# 模式识别实验报告

## 实验二 GMM分类器

学院：计算学部

姓名： 杨富祥

学号：21S003071

1. **实验内容**
2. 使用Python或Matlab编程实现GMM算法：要求独立完成算法编程，禁止调用已有函数库或工具箱中的函数
3. 使用仿真数据测试程序的正确性：
   1. 两类2维各1000个训练样本保存在Emu-Train-Samples.csv和Emu-Train-Labels.csv文件中，各1000个测试样本保存Emu-Test-Samples.csv和Emu-Test-Labels.csv文件中
   2. 每个类别的样本分别采样自包含2个分量的高斯混合模型，分别使用两个类别的训练样本学习模型化每个类别条件概率密度的GMM模型参数，并与模型的真实值比较；

类别0的GMM： ，，

，，

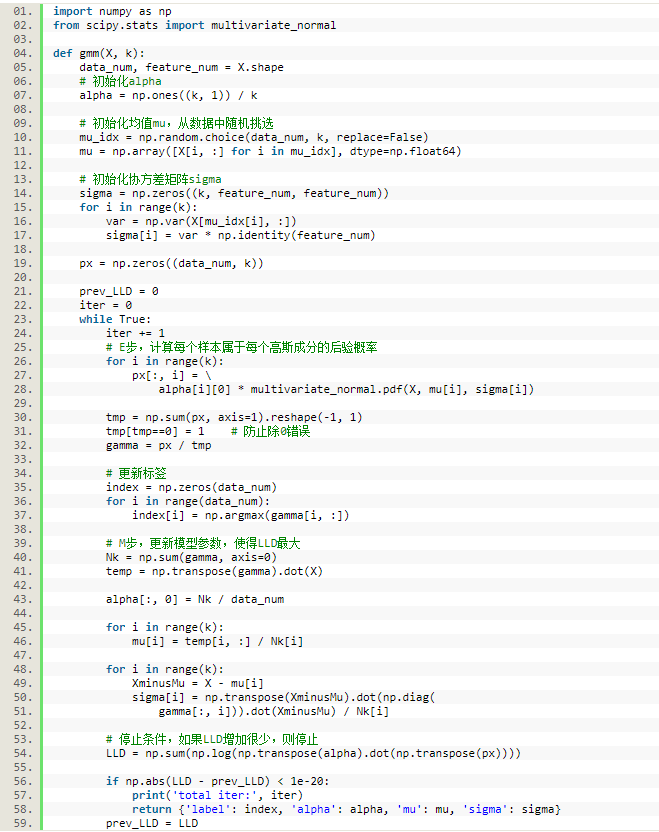
类别1的GMM： ，，

，，

* 1. 假设两个类别的先验概率相等，使用学习到的模型参数分类测试样本数据，统计分类的正确率；

1. MNIST数据集测试：
   1. MNIST-Train-Samples.csv中包含30000个17维特征手写数字样本训练，MNIST-Train-Labels.csv中包含训练样本的标签；
   2. 分别使用每个数字的样本集学习该类别的GMM模型参数；
   3. 使用10个类别GMM模型构成的分类器，分类MNIST-Test-Samples.csv中的10000个样本，与MNIST-Test-Labels.csv的类别标记比对，计算识别的正确率；
   4. 尝试设置不同的GMM模型超参数—高斯数，测试不同高斯数的GMM分类器的识别正确率；
2. **程序代码**

（GMM参数估计部分和GMM分类器部分代码）





1. **实验结果**
2. 仿真数据实验结果：给出估计出的两个GMM 模型参数，以及测试样本的识别结果。

GMM估计模型参数

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 类别0，高斯1 | 0.65890609 |  |  |
| 类别0，高斯2 | 0.34109391 |  |  |
| 类别1，高斯1 | 0.66799982 |  |  |
| 类别1，高斯2 | 0.33200018 |  |  |

**GMM分类器识别结果**

|  |  |  |
| --- | --- | --- |
|  | 正确识别数 | 正确识别率 |
| 类别0 | 1000 | 100% |
| 类别1 | 1000 | 100% |

1. MNIST数据集实验结果：

**GMM分类器识别正确率**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 高斯数 | 1 | 2 | 3 | 4 | 5 |
| 正确识别数 | 9332 | 9340 | 9486 | 9496 | 9526 |
| 正确识别率 | 93.32% | 93.40% | 94.86% | 94.96% | 95.26% |