机器学习考试范围

**4-5个简答/简述题目，计算题3-4个, 不会太难，总共9-10个大题。（最新）**

**黄色部分为必考计算题**

**红色字体部分是最后一次课的粗略划范围**

**蓝色字体部分是最新（出完卷子）之后的划范围**

**不会考Python编程题，只考伪代码编程题。**

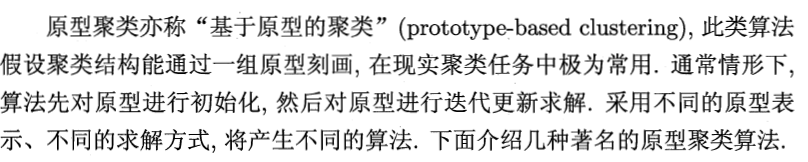
**期末占比：50%**

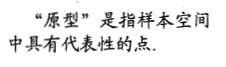
**题型参考周志华《机器学习》每个章节的题目，但是不会考类似题目。**

1. 引言
   1. **什么是监督学习？什么是非监督学习？什么是训练集？什么是测试集？各个指标的含义？什么是回归任务？什么是分类任务？其他机器学习概念。**
   2. 监督、无监督、半监督、强化—**计算或者问答（同时可能考概念，和各种方法怎么做，举例）。**
2. 机器学习概念流程
   1. **模型怎么评估？评估指标的概念（TN,TP等，ROC曲线）、**
   2. 基本流程（重点）部分：**考损失函数、优化方法、训练、测试**等
   3. 模型评估部分：经验误差与过拟合、性能度量**（考公式） .**
3. 线性模型
4. **逻辑回归/线性回归怎么计算 是什么？ 是否涉及阈值 简答/计算/应用。线性回归/逻辑回归的区别？评价模型的好坏，反差/偏差是什么？有什么用？**
5. **回归和逻辑回归的区别，逻辑回归是什么？**
6. **逻辑回归/线性回归的计算题**
7. **概念&&计算**
8. 基本分类方法：
9. **KNN 计算题必考**
10. **KNN算法步骤（伪代码）**

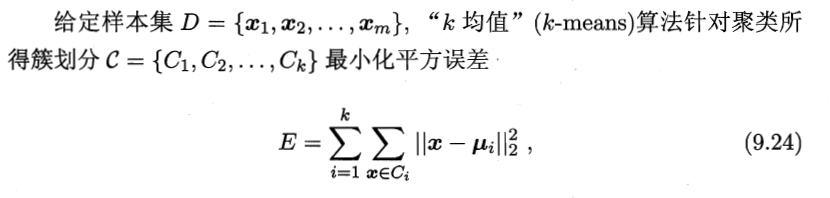
**贝叶斯：可能结合后面考，或者一个问答题，但是不会单独出贝叶斯计算题，贝叶斯不考很深，考开放性题目。举例，写方法.**

1. 神经网络
   1. **激活函数？链式公式是什么？**
   2. **权重调整表达式。反向传播必考**
   3. **常见激活函数（考简答或者问答，概念，定义，常见图像，优缺点）**
   4. **全连接神经网络的推到计算（会给定数据和激活函数）**
2. 支持向量机
3. **间隔，对偶，核函数，软/硬间隔 支持向量机距离的计算，支持向量 核心思想，哪些是支持向量，怎么求？考SVR的时候结合SVM考，不会单独考SVR。拉格朗日和对偶问题必考。**
4. 核函数的**概念**
5. **对偶问题的求解**（给定几个点把支持向量机以及目标函数写出，**求出分割平面，支持向量**，给定7，8个点或者3，4个点）
6. **单独考SVM，或者结合TSVM，考SVR。**
7. 集成学习
   1. **Adaboost算法怎么用，要考计算题，集成学习核心思想**
   2. boosting and adboast 计算题：**弱分类器迭代达到强分类器**
8. 聚类
   1. **K均值 必考 考算法/推导/怎么计算 不考概念**
   2. 原型、密度、层次聚类 **特点、算法流程**。
   3. **给定数据按照算法求解**（过程求解，10几分钟的题量）**或者问答**
9. **原型、密度、层次聚类 特点、算法流程。**

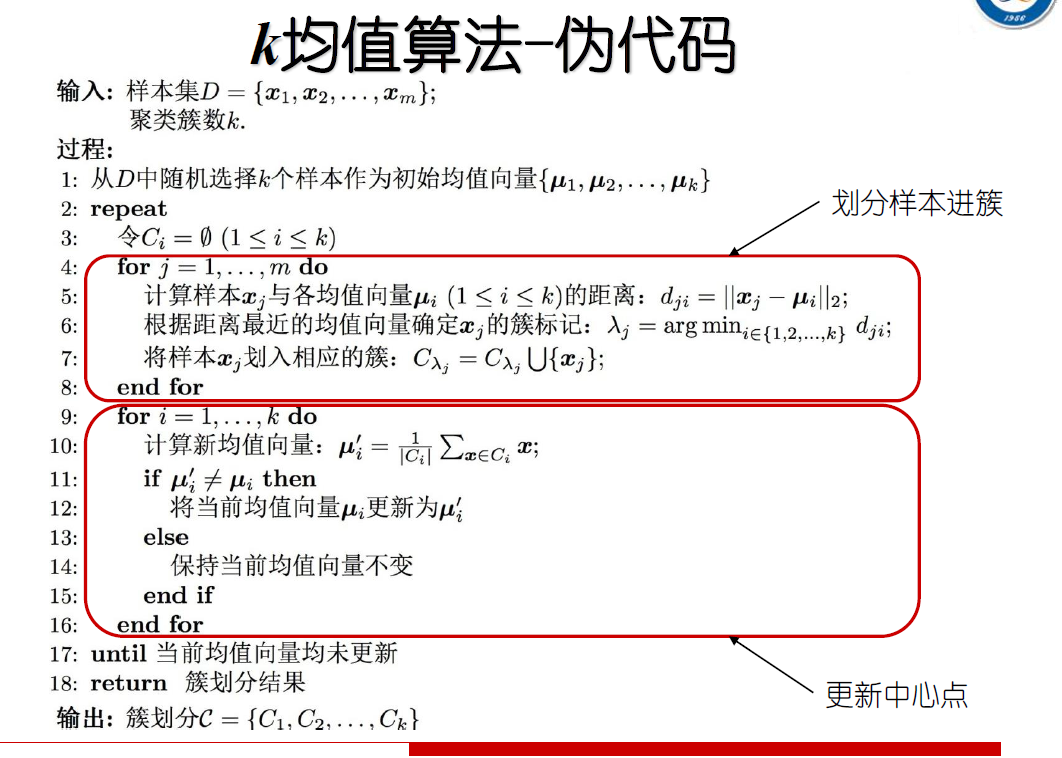




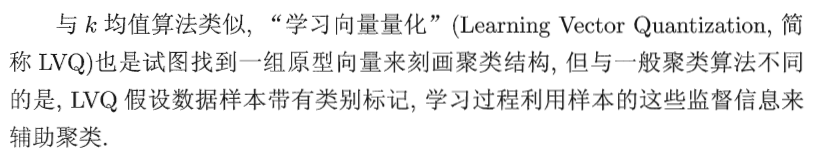
**原型聚类一-K均值聚类**

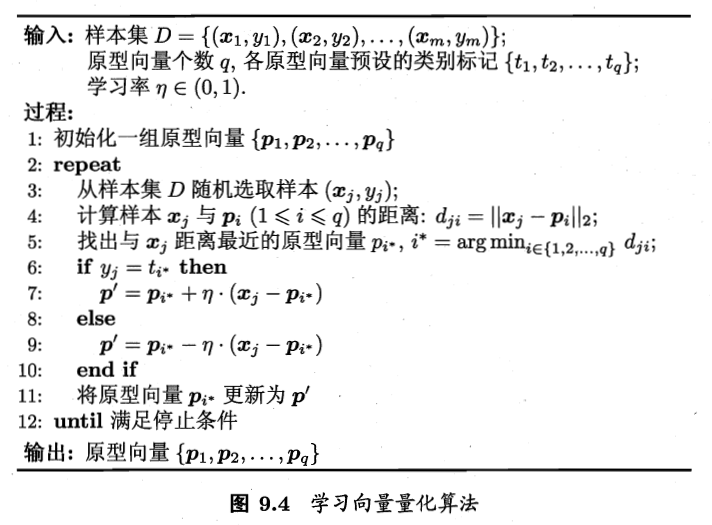
****

**算法流程**

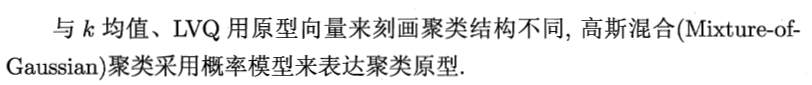
****

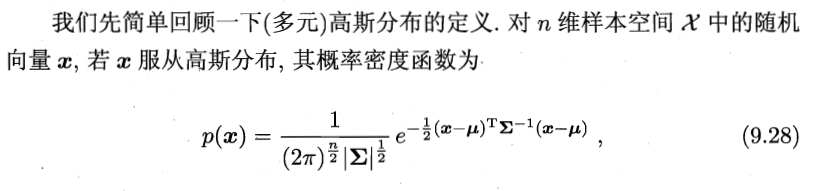
**学习向量量化**

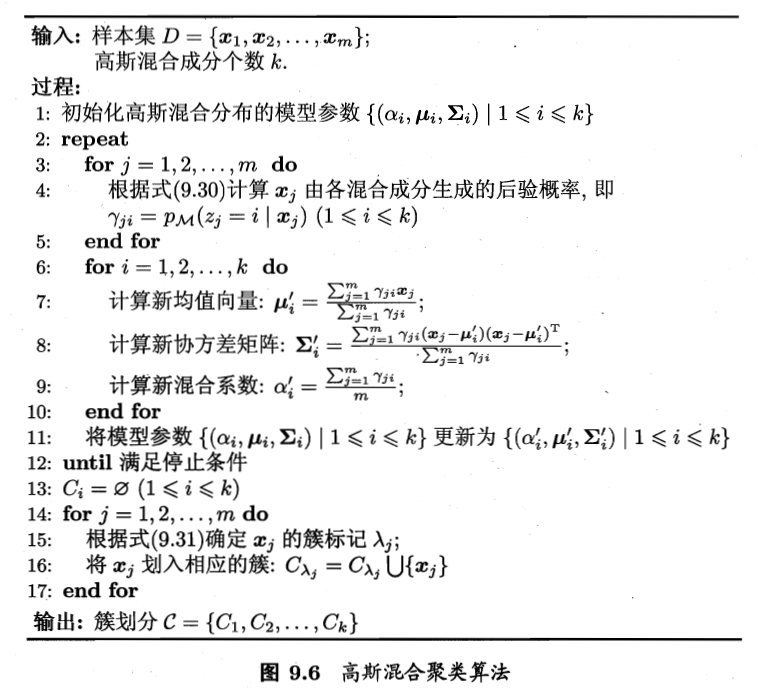
****

****

**原型聚类-高斯混合聚类**

****

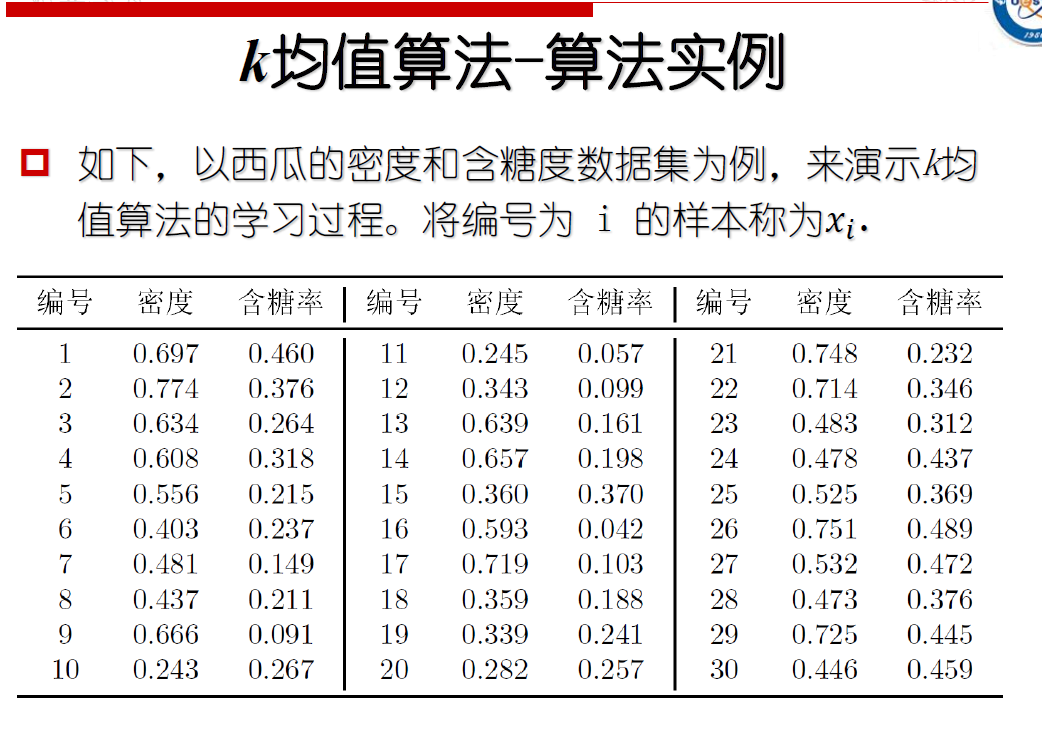
****

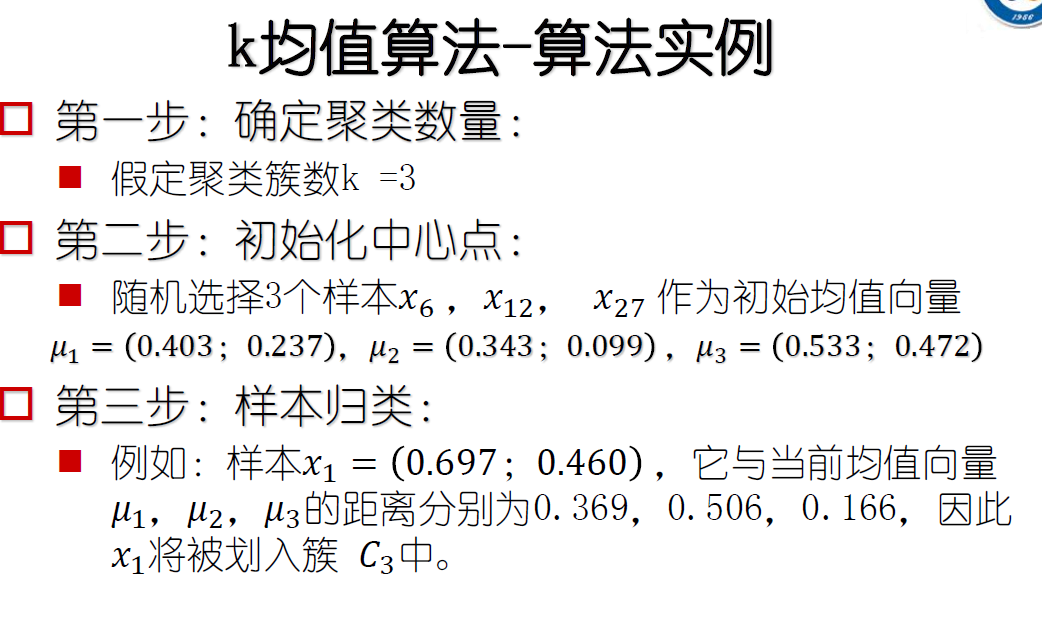
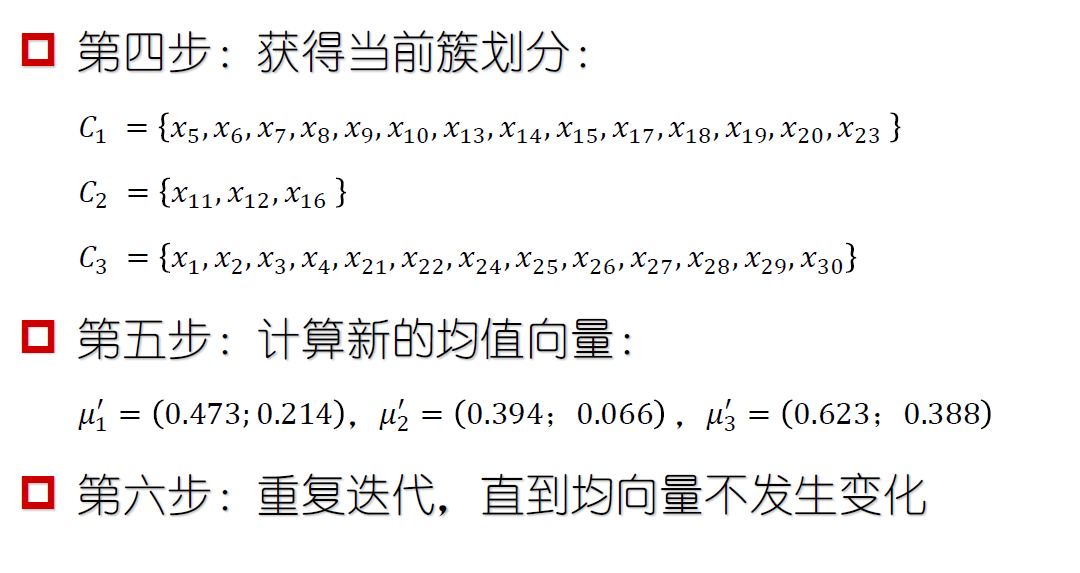
****

1. **给定数据按照算法求解**（过程求解，10几分钟的题量）**或者问答**

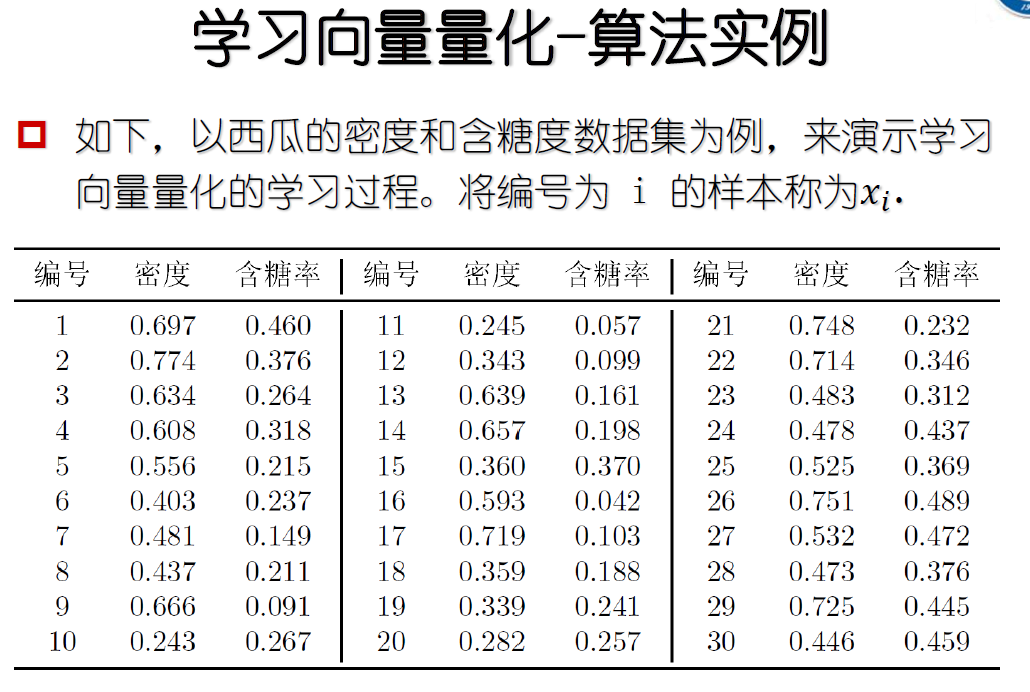
**原型聚类例题:**

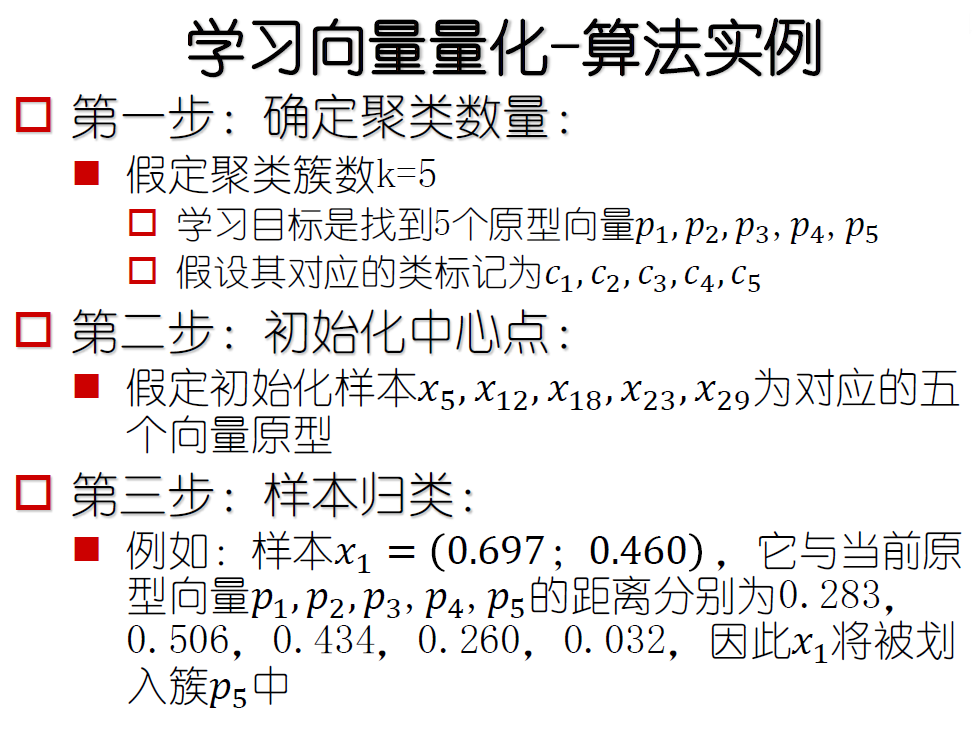
**k-均值:**

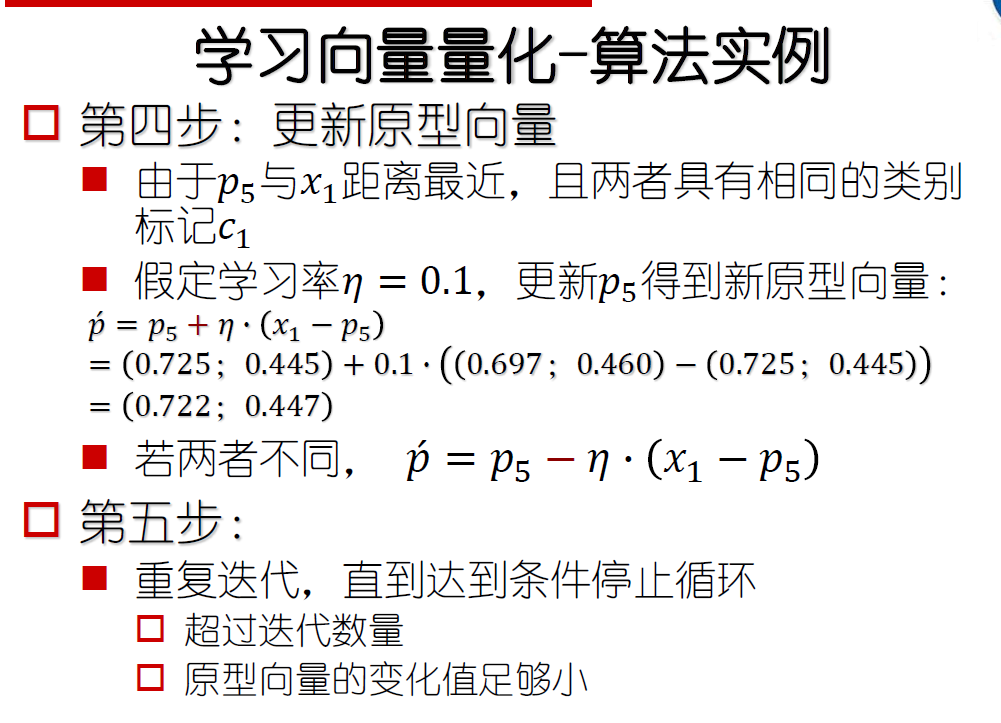
****

****

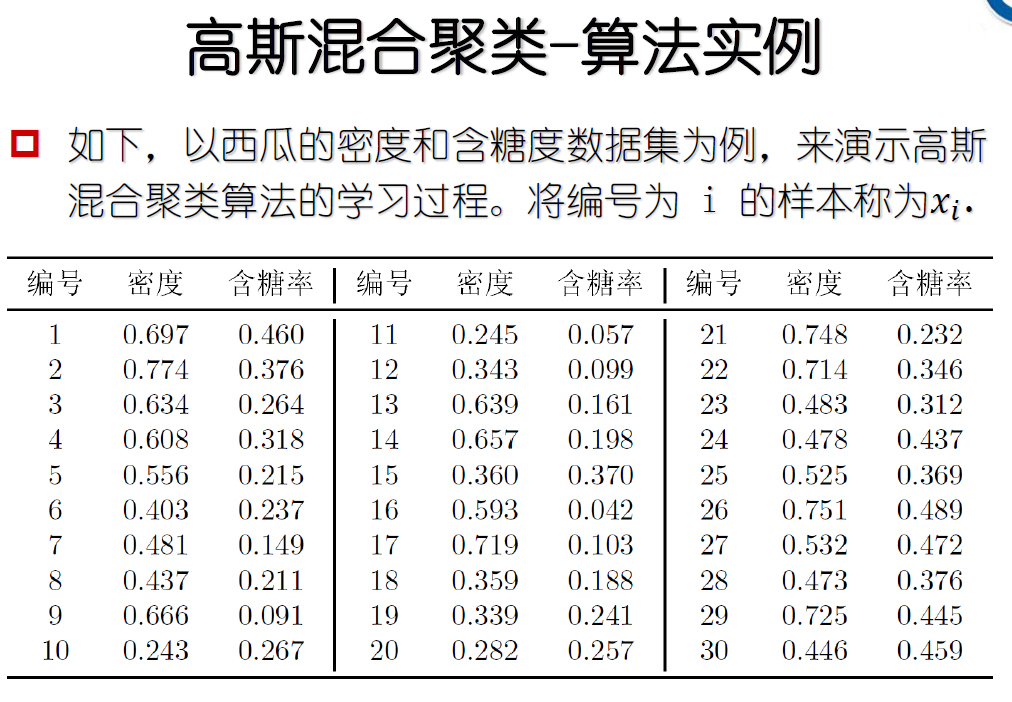
**学习向量量化**

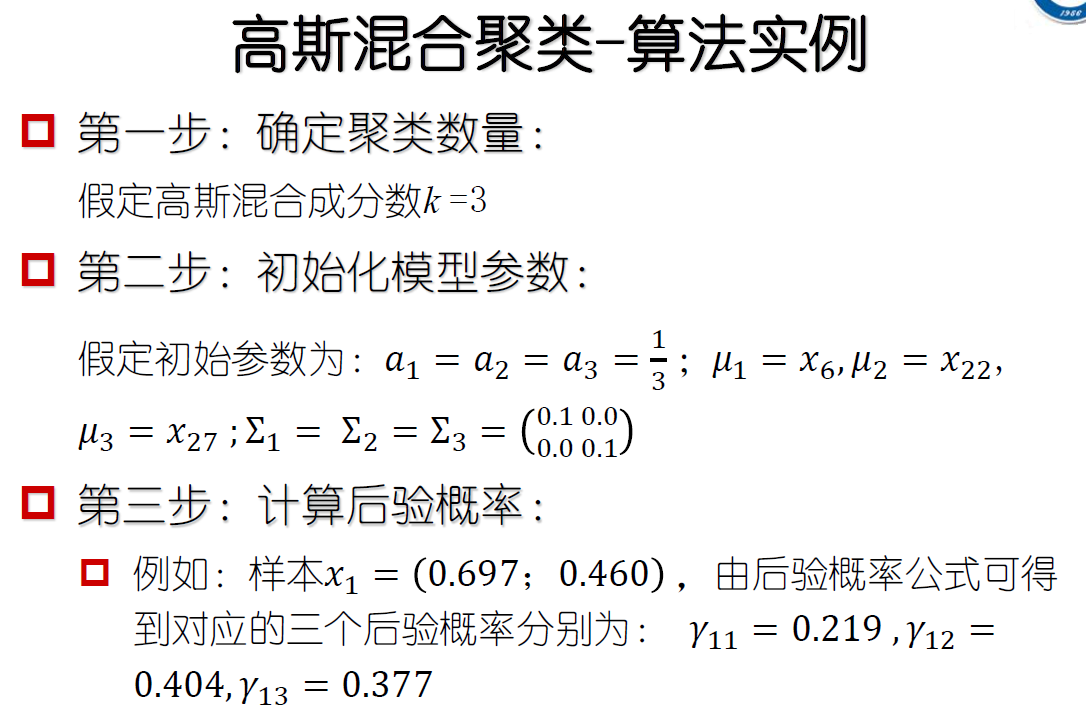
****

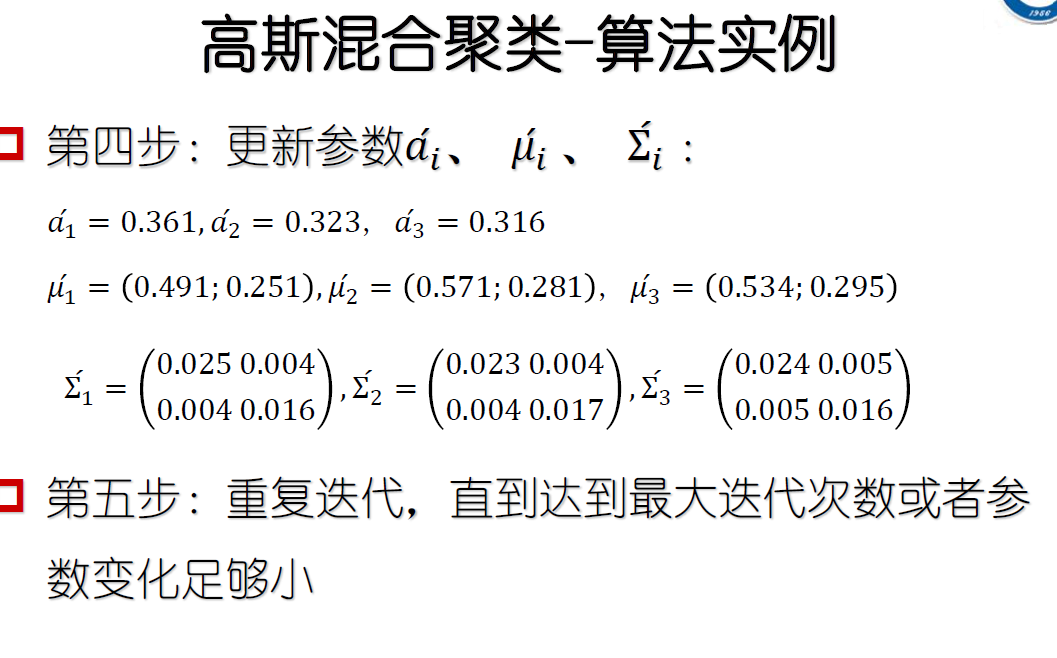
****

****

**高斯混合聚类**

****

****

****

**密度聚类例题:**

**层次聚类例题:**

1. 半监督学习
   1. **TSVM 必考 计算题/问答题/公式推导 考怎么做？计算方式？核心/基本思想 算法流程（重点） 伪代码（重点）track在哪里？SVM与TSVM 的区别在哪里，SVM相比TSVM少了什么步骤，差的点在哪里？半监督学习的流程。**
2. 降维和度量
   1. 考**动机 距离和度量 什么是度量？考PCA，但是分值不高。**
   2. 计算&&简答
3. 专题部分：
   1. **只考一个问答题或者不考，知识工程？但是由于安全部分没发PPT，因此不考安全部分，不考其他**