东南大学信息科学与工程学院研究生试题纸

课 程 ____模式识别____ 2011～2012 学年第一学期

（A）卷 （闭√ 开）卷 考试时间 2011 年 12 月 14 日 14:00-16:00

班级学号 ____22015088o____ 姓名 ____丁杰斌____ 得分 ____[redacted]____。

一、 填空题（20 分，每空 2 分）

1．模式识别系统的基本构成有：数据获取，____预处理____，____特征选择/提取____，____分类规则训练____，分类决策。

2．符合多元正态密度函数分布的样本，其等密度点的轨迹为超椭圆，椭圆的主轴与____C的特征向量方向____一致，主轴的长度由____C的特征值大小____所决定。

3．神经网络按照结构可以分为____~~前馈~~分层结构____，____相互连接结构____两种。

4．贝叶斯判别准则一般可分为：____最小错误率____，____最小误差____。

5．K-L 变换是在____最小均方误差____的意义下获得数据压缩的最佳变换。

解：① 欧氏距离 $D(x_i,x_j)=\|x_i-x_j\|=\sqrt{(x_i-x_j)^T(x_i-x_j)}$ ② 马氏距离 $D^2=(X-M)^T C^{-1}(X-M)$ ③ 明氏距离 $D_m(x_i,x_j)=\left[\sum_{k=1}^{n}|x_{ik}-x_{jk}|^m\right]^{\frac{1}{m}}$

二、 简答题（共 40 分，每题 8 分） ④ 汉明距离 $D_H=\frac{1}{2}(n-\sum_{k=1}^{n}x_{ik}\cdot x_{jk})$

解：

$(h|\omega_1)\sim N(\frac{1}{2}r_{ij}^2, r_{ij}^2)$

$h|\omega_2)\sim N(-\frac{1}{2}r_{ij}^2, r_{ij}^2)$

$(e)=\int_{-\infty}^{t} p(h|\omega_1)dh$

$(e)=\int_{t}^{+\infty} p(h|\omega_2)dh$

~~最小~~贝叶斯分类的错误概率随着 r_{ij} 的增大而减少

1．模式相似性的测度有哪些？写出其定义公式。 ⑤ $S(x_i,x_j)=\frac{x_i\cdot x_j}{\|x_i\|\|x_j\|}$

2．简述利用 K-L 变换进行特征提取的基本过程。

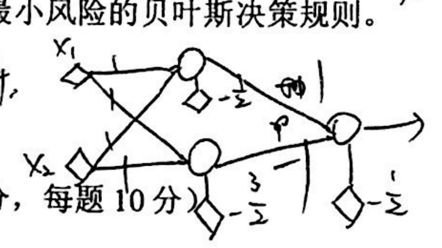
2 解：① 求出样本的自相关矩阵 $R=E[XX^T]$ ② 求矩阵 R 的特征值，按从大到小排列 ③ 选取前 m 个特征值，求出相应的特征向量并归一化处理，得到矩阵 $U=[u_1,u_2\cdots u_m]$ ④ 计算 $X^*=U^TX$ 得到 K-L 变换后的样本

3．简述贝叶斯分类器的错误概率与模式样本的马氏距离之间的关系。

4．画出解决 XOR 分类问题的多层感知器结构图。

5．简述最小错误率与最小风险的贝叶斯决策规则。

4 解，



5，解：最小错误率贝叶斯规则

~~P(x)~~ $=\max\{P(\omega_j|X)\}$ $P(\omega_i|X)$ $j=1,2,\cdots n$

最小风险 $r_i=\sum_{j=1}^{m}L_{ij}P(x|\omega_j)P(\omega_j)$ 取最小 $i=1,2,\cdots n$

三、 计算题（共 40 分，每题 10 分）

1．贝叶斯分类。设在一维特征空间中两类样本服从正态分布，$\sigma_1=\sigma_2=1$，$\mu_1=1,\mu_2=-1$，两类先验概率之比 $P(\omega_1)/P(\omega_2)=e$，试求按基于最小错误率贝叶斯决策原则的分界面。

$l_{12}=\frac{P(x|\omega_1)}{P(x|\omega_2)}$

$\therefore l_{12}=\exp\frac{-(x-1)^2+(x+1)^2}{2}=e$

$P(x|\omega_1)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\frac{-(x-\mu_1)^2}{2\sigma_1^2}$

$\therefore -1\cdot 2x=1$

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

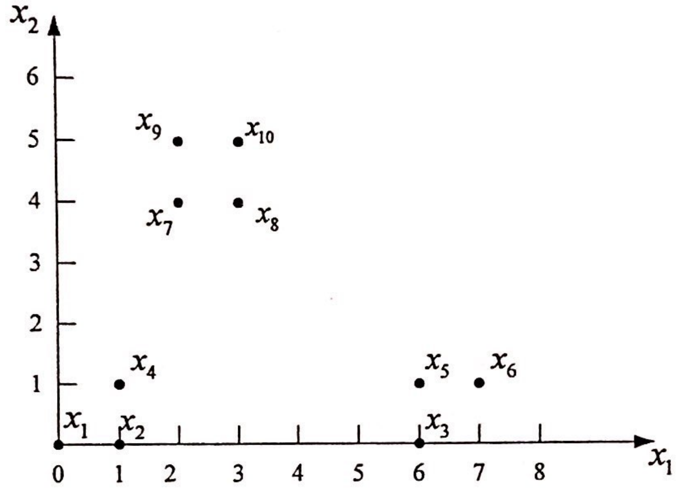
$z_1 = x_1 = (0,0)$ 分别计算各点到 z_1, z_2, z_3 的距离，距离 计算新的 返回 由新的中心值

$z_2 = x_2 = (1,0)$ x_1：D_1 D_2 D_3 $z_1(2)$ 不同 $S_1(2)$ 计算

$z_3 = x_3 = (6,0)$ 得 $S_1(1) = \{ \cdots$ $z_2(2)$ $S_2(2)$

$S_2(1)$ $S_3(1)\}$ $z_3(2)$ $S_3(2)$

2．样本分布如下图所示(样本取值均为整数点)，试利用 K 均值算法进行模式分类，取 K=3。



3．设有一个三类问题，其判别式为 $d_1(X) = x_1 + 2x_2 - 4$，$d_2(X) = x_1 - 4x_2 + 4$，$d_3(X) = -x_1 + 3$。用 $\omega_i / \overline{\omega_i}$ 两分法对模式空间的所有区域进行分类处理（即说明该区域属于哪个类别，或可能属于哪些类别，或不属于任何类别）。现有一模式 $X = [5,2]^{\mathrm{T}}$，用 $\omega_i / \overline{\omega_i}$ 两分法判定该模式属于哪一个类别。

$d_3(X) = -2$

不确定区 $d_1(X) = 5+4-4=5$ 不属于

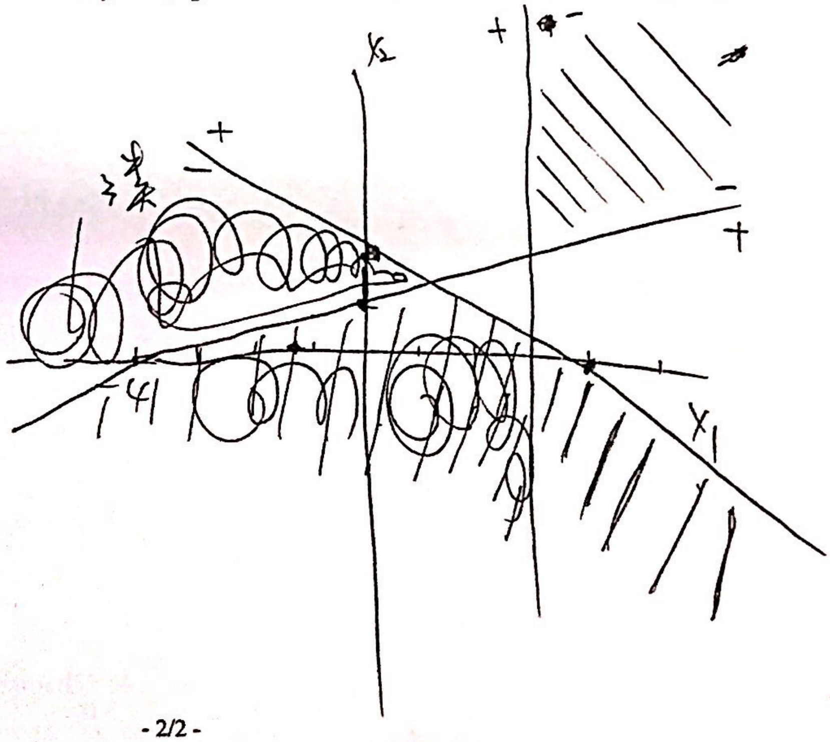
$d_2(X) = 5-8+4=1$ Ⅲ一类

4．贝叶斯分类。设在一维特征空间中两类样本服从正态分布，$\sigma_1 = \sigma_2 = 1$，$\mu_1 = 1, \mu_2 = -1$，两类先验概率之比 $P(\omega_1)/P(\omega_2) = e$，试求按基于最小错误率贝叶斯决策原则的分界面。

似然比阈值 e

似然比 $\dfrac{p(x|\omega_1)}{p(x|\omega_2)} = \dfrac{\frac{1}{\sqrt{2\pi}\sigma_1} \exp\frac{-(x-\mu_1)^2}{2\sigma_1^2}}{\frac{1}{\sqrt{2\pi}\sigma_2} \exp\frac{-(x-\mu_2)^2}{2\sigma_2^2}}$

$= \dfrac{\exp\frac{-(x-1)^2}{2}}{\exp\frac{-(x+1)^2}{2}}$



$\therefore \exp\frac{-(x-1)^2}{2} = \exp\frac{-(x+1)^2}{2} + 1$

$\cancel{-x^2}+2x\cancel{-1} = \cancel{-x^2}-2x\cancel{-1}+2$

$x = \frac{1}{2}$ 为分界

东南大学信息科学与工程学院研究生试题纸

课 程 模式识别 2012～2013 学年第一学期

（A）卷 （闭√ 开）卷 考试时间

班级学号 ______ 姓名 ______ 得分______。

一、 填空题（20 分，每空 2 分）

1．模式识别系统的基本构成有：数据获取，预处理，特征选择/提取，分类规则训练，分类决策。

2．贝叶斯判别准则一般可分为：最小错误率，最小风险。

3．模式识别按照实现方法可以分为监督分类和非监督分类。

4． K-L 变换是在最小均方误差的意义下获得数据压缩的最佳变换。

5．神经网络按照结构可以分为分层结构，相互连接结构两种。

欧氏距离 $D(x_i,x_j)=\|x_i-x_j\|=\sqrt{(x_i-x_j)^T(x_i-x_j)}$

3. 明氏距离 $D_m(x_i,x_j)=\left[\sum_{k=1}^{n}|x_{ik}-x_{jk}|^m\right]^{\frac{1}{m}}$

马氏距离 $D^2=(x-\mu)^T C^{-1}(x-\mu)$

4. 汉明距离 $D_h(x_i,x_j)=\frac{1}{2}\left(n-\sum_{k=1}^{n}x_{ik}\cdot x_{jk}\right)$

二、 简答题（共 40 分，每题 8 分）

1．模式相似性的测度有哪些？写出其定义公式（至少 4 个）。 5. 角度相似性函数 $S(x_i,x_j)=\frac{x_i^Tx_j}{\|x_i\|\|x_j\|}$

$p(h|\omega_1)\sim N(\frac{1}{2}r_{12}^2, r_{12}^2)$

2．简述利用 K-L 变换进行特征提取的基本过程。

① 计算两类样本总体自相关矩阵 $R=E[x^*x^T]$

② 求特征值 $|R-\lambda I|=0$

③ 选取最大的前n个特征值

④ 计算对应的特征向量

⑤ 将特征向量归一化 得变换矩阵 V

⑥ $X^*=V^TX$ 进行K-L变换

$p(h|\omega_2)\sim N(-\frac{1}{2}r_{12}^2, r_{12}^2)$

$P(e)=P(\omega_1)P_1(e)+P(\omega_2)P_2(e)$

3．简述感知器算法的步骤。

$P_1(e)=\int_{-\infty}^{t}p(h|\omega_1)dh$

P(e) 随着 r_{12}^2 增大而单调递减

4．简述贝叶斯分类器的错误概率与模式样本的马氏距离之间的关系。

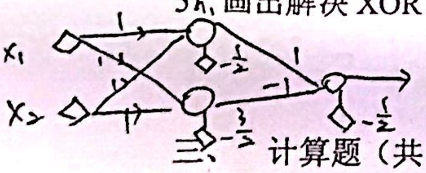
将训练样本写成增广向量模式并规范化

① 选取 $w_1=(0,0,0)^T$ 和 $c=1$

② $w(k+1)=\begin{cases}w(k) & w(k)^Tx>0\\ w(k)+cx & w(k)^Tx\le 0\end{cases}$ 对全部样本迭代

③ 一轮迭代中只要有一个样本分类出错，重新迭代，直到每个样本正确，算法收敛

5．画出解决 XOR 分类问题的多层感知器结构图。



三、 计算题（共 40 分，每题 10 分）

1．假设在某个地区的疾病普查中，正常细胞（ω_1）和异常细胞（ω_2）的先验概率分别为 $P(\omega_1)=0.9$，$P(\omega_2)=0.1$。现有一待识别细胞，其观察值为 X，从类概率密度分布曲线上查得 $p(X|\omega_1)=0.2$，$p(X|\omega_2)=0.4$。

a. 试对该细胞利用最小错误率贝叶斯决策规则进行分类。

b. 若损失函数的值分别为 $L_{11}=0$，$L_{12}=6$，$L_{21}=1$，$L_{22}=0$。试用最小风险贝叶斯决策规则对细胞进行分类。

$\frac{p(x|\omega_1)}{p(x|\omega_2)}=\frac{1}{2}$

$\frac{P(\omega_2)}{P(\omega_1)}=\frac{1}{9}$

$\frac{1}{2}>\frac{1}{9}$ ∴属于 ω_1，正常

$\theta=\frac{L_{12}-L_{22}}{L_{21}-L_{11}}\cdot\frac{P(\omega_2)}{P(\omega_1)}=\frac{6-0}{1-0}\cdot\frac{1}{9}=\frac{2}{3}$

$\frac{1}{2}<\frac{2}{3}$ ∴异常

2．样本分布如下图所示(样本取直均为整数点)，试利用K均值算法进行模式分类，取K=3。

$S_W = \frac{1}{2}(S_1+S_2) = \begin{bmatrix}1&0\\0&1\end{bmatrix}$

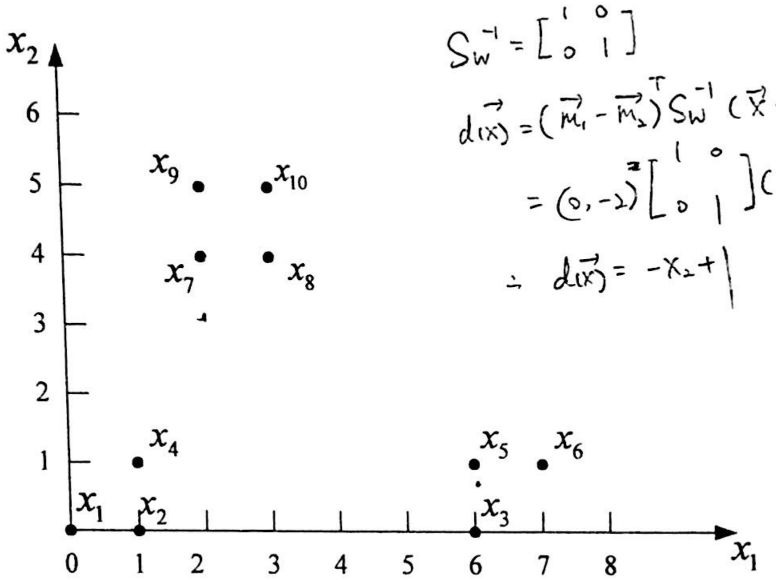
$S_W^{-1} = \begin{bmatrix}1&0\\0&1\end{bmatrix}$

$d(\vec{x}) = (\vec{m}_1-\vec{m}_2)^T S_W^{-1}(\vec{X}-\frac{\vec{m}_1+\vec{m}_2}{2})$

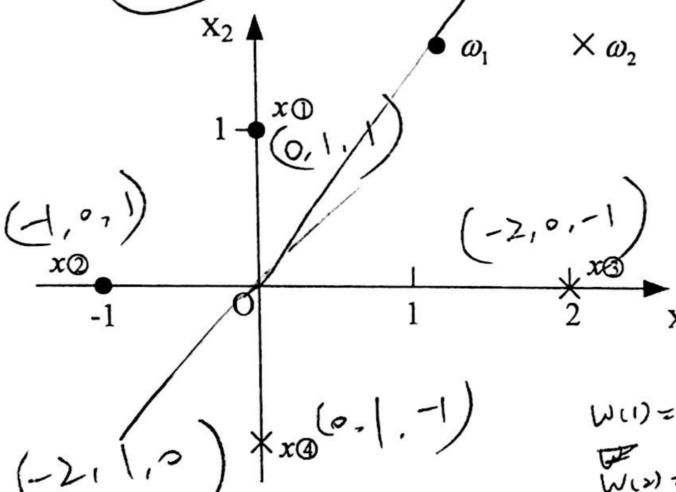
$= (0,-2)\begin{bmatrix}1&0\\0&1\end{bmatrix}(x_1-2, x_2-1)^T = -2x_2+2$

$\therefore d(\vec{x}) = -x_2+1$ (决策面方程)

$\therefore d(\vec{x}) \gtrless 0 \Rightarrow \vec{X}\in \begin{cases}\omega_1\\ \omega_2\end{cases}$



3．利用感知器算法求下列模式分类的解向量。取校正增量C=1, w(1)=0。



ω_1: $X_1=(0,1)^T$, $X_2=(-1,0)^T$ ω_2: $X_3=(2,0)^T$, $X_4=(0,-1)^T$

将向量写成增广向量形式并规范化

得 $X_1=(0,1,1)^T$ $X_2=(-1,0,1)^T$

$X_3=(-2,0,-1)^T$ $X_4=(0,-1,-1)^T$

取 $C=1$, $w(1)=0$ 分别迭代

$w(1)=0$ $w(1)X_1 = 0 < 0$

$w(2) = w(1)+X_1 = (0,1,1)^T$

$w(2)^T X_2 = 1 > 0$ ✓

$w(3) = (0,1,1)^T$

$w(3)^T X_3 = -1 < 0$

$w(4) = (-2,1,0)^T$

$w(4)^T X_4 = 1 > 0$ ✓

$w(5)X_1 = 1 > 0$ ✓

$w(6)X_2 = 2 > 0$ ✓

$w(7)X_3 = 4 > 0$ ✓

$w(8) = 1 > 0$ ✓

$\therefore$ 判决函数

$-2x_1+x_2=0$

$x_2=2x_1$

4． 设两类样本的类内离散矩阵分别为

$$S_1=\begin{pmatrix}1&1/2\\1/2&1\end{pmatrix},\quad S_2=\begin{pmatrix}1&-1/2\\-1/2&1\end{pmatrix}$$

$m_1=(2,0)^t$, $m_2=(2,2)^t$

试用 fisher 准则求其决策面方程的法线方向。

$S_W = \frac{1}{2}(S_1+S_2) = \begin{bmatrix}1&0\\0&1\end{bmatrix}$

$S_W^{-1} = \begin{bmatrix}1&0\\0&1\end{bmatrix}$

$d(\vec{x}) = (\vec{m}_1-\vec{m}_2)^T S_W^{-1}(\vec{X}-\frac{\vec{m}_1+\vec{m}_2}{2})$

$= (0,-2)\begin{pmatrix}1&0\\0&1\end{pmatrix}(x_1-2, x_2-1)^T = -2x_2$

$d(\vec{x}) = -x_2+1$