

中国科学院自动化研究所
2011 年招收攻读博士学位研究生入学统一考试试题
科目名称：模式识别

考生须知：

1. 本试卷满分为 100 分，全部考试时间总计 180 分钟。
 2. 所有答案必须写在答题纸上，写在试题纸上或草稿纸上一律无效。
-

1. (10 分) 简述模式的概念和它的直观特性，同时绘出模式识别系统的组成框图并说明各部分的主要功能特性。

2. (12 分) 试阐述线性判别函数的基本概念，并说明 Fisher 线性判别的基本原理。

3. (12 分) 对如下 10 个模式样本用 K 均值算法进行分类

$$\{x_1(0\ 0), x_2(1\ 0), x_3(0\ 1), x_4(1\ 1), x_5(8\ 7), \\ x_6(9\ 7), x_7(8\ 8), x_8(9\ 8), x_9(8\ 9), x_{10}(9\ 9)\}$$

其中取 $K=2$ ，初始聚类中心选为 x_1 和 x_2

4. (18 分) 假定对一类特定人群进行某种疾病检查，正常人以 ω_1 类代表，患病者以 ω_2 类代表。设被检查的人中正常者和患病者的先验概率分别为

$$\text{正常人: } P(\omega_1)=0.9$$

$$\text{患病者: } P(\omega_2)=0.1$$

现有一被检查者，其观察值为 x ，从类条件概率密度分布曲线上查得

$$P(x|\omega_1)=0.2, P(x|\omega_2)=0.4$$

(续下页)

同时已知风险损失函数为

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 0 & 6 \\ 1 & 0 \end{pmatrix}$$

其中 λ_{ij} 表示将本应属于第 j 类的模式判为属于第 i 类所带来的风险损失。试对该被检查者用以下两种方法进行分类：

- (1) 基于最小错误率的贝叶斯决策，并写出其判别函数和决策面方程；
- (2) 基于最小风险的贝叶斯决策，并写出其判别函数和决策面方程。

5. (16 分) 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 $\mu_2 = [-1, -1, 1]^T$ ，协方差矩阵如下

$$\Sigma_1 = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

求用 $J_s = \frac{|S_w + S_b|}{|S_w|}$ 判据的最优特征选择。

6. (16 分) 给出两类模式样本集

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

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

设其先验概率相等，即

$$P(\omega_1) = P(\omega_2) = 0.5$$

试用 K-L 变换把特征空间的维数降到一维，并用图画出样本在特征空间中的位置。

7. (16 分) 什么是人工神经网络？其主要特点有哪些？试给出误差反传训练算法（即 BP 训练算法）的原理并给出其训练步骤（转移函数取 Sigmoid 函数）。