

计算机与信息学院实验报告

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验课程： | 模式识别 | | | | |
| 实验编号： | 7 | | | | |
| 实验名称： | 实验7：类别可分性实验 | | | | |
| 实验人员： | 年级 | | 2018级 | | |
| 专业 | | 计算机科学与技术 | | |
| 学号 | | 18111207248 | | |
| 姓名 | | 吴钰 | | |
| 实验日期： | 2021.06.24 | | | | |
| 上交日期： | 2021.06.27 | | | | |
| 实 验 室： | 2060302 | | | | |
| 实验评价： |  | | | | |
|  | 实验成绩： |  | | 评定日期： |  |
|  | 指导教师： | 郑明 | | | |

一、实验目的

使用MATLAB实现类别可分性实验

二、实验环境

MATLAB2012

三、实验内容

（1）类内判别实验：

1. 分别导入4组数据x1（1,2,3,4,5,6,7,8,9,10），x2（2,3,4,5,6,7,8,9,10,11），x3（3,4,5,6,7,8,9,10,11,12），x4（4,5,6,7,8,9,10,11,12,13），1组模式样本分类数据y（1,1,1,1,1,1,1,1,1,1）T；
2. 分别计算这4组样本数据在y分类情况下的类内散布矩阵（Sw1, Sw2, Sw3, Sw4）；
3. 比较这4组值，最大值对应的分类数据是最差分类，最小值对应的分类数据是最优分类。

（2）类间判别实验方法：

1. 分别导入1组数据x1（1,2,3,4,5,6,7,8,9,10），4组模式样本分类数据y（1,2,3,4,5,6,5,4,3,2）T，（1,2,3,4,5,4,3,6,1,2）T，（1,2,2,4,3,4,1,2,6,1）T，（2,1,4,6,5,2,3,2,4,2）T；
2. 分别计算这1组样本数据在4组分类情况下的类间散布矩阵（Sb1, Sb2, Sb3, Sb4）；
3. 比较这4组值，最大值对应的分类数据是最优分类，最小值对应的分类数据是最差分类。

四、实验设计

根据题意，本文的主要任务是：实现类内判别实验和类间判别实验，主要根据数据和分类计算类间散度矩阵和类内散度矩阵。其中第1题的解决思路是计算样本均值，根据相应公式计算出类内散度矩阵。其中第2题的解决思路是先按照某一种标签分类结果来算，在确定分类结果下先以最大最小的分类标签值来进行比较，取出对应标签的样本值，再计算均值M和先验概率P，最后由公式算出不同分类结果下的类间散度矩阵。

五、实验结果

5.1 实验代码：

要求必要时对代码进行注释。

T1\_1.m

%类内判别实验

x1 = [1,2,3,4,5,6,7,8,9,10];

x2 = [2,3,4,5,6,7,8,9,10,11];

x3 = [3,4,5,6,7,8,9,10,11,12];

x4 = [4,5,6,7,8,9,10,11,12,13];

X = [x1;x2;x3;x4]';

Y = [1,1,1,1,1,1,1,1,1,1]';

%假设X为Wi类的样本集合 X为该样本的均值向量

for i = 1:size(X,2)

data = X(:,i);

M = mean(X(:,i));

k(i) = (data - M)'\*(data - M)/size(X,1); %公式除以N

%matlab自带函数Cov除以的是N-1

end

disp('该4组样本数据在y分类情况下的类内散布矩阵分别为：')

disp(k);

disp('计算各个样本数据的协方差矩阵（类内散度矩阵）,其值最后除以N-1')

disp([cov(x1),cov(x2),cov(x3),cov(x4)])

T1\_2.m

x = [1,2,3,4,5,6,7,8,9,10];

y = [1,2,3,4,5,6,5,4,3,2;1,2,3,4,5,4,3,6,1,2;1,2,2,4,3,4,1,2,6,1;2,1,4,6,5,2,3,2,4,2];

[m,n] = size(y);

for i = 1:m %对每种类别进行遍历

cnt = 1;

M = [];

P = [];

for index = min(y(i,:)):max(y(i,:)) %以最小分类数值和最大分类数值进行分别

temp\_x = x(y(i,:)==index);

if (isempty(temp\_x))

continue;

end

M(cnt) = mean(temp\_x);

P(cnt) = (length(temp\_x)/length(y(i,:)));

cnt = cnt+1;

end

M\_mean = sum(P.\*M);

sk(i) = sum(P.\*((M-M\_mean).^2));

end

disp('1组样本数据在4组分类情况下的类间散布矩阵:');

disp(sk);

5.2 结果展示：

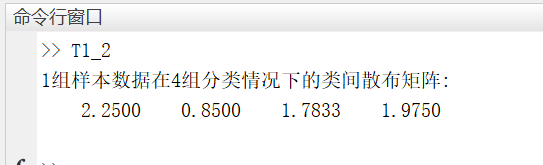
实验结果包括输出数据、截图等，将实验文档截图附在此处。

T1\_1



确定标签后，该类内部的数据的差别应该足够的小，所以其类内散度矩阵值越小越好。该题所给四种不同的样本数据，分类的结果值的优异程度一致，在用协方差矩阵计算时，结果也为一致。

T1\_2



确定标签后，该类外部的数据的差别应该足够的大，所以其类间散度矩阵值越大越好。该题所给四种不同的样本数据中，可得其四种不同的类间散度矩阵，第二类分类结果为最优分类。

六、实验总结

本次实验主要涉及计算类间散度矩阵和类内散度矩阵的操作/内容。

6.1） 在计算先验概率的时候发现先验概率不会计算的问题，在询问同学后选择直接通过该类别样本数计算。

6.2）在做实验时翻阅书本和上网查找资料对计算类间散度矩阵和类内散度矩阵这部分知识进行熟悉之后完成了实验。