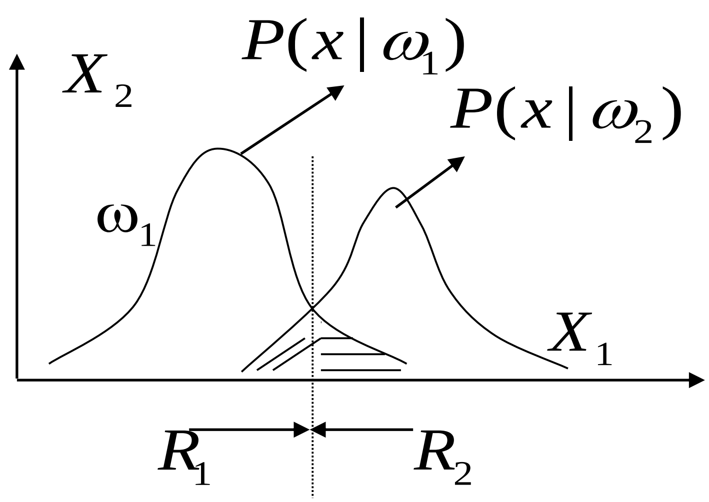
一旦R_1，R_2被确定，风险R就是$P(\omega_1)$的线性函数。

$$\therefore R = a + bP(\omega_1)$$

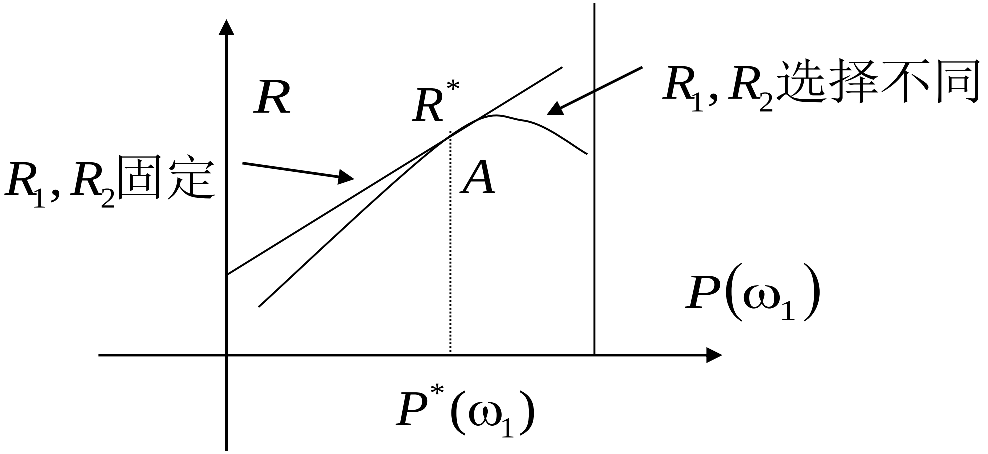
其中：

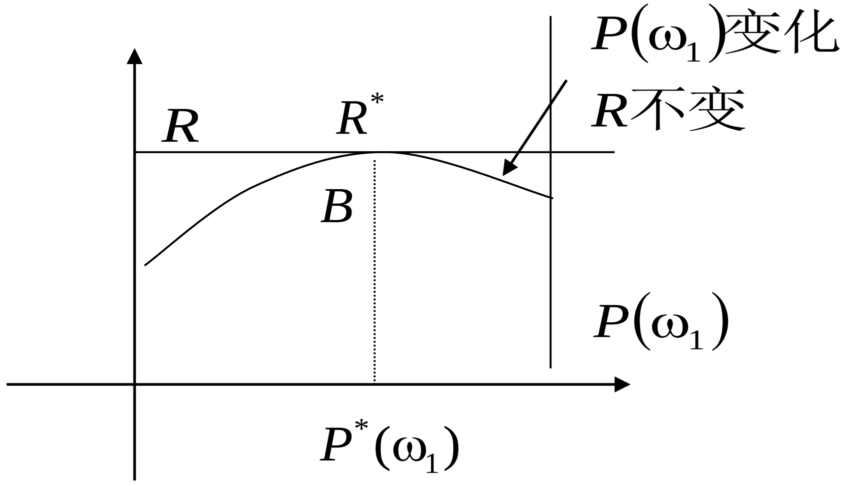
$$a = \lambda_{22} + (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x \mid \omega_2) dx$$

$$b = (\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11}) \int_{R_2} P(x \mid \omega_1) dx - (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x \mid \omega_2) dx$$



这样，就得出最小风险与先验概率的关系曲线，如图所示：





讨论：

(1)当R_1, R_2区间固定时，R与$P(\omega_1)$关系为直线关系, $P(\omega_1)\uparrow, R\uparrow$;

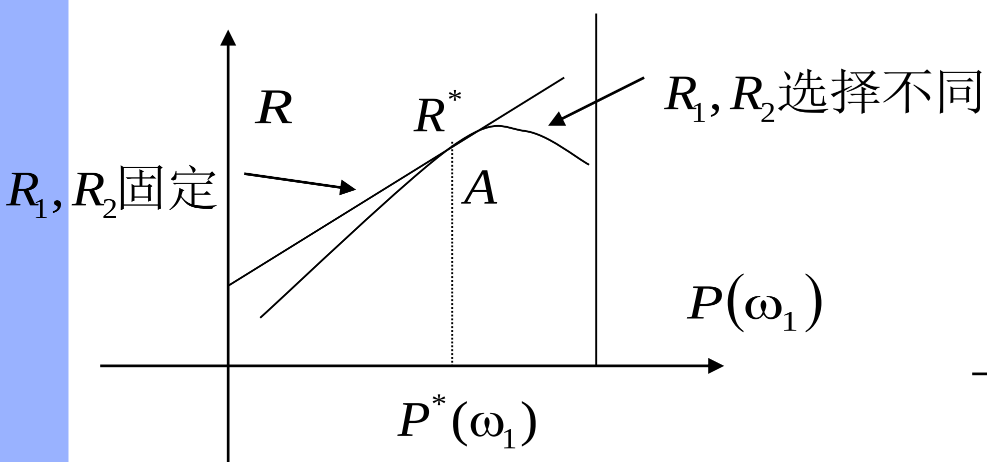
(2)当R_1, R_2选择不同时，R与$P(\omega_1)$关系为一条曲线;

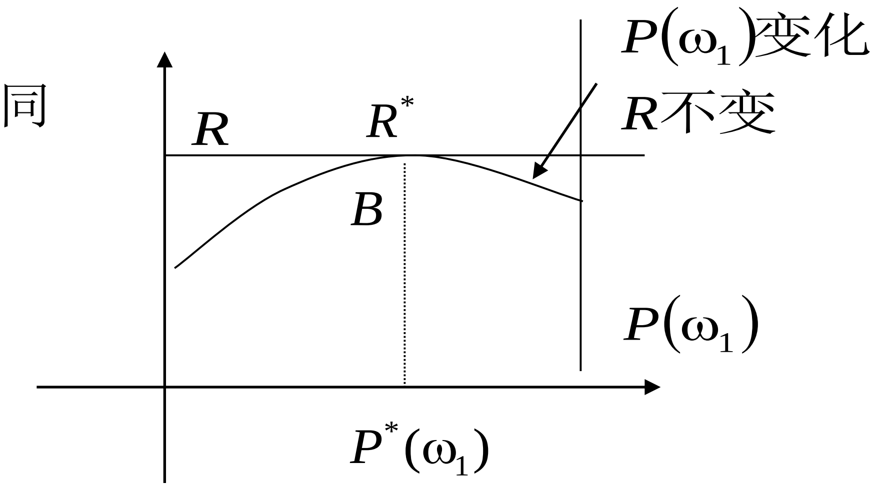
(3)如果选择R_1, R_2使$b=0$, R与$P(\omega_1)$无关.

即$(\lambda_{11}-\lambda_{22})+(\lambda_{21}-\lambda_{11})\int_{R_2} P(x/\omega_1)dx-(\lambda_{12}-\lambda_{22})\int_{R_1} P(x|\omega_2)dx=0$

这时候最大风险为最小，$R=a=\lambda_{22}+(\lambda_{12}-\lambda_{22})\int_{R_2} P(x|\omega_2)dx$

如图所示:这时直线R与横坐标$P(\omega_1)$平行，$P(\omega_1)$变化，则R不变，使最大风险为a。





所以在最大最小判别中，应该使边界R_1，R_2满足$b=0$。

若选取损失为$\begin{cases}\lambda_{11}=\lambda_{22}=0\\ \lambda_{12}=\lambda_{21}\end{cases}$，

则$\int_{R_2}P(x|\omega_1)dx=\int_{R_1}P(x|\omega_2)dx$，$P_1(e)=P_2(e)$，

两类错误概率相等。

- 上式表明，所选的判别边界，使两类的概率相等：
- 这时可使最大可能的风险为最小，这时先验概率变化，其风险不变。

一旦$R1$、 $R2$确定，R就是先验概率$P(\omega 1)$的线性函数：

$$R = a + bP(\omega_1)$$

$$a = \lambda_{22} + (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x \mid \omega_2) dx$$

$$b = (\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11}) \int_{R_2} P(x \mid \omega_1) dx - (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x \mid \omega_2) dx$$

在已知类概率密度函数、代价(损失)函数和某个先验概率$P(\omega 1)$时，按最小风险Bayes判决准则，可以确定$R1$和$R2$：

$$R_1 = \{\mathbf{x} \mid L(\mathbf{x}) > V_T\} \qquad R_2 = \{\mathbf{x} \mid L(\mathbf{x}) < V_T\}$$

其中

$$L(\mathbf{x}) = \frac{p(\mathbf{x} \mid \omega_1)}{p(\mathbf{x} \mid \omega_2)}$$

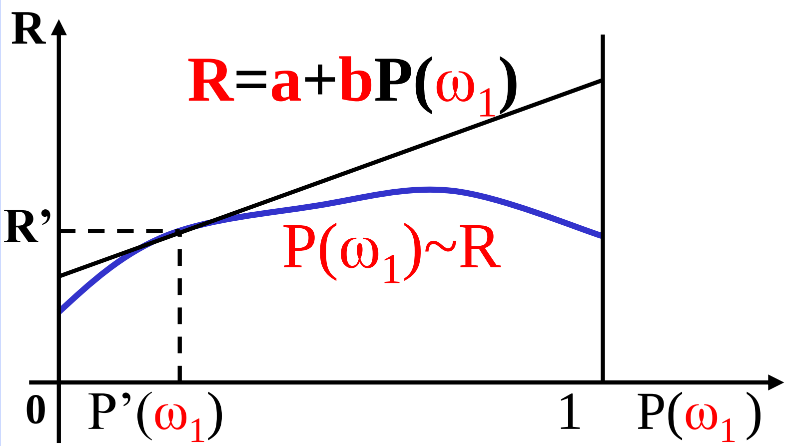
$$V_T = \frac{1 - P(\omega_1)}{P(\omega_1)} \cdot \frac{\lambda(\alpha_1, \omega_2) - \lambda(\alpha_2, \omega_2)}{\lambda(\alpha_2, \omega_1) - \lambda(\alpha_1, \omega_1)}$$

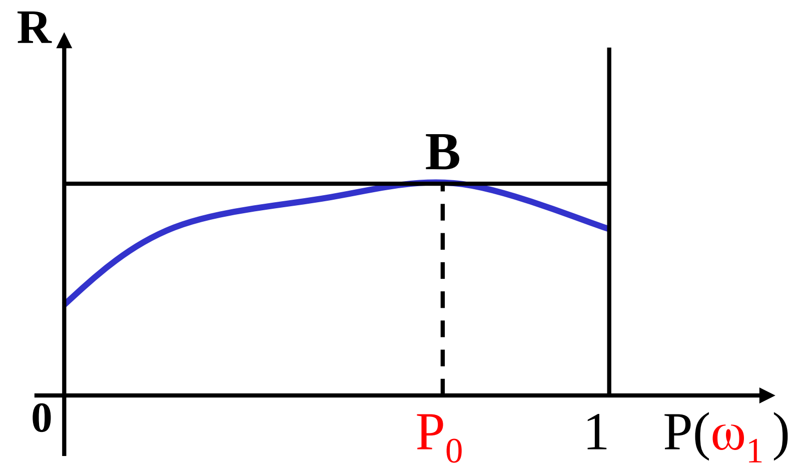
对应不同的先验概率$P(\omega 1)$，可以得到相应的最小Bayes风险，当$P(\omega 1)$遍取［0, 1］时, 就得到$P(\omega 1)\sim R$曲线。

如何选择先验概率使最大可能的平均风险(即总风险：Overall risk)最小呢？由**R**=**a**+**b**P(ω1)式知：

$$\frac{\partial R}{\partial P(\omega_1)} = b$$

若b=0，则**R**与P(ω1)无关，且恒等于a，这时最大可能的平均风险达到最小值。但b=0意味着此时的平均风险直线与最小Bayes风险曲线相切于**B**点， 平均风险**R**达到最小Bayes风险曲线的最大值。如右下图，**B**点对应的先验概率满足使最大可能的平均风险最小的要求。





$$R = a + bP(\omega_1)$$

$$a = \lambda_{22} + (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x|\omega_2)dx$$

$$b = (\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11}) \int_{R_2} P(x|\omega_1)dx - (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x|\omega_2)dx$$

$\therefore$ 最小最大风险决策为

$$(\lambda_{11} - \lambda_{22}) + (\lambda_{21} - \lambda_{11}) \int_{R_2} P(x|\omega_1)dx - (\lambda_{12} - \lambda_{22}) \int_{R_1} P(x|\omega_2)dx = 0 \quad \textbf{(1)}$$

由此来确定 $R1$ 和 $R2$.

因此, 最小最大风险决策的任务就是寻找使Bayes风险为最大时的决策域R2和R1, 它对应于方程(1)解. 在求出使Bayes风险最大时的决策域R2和R1以及相对应的先验概率P_0后,最小最大决策就和最小风险Bayes决策规则相似.

最小最大风险决策在博弈论(Game theory)中的作用比模式识别中的作用更大。在博弈论中，你会有一个对手以对你最不利的方式与你竞争。因此，对于你来说，如何采取一种行为（如做出一个分类）使你所付出的代价（由你对手的对策行为所产生的）最小，具有十分重要的意义。

3.7 正态分布Bayes决策

最小错误率Bayes决策(最大后验概率Bayes决策)和最小风险Bayes决策应用范围很广泛，但事先必须求出 $p(X|\omega i)$和$P(\omega i)$才能做出判决，这一工作一般做起来比较麻烦。当$p(X|\omega i)$呈现正态分布时，将会使决策简化，这时不再需要求出$p(X|\omega i)$的具体形式，只需要知道它的均值向量μ_i和协方差矩阵$\sum_i$这两个参数即可。

本节要用到二次型和多元正态分布的相关数学知识（见补充课件）。

3.7 正态分布Bayes决策

一、正态分布判别函数

1、为什么采用正态分布

a、正态分布在物理上是合理的、广泛的。

b、正态分布数学上简单，N(μ, σ ²) 只有均值和方差两个参数。

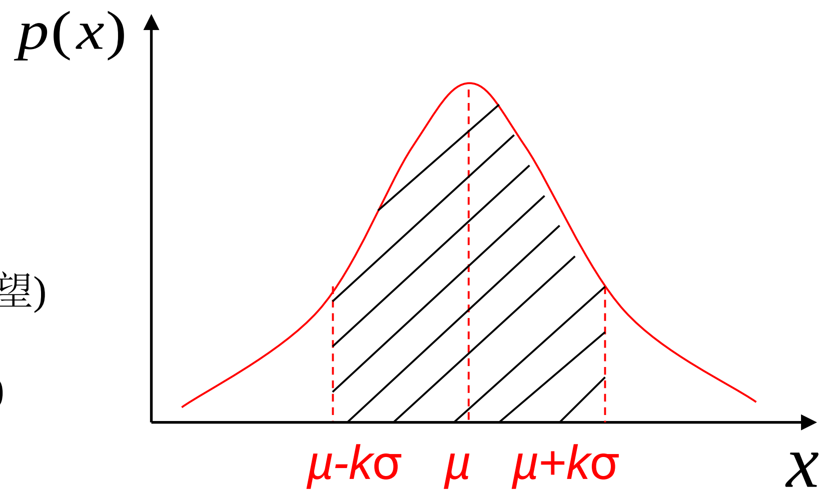
2、单变量正态分布

概率密度函数形式:

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]=N(\mu,\sigma^2)$$

其中: $\mu=E(x)=\int_{-\infty}^{\infty}xp(x)dx$(均值或数学期望)

$$\sigma^2=E\left[(x-\mu)^2\right]=\int_{-\infty}^{\infty}(x-\mu)^2p(x)dx\text{(方差)}$$



概率密度函数应满足下列关系:

$$\begin{cases}p(x)\geq 0,(-\infty<x<\infty)\\ \int_{-\infty}^{\infty}p(x)dx=1\end{cases}$$

3、(多元随机变量)多维正态分布

（1）概率密度函数形式：

$$p(x)=\frac{1}{(2\pi)^{n/2}\left|\sum\right|^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{T}\sum{}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

$$\text{其中}:\mathbf{x}=(x_1,x_2,\ldots,x_n)^{T},n\text{维特征向量}$$

$$\boldsymbol{\mu}=(\mu_1,\mu_2,\ldots,\mu_n)^{T}，\ n\text{维均值向量}$$

$\sum$为$n\times n$维协方差矩阵,$\sum^{-1}$为$\sum$的逆阵,$|\Sigma|$为$\sum$的行列式

$$\mu_i=E(x_i)=\int_{-\infty}^{\infty}x_i\,p(x_i)dx_i$$

$$\sum=E\left[(x-\mu)(x-\mu)^{T}\right]$$

$$=E\left\{\begin{bmatrix}(x_1-\mu_1)\\ \cdots\cdots\\ (x_n-\mu_n)\end{bmatrix}[(x_1-\mu_1),\ldots,(x_n-\mu_n)]\right\}$$

$$=E\left\{\begin{bmatrix}(x_1-\mu_1)(x_1-\mu_1)\ldots(x_1-\mu_1)(x_n-\mu_n)\\ \cdots\cdots\\ (x_n-\mu_n)(x_1-\mu_1)\ldots(x_n-\mu_n)(x_n-\mu_n)\end{bmatrix}\right\}$$

$$=\begin{bmatrix} E\left[\left(x_1-\mu_1\right)\left(x_1-\mu_1\right)\right] \ldots E\left[\left(x_1-\mu_1\right)\left(x_n-\mu_n\right)\right] \\ \ldots\ldots \\ E\left[\left(x_n-\mu_n\right)\left(x_1-\mu_1\right)\right] \ldots E\left[\left(x_n-\mu_n\right)\left(x_n-\mu_n\right)\right] \end{bmatrix}$$

$$=\begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \ldots & \sigma_{1n}^2 \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \ldots & \sigma_{nn}^2 \end{bmatrix}, \begin{pmatrix} \text{对角线}\sigma_{ij}^2, i=j\text{是方差} \\ \text{非对角线}\sigma_{ij}^2,\ i\neq j\text{是协方差} \\ \text{(注:协方差}\sigma_{ij}^2\text{可能为负)} \end{pmatrix}$$

(2) 性质：

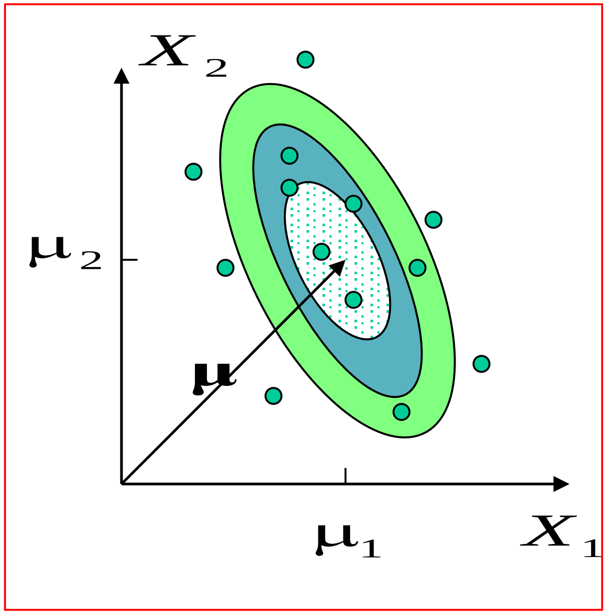
①$\boldsymbol{\mu}$与$\sum$对分布起决定作用$P(\chi)=N(\boldsymbol{\mu}, \sum)$, μ由n个分量组成，$\sum$由n(n+1)/2(因$\sum_i$对称)个元素组成。因此，多维正态分布由n+n(n+1)/2个参数组成。

②等密度点的轨迹是一个超椭球面。区域中心由μ决定，区域形状由$\sum$决定，椭球的主轴与$\sum$的特征值平方成正比。

③不相关性等价于独立性。若x_i与x_j互不相关,则x_i与x_j一定独立。

④线性变换的正态性Y=AX，A为线性变换矩阵。若X为正态分布，则Y也是正态分布。

⑤线性组合的正态性。



➢ 判别函数$g_i(\mathbf{x})$:

类条件概率密度用正态来表示

$$g(\mathbf{x}) = P(\mathbf{x}|\omega_i)P(\omega_i)$$

$$= \frac{1}{(2\pi)^{n/2}\left|\Sigma_i\right|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i)\right] P(\omega_i)$$

得到正态分布的最小错误率Bayes判别函数

$$\boxed{g_i(\mathbf{x})} = \ln g(\mathbf{x})$$

$$= \ln\left\{\frac{1}{(2\pi)^{n/2}\left|\Sigma_i\right|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)\right]\right\} + \ln P(\omega_i)$$

$$\boxed{= -\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i) - \frac{n}{2}\ln 2\pi - \frac{1}{2}\ln\left|\Sigma_i\right| + \ln P(\omega_i)} \quad \textbf{(a)}$$

➢ 决策面方程: $g_i(\mathbf{x}) - g_j(\mathbf{x}) = 0$

$$g_i(\mathbf{x}) - g_j(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{1}{2}\ln\frac{|\Sigma_i|}{|\Sigma_j|}$$

$$+ \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^T \Sigma_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j) + \ln\frac{P(\omega_i)}{P(\omega_j)}$$

$$= 0$$

- 判别规则：

$$若 g_i(\mathbf{x}) > g_j(\mathbf{x}),\quad i=1,2,\ldots,M; i \neq j, 则 \mathbf{x} \in \omega_i$$

- **二、最小错误率(Bayes)分类器：从最小错误率这个角度来分析Bayes 分类器**

1.第一种情况： $\boldsymbol{\Sigma_i = \sigma^2}$

(最简单情况:各个特征统计独立，且同方差情况)

$$即: \sum\nolimits_i = \sigma^2 I = \begin{bmatrix} \sigma_{11}^2 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \sigma_{nn}^2 \end{bmatrix}，只有方差(协方差为零)。$$

➢ 判别函数:

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{n}{2}\ln 2\pi - \frac{1}{2}\ln\left|\Sigma_i\right| + \ln P(\omega_i)$$

$\because \Sigma_i = \sigma^2 I, \Sigma_i^{-1} = \left(1/\sigma^2\right)I; \left|\Sigma_i\right| = \sigma^2 I$和$\frac{n}{2}\ln 2\pi$与$i$无关, 对分类无影响, 可以省略。

$$\therefore g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln P(\omega_i)$$

$$\boxed{\therefore g_i(\mathbf{x}) = -\frac{\left\|\mathbf{x}-\boldsymbol{\mu}_i\right\|^2}{2\sigma^2} + \ln P(\omega_i),\text{其中}\left\|\mathbf{x}-\boldsymbol{\mu}_i\right\|^2 = (\mathbf{x}-\boldsymbol{\mu}_i)^T(\mathbf{x}-\boldsymbol{\mu}_i)} \quad \textbf{(b)}$$

❖ 若M类先验概率相等:

$$P(\omega_1) = P(\omega_2) = \ldots = P(\omega_M)$$

则有:$g(\mathbf{x}) = \left\|\mathbf{x}-\boldsymbol{\mu}_i\right\|^2 \big/ 2\sigma^2$,(欧氏距离)

➢ 最小距离分类器:未知$\mathbf{x}$与$\boldsymbol{\mu}_\mathbf{i}$相减,找最近的$\boldsymbol{\mu}_\mathbf{i}$把$\mathbf{x}$归类。

$(\mathbf{x}-\boldsymbol{\mu}_i)^T(\mathbf{x}-\boldsymbol{\mu}_i)=\mathbf{x}^T\mathbf{x}-2\boldsymbol{\mu}_i^T\mathbf{x}+\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i$,因为(b)式中二次项$\mathbf{x}^T\mathbf{x}$与$i$无关

$\therefore$ 简化可得：$g_i(\mathbf{x})=\mathbf{w}_i^T\mathbf{x}+w_{i0}$,(线性判别函数)

其中：$\mathbf{w}_i=\dfrac{1}{\sigma^2}\boldsymbol{\mu}_i, w_{i0}=-\dfrac{1}{2\sigma^2}\boldsymbol{\mu}_i^T\boldsymbol{\mu}_i+\ln P(\omega_i)$

w_{i0}**称为第i个方向的阈值(threshold)或偏置(bias)**

- 判别规则：$g_i(\mathbf{x})=\mathbf{w}_i^T\mathbf{x}+w_{i0}=\max\limits_{1\le j\le M}\{\mathbf{w}_j^T\mathbf{x}+w_{j0}\}\Rightarrow\mathbf{x}\in\omega_i$

对于二类情况$g(\mathbf{x})=g_2(\mathbf{x})-g_1(\mathbf{x})$

$$=\frac{1}{\sigma^2}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^T\mathbf{x}+\frac{1}{2\sigma^2}(\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1-\boldsymbol{\mu}_2^T\boldsymbol{\mu}_2)\mathop{\gtrless}^{<}_{>}\ln\frac{P(\omega_1)}{P(\omega_2)}\Rightarrow\mathbf{x}\in\begin{matrix}\omega_1\\ \omega_2\end{matrix}$$

- 决策面方程：$g_i(\mathbf{x})-g_j(\mathbf{x})=0$

在$\Sigma_i=\sigma^2 I$的特殊情况下，该方程可表示为:

$$\boxed{\mathbf{w}^T(\mathbf{x}-\mathbf{x}_0)=0}$$

其中$\mathbf{w}=\boldsymbol{\mu}_i-\boldsymbol{\mu}_j$

且$\mathbf{x}_0=\dfrac{1}{2}(\boldsymbol{\mu}_i+\boldsymbol{\mu}_j)-\dfrac{\sigma^2(\boldsymbol{\mu}_i-\boldsymbol{\mu}_j)}{\|\boldsymbol{\mu}_i-\boldsymbol{\mu}_j\|}\ln\dfrac{P(\omega_i)}{P(\omega_j)}$

讨论：

二类情况下$\omega_i = \omega_1\omega_2$

(a)因为$\sum_i = \sigma^2 I$，协方差为零。所以等概率面是一个圆形。

$(b) \because W$与$(x - x_0)$内积为0，因此分界面H与W垂直

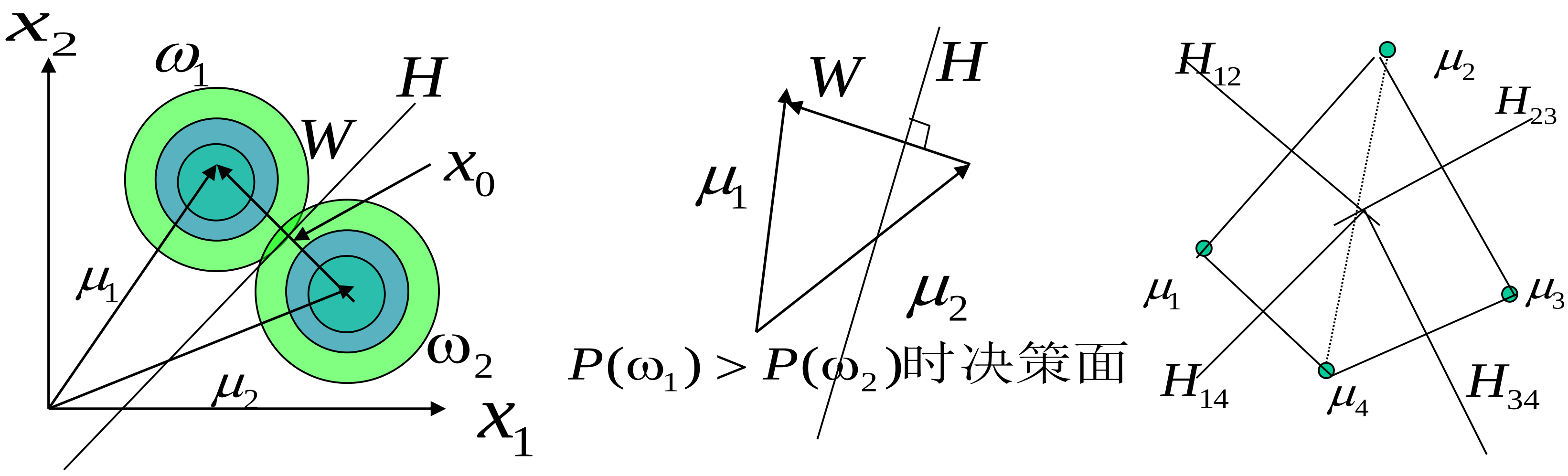
又$\because W = \mu_i - \mu_j = \mu_1 - \mu_2$，所以$W$与$\mu_1 - \mu_2$同相(同方向)

$\therefore$决策面H垂直于μ的联线。

(c)如果先验概率相等$P(\omega_1) = P(\omega_2)$，$H$通过$\mu$联线的中点。

否则就是$P(\omega_1) \neq P(\omega_2)$，$H$离开先验概率大的一类。

(d)对多类情况，用各类的均值联线的垂直线作为分界面。



➢ **2、第二种情况：$\Sigma_i = \Sigma$相等，**即各类协方差相等。

因为(a)式中$\sum_1 = \sum_2 = \ldots = \sum_M = \sum$与$i$无关

$$\therefore \boxed{g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \sum{}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) + \ln P(\omega_i)}$$

若先验概率相等$P(\omega_1) = P(\omega_2) = P(\omega_3) = \ldots = P(\omega_i)$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \sum{}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) = r^2(\text{马氏距离})$$

➢ 最小距离分类器：未知$\mathbf{x}$，把$\mathbf{x}$与各类均值相减，把$\mathbf{x}$归于最近的一类。

把$(\mathbf{x}-\boldsymbol{\mu}_i)^T \sum{}^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)$展开；$\mathbf{x}^T \sum{}^{-1}\mathbf{x}$与$i$无关。

$$\boxed{\therefore g_i(\mathbf{x}) = (\boldsymbol{\mu}_i)^T \sum{}^{-1}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_i^T \sum{}^{-1}\boldsymbol{\mu}_i + \ln P(\omega_i)}$$

$$= \mathbf{w}_i^T \mathbf{x} + w_{i0}\text{（线性函数）}$$

其中$\mathbf{w}_i = \sum{}^{-1}\mu_i$

且$w_{i0} = -\frac{1}{2}\boldsymbol{\mu}_i^T \sum{}^{-1}\boldsymbol{\mu}_i + \ln P(\omega_i)$

- 决策规则：$g_i(\mathbf{x}) = \mathbf{w}_i^T x + w_{i0} = \max\limits_{1\le j\le M}\left\{\mathbf{w}_j^T\mathbf{x} + w_{j0}\right\} \Rightarrow \mathbf{x}\in\omega_i$

$$\boxed{\begin{aligned}&\text{对于二类情况}g(\mathbf{x}) = g_2(\mathbf{x}) - g_1(\mathbf{x}) = (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \sum{}^{-1}\mathbf{x}\\&+\frac{1}{2}\sum{}^{-1}(\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^T\boldsymbol{\mu}_2) \mathop{\gtrless}\limits ^{<}_{>} \ln\frac{P(\omega_1)}{P(\omega_2)} \Rightarrow \mathbf{x}\in \begin{matrix}\omega_1\\ \omega_2\end{matrix}\end{aligned}}$$

- 决策界面：若ω_i与ω_j相邻, $g_i(\mathbf{x}) - g_j(\mathbf{x}) = 0$

$$\therefore \mathbf{w}^T(\mathbf{x}-\mathbf{x}_0) = 0$$

$$\text{其中}\mathbf{w} = \sum{}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$$

$$\text{且}\mathbf{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) - \frac{\ln\dfrac{P(\omega_i)}{P(\omega_j)}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}{(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^T\sum{}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}$$

➢ 讨论：针对ω_1, ω_2二类情况，如图：

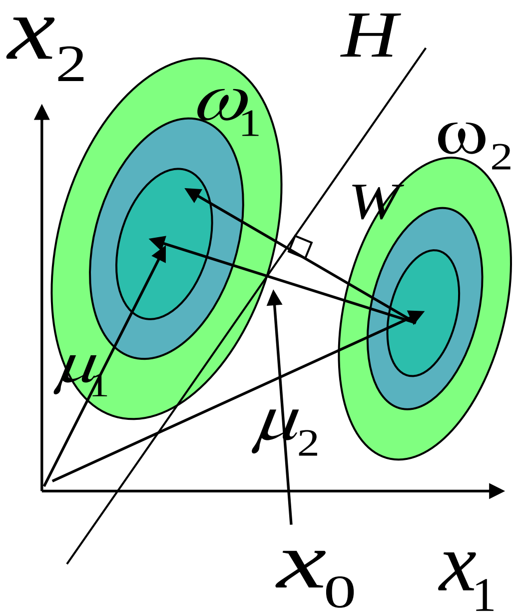
(a)因为$\sum_i = \sum \neq \sigma I$，所以等概率面是椭圆，长轴由$\sum_i$特征值决定

(b)因为W与$(x-x_0)$内积为0，所以W与$(x-x_0)$正交，$\therefore H$通过x_0点。

(c)因为$W = \sum^{-1}(\mu_i - \mu_j)$；所以$W$与$(\mu_i - \mu_j)$不同相；$\therefore H$不垂直于$\mu$值联线

(d)若各类先验概率相等，则$x_0 = \frac{1}{2}(\mu_i - \mu_j)$，则$H$通过均值联线中点；

否则H离开先验概率大的一类。



➢ 3、第三种情况(一般情况)：Σ_i任意，各类协方差矩阵不等，（a)式中二次项$\mathbf{x}^T\Sigma_i\mathbf{x}$与$i$有关，因此判别函数为二次型。

- 判别函数：$g_i(\mathbf{x})=\mathbf{x}^T\mathbf{W}_i\mathbf{x}+\mathbf{w}_i^T\mathbf{x}+w_{i0}$，其中$\mathbf{W}_i=-\frac{1}{2}\sum_i^{-1}(n\times n\text{矩阵})$，

$$\mathbf{w}_i=\sum_i^{-1}\boldsymbol{\mu}_i(n\text{维列向量}),\quad w_{i0}=-\frac{1}{2}\boldsymbol{\mu}_i^T\sum_i^{-1}\boldsymbol{\mu}_i-\frac{1}{2}\ln\left|\sum_i\right|+\ln P(\omega_i)$$

- 决策规则：$g_i(\mathbf{x})=\mathbf{x}^T\mathbf{W}_i\mathbf{x}+\mathbf{w}_i^T\mathbf{x}+w_{i0}$

$$=\max_{1\le j\le M}\left\{\mathbf{x}^T\mathbf{W}_j\mathbf{x}+\mathbf{w}_j^T\mathbf{x}+w_{j0}\right\}\Rightarrow\mathbf{x}\in\omega_i$$

对于二类情况：

$$g(\mathbf{x})=g_2(\mathbf{x})-g_1(\mathbf{x})$$

$$=\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T\sum_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^T\sum_2^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)$$

$$+\frac{1}{2}\ln\frac{\left|\sum_1\right|}{\left|\sum_2\right|}\begin{matrix}<\\>\end{matrix}\ln\frac{P(\omega_1)}{P(\omega_2)}\Rightarrow\mathbf{x}\in\begin{matrix}\omega_1\\\omega_2\end{matrix}$$

协方差矩阵不等，决策面是二次曲面(可能是超球面、超椭球面、超双曲面、超抛物面)。当**x**是二维模式时，判决面为二次曲线。

协方差矩阵相等，决策面是超平面。当**x**是二维模式时，决策面为直线。

决策面方程：$g_i(x) - g_j(x) = 0$

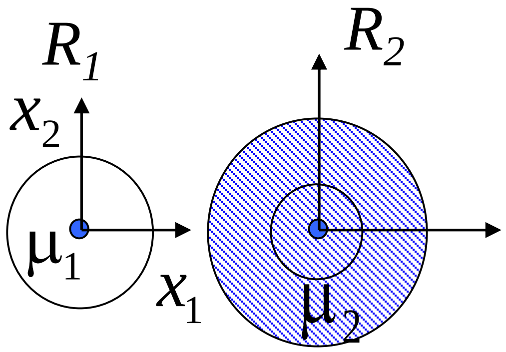
下面看看协方差矩阵不等时，二维特征空间二类问题决策面的各种图形。对于二类问题，有如下条件：

a、二类情况$\omega_1\omega_2$;

b、x_1x_2为条件独立；

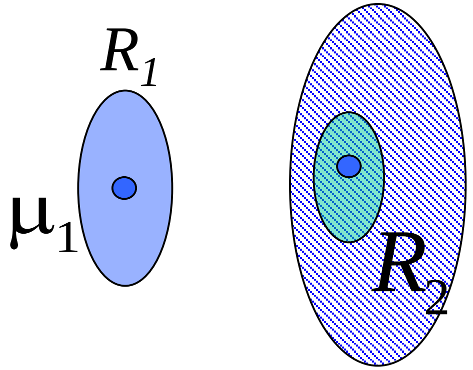
c、先验概率相等。

下面各图示出了二维特征空间两类问题的决策面的各种形式，图中的圆、椭圆表示等概密点轨迹。



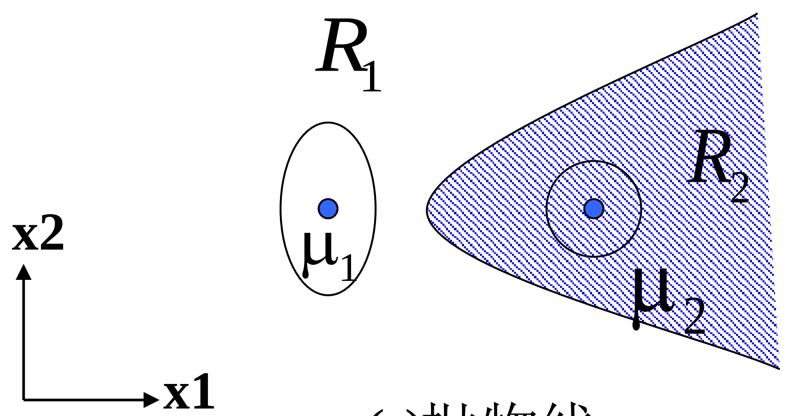
(*a*)圆

(a)图：由于**ω2**类的方差小，**ω2**类的模式更集中于**μ2**，决策面是包围**μ2**的一个圆



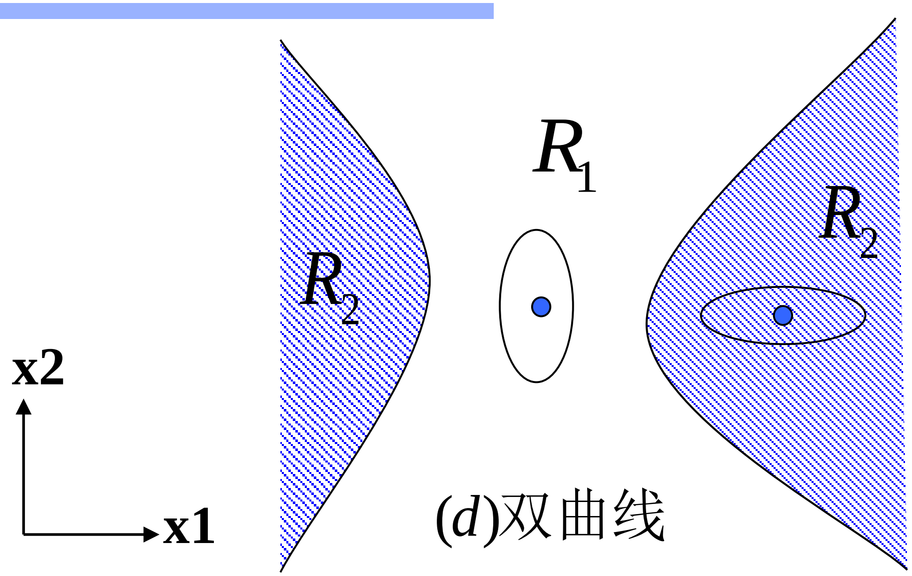
(*b*)椭圆

（b)图：决策面是包围**μ2**一个椭圆

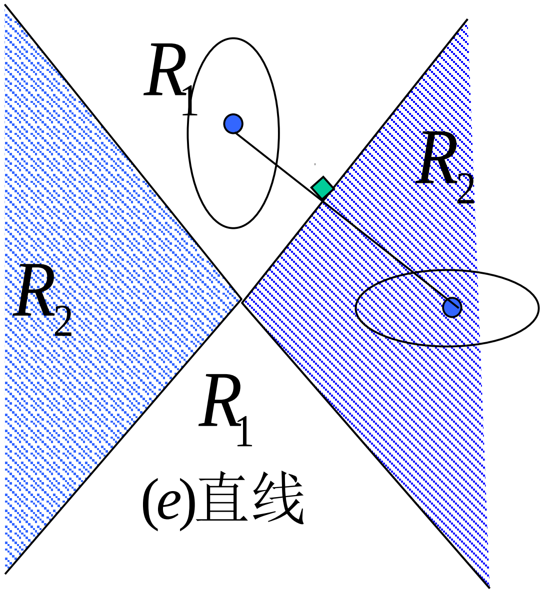


(*c*)抛物线

(c)图： 类ω1和ω2的x1的方差相同，但ω1的x2方差较ω2的x2方差大，从而x2值较大的模式更可能来自ω1类，因此决策面向右弯，呈抛物线状



(d)图：由于**ω2**类的**x1**的方差大于**ω1**类的**x1**的方差，而两类**x2**的方差情况正相反，因此决策面呈双曲线状



(e)直线

(e)图：由于两类的分布关于一直线对称，因此双曲线退化为相交直线。

3.8 正态分布Bayes分类器算法和例题

Bayes分类算法(假定各类样本服从正态分布)

1.输入类数M；特征数n，待分样本数m。

2.输入训练样本数N和训练集资料矩阵X(N×n)。并计算有关参数。

3.计算矩阵X中各类的后验概率。

4.若按最小错误率原则分类，则可根据 3 的结果判定X中各类样本的类别。

5.若按最小风险原则分类，则输入各值，并计算X中各样本属于各类时的风险并判定各样本类别。

❖ **例1**、有训练集资料矩阵如下表所示，现已知，N=9、N_1=5、N_2=4、n=2、M=2，试问，$X=(0,0)^T$应属于哪一类？

训练样本号k	1	2	3	4	5	1	2	3	4
特征 x_1	1	1	0	-1	-1	0	1	0	-1
特征 x_2	0	1	1	1	0	-1	-2	-2	-2
类别	ω_1					ω_2			

❖ **解1**、假定二类协方差 矩阵不等（$\sum_1 \neq \sum_2$），则均值:

$$\overline{X}_{11}=\frac{1}{5}(1+1+0-1-1)=0, \overline{X}_{12}=\frac{3}{5}; \overline{X}_{21}=\frac{1}{4}(0+1+0-1)=0, \overline{X}_{22}=\frac{1}{4}(-1-2-2-2)=-\frac{7}{4}$$

$$\overline{X_1}=\left(\overline{X_{11}},\overline{X_{12}}\right)^T=(0,\frac{3}{5})^T, \overline{X_2}=\left(\overline{X_{21}},\overline{X_{22}}\right)^T=(0,-\frac{7}{4})^T$$.协方差矩阵为:

$$\sum_1=\begin{pmatrix}1 & 0\\ 0 & \frac{3}{10}\end{pmatrix}, \sum_2=\begin{pmatrix}\frac{2}{3} & 0\\ 0 & \frac{1}{4}\end{pmatrix}, \sum_1=\begin{pmatrix}C_{11} & C_{12}\\ C_{21} & C_{22}\end{pmatrix}$$(见前面协方差$\sum_1$的计算方法）

$$C_{11}=\frac{1}{5-1}\sum_{k=1}^{5}(x_{1k}-\bar{x}_{11})^{T}(x_{1k}-\bar{x}_{11})$$

$$=\frac{1}{4}\left[(1-0)^{2}+(1-0)^{2}+(0-0)^{2}+(-1-0)^{2}+(-1-0)^{2}\right]=1$$

$$C_{12}=\frac{1}{4}\sum_{k=1}^{5}(x_{1k}-\bar{x}_{11})(x_{2k}-\bar{x}_{12})^{T}=0$$

$$C_{12}=C_{21}$$

$$C_{22}=\frac{1}{4}\sum_{k=1}^{5}(x_{2k}-\bar{x}_{12})(x_{2k}-\bar{x}_{12})^{T}=\frac{3}{10}$$

协方差矩阵为$\sum_{1}=\begin{pmatrix}1 & 0\\ 0 & \frac{3}{10}\end{pmatrix},\sum_{2}=\begin{pmatrix}\frac{2}{3} & 0\\ 0 & \frac{1}{4}\end{pmatrix}$(计算方法同上）

利用公式：$g(x)=g_2(x)-g_1(x)$

$$=\frac{1}{2}(\mathrm{x}-\overline{\mathrm{x}_1})^T\sum_1^{-1}(\mathrm{x}-\overline{\mathrm{x}_1})-\frac{1}{2}(\mathrm{x}-\overline{\mathrm{x}_2})^T\sum_2^{-1}(\mathrm{x}-\overline{\mathrm{x}_2})$$

$$+\frac{1}{2}\ln\frac{\left|\sum_1\right|}{\left|\sum_2\right|}-\ln\frac{P(\omega_1)}{P(\omega_2)}\begin{matrix}<\\>\end{matrix}0\Rightarrow x\in\begin{matrix}\omega_1\\\omega_2\end{matrix}$$

$\mathrm{x}=(x_1,x_2)^T$，将$\mathrm{x}=(x_1,x_2)^T=(0,0)^T$代入得：$g(x)=-10.91<0$

所以判$X=(0,0)^T$属于ω_1类。

令$g(x)=0$得分界线方程为：$\frac{1}{2}x_1^2+\frac{2}{3}x_2^2+18x_2+10.91=0$

这是一个非线性椭圆方程：$\frac{x_1^2}{14.81^2}+\frac{(x_2+13.5)^2}{12.88^2}=1$

❖ 解2、假定两类协方差矩阵相等$\sum=\sum_1+\sum_2$

$$\sum=\sum{}_1+\sum{}_2=\begin{pmatrix}\frac{5}{3} & 0\\ 0 & \frac{11}{20}\end{pmatrix},\sum{}^{-1}=\begin{pmatrix}\frac{3}{5} & 0\\ 0 & \frac{20}{11}\end{pmatrix},$$

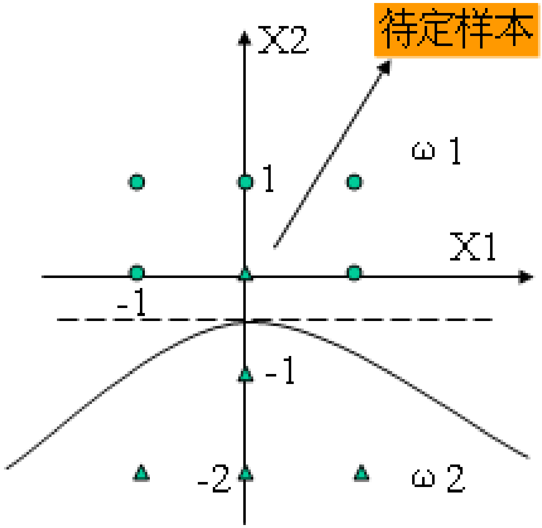
所以代入$x=(0,0)^T$得：

$$g(x)=(\bar{x}_2-\bar{x}_1)^T\sum{}_1^{-1}x+\frac{1}{2}(\bar{x}_1^T\sum{}^{-1}\bar{x}_1-\bar{x}_2^T\sum{}^{-1}\bar{x}_2)-\ln\frac{P(\omega_1)}{P(\omega_2)}=-2.68<0$$

故应把$x=(0,0)^T$判为ω_1类，

分界线方程为$g(x)=\dfrac{-47}{11}x_2-2.68=0$

从而得$x_2=-0.61$为一直线，如图中虚线所示。



两种解得分界线

例2： 有训练集资料矩阵如下表所示，现已知， $N=9$、$N_1=N_2=3$、$n=2$、$M=3$，试问，未知样本 $X=(0,0)^T$应属于哪一类？

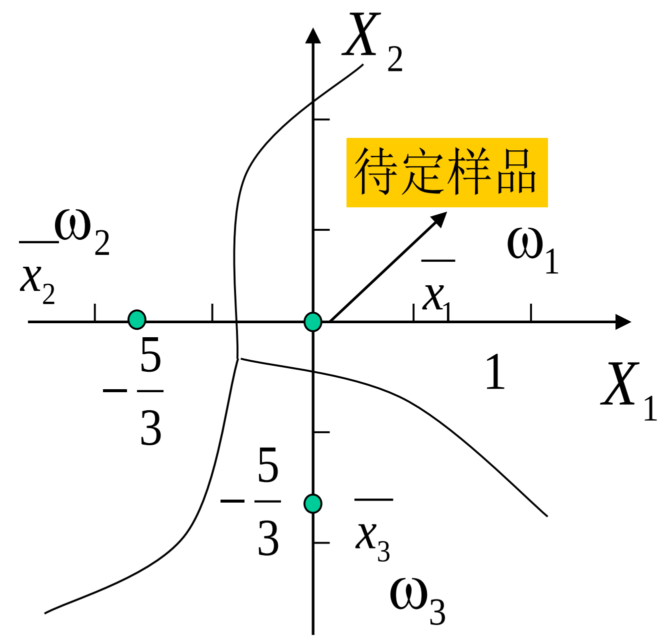
训练样本号k	1 2 3	1 2 3	1 2 3
特征 x_1	0 1 2	-2 -1 -2	0 1 -1
特征 x_2	1 0 -1	1 0 -1	-1 -2 -2
类别	ω_1	ω_2	ω_3

解1、 假定三类协方差不等；

$$\text{均值}\,\overline{x}_1=(1,0)^T,\overline{x}_2=(-\frac{5}{3},0)^T,\overline{x}_3=(0,-\frac{5}{3})^T$$

$$\text{协方差矩阵为:}\sum_1=\begin{pmatrix}1&0\\0&1\end{pmatrix},\sum_2=\begin{pmatrix}\frac{1}{3}&0\\0&1\end{pmatrix},\sum_3=\begin{pmatrix}1&0\\0&\frac{1}{3}\end{pmatrix}$$

$$\text{所以}\sum_1^{-1}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\sum_2^{-1}=\begin{pmatrix}3&0\\0&1\end{pmatrix},\sum_3^{-1}=\begin{pmatrix}1&0\\0&3\end{pmatrix}$$



$$\left|\sum_1\right| = 1, \left|\sum_2\right| = \frac{1}{3}, \left|\sum_3\right| = \frac{1}{3}$$

先验概率$P(\omega_1) = P(\omega_2) = P(\omega_3) = \frac{1}{3}$

代入多类判别函数$g_i(x) = x^T W_i x + w_i^T x + w_{io}$,

其中$W_i = -\frac{1}{2}\sum_i^{-1}, w_i = \sum_i^{-1}\mu_i, w_{io} = -\frac{1}{2}\mu_i^T \sum_i^{-1}\mu_i - \frac{1}{2}\ln\left|\sum_i\right| + \ln P(\omega_i)$.

所以$g_1(x) = -\frac{1}{2}\left(x_1^2 + x_2^2 - 2x_1 + 1\right)$

$$g_2(x) = -\frac{1}{2}\left(3x_1^2 + x_2^2 + 10x_1 + 7.2\right)$$

$$g_3(x) = -\frac{1}{2}\left(x_1^2 + 3x_2^2 + 10x_2 + 7.2\right)$$

将$X = (0,0)^T$代入得:

$$g_1(x) = -0.5, g_2(x) = g_3(x) = -3.6$$

故应判样品$X=(0,0)^T$为ω_1类

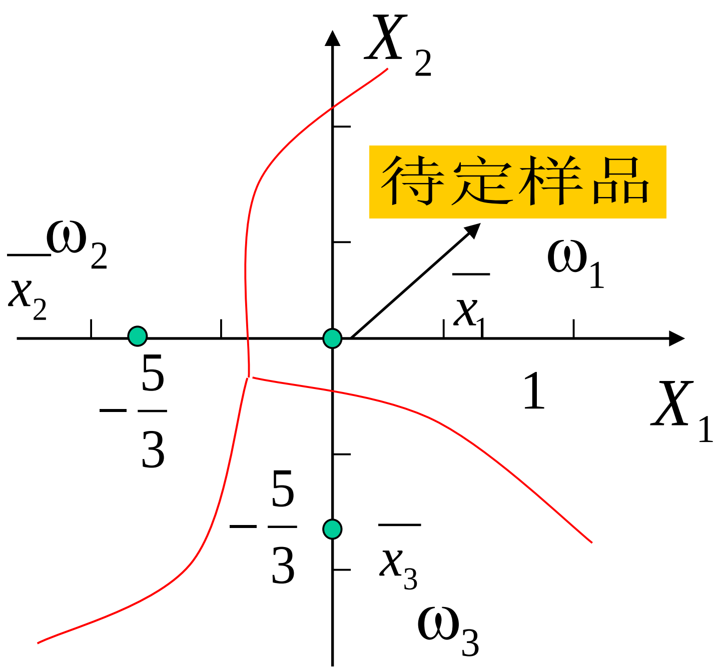
分别令$g_1(x)=g_2(x), g_2(x)=g_3(x), g_3(x)=g_1(x)$

$$g_1(x)-g_2(x)=x_1^2+6x_1+3.1=0$$

$$g_2(x)-g_3(x)=-x_1^2+x_2^2-5x_1+5x_2=0$$

$$g_3(x)-g_1(x)=-x_2^2-2x_1-5x_2-2.6=0$$

❖ 可得三类分界线如图所示：



➢ 解2、设三类协方差矩阵相等

$$\sum=\sum_1+\sum_2+\sum_3=\begin{pmatrix}\frac{7}{3} & 0\\ 0 & \frac{7}{3}\end{pmatrix},\sum^{-1}=\begin{pmatrix}\frac{3}{7} & 0\\ 0 & \frac{3}{7}\end{pmatrix}$$

代入多类时判别函数 : $g_i(x)=w_i^T x+w_{i0}$

所以$g_1(x)=\frac{3}{7}x_1-\frac{3}{14},g_2(x)=-\frac{5}{7}x_1-\frac{25}{42}$

$$g_3(x)=-\frac{5}{7}x_2-\frac{25}{42}$$

将$X=(0,0)^T$代入得 :

$$g_1(x)=-\frac{3}{14},g_2(x)=g_3(x)=-\frac{25}{42}$$

故应判样品$X=(0,0)^T$为ω_1类

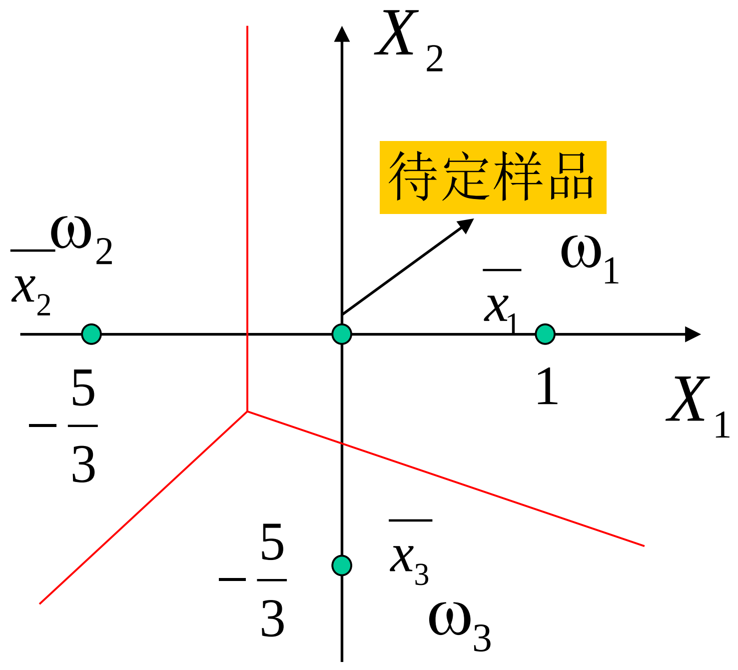
分别令$g_1(x)=g_2(x), g_2(x)=g_3(x), g_3(x)=g_1(x)$

$$g_1(x)-g_2(x)=\frac{8}{7}x_1+\frac{8}{21}$$

$$g_2(x)-g_3(x)=-\frac{5}{7}x_1+\frac{5}{7}x_2$$

$$g_3(x)-g_1(x)=-\frac{3}{7}x_1-\frac{5}{7}x_2-\frac{8}{21}$$

- 可得三类分界线如图所示：



3.9 序惯分类

(Sequential classification)

迄今为止所讨论的分类问题，关于待分类样本的所有信息都是一次性提供的。但是，在许多实际问题中，观察实际上是序贯的。随着时间的推移可以得到越来越多的信息。

假设对样品进行第 i 次观察获取一序列特征为：

$X=(x_1,x_2,\ldots,x_i)^T$，则对于ω_1、ω_2两类问题，

- 若$X\in\omega_1$，则判决完毕
- 若$X\in\omega_2$，则判决完毕
- 若X不属ω_1也不属ω_2，则不能判决，进行第i+1次观察，得$X=(x_1,x_2,\ldots,x_i,x_{i+1})^T$，再重复上面的判决，直到所有的样品分类完毕为止。

这样做的好处是使那些在二类边界附近的样本不会因某种偶然的微小变化而误判，当然这是以多次观察为代价的。

❖由最小错误概率的Bayes判决，对于两类问题，似然比为：

$$l_i(x) = \frac{P(X/\omega_1)}{P(X/\omega_2)} \begin{matrix} > \\ < \end{matrix} \frac{P(\omega_2)}{P(\omega_1)} \Rightarrow X \in \begin{matrix} \omega_1 \\ \omega_2 \end{matrix}$$

其中，特征向量$X = (x_1, x_2, ..., x_N)^T$

当测得第一个特征参数x_1时可计算其似然比：

$$l_1(x_1) = \frac{P_1(x|\omega_1)}{P_1(x|\omega_2)} = \frac{P(x_1/\omega_1)}{P(x_1/\omega_2)}$$

如果$l_1(x_1) \geq A$, 则$X = x_1 \in \omega_1$

如果$l_1(x_1) \leq B$, 则$X = x_1 \in \omega_2$

如果$B < l_1(x_1) < A$, 则测量下一个特征参数x_2并计算似然比：

$$l_2(x_1,\ x_2) = \frac{P(x_1,\ x_2|\omega_1)}{P(x_1,\ x_2|\omega_2)} \stackrel{def}{=} l_2(x) = \frac{P_2(x|\omega_1)}{P_2(x|\omega_2)}$$

若$l_2(x_1, x_2) \geq A$,则$X = (x_1, x_2)^T \in \omega_1$

若$l_2(x_1, x_2) \leq B$,则$X = (x_1, x_2)^T \in \omega_2$

若$B < l_2(x_1, x_2) < A$，再测第三个特征参数x_3，为重复以上过程直到所有样品的类别全部确定为止（其中A，B是上下门限）。

- 现在来确定A、B的值。
- 因为

$$l_N(x) = \frac{P_N(X/\omega_1)}{P_N(X/\omega_2)} \geq A, N\text{表示第}N\text{次测量}$$

$$\therefore P_N(X/\omega_1) \geq AP_N(X/\omega_2)$$

对上式两边对应于ω_1的特征空间内取积分可得：

$\int_{R_1} P_N(X/\omega_1)dX \geq A\int_{R_1} P_N(X/\omega_2)dX$，左边的积分代表模式X属于$\omega_1$类而判决为$\omega_1$类的概率。

即：$\int_{R_1} P_N(X/\omega_1)dX = 1-P_1(e)$

$P_1(e)=\int_{R_2} P_N(X/\omega_1)dX$表示本来属于$\omega_1$类而错判为$\omega_2$类的分类错误概率

而积分$\int_{R_1} P_N(X/\omega_2)dX = P_2(e)$

为本来属于ω_2类而错判为ω_1类的分类错误概率

$\therefore 1-P_1(e)\ge AP_2(e)$或$A\le {[1-P_1(e)]}\big/{P_2(e)}$

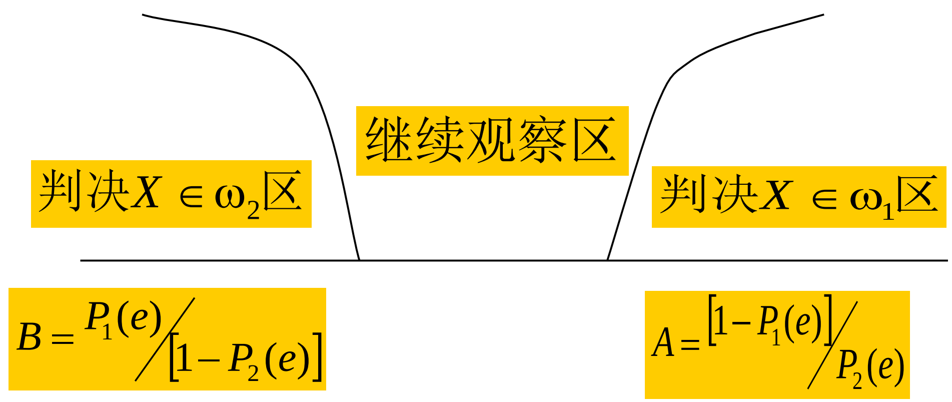
同理，因为$l_N(x)=\dfrac{P_N(X/\omega_1)}{P_N(X/\omega_2)}\le B$,所以$P_N(X/\omega_1)\le BP_N(X/\omega_1)$

即$\int_{R_2} P_N(X/\omega_1)dX \le B\int_{R_2} P_N(X/\omega_2)dX$

用错误概率表示为$P_1(e)\le B(1-P_2(e)),\therefore B\ge {P_1(e)}\big/{[1-P_2(e)]}$

$$A\le {[1-P_1(e)]}\big/{P_2(e)} \Rightarrow X\in\omega_1$$

$$B\ge {P_1(e)}\big/{[1-P_2(e)]} \Rightarrow X\in\omega_2$$



一般地，若取n个特征中的k个特征$x_1, x_2, \ldots, x_k$作特征向量，

令$\mathbf{x}^{(k)} = (x_1, x_2, \ldots, x_k)^{\mathrm{T}}$，定义关于这k个特征的似然比，称为序贯似然比：

$$l_k(\mathbf{x}^{(k)}) = \frac{P(x_1,\ x_2, \ldots, x_k \mid \omega_1)}{P(x_1,\ x_2, \ldots, x_k \mid \omega_2)} \stackrel{def}{=} \frac{P(\mathbf{x}^{(k)} \mid \omega_1)}{P(\mathbf{x}^{(k)} \mid \omega_2)}$$

❖ 序贯分类决策规则：

$$
\begin{cases}
l_k(\mathbf{x}^{(k)}) = \dfrac{P(\mathbf{x}^{(k)} \mid \omega_1)}{P(\mathbf{x}^{(k)} \mid \omega_2)} \geq A = \dfrac{1-P_1(e)}{P_2(e)} \text{时} \mathbf{x} \in \omega_1 \\
l_k(\mathbf{x}^{(k)}) = \dfrac{P(\mathbf{x}^{(k)} \mid \omega_1)}{P(\mathbf{x}^{(k)} \mid \omega_2)} \leq B = \dfrac{P_1(e)}{1-P_2(e)} \text{时} X \in \omega_2 \\
B < \dfrac{P(\mathbf{x}^{(k)} \mid \omega_1)}{P(\mathbf{x}^{(k)} \mid \omega_2)} < A \text{时，增加一个特征观察，继续识别}
\end{cases}
$$

- 上下门限A、B是由设计给定的错误概率$P_1(e)$，$P_2(e)$来确定的，Wald 已证明（此略），观察次数不会很大，它收敛的很快。

课堂思考题：

1.二类问题最小错误概率的Bayes决策的内容是什么？

2.二类问题最小风险Bayes决策的内容是什么？

3.对一批病人进行癌症普查，患癌症者定为属ω_1类，正常者属ω_2类。统计资料表明人们患癌症的概率为0.005。设有一种诊断此病的实验，其结果有阳性反应和阴性反应之分，依其作为诊断。化验结果是一维离散模式特征。统计资料表明：癌症有阳性反应的概率为0.95，正常人阳性反应的概率为0.01。若采用最小错误概率Bayes决策，问有阳性反映的人患癌症的概率有多大？

参考答案

1.答：

若对待分类模式的特征得到一个观察值x，利用下面公式：

$$P(\omega_i \mid X) = \frac{P(X \mid \omega_i)P(\omega_i)}{P(X)}$$

(1)若$P(X|\omega 1)P(\omega 1)>P(X|\omega 2)P(\omega 2)$，那么模式x真正类别状态是ω1；

(2)若$P(X|\omega 1)P(\omega 1)<P(X|\omega 2)P(\omega 2)$，那么模式x真正类别状态是ω2。

2.答：

对于二类问题最小风险Bayes决策，行动a1相当于决策“模式x的真正状态为ω1”，行动a2相当于决策“模式x的真正状态为ω2”。记 $\lambda_{ij} = \lambda(a_i, \omega j)$表示模式$x$本来属于ω$j$类而错判为ω$i$所受损失。将多类条件风险公式按二类问题展开可得到：

$$R(a_1 \mid X) = \lambda_{11} P(\omega_1 \mid X) + \lambda_{12} P(\omega_2 \mid X)$$

$$R(a_2 \mid X) = \lambda_{21} P(\omega_1 \mid X) + \lambda_{22} P(\omega_2 \mid X)$$

二类问题最小风险Bayes决策法则就是：

若R(a1|X)<R(a2|X)，则判定模式x的真正状态为ω_1；否则模式x的真正状态为ω_2。

3.解：

由题意：

(1)患癌症的概率为0.005，即$P(\omega 1)=0.005$，故$P(\omega 2)=0.995$；

(2)癌症有阳性反应的概率为0.95，即$P(x=阳|\omega 1)=0.95$，故$P(x=阴|\omega 1)=0.05$；

(3)正常人阳性反应的概率为0.01，即$P(x=阳|\omega 2)=0.01$，故$P(x=阴|\omega 2)=0.99$。

$$P(\omega_1 \mid x=\text{阳}) = \frac{P(x=\text{阳} \mid \omega_1)P(\omega_1)}{P(x=\text{阳})}$$

$$= \frac{P(x=\text{阳} \mid \omega_1)P(\omega_1)}{P(x=\text{阳} \mid \omega_1)P(\omega_1) + P(x=\text{阳} \mid \omega_2)P(\omega_2)}$$

$$= \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.01 \times 0.995}$$

$$= 32.3\%$$

说明有阳性反应的人其患癌症的概率为32.3%。

当医生必须回答其属于哪个类时，因为：

$P(x=阳|\omega 1)P(\omega 1)=0.95\times 0.005=0.00475;$

$P(x=阳|\omega 2)P(\omega 2)=0.01\times 0.995=0.00995$

$P(x=阳|\omega 1)P(\omega 1)<P(x=阳|\omega 2)P(\omega 2)$

故：$x=阳\in\omega 2$(即把有阳性反应的人判属为正常人)。