

《模式识别》试题库

一、基本概念题

1.1 模式识别的三大核心问题是：_____、_____、_____。

1.2、模式分布为团状时，选用_____聚类算法较好。

1.3 欧式距离具有_____。 马式距离具有_____。

(1) 平移不变性 (2) 旋转不变性 (3) 尺度缩放不变性 (4) 不受量纲影响的特性

1.4 描述模式相似的测度有：_____。

(1) 距离测度 (2) 模糊测度 (3) 相似测度 (4) 匹配测度

1.5 利用两类方法处理多类问题的技术途径有：(1) _____；(2) _____；

(3) _____。其中最常用的是第_____个技术途径。

1.6 判别函数的正负和数值大小在分类中的意义

是：_____，
_____。

1.7 感知器算法_____。

(1) 只适用于线性可分的情况；(2) 线性可分、不可分都适用。

1.8 积累位势函数法的判别界面一般为_____。

(1) 线性界面；(2) 非线性界面。

1.9 基于距离的类别可分性判据有：_____。

$$(1) \operatorname{Tr}[S_w^{-1} S_B] \quad (2) \frac{S_B}{S_W} \quad (3) \left| \frac{S_B}{S_W + S_B} \right|$$

1.10 作为统计判别问题的模式分类，在()情况下，可使用聂曼-皮尔逊判决准则。

1.11 确定性模式非线性分类的势函数法中，位势函数 $K(x, x_k)$ 与积累位势函数 $K(x)$ 的关系为

()。

1.12 用作确定性模式非线性分类的势函数法，通常，两个 n 维向量 x 和 x_k 的函数 $K(x, x_k)$ 若同时满足下列三个条件，都可作为势函数。

① ()；

② () ;

③ $K(x, x_k)$ 是光滑函数, 且是 x 和 x_k 之间距离的单调下降函数。

1.13 散度 J_{ij} 越大, 说明 ω_i 类模式与 ω_j 类模式的分布 ()。当 ω_i 类模式与 ω_j 类模式的分布相同时, $J_{ij} = ()$ 。

1.14 若用 Parzen 窗法估计模式的类概率密度函数, 窗口尺寸 h_1 过小可能产生的问题是 (), h_1 过大可能产生的问题是 ()。

1.15 信息熵可以作为一种可分性判据的原因是: _____。

1.16 作为统计判别问题的模式分类, 在 () 条件下, 最小损失判决规则与最小错误判决规则是等价的。

1.17 随机变量 $l(\vec{x}) = p(\vec{x}|\omega_1)/p(\vec{x}|\omega_2)$, $l(\vec{x})$ 又称似然比, 则 $E\{l(\vec{x})|\omega_2\} = ()$ 。在最小误判概率准则下, 对数似然比 Bayes 判决规则为 ()。

1.18 影响类概率密度估计质量的最重要因素是 ()。

1.19 基于熵的可分性判据定义为 $J_H = E_x[-\sum_{i=1}^c P(\omega_i | \vec{x}) \log P(\omega_i | \vec{x})]$, J_H 越 (), 说明模式的可分性越强。当 $P(\omega_i | \vec{x}) = ()$ ($i=1, 2, \dots, c$) 时, J_H 取极大值。

1.20 K_n 近邻元法较之于 Parzen 窗法的优势在于 ()。

上述两种算法的共同弱点主要是 ()。

1.21 已知有限状态自动机 $A_f = (\Sigma, Q, \delta, q_0, F)$, $\Sigma = \{0, 1\}$; $Q = \{q_0, q_1\}$;

δ : $\delta(q_0, 0) = q_1$, $\delta(q_0, 1) = q_1$, $\delta(q_1, 0) = q_0$, $\delta(q_1, 1) = q_0$; $q_0 = q_0$; $F = \{q_0\}$ 。现有输入字符串: (a) 00011101011, (b) 1100110011, (c) 101100111000, (d) 0010011, 试问, 用 A_f 对上述字符串进行分类的结果为 ()。

1.22 句法模式识别中模式描述方法有: _____。

(1) 符号串 (2) 树 (3) 图 (4) 特征向量

1.23 设集合 $X=\{a, b, c, d\}$ 上的关系,

$R=\{(a, a), (a, b), (a, d), (b, b), (b, a), (b, d), (c, c), (d, d), (d, a), (d, b)\}$, 则 a, b, c, d 生成的 R 等价类分别为 ($[a]_R=$, $[b]_R=$, $[c]_R=$, $[d]_R=$)。

1.24 如果集合 X 上的关系 R 是传递的、() 和 () 的, 则称 R 是一个等价关系。

1.25 一个模式识别系统由那几部分组成? 画出其原理框图。

1.26 统计模式识别中, 模式是如何描述的。

1.27 简述随机矢量之间的统计关系: 不相关, 正交, 独立的定义及它们之间的关系。

1.28 试证明, 对于正态分布, 不相关与独立是等价的。

1.29 试证明, 多元正态随机矢量的线性变换仍为多元正态随机矢量。

1.30 试证明, 多元正态随机矢量 \vec{X} 的分量的线性组合是一正态随机变量。

第二部分 分析、证明、计算题

第二章 聚类分析

2.1 影响聚类结果的主要因素有那些?

2.2 马氏距离有那些优点?

2.3 如果各模式类呈现链状分布, 衡量其类间距离用最小距离还是用最大距离? 为什么?

2.4 动态聚类算法较之于简单聚类算法的改进之处何在? 层次聚类算法是动态聚类算法吗? 比较层次聚类算法与 c -均值算法的优劣。

2.5 ISODATA 算法较之于 c -均值算法的优势何在?

2.6 简述最小张树算法的优点。

2.7 证明马氏距离是平移不变的、非奇异线性变换不变的。

2.8 设, 类 ω_p 、 ω_q 的重心分别为 \vec{x}_p 、 \vec{x}_q , 它们分别有样本 n_p 、 n_q 个。将和 ω_q 合并为 ω_l , 则 ω_l 有 $n_l = n_p + n_q$ 个样本。另一类 ω_k 的重心为 \vec{x}_k 。试证明 ω_k 与 ω_l 的距离平方是

$$D_{kl}^2 = \frac{n_p}{n_k + n_l} D_{kp}^2 + \frac{n_q}{n_k + n_l} D_{kq}^2 - \frac{n_p n_q}{n_k + n_l} D_{pq}^2$$

2.9 (1) 设有 M 类模式 ω_i , $i=1, 2, \dots, M$, 试证明总体散布矩阵 S_T 是总类内散布矩阵 S_W 与类间散布矩阵 S_B 之和, 即 $S_T = S_W + S_B$ 。

(2) 设有二维样本: $x_1=(-1, 0)^T$, $x_2=(0, -1)^T$, $x_3=(0, 0)^T$, $x_4=(2, 0)^T$ 和 $x_5=(0, 2)^T$ 。试选用一种合适的方法进行一维特征提取 $y_i = W^T x_i$ 。要求求出变换矩阵 W , 并求出变换结果 y_i , ($i=1, 2, 3, 4, 5$)。

(3) 根据(2)特征提取后的一维特征, 选用一种合适的聚类算法将这些样本分为两类, 要求每类样本个数不少于两个, 并写出聚类过程。

2.10 (1) 试给出 c-均值算法的算法流程图;

$$J^{(k)} = \sum_{j=1}^c \sum_{\bar{x}_i \in \omega_j^{(k)}} (\bar{x}_i - \bar{z}_j^{(k)})^T (\bar{x}_i - \bar{z}_j^{(k)})$$

(2) 试证明 c-均值算法可使误差平方和准则

其中, k 是迭代次数; $\bar{z}_j^{(k)}$ 是 $\omega_j^{(k)}$ 的样本均值。

2.11 现有 $2k+1$ 个一维样本, 其中 k 个样本在 $x=-2$ 处重合, 另 k 个样本在 $x=0$ 处重合, 只有 1 个在 $x=a>0$ 处。若 $a=2(k+1)$, 证明, 使误差平方和准则 J_c 最小的两类划分是 $x=0$ 处的 k 个样本与 $x=a$ 处的 1 个样本为一类, 其余为另一类。这里,

$$J_c = \sum_{j=1}^c \sum_{i=1}^{N_j} (x_i - m_j)^2$$

其中, c 为类别数, N_j 是第 j 类的样本个数, $x_i \in \omega_j$, $i=1, 2, \dots, N_j$, m_j 是第 j 类的样本均值。

2.12 有样本集 $\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$, 试用谱系聚类算法对其分类。

2.13 设有样本集 $S = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$, 证明类心 \bar{z} 到 S 中各样本点距离平方和 $\sum_{i=1}^n (\bar{x}_i - \bar{z})^T (\bar{x}_i - \bar{z})$ 为最

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n \bar{x}_i$$

小时, 有

2.14 假设 s 为模式矢量集 X 上的距离相似侧度, 有 $\forall x, y \in X, s(x, y) > 0$ 且当 $a > 0$ 时,

$$d(x, y) = a / s(x, y)$$

。证明 d 是距离差异性测度。

2.15 证明欧氏距离满足旋转不变性。

提示: 运用 Minkowski 不等式, 对于两矢量 $x = [x_1, \dots, x_l]^T$ 和

$d_{\min}^{ss}(s_{\min}^{ss}), d_{\max}^{ss}(s_{\max}^{ss}), d_{\text{avg}}^{ss}(s_{\text{avg}}^{ss}), d_{\text{mean}}^{ss}(s_{\text{mean}}^{ss})d(s)$, 满足

$$\left(\sum_{i=1}^l |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^l |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^l |y_i|^p\right)^{1/p}$$

2.16 证明:

(a) 如果 s 是类 X 上的距离相似测度, $\forall x, y > 0, s(x, y) > 0$, 那么对于 $\forall a > 0$, $s(x, y) + a$ 也是类 X 上的距离测度。

(b) 如果 d 是类 X 上的距离差异性测度, 那么对于 $\forall a > 0$, $d + a$ 也是类 X 上的距离差异性测度

2.17 假设 $f: R^+ \rightarrow R^+$ 是连续单调递增函数, 满足

$$f(x) + f(y) \geq f(x + y), \forall x, y \in R^+$$

d 是类 X 上的距离差异性测度且 $d_0 \geq 0$ 。证明 $f(d)$ 也是类 X 上的距离差异性测度。

2.18 假设 s 为类 X 上的距离相似测度, 有 $\forall x, y > 0, s(x, y) > 0$, $f: R^+ \rightarrow R^+$ 是连续单调递增函数, 满足

$$f(x) + f(y) \geq f\left(\frac{1}{\frac{1}{x} + \frac{1}{y}}\right), \forall x, y \in R^+$$

证明 $f(x)$ 是 X 上的距离相似测度。

2.19 证明: 对于模式矢量集 X 上任意两个矢量 \vec{x} 和 \vec{y} 有

$$d_{\infty}(\vec{x}, \vec{y}) \leq d_2(\vec{x}, \vec{y}) \leq d_1(\vec{x}, \vec{y})$$

2.20 (a) 证明公式 $s_F^q(\vec{x}, \vec{y}) = \left(\sum_{i=1}^l s(\vec{x}_i, \vec{y}_i)^q\right)^{1/q}$ 中 $s_F(\vec{x}, \vec{y})$ 的最大最小值分别是和 $0.5l^{1/q}$ 。

(b) 证明当 $q \rightarrow +\infty$ 时, 公式 $s_F^q(\vec{x}, \vec{y}) = \left(\sum_{i=1}^l s(\vec{x}_i, \vec{y}_i)^q\right)^{1/q}$ 中 $s_F(\vec{x}, \vec{y}) = \max_{1 \leq i \leq l} s(\vec{x}_i, \vec{y}_i)$

2.21 假设 d 是模式矢量集 X 上的差异性测度, $s = d_{\max} - d$ 是相应相似测度。

证明 $s_{\text{avg}}^{ps}(x, C) = d_{\max} - d_{\text{avg}}^{ps}(x, C), \forall x \in X, C \subset X$

其中 s_{avg}^{ps} 和 d_{avg}^{ps} 是分别根据 s 和 d 所定义的。 Ψ_{avg}^{ps} 的定义来自于下面公式，其中第一个集合只含有一个矢量。

提示：平均亲近函数

$$\Psi_{avg}^{ps}(D_i, D_j) = \frac{1}{n_{D_i} n_{D_j}} \sum_{x \in D_i} \sum_{y \in D_j} \Psi(x, y), \text{ 其中 } n_{D_i} \text{ 和 } n_{D_j} \text{ 分别是集合 } D_i \text{ 和 } D_j \text{ 的}$$

势。即使 Ψ 是测度，显然 Ψ_{avg}^{ps} 不是测度。在公式中， D_i 和 D_j 中的所有矢量都参与计算。

2.22 假设 $x, y \in \{0, 1\}^l$ 。证明 $d_2(x, y) = \sqrt{d_{Hamming}(x, y)}$ 。

2.23 考虑一维空间的两矢量， $x = [x_1, \dots, x_l]^T$ 和 $y = [y_1, \dots, y_l]^T$,

$$|x_i - y_i| = \max_{j=1, \dots, l} \{|x_j - y_j|\}, \text{ 定义距离 } d_n(x, y) \text{ 为}$$

$$d_n(x, y) = |x_i - y_i| + \frac{1}{l - [(l-2)/2]} \sum_{j=1, j \neq i}^l |x_i - y_i|$$

这个距离曾被提议作为欧氏距离的近似值。

(a) 证明 d_n 是距离。

(b) 比较 d_n 和 d_2 的计算复杂度。

2.24 若定义下列准则函数

$$J_T = \sum_{i=1}^c \sum_{x \in X_i} (x - m_i)^T S_T^{-1} (x - m_i)$$

其中 m_i 是 X_i 中 N_i 个样本的均值向量， S_T 是总散布矩阵，

(1) 证明 J_T 对数据的非奇异线形变换具有不变性。

(2) 证明把 X_i 中的样本 \hat{x} 转移到 X_j 中去，则使 J_T 改变为

$$J_T^* = J_T - \left[\frac{N_j}{N_j + 1} (\hat{x} - m_j)^T S_T^{-1} (\hat{x} - m_j) - \frac{N_i}{N_i - 1} (\hat{x} - m_i)^T S_T^{-1} (\hat{x} - m_i) \right]$$

(3) 写出使 J_T 最小化的迭代程序。

2.25 证明对于 C-均值算法, 聚类准则函数满足使算法收敛的条件。(即若 $J(\Gamma, \tilde{K}) \leq J(\Gamma, K)$, 则有

$$J(\tilde{\Gamma}, \tilde{K}) \leq J(\Gamma, \tilde{K})$$

2.26 令 $\Delta(y, K_j) = \frac{1}{2}(y - m_j)^T \Sigma_j^{-1}(y - m_j) + \frac{1}{2} \log |\Sigma_j|$ 是点到聚类的相似性度量, 式中 m_j 和 Σ_j 是

聚类 Γ_j 的均值和协方差矩阵, 若把一点从 Γ_i 转移到 Γ_j 中去, 计算由公式

$$J_K = \sum_{i=1}^c \sum_{y \in \Gamma_j} \Delta(y, K_j)$$

所示 J_K 的变化值。

第三章 判别域代数界面方程法

3.1 证明感知器算法在训练模式是线性可分的情况下, 经过有限次迭代后可以收敛到正确的解矢量 w^* 。

3.2

(1) 试给出 LMSE 算法 (H-K 算法) 的算法流程图;

(2) 试证明 $X^*e(k)=0$, 这里, X^* 是伪逆矩阵; $e(k)$ 为第 k 次迭代的误差向量;

(3) 已知两类模式样本 ω_1 : $x_1=(-1, 0)^T$, $x_2=(1, 0)^T$; ω_2 : $x_3=(0, 0)^T$, $x_4=(0, -1)^T$ 。

试用 LMSE 算法判断其线性可分性。

3.3 设等式方程组 $X\bar{w} = \bar{b}$, 其中: 属于 ω_1 的样本作为 X 的前 N_1 行, 属于 ω_2 的样本作为 X 的后 N_2

$$b = (\underbrace{\frac{N}{N_1}, \dots, \frac{N}{N_1}}_{N_1}, \underbrace{\frac{N}{N_2}, \dots, \frac{N}{N_2}}_{N_2})$$

行。证明: 当余量矢量

时, MSE 解等价于 Fisher 解。

3.4 已知二维样本: $\vec{x}_1=(-1, 0)^T$, $\vec{x}_2=(0, -1)^T$, $\vec{x}_3=(0, 0)^T$, $\vec{x}_4=(2, 0)^T$ 和 $\vec{x}_5=(0, 2)^T$, $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\} \in \omega_1$,

$\{\vec{x}_4, \vec{x}_5\} \in \omega_2$ 。试用感知器算法求出分类决策函数, 并判断 $\vec{x}_6=(1, 1)^T$ 属于哪一类?

3.4. 已知模式样本 $x_1=(0, 0)^T$, $x_2=(1, 0)^T$, $x_3=(-1, 1)^T$ 分别属于三个模式类别, 即, $x_1 \in \omega_1$, $x_2 \in \omega_2$, $x_3 \in \omega_3$,

(1) 试用感知器算法求判别函数 $g_i(x)$, 使之满足, 若 $x_i \in \omega_i$ 则 $g_i(x) > 0$, $i=1, 2, 3$;

(2) 求出相应的判决界面方程, 并画出解区域的示意图。

给定校正增量因子 $C=1$, 初始值可以取:

$$w_1(1) = (4, -9, -4)^T, w_2(1) = (4, 1, -4)^T, w_3(1) = (-4, -1, -6)^T.$$

3.5 已知 $\omega_1: \{(0, 0)^T\}$, $\omega_2: \{(1, 1)^T\}$, $\omega_3: \{(-1, 1)^T\}$ 。用感知器算法求该三类问题的判别函数，并画出解区域。

3.6 试证明：

(1) 从 \vec{x} 到超平面 $g(\vec{x}) = \vec{w}^T \vec{x} + w_0 = 0$ 的距离 $r = \frac{|g(\vec{x})|}{\|\vec{w}\|}$ 是在 $g(\vec{x}_q) = 0$ 的约束条件下，使 $\|\vec{x} - \vec{x}_q\|^2$ 达到极小的解。

(2) \vec{x} 在超平面上的投影是 $\vec{x}_p = \vec{x} - \frac{g(\vec{x})}{\|\vec{w}\|^2} \vec{w}$ 。

3.7 设有一维空间二次判别函数 $g(x) = 5 + 7x + 9x^2$ ，试将其映射成广义齐次线性判别函数 $g(\vec{x}) = \vec{a}^T \vec{y}$ 。

3.8 对二维线性判别函数 $g(\vec{x}) = x_1 + 2x_2 - 2$

(1) 将判别函数写成 $g(\vec{x}) = \vec{w}^T \vec{x} + w_0$ 的形式，并画出 $g(\vec{x}) = 0$ 的几何图形；

(2) 将其映射成广义齐次线性判别函数 $g(\vec{x}) = \vec{a}^T \vec{y}$ ；

(3) 指出上述 X 空间实际是 Y 空间的一个子空间，且 $\vec{a}^T \vec{y} = 0$ 对 X 子空间的划分与原空间中 $\vec{w}^T \vec{x} + w_0 = 0$ 对原 X 空间的划分相同，并在图上表示出来。

3.9 指出在Fisher线性判别中， \vec{w} 的比例因子对Fisher判别结果无影响的原因。

3.10 证明两向量外积组成的矩阵一般是奇异的。

3.11 证明，在几何上，感知器准则函数值正比于被错分类样本到决策面的距离之和。

3.12 解释为什么感知器函数是一个连续分段的线性分类器。

3.13 如果在感知器算法中 $\rho_k = \rho$ ，那么在 $k_0 = \frac{\|w(0) - \alpha w^*\|}{\beta^2 \rho (2 - \rho)}$ 步之后，这个算法收敛，其中

$$\alpha = \frac{\beta^2}{|\gamma|}, \quad \rho < 2。$$

3.14 证明感知器算法的正确分类和错误分类在有限个反复的运算以后是收敛的

3.15 考虑一种情况，在类 ω_1 中包含两个特征向量， $[0,1]^T$ 。类 ω_2 中包含 $[1,0]^T$ 和 $[1,1]^T$ 两

个向量。根据感知器算法，其中 $\rho = 1$ ， $w(0) = [0.5, 0.5]^T$ ，设计一个线性分离器来区分这两类

3.16 在上一章 2.12 问题中两分类问题中，取 $\mu_1^T = [1,1]$ ， $\mu_2^T = [0,0]$ ， $\sigma_1^2 = \sigma_2^2 = 0.2$ 。对于每一类产生 50 个向量。为了确保对于这两类的线性分离，对于向量 $[1, 1]$ 类确保 $x_1 + x_2 < 1$ ，

对于 $[0, 0]$ 向量类 $x_1 + x_2 > 1$ 。下面的步骤就是使用这些向量去设计一个线性分类器使用 (3.21) 中的感知器算法。在收敛以后，画出相关的判定线

3.17 假如 2.12 问题中是多类分类问题，每一类有 100 个样本点。根据 LMS 算法使用这些数据去设计一个线性分类器。当所有的点被带入这个算法中进行计算的时候，画出这个算法收敛的相关超平面。其中 $\rho_k = \rho = 0.01$ ，然后使用 $\rho = 0.01$ 。观察这个结果

3.18 证明，使用 KESLER 构造器，经过前面 3.21 感知器算法的有限步正确与错误分类计算后，对于一个 $x_{(t)} \in \omega_i$ ，变为

$$\begin{aligned} \omega_i(t+1) &= \omega_i(t) + \rho x_{(t)} & \text{if } \omega_i^T x_{(t)} < \omega_j^T(t) x_{(t)} & \quad j \neq i \\ \omega_i(t+1) &= \omega_i(t) - \rho x_{(t)} & \text{if } \omega_i^T x_{(t)} < \omega_j^T(t) x_{(t)} & \quad j \neq i \\ \omega_k(t+1) &= \omega_k(t) & \forall k \neq j & \quad \text{and } k \neq i \end{aligned}$$

3.19 证明理想权重向量的误差平方和趋渐进于 MSE 的解。

3.20 使用均方误差和的原则解问题 3.6 并设计一个线性分类器。

3.21 证明设计一个 M 类的线性分类器，有最佳误差平方和。分类器减少到 M 等价个有相应的效果。

3.22 证明，假如 x, y 服从联合高斯分布，对于 x 条件下 y 的分布是

$$E[y|x] = \frac{\alpha \sigma_y x}{\sigma_x} + \mu_y - \frac{\alpha \sigma_y \mu_x}{\sigma_x},$$

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \alpha \sigma_x \sigma_y \\ \alpha \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}$$

3.23 取 M 类分类器按照参数函数 $g(x; \omega_k)$ 的形式存在，目的是估计参数 ω_k ，使得分类器根据输入向量 x 能够产生期望的响应输出值。假设在每一类中 x 是随机分布，分类器的输出根据相关期望响应值的不同而不同。按照高斯已知变量的一个高斯分布，假设所有的输出都是相同的。证明按照误差平方和的原则，ML 估计是产生一个等价的估计值。

提示：在已知的类别当中取出 N 个训练样本值。对于他们中的每一个形成 $y_i = g(x_i; \omega_k) - d_k^i$ 。 d_k^i 是第 k 类中第 i 个样本点的期望响应值。 y_i^s 服从正态 0 均值，方差为 σ^2 的分布。这个似然函数使用 y_i^s

3.24 在二类分类问题中，贝叶斯最佳判定截面是通过 $g(x) = P(\omega_1|x) - P(\omega_2|x) = 0$ 给出，证明 MSE 中训练一个判定界面 $f(x; \omega)$ ，目的是对两类进行有效判别，相关的，它等价于在 MSE 最优感知中，它等价于 $f(x; \omega)$ 的渐进函数形式 $g(\cdot)$ 。

3.25 假设在两类分类问题中有服从联合分布的特征向量，他们在有共同的方差 Σ 。设计一个线性 MSE 分类器，证明在 2.11 问题中的贝叶斯分类器和这个结果的 MSE 分类器仅仅通过一个阈值就可以区分。简化起见，仅仅考虑等概率的类的情况。

提示：计算 MSE 超平面 $\omega^T x + \omega_0 = 0$ ，增加 x 的维数，它的解按照下列方式提供，

$$\begin{bmatrix} R & E[x] \\ E[x]^T & 1 \end{bmatrix} \begin{bmatrix} w \\ w_0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(\mu_1 - \mu_2) \\ 0 \end{bmatrix}$$

相关的 R 和 Σ 在 MSE 分类器中按照下列的形式给出

$$(\mu_1 - \mu_2)^T \Sigma^{-1} \left(x - \frac{1}{2}(\mu_1 + \mu_2) \right) > (<) 0$$

第四章 统计判决

4.1 使用最小最大损失判决规则的错分概率是最小吗？为什么？

4.2 当 $\Sigma i = \sigma^2 I$ 时，先验概率对决策超平面的位置影响如何？

4.3 假设在某个地区的细胞识别中正常 ω_1 和异常 ω_2 两类的先验概率分别为

$$\text{正常状态 : } P(\omega_1) = 0.9$$

$$\text{异常状态: } P(\omega_2) = 0.1$$

现有一待识的细胞，其观测值为 x ，从类条件概率密度分布曲线上查得 $p(x/\omega_1) = 0.2$, $p(x/\omega_2) = 0.4$ 并且已知损失系数为 $\lambda_{11}=0$, $\lambda_{12}=1$, $\lambda_{21}=6$, $\lambda_{22}=0$ 。

试对该细胞以以下两种方法进行分类：①基于最小错误概率准则的贝叶斯判决；②基于最小损失准则的贝叶斯判决。请分析两种分类结果的异同及原因。

4.4 试用最大似然估计的方法估计单变量正态分布的均值 μ 和方差 σ^2 。

4.5 已知两个一维模式类别的类概率密度函数为

$$\begin{aligned} p(x|\omega_1) &= \begin{cases} x & 0 \leq x < 1 \\ 2-x & 1 \leq x \leq 2 \\ 0 & \text{其它} \end{cases} \\ p(x|\omega_2) &= \begin{cases} x-1 & 1 \leq x < 2 \\ 3-x & 2 \leq x \leq 3 \\ 0 & \text{其它} \end{cases} \end{aligned}$$

先验概率 $P(\omega_1)=0.6$, $P(\omega_2)=0.4$,

(1) 求 0-1 代价 Bayes 判决函数；

(2) 求总错误概率 $P(e)$ ；

(3) 判断样本 $\{x_1=1.35, x_2=1.45, x_3=1.55, x_4=1.65\}$ 各属于哪一类别。

4.6 在目标识别中，假定有农田和装甲车两种类型，类型 ω_1 和类型 ω_2 分别代表农田和装甲车，它们的先验概率分别为 0.8 和 0.2，损失函数如表 1 所示。现在做了三次试验，获得三个样本的类概率密度如下：

$$p(x/\omega_1): 0.3, 0.1, 0.6$$

$$p(x/\omega_2): 0.7, 0.8, 0.3$$

- (1) 试用贝叶斯最小误判概率准则判决三个样本各属于哪一个类型；
- (2) 假定只考虑前两种判决，试用贝叶斯最小风险准则判决三个样本各属于哪一个类型；
- (3) 把拒绝判决考虑在内，重新考核三次试验的结果。

表 1

类型 损失 判决	ω_1	ω_2
α_1	0.5	3
α_2	5	1
α_3	2	2

4.7 已知两个一维模式类别的类概率密度函数为

$$p(x|\omega_1) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{其它} \end{cases}$$

$$p(x|\omega_2) = \begin{cases} 2-2x, & 0 \leq x \leq 1 \\ 0, & \text{其它} \end{cases}$$

先验概率 $P(\omega_1)=P(\omega_2)$ ，损失函数， $\lambda_{11}=\lambda_{22}=0$ ， $\lambda_{12}=0.6$ ， $\lambda_{21}=0.4$ 。

- (1) 求最小平均损失 Bayes 判决函数；
- (2) 求总的误判概率 $P(e)$ ；
- (3) 对于一个两类一维问题，若这两类的类概率密度分别服从正态分布 $N(0, \sigma^2)$ 和 $N(1, \sigma^2)$ ，证明使平

均决策风险最小的决策门限为

$$x_0 = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{21}P(\omega_2)}{\lambda_{12}P(\omega_1)}$$

这里，假设风险函数 $\lambda_{11}=\lambda_{22}=0$ 。一维正态分布：
$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

4.8 设 $\hat{C}(N) = \frac{1}{N} \sum_{j=1}^N (\bar{x}_j - \hat{\bar{m}}(N))(\bar{x}_j - \hat{\bar{m}}(N))^T$ 是基于样本集 $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N\}$ 对总体 $\bar{x} \sim N(\bar{m}, C)$

的协方差矩阵的最大似然估计。试推导由 $\hat{C}(N)$ 求增加一个样本 \bar{x}_{N+1} 后协方差矩阵的估计 $\hat{C}(N+1)$

的递推公式。其中， $\hat{\bar{m}}(N)$ 是基于样本集 $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N\}$ 对总体 \bar{x} 的均值向量 \bar{m} 的最大似然估计

$$\hat{\bar{m}}(N) = \frac{1}{N} \sum_{j=1}^N \bar{x}_j$$

4.9 设以下两类模式均为正态分布

$$\omega_1: \{(0, 0)^T, (2, 0)^T, (2, 2)^T, (0, 2)^T\}$$

$$\omega_2: \{(4, 4)^T, (6, 4)^T, (6, 6)^T, (4, 6)^T\}$$

(1) 设 $P(\omega_1) = P(\omega_2) = 1/2$ ，求该两类模式之间的 Bayes 判别界面的方程。

(2) 绘出判别界面。

4.10 设以下两类模式均为正态分布

$$\omega_1: \{(-5, -5)^T, (-5, -4)^T, (-4, -5)^T, (-6, -5)^T, (-5, -6)^T\}$$

$$\omega_2: \{(5, 5)^T, (5, 6)^T, (6, 5)^T, (5, 4)^T, (4, 5)^T\}$$

(1) 试用正交函数逼近法求类概率密度的估计 $\hat{p}(\bar{x}|\omega_1)$ 和 $\hat{p}(\bar{x}|\omega_2)$ ，可选用 Hermite 正交多项式前

四项低阶基函数： $H_0(x)=1$ ， $H_1(x)=2x$ ， $H_2(x)=4x^2-2$ ， $H_3(x)=8x^3-12x$ ；

(2) 设 $P(\omega_1) = P(\omega_2) = 1/2$ ，求 Bayes 判决函数；

(3) 给出判别界面方程和图示。

4.11 证明在多元问题中，贝叶斯决策准则使错误分类概率最小。

提示：使用正确分类概率来证明要方便一些。

4.12 在一个两类一维问题中，两类的概率分布密度函数分别为高斯分布 $N(0, \sigma^2)$ 和 $N(1, \sigma^2)$ ，证明

使平均风险最小的门限 x_0 为：

$$x_0 = \frac{1}{2} - \sigma^2 \ln \frac{\lambda_{21} P(\omega_2)}{\lambda_{12} P(\omega_1)}$$

其 中

$$\lambda_{11} = \lambda_{22} = 0。$$

- 4.13 假设两类分类问题中损失矩阵为 $L = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}$, ϵ_1 是将本来属于 ω_1 类的样本错分为 ω_2 的概率, ϵ_2 是将本来属于 ω_2 类的样本错分为 ω_1 的概率。试证明平均风险为

- 4.14 证明在多元分类问题中, M 类的分类错误概率上限为 $P_e = (M-1)/M$ 。

提示: 对于每一个向量 x 最大后验概率密度函数 $P(\omega_i | \bar{x})$, $i=1, 2, \dots, M$, 大于或等于 $1/M$ 。这等于每一个 $P(\omega_i | \bar{x})$ 都是相等的。

- 4.15 假设在一维两类分类当中样本点符合 Rayleigh 概率密度函数分布:

$$p(x | \omega_i) = \begin{cases} \frac{x}{\sigma_i^2} \exp\left(-\frac{x^2}{2\sigma_i^2}\right) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \text{试求判决边界 } g(x) = 0。$$

- 4.16 在两类分类问题中, 限定其中一类的错分误概率为 $\epsilon_1 = \epsilon$, 证明, 使另一类的错分概率 ϵ_2 最小等价于似然比判决: 如果 $P(\omega_1)/P(\omega_2) > \theta$, 则判 $x \in \omega_1$, 这里, θ 是使 $\epsilon_1 = \epsilon$ 成立的似然比判决门限。

注: 这就是 Neyman-Pearson 判决准则, 它类似于贝叶斯最小风险准则。

提示: 该问题等价于用 Lagrange 乘子法, 使 $q = \theta(\epsilon_1 - \epsilon) + \epsilon_2$ 最小化。

- 4.17. 二维三类问题, 假设每一类都服从同一正态分布, 且特征向量的协方差矩阵为

$$\Sigma = \begin{bmatrix} 1.2 & 0.4 \\ 0.4 & 1.8 \end{bmatrix}$$

各类的均值向量分别是 $[0.1, 0.1]^T$, $[2.1, 1.9]^T$, $[-1.5, 2.0]^T$ 。

(1) 用贝叶斯最小错误概率分类器将向量 $[1.6, 1.5]^T$ 分类。

(2) 画出距离向量 $[2.1, 1.9]^T$ 的等马氏距离曲线图 (略图)。

4. 18. 在两类三维空间分类问题中, 每一类中的特征向量都服从正态分布, 协方差矩阵为

$$\Sigma = \begin{bmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & 0.3 & -0.1 \\ 0.1 & -0.1 & 0.3 \end{bmatrix}$$

这两类的各自的均值向量分别为 $[0, 0, 0]^T$ 和 $[0.5, 0.5, 0.5]^T$ 。试推导相应的线性决策函数和决策界面方程。

4. 19. 在两类等概率分类问题中, 每一类中的特征向量的协方差矩阵均为 Σ , 相关的均值向量为 μ_1, μ_2 , 证明对于贝叶斯最小错误概率分类器, 错误概率分布是

$$P_B = \int_{(1/2)d_m}^{+\infty} \frac{1}{2\pi} \exp(-z^2/2) dz$$

其中, d_m 是这两个均值向量之间的马氏距离。该函数是 d_m 的增函数。

提示: 对数似然比 $u = \ln p(x | \omega_1) - \ln p(x | \omega_2)$ 是一个随机变量, 且服从高斯分布:

$$N\left(\frac{1}{2}d_m^2, d_m^2\right), \forall x \in \omega_1; \text{ 和 } N\left(-\frac{1}{2}d_m^2, d_m^2\right), \forall x \in \omega_2。 \text{ 据此计算错误概率。}$$

4. 20. 证明假设每个向量都遵循高斯概率密度函数分布, 在 (2. 19) 的最大似然概率检测

$$x \in \omega_1(\omega_2) \quad \text{if} \quad l_{12} = \frac{p(x | \omega_1)}{p(x | \omega_2)} > (<) \theta$$

等价于

$$d_m^2(\mu_1, x | \Sigma_1) - d_m^2(\mu_2, x | \Sigma_2) + \ln \frac{|\Sigma_1|}{|\Sigma_2|} < (>) - 2 \ln \theta$$

这里 $d_m^2(\mu_i, x | \Sigma_i)$ 是 μ_i 和 x 之间关于 Σ_i 矩阵的的马氏距离。

4.21. 如果 $\Sigma_1 = \Sigma_2 = \Sigma$, 证明上个问题成为 $(\mu_2 - \mu_1)^T \Sigma^{-1} x > (<) \Theta$, 这里

$$\Theta = \ln \theta + 1/2 \left(\|\mu_1\| \Sigma - \|\mu_2\| \Sigma \right) .$$

4.22. 在二维两类问题中, 每一类 ω_1, ω_2 都服从以下分布:

$$p(\vec{x} | \omega_1) = \frac{1}{2\pi\sigma_1^2} \exp\left(-\frac{1}{2\sigma_1^2} (\vec{x} - \vec{\mu}_1)^T (\vec{x} - \vec{\mu}_1)\right)$$

$$p(\vec{x} | \omega_2) = \frac{1}{2\pi\sigma_2^2} \exp\left(-\frac{1}{2\sigma_2^2} (\vec{x} - \vec{\mu}_2)^T (\vec{x} - \vec{\mu}_2)\right)$$

其中 $\vec{\mu}_1 = (1, 1)^T$, $\vec{\mu}_2 = (1.5, 1.5)^T$, $\sigma_1^2 = \sigma_2^2 = 0.2$ 假设 $P(\omega_1) = P(\omega_2)$, 设计一个贝叶斯分类器, 满足

(a) 错误分类概率最小

$$\Lambda = \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix}$$

(b) 具有损失矩阵 Λ 的平均风险最小

使用一个伪随机的数值产生器, 从每一个类中得到 100 个特征向量。按照上面的概率密度函数。使用这个分类器去分类已经产生的向量。对于每个事例中的错误概率是多少? 用 $\vec{\mu}_2 = (3.0, 3.0)^T$ 重复这个实验。

4.23. 重复上面的实验, 特征向量服从以下分布:

$$p(\vec{x} | \omega_i) = \frac{1}{2\pi |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\vec{x} - \vec{\mu}_i)^T \Sigma^{-1} (\vec{x} - \vec{\mu}_i)\right)$$

而且

$$\Sigma = \begin{bmatrix} 1.01 & 0.2 \\ 0.2 & 1.01 \end{bmatrix}$$

并且 $\mu_1 = [1, 1]^T$, $\mu_2 = [1.5, 1.5]^T$

提示: 一个高斯随机向量的线性变换仍然是一个高斯随机向量。注意

$$\begin{bmatrix} 1.01 & 0.2 \\ 0.2 & 1.01 \end{bmatrix} = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$$

4. 24. 二维两类问题, 假设两类服从同一正态分布, 其协方差矩阵为 $\Sigma = \begin{pmatrix} 1.1 & 0.3 \\ 0.3 & 1.9 \end{pmatrix}$, 均值向量分别为 $\vec{\mu}_1 = (0, 0)^T, \vec{\mu}_2 = (3, 3)^T$ 。试用贝叶斯分类器对向量 $\vec{x} = (1.0, 2.2)^T$ 进行分类。

4. 25. 假设在两类一维问题中 $p(x|\omega_1)$ 服从高斯分布 $N(\mu, \sigma^2)$, $p(x|\omega_2)$ 服从 a 到 b 之间的均匀分布。证明贝叶斯错误概率的上限为 $G\left(\frac{b-\mu}{\sigma}\right) - G\left(\frac{a-\mu}{\sigma}\right)$, 其中 $G(x) \equiv P(y \leq x)$, 并且 y 服从

高斯 $N(0, 1)$ 分布。

4. 26. 证明随机向量 $\frac{\partial \ln(p(x_k; \vec{\theta}))}{\partial \vec{\theta}}$ 的均值是 $\vec{0}$ 。

4. 27. 在掷硬币的游戏实验中, 正面 (1) 出现的概率是 q, 反面出现的概率是 (1-q)。设 $x_i, i=1, 2, \dots,$

N 是这个实验的结果, $x_i \in \{0, 1\}$, 证明 q 的最大似然估计是 $q_{ML} = \frac{1}{N} \sum_{i=1}^N x_i$ 提示:

似然函数是: $P(X; q) = \prod_{i=1}^N q^{x_i} (1-q)^{(1-x_i)}$

证明 ML 结果是下列方程的解 $q^{\sum_i x_i} (1-q)^{(N-\sum_i x_i)} \left(\frac{\sum_i x_i}{q} - \frac{N - \sum_i x_i}{1-q} \right) = 0$

4. 28. 随机变量 x 服从高斯 $N(\mu, \sigma^2)$ 分布, μ 未知。给定该变量的 N 个观测值, 设 $L(\mu)$ 为 μ 的对数

似然函数: $L(\mu) = \ln(p(x; \mu))$ 。试求该随机变量的 Cramer-Rao 界: $-E \left[\frac{\partial^2 L(\mu)}{\partial^2 \mu} \right]$ 。将该结果与 μ 的 ML 估计值的方差进行比较, 有何结论? 假如这个未知参数是方差 σ^2 , 结论又如何?。

4. 29. 证明假如似然函数是高斯函数有未知的均值 μ , 和协方差矩阵 Σ , 然后 ML 估计如下给出

$$\tilde{\mu} = \frac{1}{N} \sum_{k=1}^N x_k$$

$$\tilde{\Sigma} = \frac{1}{N} \sum_{k=1}^N \left(x_k - \tilde{\mu} \right) \left(x_k - \tilde{\mu} \right)^T$$

4.30. 随机变量 x 服从 Erlang 分布, 概率密度函数为

$$p(x; \theta) = \theta^2 x \exp(-\theta x) u(x)$$

其中 $u(x)$ 是一个阶跃函数

$$u(x) = \begin{cases} 1 & , \quad x > 0 \\ 0 & , \quad x < 0 \end{cases}$$

假设 x 的 N 个观测值 x_1, \dots, x_N , 证明 θ 的最大似然估计为

$$\hat{\theta}_{ML} = \frac{2N}{\sum_{k=1}^N x_k}$$

4.31. 随机变量 x 是服从正态分布 $N(\mu, \sigma^2)$, 其中未知参数 μ 服从 Rayleigh 分布, 其概率密度函数为

$$p(\mu) = \frac{\mu \exp(-\mu^2 / 2\sigma_\mu^2)}{\sigma_\mu^2}$$

试证明 μ 的最大后验概率估计为

$$\hat{\mu}_{MAP} = \frac{Z}{2R} (1 + \sqrt{1 - 4R / Z^2})$$

$$Z = \frac{1}{\sigma^2} \sum_{k=1}^N x_k, \quad R = \frac{N}{\sigma^2} + \frac{1}{\sigma_\mu^2}$$

其中,

4.32. 证明对于对数正态分布

$$p(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left(-\frac{(\ln x - \theta)^2}{2\sigma^2} \right), x > 0$$

最大似然估计为

$$\hat{\theta}_{ML} = \frac{1}{N} \sum_{k=1}^N \ln x_k$$

4.33. 若已知一个随机变量 x 的均值和方差:

$$\mu = \int_{-\infty}^{+\infty} xp(x)dx,$$

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 p(x)dx$$

试证明, 该随机变量概率密度函数的最大熵估计服从高斯分布 $N(\mu, \sigma^2)$

4.34. P 为一个随机点 x 位于某区间 h 的概率。给定 x 的 N 个观测值, 其中有 k 个落入区间 h 的概率服从二项式分布:

$$\text{prob}\{k\} = \frac{N!}{k!(N-k)!} P^k (1-P)^{N-k}$$

证明 $E[k/N] = P$, 并且它的方差为 $\sigma^2 = E[(k/N - P)^2] = P(1-P)/N$

而且这个概率估计 $P=k/N$ 是无偏的和渐进一致的。

第五章 特征提取与选择

5.1 设有 M 类模式 ω_i , $i=1, 2, \dots, M$, 试证明总体散布矩阵 S_t 是总类内散布矩阵 S_w 与类间散布矩阵 S_b 之和, 即 $S_t = S_w + S_b$ 。

5.2 下面哪个矩阵可以用在二维空间线性变换中, 并保持马氏距离的特性? 请解释原因。

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & -1 \end{pmatrix}$$

5.3 Bhattacharyya 可分性判据定义为

$$J_B = -\ln \int_{\Omega} [p(\bar{x} | \omega_1) p(\bar{x} | \omega_2)]^{1/2} d\bar{x}$$

式中 Ω 表示特征空间。试证明，在最小误判概率准则判决下，最小最小误判概率有

$$P(e) \leq [P(\omega_1)P(\omega_2)]^{1/2} \exp(-J_B)。$$

5.4 令 x_i , $i=1, 2, 3$ 为独立的二值特征，且 $p(x_i=1 | \omega_1) = \alpha_i$, $p(x_i=1 | \omega_2) = \beta_i$, 两类的先验概率相等，且 α_i, β_i 满足以下条件：

$$(1) \alpha_i < \beta_i, \forall i, (2) \beta_1 - \alpha_1 > \beta_2 - \alpha_2 > \beta_3 - \alpha_3。$$

试证明各特征分别使用时之错误概率 $e(x_i)$ 满足： $e(x_1) < e(x_2) < e(x_3)$ 。

5.5 按上题条件，试证明当两个特征合用时，其错误概率为

$$e(x_i, x_j) = \frac{1}{2} [e(x_i) + e(x_j) - (\beta_i - \alpha_i) | e(x_j) - \alpha_j | - (\beta_j - \alpha_j) | e(x_i) - \alpha_i |]$$

请找出使 $e(x_1, x_2) < e(x_2, x_3)$ 之条件。

5.6 同上题，如果给定

$$\alpha_1 = 0.10, \alpha_2 = 0.05, \alpha_3 = 0.01, \\ \beta_1 = 0.90, \beta_2 = 0.80, \beta_3 = 0.70$$

试计算 $e(x_1), e(x_2), e(x_3), e(x_1, x_2), e(x_1, x_3), e(x_2, x_3)$ 。

5.7 已知以下两类模式

$$\omega_1: \{(0, 0, 0)^T, (1, 0, 0)^T, (1, 0, 1)^T, (1, 1, 0)^T\}$$

$$\omega_2: \{(0, 0, 1)^T, (0, 1, 0)^T, (0, 1, 1)^T, (1, 1, 1)^T\}$$

试用 K-L 变换分别把特征空间维数降到 $d=2$ 和 $d=1$ ，并作图画出样本在该特征空间中的位置。

5.8 令 Σ_i 和 P_i 分别是 ω_i 类 ($i=1, 2$) 的协方差矩阵和先验概率。假设对数据进行了白化变换，即，使

$$B^T S_w B = I。这里，S_w = \sum_i P_i \Sigma_i, I \text{ 是单位矩阵。试证明矩阵 } P_1 B^T \Sigma_1 B \text{ 和 } P_2 B^T \Sigma_2 B \text{ 所产生的 K-L}$$

坐标轴是相同的。若用 Λ_i 表示矩阵 $P_i B^T \Sigma_i B$ 的本征值矩阵，求证 $\Lambda_1 = I - \Lambda_2$ 。

8. 1 三类 $\omega_1, \omega_2, \omega_3$, 其中 $\omega_1: \{(1,0)^T, (2,0)^T, (1,1)^T\}$, $\omega_2: \{(-1,0)^T, (-1,1)^T, (0,1)^T\}$, 求 S_w 及 S_b 。

8. 3 令 $p(x|\omega_i) \sim N(\mu_i, \Sigma_i)$, $i=1,2$, 假定各特征分量 $x_j, j=1,2,\dots,D$, 相互独立, 试证按式

$$J_D = I_{ij} + I_{ji} = \int_X [p(x|\omega_i) - p(x|\omega_j)] \ln \frac{p(x|\omega_i)}{p(x|\omega_j)} dx$$

定义的散度 J_D 可写为 $J_D = \sum_{j=1}^D J_{Dj}$

8. 6 两个一维正态分布, 其期望与方差如下:

第一组 $\mu_1 = 0, \mu_2 = 2, \sigma_1^2 = 4, \sigma_2^2 = 0.25$; 第二组 $\mu_1 = 0, \mu_2 = 2, \sigma_1^2 = 1, \sigma_2^2 = 1$ 。

求 Bhattacharyya 距离及散度。

8. 10 令 $x_i, i=1,2,3$ 为独立的二值特征, 且 $p(x_i=1|\omega_1) = \alpha_i, p(x_i=1|\omega_2) = \beta_i$, 二类先验概率相等, 且 α_i, β_i 满足以下条件:

$$\textcircled{1} \alpha_i < \beta_i, \forall i, \textcircled{2} \beta_1 - \alpha_1 > \beta_2 - \alpha_2 > \beta_3 - \alpha_3。$$

试证各特征分别使用时之错误概率 $e(x_i)$ 满足: $e(x_1) < e(x_2) < e(x_3)$ 。

8. 11 令 $x_i, i=1,2,3$ 为独立的二值特征, 且 $p(x_i=1|\omega_1) = \alpha_i, p(x_i=1|\omega_2) = \beta_i$, 二类先验概率相等, 且 α_i, β_i 满足以下条件:

$$\textcircled{1} \alpha_i < \beta_i, \forall i, \textcircled{2} \beta_1 - \alpha_1 > \beta_2 - \alpha_2 > \beta_3 - \alpha_3。$$

试证当两个特征合用时其错误概率为

$$e(x_i, x_j) = \frac{1}{2} [e(x_i) + e(x_j) - (\beta_i - \alpha_i) | e(x_j) - \alpha_j | - (\beta_j - \alpha_j) | e(x_i) - \alpha_i |]$$

找出使 $e(x_1, x_2) < e(x_2, x_3)$ 之条件。

8. 12 令 $x_i, i=1,2,3$ 为独立的二值特征, 且 $p(x_i=1|\omega_1) = \alpha_i, p(x_i=1|\omega_2) = \beta_i$, 二类先验概率相等, 且 α_i, β_i 满足以下条件:

$$\textcircled{1} \alpha_i < \beta_i, \forall i, \textcircled{2} \beta_1 - \alpha_1 > \beta_2 - \alpha_2 > \beta_3 - \alpha_3。$$

当 两 个 特 征 合 用 时 其 错 误 概 率 为

$$e(x_i, x_j) = \frac{1}{2} [e(x_i) + e(x_j) - (\beta_i - \alpha_i) | e(x_j) - \alpha_j | - (\beta_j - \alpha_j) | e(x_i) - \alpha_i |]$$
, 如 果 给 定
 $\alpha_1 = 0.10, \alpha_2 = 0.05, \alpha_3 = 0.01, \beta_1 = 0.90, \beta_2 = 0.80, \beta_3 = 0.70$
。 试 计 算
 $e(x_1), e(x_2), e(x_3); e(x_1, x_2), e(x_1, x_3), e(x_2, x_3)$ 。

第六章 句法模式识别

6.1 句法模式识别中，模式类是如何描述的。

6.2 文法 $G=(V_N, V_T, P, S)$ 中起始符 S 可以是非终止符吗？

6.3 试给出能产生不含子串 011 的 0, 1 字符串的全体的正则文法 G_n ，画出 G_n 对应的有限状态自动机的状态转移图。设 $x=10010$ ，试根据 CYK 算法导出对 x 的最左剖析。

6.4 已知文法 $G=(V_N, V_T, P, S)$ ， $V_N=\{S, \Delta\}$ ， $V_T=\{a, d, (,), +, \times, -, *, \sim\}$ ，其中， a, d 为图形描述语言 (PDL) 的基元： $a: \rightarrow$ ； $d: \downarrow$ ；“ $+$ ， \times ， $-$ ， $*$ ， \sim ”为基元链接运算：“ $+$ ”为头接尾；“ \times ”为尾尾相接；“ $-$ ”为头头相接；“ $*$ ”为头接和尾接；“ \sim ”为反向；“ $(,)$ ”为运算顺序控制符。试写出描述字母“E”的 PDL 句子及相应的产生式 P 。

第三部分 综合分析题

1、在描述一个完整的模式识别系统的基础上，给出你所了解的一种模式识别应用实例，并对其中可能用到的模式识别典型理论和方法进行简要说明。

2、举出日常生活或技术、学术领域中应用模式识别理论解决问题的实例（包括问题模型，解决的方法，体会）

3、若要实现手机上的语音拨号功能，你认为应该有哪些处理步骤？分别需要哪些模式识别方法？试用流程图予以说明。

4、若要实现机器（计算机、手机等）的手写字符输入功能，你认为应该有哪些处理步骤？分别需要哪些模式识别方法？试用流程图予以说明。

5、若要实现计算机指纹识别功能，你认为应该有哪些处理步骤？分别需要哪些模式识别方法？试用流程图予以说明。

6、若要实现汽车车牌自动识别，你认为应该有哪些处理步骤？分别需要哪些模式识别方法？试用流程图予以说明。



妈妈新开了个淘宝店，欢迎前来捧场

妈妈的淘宝点开了快半年了，主要卖的是毛绒玩具、坐垫、抱枕之类的，但生意一直不是很好，感觉妈妈还是很用心的，花了不少功夫，但是就是没有人气，所以我也来出自己的一份力，帮忙宣传一下。

并且妈妈总是去五亭龙挑最好的玩具整理、发货，质量绝对有保证。

另外我家就在扬州五亭龙玩具城旁边，货源丰富，质量可靠，价格便宜。

欢迎大家来逛逛【扬州五亭龙玩具总动员】 99toy.taobao.com