# 模式识别实验报告

## 实验二 GMM分类器

学院：计算机科学与技术学院

姓名：李星晨

学号：16S003084

1. **实验内容**
2. 使用C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. 使用仿真数据测试算法的正确性：两类2维各1000个训练样本Train1和Train2分别采样自如下两个GMM，使用训练样本分别估计包含2个分量高斯的GMM参数。

GMM1： ，，

，，

GMM2： ，，

，，

构造区分两类的GMM分类器，测试采样自同样GMM的测试样本Test1和Test2。

1. MNIST数据集测试：使用TrainSamples中的10000个17维特征手写数字样本训练GMM分类器区分10个类别，TrainLabels中包含训练样本的标签；测试设置不同高斯数量GMM分类器对TestSamples中10000个样本的识别正确率。
2. **程序代码**

（GMM参数估计部分和GMM分类器部分代码）

参数估计部分：

Epsilon = 0.0001;

while 1

Mu\_old = Mu;

Alpha\_old = Alpha;

Sigma\_old = Sigma;

for i=1:M

Pxi(:,i) = Gauss(Data, Mu(i,:), Sigma(:,:,i));

end

Pix\_tmp = repmat(Alpha,m,1).\*Pxi;

Pix = Pix\_tmp./(repmat(sum(Pix\_tmp,M),1,M));

Beta = sum(Pix);

for i=1:M

Alpha(i) = Beta(i)/m;

Mu(i,:) = Pix(:,i)'\*Data / Beta(i);

Data\_tmp = Data - repmat(Mu(i,:),m,1);

Sigma(:,:,i) = (repmat(Pix(:,i)',n, 1) .\* Data\_tmp'\*Data\_tmp) / Beta(i);

end

v = [sum(abs(Mu - Mu\_old)), abs(Alpha - Alpha\_old)];

s = abs(Sigma-Sigma\_old);

v2 = 0;

for i=1:M

v2 = v2 + det(s(:,:,i));

end

if ((sum(v) + v2) < Epsilon)

break;

end

end

高斯计算过程：

function prob = Gauss( Data, Mu, Sigma )

%UNTITLED Summary of this function goes here

% Detailed explanation goes here

[m n] = size(Data);

Data = Data - repmat(Mu,m,1);

prob = sum((Data\*inv(Sigma)).\*Data,2);

prob = exp(-0.5\*prob)/((2\*pi)^(n/2)\*(abs(det(Sigma))^(1/2)));

end

分类器部分：

[Alpha1, Mu1, Sigma1] = em( Data1, Alpha1, Mu1, Sigma1 )

[Alpha2, Mu2, Sigma2] = em( Data2, Alpha2, Mu2, Sigma2 )

M = size(Mu1,2);

[m n] = size(Data);

for i=1:M

Pxi1(:,i) = Gauss(Data, Mu1(i,:), Sigma1(:,:,i));

end

result1 = sum(Pxi1,2);

for i=1:M

Pxi2(:,i) = Gauss(Data, Mu2(i,:), Sigma2(:,:,i));

end

result2 = sum(Pxi2,2);

count = 0;

for i=1:m

if result1(i,1) >= result2(i,1)

count = count +1;

end

end

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

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| GMM1-Gauss1 | 0.6589 | （-0.0488，-0.0349） | （2.85162，0.97072 0.97072，0.96895） |
| GMM1-Gauss2 | 0.3411 | （9.9703，9.9535） | （2.01632， 2.35543  2.35543，5.31821） |
| GMM2-Gauss1 | 0.6680 | （2.0221，10.1670） | （0.96728， 0.91060 0.91060， 2.74893） |
| GMM2-Gauss2 | 0.3320 | （14.9710，19.9925） | （5.28810，2.18045 2.18045， 1.12283） |

**GMM分类器识别结果**

|  |  |  |
| --- | --- | --- |
|  | 正确识别数 | 正确识别率 |
| Test1 | 1000 | 1000 |
| Test2 | 100% | 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 | 9215 | 9243 | 9301 | 9324 | 9349 |
| 正确识别率 | 92.15% | 92.43% | 93.01% | 93.24% | 93.49% |