中国科学院自动化研究所

2001年　　博士生入学考试
"模式识别"试题

一、　论述以下概念并分析其解决问题的思想方法

(1)、基于最小错误率的 Bayes 决策;

(2)、最小最大决策;

(3)、Bayes 估计;

(4)、Fisher 线性判别;

(5)、最小平方误差准则函数;

(6)、特征选择;

二、　证明：二元正态分布的边缘分布和条件分布仍然是正态分布。

三、　在 MSE（最小平方误差）准则函数

$$J_s(a)=\|Ya-b\|^2=\sum_{n=1}^{N}(a^T y_n-b_n)^2$$

中，取 $b=U_N=(\underbrace{1,1,\ldots,1}_{N个})^T$，证明，当 $N\to\infty$ 时，MSE 的解以最小均方误差逼近 Bayes 判别函数

$$g_0(x)=P(\omega_1|x)-P(\omega_2|x)$$

在此，假定样本按照概率 $p(x)=p(x|\omega_1)P(\omega_1)+p(x|\omega_2)P(\omega_2)$ 独立抽取。

四、 给定先验概率相等的两类，其均值向量分别为：$\mu_1=[1,3,-1]^T$ 和 $\mu_2=[-1,-1,1]^T$，协方差矩阵是

$$S_1=\begin{bmatrix}4 & 1 & 0\\ 1 & 4 & 0\\ 0 & 0 & 1\end{bmatrix},\quad S_2=\begin{bmatrix}2 & 1 & 0\\ 1 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}$$

求用 J_5 判据的最优特征提取。

五、 利用最大似然估计方法来估计单变量正态分布的均值 μ 和方差 σ^2。

六、 设两类样本 ω_1 和 ω_2，如下图

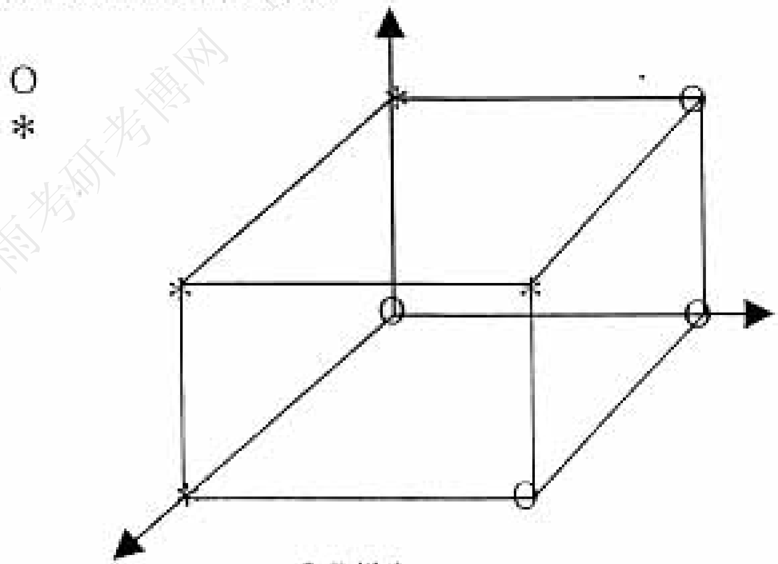
ω_1（用 O 表示）包含四个样本点：

$$x_{1,1}=(0,0,0)^T \qquad x_{1,2}=(1,0,0)^T$$
$$x_{1,3}=(1,0,1)^T \qquad x_{1,4}=(1,1,0)^T$$

ω_2（用*表示）包含四个样本点：

$$x_{2,1}=(0,0,1)^T, \qquad x_{2,2}=(0,1,0)^T,$$
$$x_{2,3}=(0,1,1)^T, \qquad x_{2,4}=(1,1,1)^T$$

试利用散度判据降低维数。



两类样本 ω_1 和 ω_1