

# 2023-2024 学年秋季学期《模式识别与机器学习》课程

## 第 10 章作业题

### 一、选择题

1. (多选题) 以下关于聚类的说法, 正确的是 ( )
  - A. 聚类是一种无监督学习方法
  - B. EM 算法的 M 步可由极大似然估计推导得到
  - C. 层次聚类只能使用自底向上的聚合策略
  - D. K 均值聚类算法的时间复杂度与聚类样本的数量成线性关系
2. (多选题) 关于数据聚类方法, 以下说法正确的是 ( )
  - A. K 均值聚类、高斯混合聚类和层次化聚类方法都需要预先给定类簇个数
  - B. K 均值聚类易受初始均值向量选取的影响
  - C. 设  $S_w$  是类内离散度矩阵,  $S_B$  是类间离散度矩阵, 则聚类目标可以是最大化  $|S_w S_B^{-1}|$
  - D. 高斯混合聚类问题可以使用 EM 算法来求解

### 二、计算题

1. (层次聚类) 现有 7 个一维数据点  $\{-7, -2, -0.5, 0.5, 2.5, 3, 3.5\}$ , 采用自底向上的层次聚类方法对其进行聚类, 其中类簇  $D_i$  和  $D_j$  之间的距离定义为下式, 并画出聚类树。

$$d_{\min}(D_i, D_j) = \min_{\substack{x \in D_i \\ x' \in D_j}} \|x - x'\|_2$$