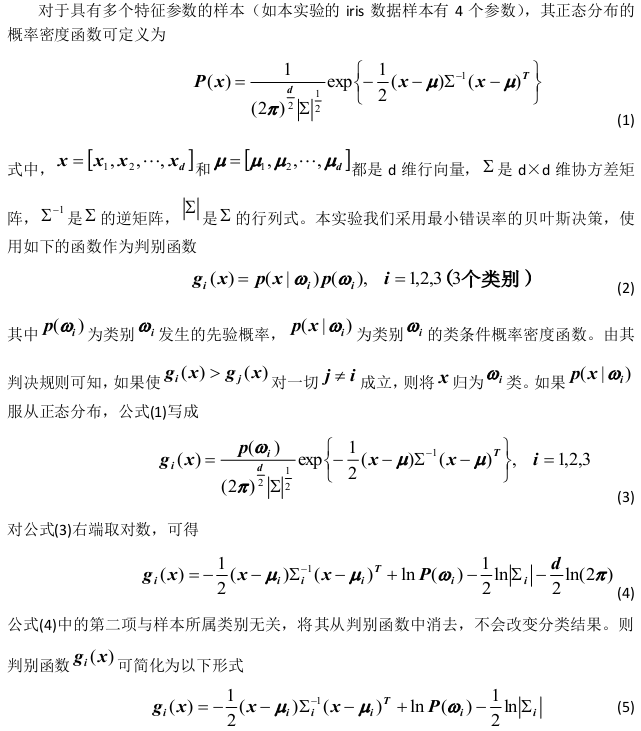
**《模式识别》实验报告**

**题目：** 最小错误率贝叶斯决策

1. **实验内容**

1，实验原理



2，实验步骤

1）从iris.txt 文件（课程邮箱-文件中心）中读取估计参数用的样本，每一类样本抽出前40个，分别求其均值；

（2）求每类样本的协方差矩阵、逆矩阵以及协方差矩阵的行列式；

（3）对三个类别，分别取每组剩下的 10个样本，每两组进行分类。由于每类样本都相等，

且每类选取用作训练的样本也相等，在每两组进行分类时，待分类样本的类先验概率为0.5。

将各个样本代入判别函数既公式(5)，进行分类。

3，实验要求

（1）复习最小错误率贝叶斯决策原理，写出实验代码，实现对三类样本的分类；

（2）计算分类的正确率，画出三维空间的样本分类图；

（3）分析实验结果，完成实验报告。

**二、实验代码**

（1），

clear

% 原始数据导入

iris=load('iris.txt');

N=40;%每组取N=40个样本

%求第一类样本均值

for i = 1:N

for j = 1:4

w1(i,j) = iris(i,j+1);

end

end

sumx1 = sum(w1,1);

for i=1:4

meanx1(1,i)=sumx1(1,i)/N;

end

%求第二类样本均值

for i = 1:N

for j = 1:4

w2(i,j) = iris(i+50,j+1);

end

end

sumx2 = sum(w2,1);

for i=1:4

meanx2(1,i)=sumx2(1,i)/N;

end

%求第三类样本均值

for i = 1:N

for j = 1:4

w3(i,j) = iris(i+100,j+1);

end

end

sumx3 = sum(w3,1);

for i=1:4

meanx3(1,i)=sumx3(1,i)/N;

end

（2），

%求第一类样本协方差矩阵

z1(4,4) = 0;

var1(4,4) = 0;

for i=1:4

for j=1:4

for k=1:N

z1(i,j)=z1(i,j)+(w1(k,i)-meanx1(1,i))\*(w1(k,j)-meanx1(1,j));

end

var1(i,j) = z1(i,j) / (N-1);

end

end

%求第二类样本协方差矩阵

z2(4,4) = 0 ;

var2(4,4) = 0;

for i=1:4

for j=1:4

for k=1:N

z2(i,j)=z2(i,j)+(w2(k,i)-meanx2(1,i))\*(w2(k,j)-meanx2(1,j));

end

ar2(i,j) = z2(i,j) / (N-1);

end

end

%求第三类样本协方差矩阵

z3(4,4) = 0 ;

var3(4,4) = 0;

for i=1:4

for j=1:4

for k=1:N

z3(i,j)=z3(i,j)+(w3(k,i)-meanx3(1,i))\*(w3(k,j)-meanx3(1,j));

end

var3(i,j) = z3(i,j) /( N-1);

end

end

%求各类的协方差矩阵逆矩阵及行列式

var1\_inv = [];var1\_det = [];

var2\_inv = [];var2\_det = [];

var3\_inv = [];var3\_det = [];

var1\_inv = inv(var1);

var2\_inv = inv(var2);

var3\_inv = inv(var3);

var1\_det = det(var1);

var2\_det = det(var2);

var3\_det = det(var3);

（3），

M=10;

for i = 1:M

for j = 1:4

test(i,j) = iris(i+50,j+1); % 取测试数据

end

end

t1=0;t2=0;t3=0;

for i = 1:M

x=test(i,1);y=test(i,2);

z=test(i,3);h=test(i,4);

g1 = (-0.5)\*([x,y,z,h]-meanx1)\*var1\_inv\*([x,y,z,h]'-meanx1') - 0.5\*log(abs(var1\_det)) +log(0.5); % p1

g2 = (-0.5)\*([x,y,z,h]-meanx2)\*var2\_inv\*([x,y,z,h]'-meanx2') - 0.5\*log(abs(var2\_det)) +log(0.5); % p2

if g1>g2

t1=t1+1; %若g1>g2,则属于第一类，否则属于第二类，并统计属于每一类的个数

else

t2=t2+1;

end

end

**三、实验结果**

（1）

第一类样本均值：5.0375 3.4525 1.46 0.235

第二类样本均值：6.01 2.78 4.3175 1.35

第三类样本均值：6.6225 2.96 5.6075 1.99

（2）每类样本的协方差矩阵、逆矩阵以及协方差矩阵的行列式

第一类样本的协方差矩阵：

0.131121794871795 0.0972115384615384 0.0133333333333333 0.0132692307692308

0.0972115384615384 0.130250000000000 0.00215384615384614 0.0119615384615385

0.0133333333333333 0.00215384615384614 0.0296410256410257 0.00502564102564103

0.0132692307692308 0.0119615384615385 0.00502564102564103 0.00951282051282051

逆矩阵：18.8146188042527 -13.3847352886752 -6.47453636800065 -5.99346924772421

-13.3847352886752 18.2504706645718 5.95329943788004 -7.42342518317483

-6.47453636800065 5.95329943788004 39.4922929094642 -19.3184124732768

-5.99346924772421 -7.42342518317483 -19.3184124732768 133.021717057711

协方差矩阵的行列式：1.56318133831812e-06

第二类样本的协方差矩阵：

0 0 0 0

0 0 0 0

0 0 0 0

0 0 0 0

逆矩阵：

Inf Inf Inf Inf

Inf Inf Inf Inf

Inf Inf Inf Inf

Inf Inf Inf Inf

协方差矩阵的行列式：0

第三类样本的协方差矩阵：

0.467942307692308 0.110410256410256 0.357775641025641 0.0512564102564103

0.110410256410256 0.113230769230769 0.0810769230769231 0.0462564102564103

0.357775641025641 0.0810769230769231 0.345326923076923 0.0593076923076923

0.0512564102564103 0.0462564102564103 0.0593076923076923 0.0742564102564103

逆矩阵：

12.0677437364511 -4.88916379648501 -12.1082128512238 4.38637019494952

-4.88916379648501 14.8211423555543 3.00369869101398 -8.25671873361937

-12.1082128512238 3.00369869101398 15.7866851329687 -6.12187764556262

4.38637019494952 -8.25671873361937 -6.12187764556262 20.4719162879598

协方差矩阵的行列式：0.000142786635162081

（3）各类样本代入公式分类得：

第一类样本：

5.11375000000000 3.79125000000000 0.520000000000000 0.517500000000000

3.79125000000000 5.07975000000000 0.0839999999999996 0.466500000000000

0.520000000000000 0.0839999999999996 1.15600000000000 0.196000000000000

0.517500000000000 0.466500000000000 0.196000000000000 0.371000000000000

第二类样本：

10.6760000000000 3.37800000000000 6.71300000000000 2.04000000000000

3.37800000000000 4.32400000000000 3.15400000000000 1.77000000000000

6.71300000000000 3.15400000000000 7.93775000000000 2.87500000000000

2.04000000000000 1.77000000000000 2.87500000000000 1.68000000000000

第三类样本：

18.2497500000000 4.30600000000000 13.9532500000000 1.99900000000000

4.30600000000000 4.41600000000000 3.16200000000000 1.80400000000000

13.9532500000000 3.16200000000000 13.4677500000000 2.31300000000000

1.99900000000000 1.80400000000000 2.31300000000000 2.89600000000000

三维空间分类图：









**四 实验心得**

首先感谢老师给予我们这次试验的机会，通过这次试验，我更深入了解了最小错误率贝叶斯决策原理，通过计算期望，可以对现实问题进行数学计算，从而获得最优解。通过自己对相关程序的理解和老师的帮助下我们完成了程序，并仔细阅读了程序,希望对以后的学习有一定的帮助。经过实验,我对模式识别的认识更深入了,希望能学好模式识别。

**实验报告成绩评定表**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **评定项目** | **内 容** | **满分** | **评分** | **总分** |
| 实验内容 | 是否明确实验目的，是否熟悉算法原理。 | 20 |  |  |
| 实验代码 | 代码是否完整、准确。 | 30 |  |
| 实验结果 | 结果是否真实、准确。 | 20 |  |
| 实验分析 | 是否撰写实验分析，是否阐述自己的观点。 | 30 |  |
| 总成绩 | 采用五级分制：优、良、中、及格、不及格 |  | | |
| **指导教师评语：** | | | | |