

计算机与信息学院实验报告

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验课程： | 模式识别 | | | | |
| 实验编号： | 5 | | | | |
| 实验名称： | 实验5：贝叶斯决策实验 | | | | |
| 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： | 2021.5.24 | | | | |
| 上交日期： | 2021.6.4 | | | | |
| 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实现贝叶斯决策相关实验

了解最小风险判决的基本原理

掌握最小风险判决模型的算法设计

二、实验环境

MATLAB2012

三、实验内容

已知存在23个细胞样本的集合*X*={-3.9847 -3.5549 -1.2401 -0.7932 -2.8531 -2.7605 -3.7287 -2.5414 -2.2692 -3.4549 -3.0752 -3.9934 2.8792 -0.9780 0.7932 1.1882 3.0682 -1.5799 -1.4885 -0.7431 -0.4221 -1.1186 4.2532}。采用最小风险贝叶斯决策对细胞进行判决分类，将其划分为正常细胞和异常细胞。已知正常细胞和异常细胞的先验概率分别为*P*（*ω*1）=0.9和*P*（*ω*2）=0.1。正常细胞和异常细胞分别服从均值μ=-2和方差σ=0.5，均值μ=2和方差σ=2的单变量正态分布。并且损失函数*r*11=0，*r*12 = *r*21=4，*r*22=0。请设计一个最小风险贝叶斯决策模型，对23个细胞样本进行判决分类，并绘制出相应的条件风险曲线和细胞分类结果。

步骤如下：

1. 给定样本*x*，计算其后验概率*P*（*ωj*|*x*），*j*=1，2，…，*c*。
2. 在已知损失函数条件下，根据公式求出两种分类的条件平均风险，*i* =1，2。
3. 根据公式，则*x*∈*ωi*。比较各自判决分类情况下的条件平均风险，把样本x判决分类到条件平均风险最小的那一个类别。

四、实验设计

根据题意，本文的主要任务是：采用最小风险贝叶斯决策对细胞进行判决分类，将其划分为正常细胞和异常细胞。其中第1题的解决思路是先对给定样本x，计算其后验概率；在已知损失函数条件下，根据公式求出两种分类的条件平均风险R1\_plot(j)和R2\_plot(j)；再根据公式比较各自判决分类情况下的条件平均风险，把样本x判决分类到条件平均风险最小的那一个类别。。

五、实验结果

5.1 实验代码：

要求必要时对代码进行注释。

Bayes.m

clear all

x=[-3.9847 -3.5549 -1.2401 -0.7932 -2.8531 -2.7605 -3.7287 -2.5414 -2.2692 -3.4549 -3.0752 -3.9934 2.8792 -0.9780 0.7932 1.1882 3.0682 -1.5799 -1.4885 -0.7431 -0.4221 -1.1186 4.2532]

disp(x);

pw1=0.9;

pw2=0.1;

[R1\_x,R2\_x,result]=danger(x,pw1,pw2);

danger.m

function [R1\_x,R2\_x,result]=danger(x,pw1,pw2)

m=numel(x);%得到待测细胞个数

R1\_x=zeros(1,m);%存放把样本X判为正常细胞所造成的整体损失

R2\_x=zeros(1,m);%存放把样本X判为异常细胞所造成的整体损失

result=zeros(1,m);%存放比较结果

u1=-2;

a1=0.5;

u2=2;

a2=2;

%类条件概率分布px\_w1: ( -2，0.25 )px\_w2 (2,4)

r11=0;

r12=4;

r21=4;

r22=0;

%风险决策表

for i=1:m %计算两类风险值

R1\_x(i)=r11\*pw1\*normpdf(x(i),u1,a1)/(pw1\*normpdf(x(i),u1,a1)+pw2\*normpdf(x(i),u2,a2))+r21\*pw2\*normpdf(x(i),u2,a2)/(pw1\*normpdf(x(i),u1,a1)+pw2\*normpdf(x(i),u2,a2));

R2\_x(i)=r12\*pw1\*normpdf(x(i),u1,a1)/(pw1\*normpdf(x(i),u1,a1)+pw2\*normpdf(x(i),u2,a2))+r22\*pw2\*normpdf(x(i),u2,a2)/(pw1\*normpdf(x(i),u1,a1)+pw2\*normpdf(x(i),u2,a2));

end

for i = 1:m

if R2\_x(i) > R1\_x(i)%第二类比第一类风险大

result(i) = 0%判为正常细胞（损失较小），用0表示

else

result(i) = 1%判为异常细胞，用1表示

end

end

a=[-5:0.05:5];%取样本点以画图

n=numel(a);

R1\_plot=zeros(1,n);

R2\_plot=zeros(1,n);

for j=1:n

R1\_plot(j)=r11\*pw1\*normpdf(a(j),u1,a1)/(pw1\*normpdf(a(j),u1,a1)+pw2\*normpdf(a(j),u2,a2))+r21\*pw2\*normpdf(a(j),u2,a2)/(pw1\*normpdf(a(j),u1,a1)+pw2\*normpdf(a(j),u2,a2));

R2\_plot(j)=r12\*pw1\*normpdf(a(j),u1,a1)/(pw1\*normpdf(a(j),u1,a1)+pw2\*normpdf(a(j),u2,a2))+r22\*pw2\*normpdf(a(j),u2,a2)/(pw1\*normpdf(a(j),u1,a1)+pw2\*normpdf(a(j),u2,a2));

%计算各样本点的风险画图

end

figure(1)

hold on

plot(a,R1\_plot,'b-',a,R2\_plot,'g\*-')

for k=1:m

if result(k)==0

plot(x(k),-0.1,'b^')%正常细胞用上三角表示

else

plot(x(k),-0.1,'go')%异常细胞用圆圈表示

end;

end;

legend('正常细胞','异常细胞','Location', 'Best')

xlabel('细胞分类结果')

ylabel('条件风险')

title('风险判决曲线')

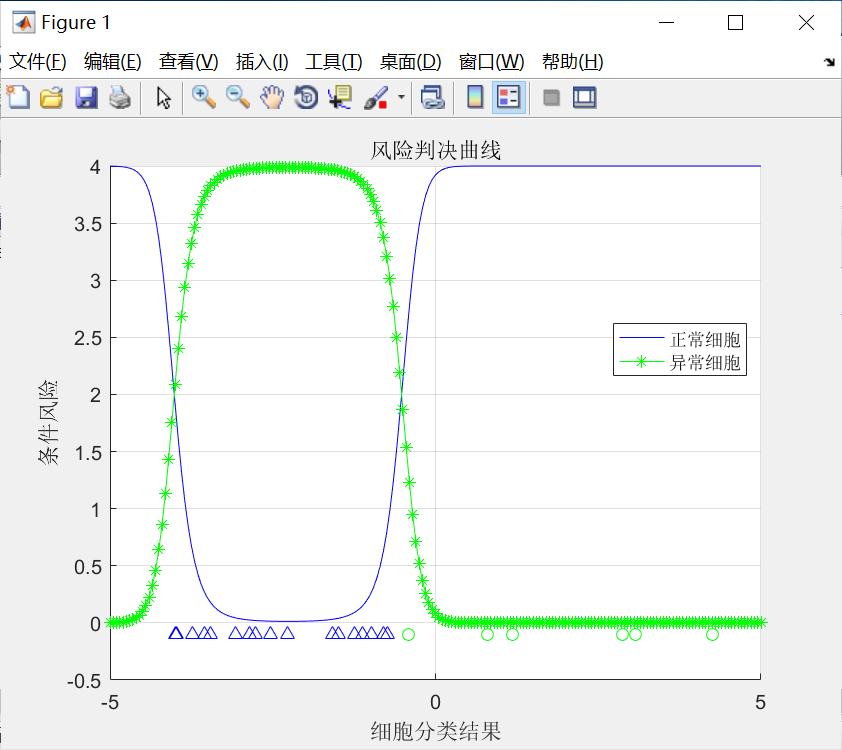
grid on

return

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。

风险判决曲线：



风险判决曲线如图所示，其中带\*的绿色曲线代表异常细胞的条件风险曲线;另一条光滑的蓝色曲线为判为正常细胞的条件风险曲线。根据贝叶斯最小风险判决准则，判决结果见曲线下方，其中“蓝三角”代表判决为正常细胞，“绿圆圈“代表异常细胞。

分类结果：0为正常细胞，1为异常细胞



六、实验总结

本次实验主要涉及贝叶斯最小风险判决模型的算法设计的操作/内容。

实验思考：

6.1) 贝叶斯分类特点：先验概率是计算后验概率的基础，是通过大量统计来得到的，这就是大数定理。许多事件的发生不具有可重复性，所以先验概率只能根据对置信度的主观判断，那么就以新获得的信息对先验概率进行修正。分类决策一定存在错误率，即使错误率很低。

6.2) 不同的贝叶斯分类器有不同的贝叶斯决策

（1）最小错误率分类器：就是把样本划分到后验概率大的那一类去。最小错误率和最大后验概率两者等价。对于最小错误率规则，确定了最小错误率也就确定了决策边界，也就是两个后验概率相等的点。

（2）最小风险贝叶斯决策：正是考虑各种错误造成损失不同而提出的一种决策规则，对待风险的态度:“宁可错杀一千，也不放走一个”。

（3）二者差别：差别在于是否考虑风险，即错误损失。最小风险决策可看作加权形式的最小错误率决策，加权值即损失函数取特定形式时二者可能等价，如损失函数取0-1形式。