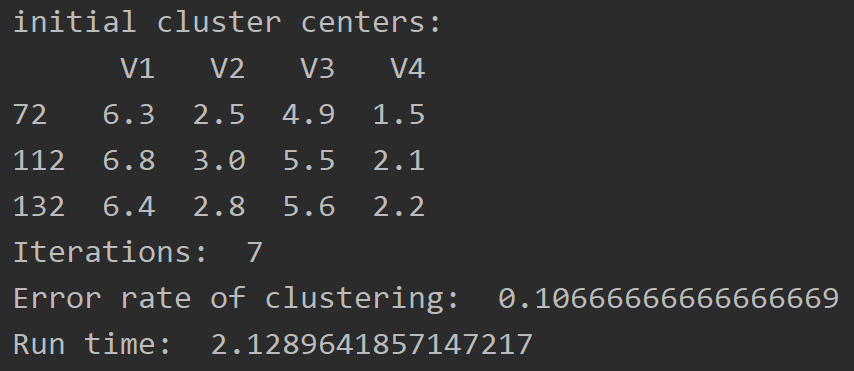
练习一

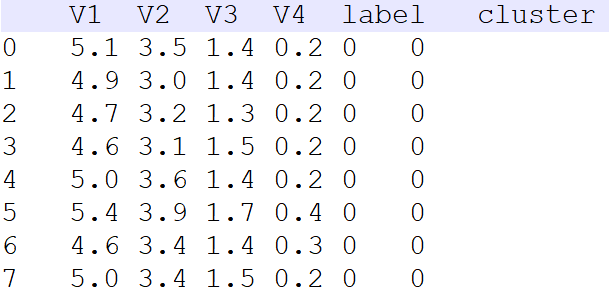
张雨 生物统计学 18110700014

2. C\_means方法聚类

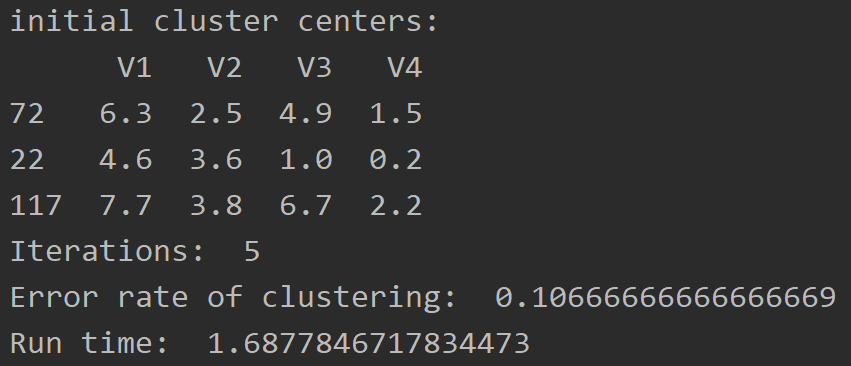
a) 初始聚类中心点，迭代次数，错误率的输出结果如下：



聚类结果储存在results\_cmean.txt文件中的cluster列，示例如下：



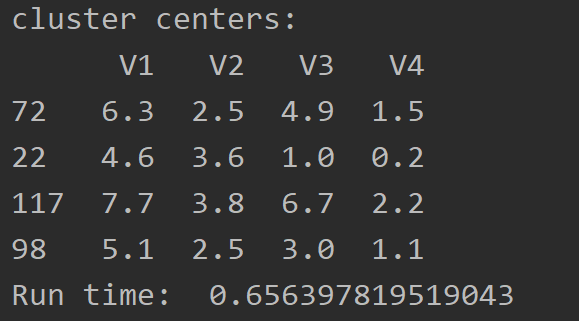
b) a)中的初始类中心点是随机选择的，接下来我们通过MMD方法先选择类中心点



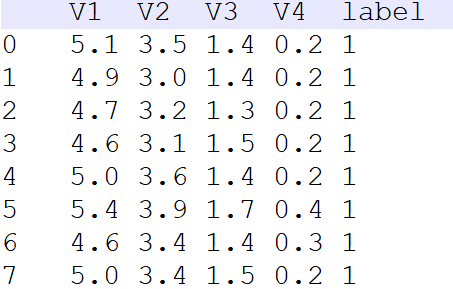
可以看出基于MMD方法选择类中心点错误率和随机选择相同，因为迭代次数变少了，运行时间有所缩短

3. 最大最小距离方法聚类

选择阈值为0.5，得到四个类中心

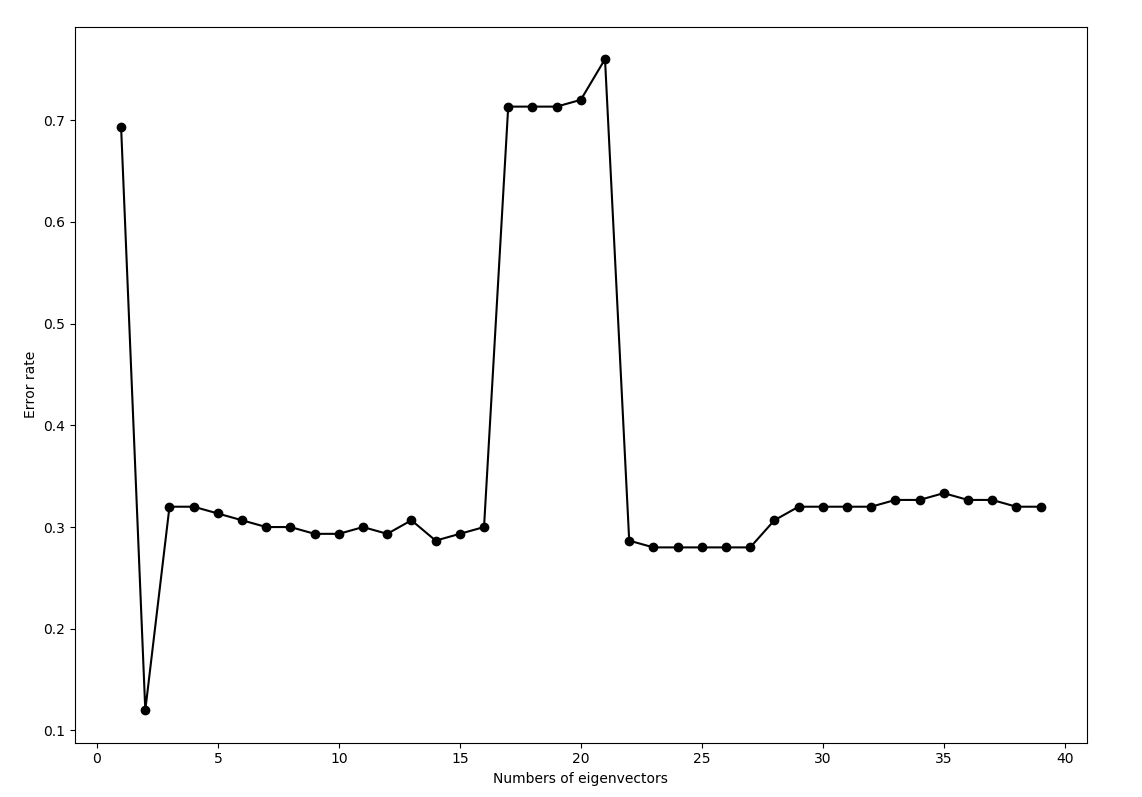


聚类结果储存在文件results\_mmd.txt中，截取前几行如下图所示

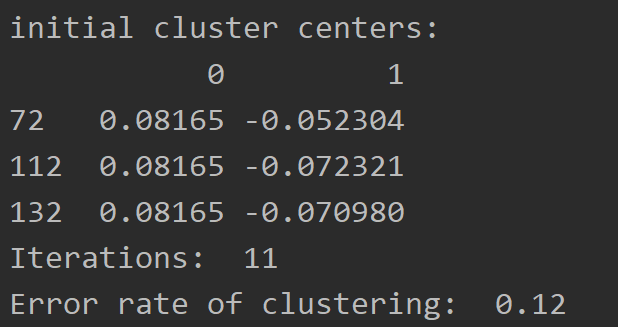


4. 谱聚类方法

谱聚类方法需要预先指定k个最小特征值对应的特征向量，这里我们尝试取1-40个特征向量进行计算，计算结果如下图：



可以看出当取两个最小特征值对应的特征向量时，聚类错误率最小，结果输出如下：



聚类结果输出在文件result\_spec\_clustering.txt中。