

考试注意事项：

1. 考试形式：开卷网络笔试。
2. 分 AB 卷，学号尾数为奇数的同学，答 A 卷，其他同学答 B 卷。
3. 直接在自备的 A4 答题纸上答题。
4. 在每张答题纸的上方区域写清楚自己的姓名和学号。
5. 答题时，写清题号（注：不必在答题纸上重新抄写题目）。
6. 答题时请保持字迹清楚，卷面整洁。
7. 完成考试后将答题纸扫描或拍照后打包发送到以下邮箱。
  - （1）A 卷：wzfu@ustc.edu.cn, cjwbdw6@mail.ustc.edu.cn
  - （2）B 卷：wzfu@ustc.edu.cn, xlh1995@mail.ustc.edu.cn
8. 考试结束后请销毁试卷。

# 中国科学技术大学

## 2019-2020 学年第二学期研究生考试试卷-A 卷

考试科目：模式识别

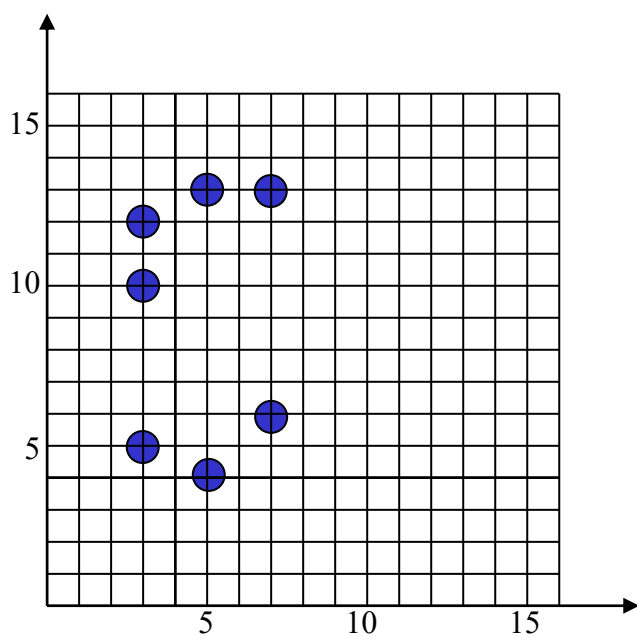
学生所在系：\_\_\_\_\_ 姓名：\_\_\_\_\_ 学号：\_\_\_\_\_

1. 是非判断题（每小题 2 分，共 30 分）

- (1) 多元正态分布概率密度函数等密度点的轨迹为超椭球面。( )
- (2) 设  $A=\{1, 4, 7\}$ ，则  $R=\{(a, b) \mid a, b \in A \wedge a \equiv b \pmod{3}\}$  为  $A$  上的等价关系。( )
- (3) 设  $a, b, c$  属于字母表，则句子  $abcb$  的所有长度为 2 的子串是： $ab, ac, bb, bc, cb$ 。( )
- (4) 两点间的欧氏距离不是平移变换下的不变量。( )
- (5) 设  $a, b, c$  属于字母表，则根据定义，串  $\lambda ab \lambda$  的长度  $|\lambda ab \lambda|=4$ 。( )
- (6) 设  $T$  是由一个非确定的有限状态自动机接受的链集，则必存在一个能接受  $T$  的确定的有限状态自动机。( )
- (7) 设  $T$  是由一个非确定的下推自动机接受的链集，则必存在一个能接受  $T$  的确定的下推自动机。( )
- (8) 对一个二分类问题而言，即使给定的训练样本集是线性可分的，势函数法也未必总是给出正确的区分超平面。( )
- (9) 设  $T$  是由一个非确定的图灵机接受的链集，则必存在一个能接受  $T$  的确定的图灵机。( )
- (10) 在一维特征空间中，用于区分两个类别的分界面退化为一个点。( )
- (11) 当一个语言由有限状态文法生成时，它能被一个有限状态自动机所识别。( )
- (12) 多元正态分布的条件分布和边缘分布仍然是正态分布。( )
- (13) 设有  $N$  个可用样本，则使用留一法设计分类器时用于检验的样本的个数为 1。( )
- (14) 最小风险判决与模式发生的后验概率无关。( )
- (15) 设  $X$  是  $d$  维的正态随机向量， $C$  是与  $X$  同维的常向量，则  $C^T X$  是一个 1 维的正态随机变量。( )

2. 设  $\mathbf{X}$  和  $\boldsymbol{\mu}$  是  $d$  维模式向量,  $\boldsymbol{\Sigma}$  为  $d \times d$  维可逆对称矩阵, 证明:  $\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}$ 。  
(共 10 分)

3. 设有如下图所示的样本集, 选择欧氏距离作为距离测度, 运用近邻函数值准则聚类算法完成聚类分析, 给出聚类结果。  
(共 15 分)



4. 设有文法  $G = (N, \Sigma, P, S)$ 。其中:  
 $N = \{S, A, B, C, D\}$ ,  $\Sigma = \{a, b\}$ , 以及  
 $P$ : (1)  $S \rightarrow B$ , (2)  $S \rightarrow BC$ , (3)  $A \rightarrow D$ , (4)  $D \rightarrow a$ , (5)  $B \rightarrow b$ , (6)  $C \rightarrow SA$ 。  
 用 CYK 算法对输入链  $bbbba$  进行句法分析以判断其是否属于  $L(G)$ 。  
 (共 15 分)

5. 设  $G = (N, \Sigma, P, S)$ 。其中:  
 $N = \{S\}$ ,  $\Sigma = \{a, b, c\}$ , 以及  
 $P$ : (1)  $S \rightarrow aSa$ , (2)  $S \rightarrow bSb$ , (3)  $S \rightarrow c$ 。  
 求  $L(G)$ 。  
 (共 10 分)

6. 设有如下分属于两个类别 $\omega_1$ 和 $\omega_2$ 的样本集合：

$$\omega_1: (1,1)^T, (-1,1)^T; \omega_2: (-1,-1)^T, (1,-1)^T$$

问：(1) 该样本集合是否是线性可分的？

(2) 如果回答是线性可分的，用感知器算法求解分类器。否则，说明不是线性可分的理由。

(共 20 分)