式(或为了实现方便的考虑将判别函数设定为某种较简单的形式),再通过训练样本确定其中的参数,就能够更简便地设计出分类器。这就是从样本出发直接设计分类器的思路,本书第 4 章、第 5 章中的方法及第 15 章中介绍的支持向量机等都属于这类方法。在很多实际问题中,这类方法往往更具实用价值。

这里提到的方法都是分两步来解决模式识别问题的,即首先根据已知数据(训练样本)设计分类器,然后用它对未知数据进行分类。能否直接从训练样本出发把未知数据分类呢?本书第 6 章将要介绍的近邻法就是采用了这种做法。

习　　题

2.1　如果只知道各类的先验概率,最小错误率贝叶斯决策规则应如何表示?

2.2　利用概率论中的乘法定理和全概率公式证明贝叶斯公式

$$P(\omega_i,\boldsymbol{x})=\frac{p(\boldsymbol{x}|\omega_i)P(\omega_i)}{p(\boldsymbol{x})}$$

2.3　证明:在两类情况下 $P(\omega_1|\boldsymbol{x})+P(\omega_2|\boldsymbol{x})=1$。

2.4　分别写出在以下两种情况

(1)$P(\boldsymbol{x}|\omega_1)=P(\boldsymbol{x}|\omega_2)$

(2)$P(\omega_1)=P(\omega_2)$

下的最小错误率贝叶斯决策规则。

2.5　(1)对 c 类情况推广最小错误率贝叶斯决策规则;

(2)指出此时使错误率最小等价于后验概率最大,即

$$P(\omega_i|\boldsymbol{x})>P(\omega_j|\boldsymbol{x})\quad 对一切\ j\neq i\ 成立时,\boldsymbol{x}\in\omega_i。$$

2.6　对两类问题,证明最小风险贝叶斯决策规则可表示为

$$若\quad \frac{p(\boldsymbol{x}|\omega_1)}{p(\boldsymbol{x}|\omega_2)}\gtrless\frac{(\lambda_{12}-\lambda_{22})P(\omega_2)}{(\lambda_{21}-\lambda_{11})P(\omega_1)},则\ \boldsymbol{x}\in\begin{cases}\omega_1\\ \omega_2。\end{cases}$$

2.7　若 $\lambda_{11}=\lambda_{22}=0,\lambda_{12}=\lambda_{21}$,证明此时最小最大决策面使来自两类的错误率相等。

2.8　对于同一个决策规则判别函数可定义成不同形式,从而有不同的决策面方程,指出决策区域是不变的。

2.9　写出两类和多类情况下最小风险贝叶斯决策判别函数和决策面方程。

2.10　随机变量 $l(\boldsymbol{x})$ 定义为

$$l(\boldsymbol{x})=\frac{p(\boldsymbol{x}|\omega_1)}{p(\boldsymbol{x}|\omega_2)},l(\boldsymbol{x})又称为似然比,试证明$$

(1)$E\{l^n(\boldsymbol{x})|\omega_1\}=E\{l^{n-1}(\boldsymbol{x})|\omega_2\}$

(2)$E\{l(\boldsymbol{x})|\omega_2\}=1$

(3)$E\{l(\boldsymbol{x})|\omega_1\}-E\{l(\boldsymbol{x})|\omega_2\}=\mathrm{Var}\{l(\boldsymbol{x})|\omega_2\}$

2.11　$x_j(j=1,2,\cdots,n)$为 n 个独立随机变量,有

$$E[x_j|\omega_i]=ij\eta\quad \mathrm{Var}[x_j|\omega_i]=i^2j^2\sigma^2$$

计算在 $\lambda_{11}=\lambda_{22}=0$ 及 $\lambda_{12}=\lambda_{21}=1$ 情况下,由贝叶斯决策引起的错误率。(提示:用中心极限定理)

2.12 写出离散情况的贝叶斯公式。

2.13 把连续情况的最小错误率贝叶斯决策推广到离散情况，并写出其判别函数。

2.14 写出离散情况条件风险 $R(a_i|\boldsymbol{x})$的定义，并指出其决策规则。

2.15 证明多元正态分布的等密度点轨迹是一个超椭球面，且其主轴方向由 Σ 的特征向量决定，轴长度由 Σ 的特征值决定。

2.16 证明 Mahalanobis 距离 r 符合距离定义三公理，即

(1) $r(a,b)=r(b,a)$

(2) 当且仅当 $a=b$ 时，$r(a,b)=0$

(3) $r(a,c)\leqslant r(a,b)+r(b,c)$

2.17 若将 Σ^{-1}阵写为

$$\Sigma^{-1}=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1d} \\ h_{12} & h_{22} & \cdots & h_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1d} & h_{2d} & \cdots & h_{dd} \end{bmatrix}$$

证明 Mahalanobis 距离平方为

$$\gamma^2=\sum_{i=1}^{d}\sum_{j=1}^{d}h_{ij}(\boldsymbol{x}_i-\mu_i)(\boldsymbol{x}_j-\mu_j)$$

2.18 分别对于 $d=2,d=3$ 证明对应于 Mahalanobis 距离 γ 的超椭球体积是

$$V=V_d|\Sigma|^{\frac{1}{2}}\gamma^d$$

其中 V_d 是 d 维单位超球体积

$$V_d=\begin{cases}\pi^{d/2}\Big/\left(\dfrac{d}{2}\right)! & d=2\\[2ex] \dfrac{2^d\pi^{(d-1)/2}\left(\dfrac{d-1}{2}\right)!}{d!} & d=3\end{cases}$$

2.19 假定 $\boldsymbol{x}$ 和 m 是两个随机变量，并设在给定 m 时，x 的条件密度为

$$p(x|m)=(2\pi)^{\frac{1}{2}}\sigma^{-1}\exp\left\{-\frac{1}{2}(x-m)^2/\sigma^2\right\}$$

再假定 m 的边缘分布是正态的，期望值为 m_0，方差为 σ_m^2，证明

$$p(m|x)=\frac{(\sigma^3+\sigma_m)^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\sigma\sigma_m}\exp\left[-\frac{1}{2}\frac{\sigma^2+\sigma_m^2}{\sigma^2\sigma_m^2}\left(m-\frac{\sigma_m^2x+m_0\sigma^2}{\sigma^2+\sigma_m^2}\right)^2\right]$$

（提示：利用贝叶斯公式）

2.20 对 $\Sigma_i=\sigma^2I$ 的特殊情况，证明

(1)若 $P(\omega_i)\neq P(\omega_j)$、则超平面靠近先验概率较小的类；

(2)在什么情况下，先验概率对超平面的位置影响不大。

2.21 对 $\Sigma_i=\Sigma$ 的特殊情况，指出在先验概率不等时，决策面沿 μ_i 点与 μ_j 点连线向先验概率小的方向移动。

2.22 似然比决策准则为

$$若\ l(\boldsymbol{x})=\frac{p(\boldsymbol{x}|\omega_1)}{p(\boldsymbol{x}|\omega_2)}\gtrless\frac{P(\omega_2)}{P(\omega_1)}\quad 则\ \boldsymbol{x}\in\begin{cases}\omega_1\\ \omega_2\end{cases}。$$

负对数似然比为 $h(\boldsymbol{x})=-\ln[l(\boldsymbol{x})]$，当 $p(\boldsymbol{x}|\omega_i)$是均值向量为 μ_i 和协方差矩阵为 Σ_i 的正态分布时：

(1)试推导出 $h(\boldsymbol{x})$，并指出其决策规则；

(2)当 $\Sigma_1=\Sigma_2=\Sigma$ 时，推导 $h(\boldsymbol{x})$及其决策规则；

(3)分析(1)，(2)两种情况下的决策面类型。

2.23 二维正态分布，$\mu_1=(-1,0)^T$，$\mu_2=(1,0)^T$，$\Sigma_1=\Sigma_2=I$，$P(\omega_1)=P(\omega_2)$。试写出负对数似然比决策规则。

2.24 习题 2.23 中，若 $\Sigma_1\neq\Sigma_2$，

$$\Sigma_1=\begin{bmatrix}1 & \frac{1}{2}\\ \frac{1}{2} & 1\end{bmatrix},\quad \Sigma_2=\begin{bmatrix}1 & -\frac{1}{2}\\ -\frac{1}{2} & 1\end{bmatrix},$$

写出负对数似然比决策规则。

2.25 在习题 2.24 的情况下，若考虑损失函数 $\lambda_{11}=\lambda_{22}=0$，$\lambda_{12}=\lambda_{21}$，画出似然比阈值与错误率间的关系。

(1)求出 $P(e)=0.05$ 时完成 Neyman-Pearson 决策时总的错误率；

(2)求出最小最大决策的域值和总的错误率。