1. 第一题
   1. 写出贝叶斯最小风险决策和最小错误率决策的决策规则

贝叶斯最小风险决策：

为了最小化总风险，对所有的i = 1，… , c 计算条件风险：



其中αi是指将分类判别为第i类的行为。

则总的条件风险为：



贝叶斯决策的规则就是将上述条件风险最小化

最小错误率决策的决策规则为：



则决策规则为：



* 1. 写出最小损失决策的决策规则（包括分类规则和拒识规则）



1. 第二题
   1. 请写出类条件概率密度函数的数学形式



* 1. 两种情况下的最小错误率决策判别函数
     1. 类协方差矩阵不等

判别函数为：



* + 1. 类协方差矩阵相等

有线性的判别函数：



* 1. 协方差矩阵奇异时的处理方法
     1. 协方差矩阵奇异说明必然有两个维度线性相关，因此可以将所有线性相关的维度，只保留一个，即将线性相关的参数去掉至保留一个。
     2. 先假设不同类别的协方差矩阵相同，计算得到共同的协方差矩阵，然后对相同的协方差和不同的协方差之间进行加权平均。

1. 第三题
   1. 此规则下的错误率：



* 1. 当错误率最小时，满足的条件：

此时P(error)的导数为0



* 1. 这个方程是否决定了θ？

此方程并不能决定θ，因为导数等于0的情况除了错误率最小的情况，当错误率最大的时候，导数也等于0。

* 1. 举出一个例子

假设样本在(0,1]之间符合均匀分布，且分布在（0, 0.5）之间的类别为w2，分布在（0.5, 1]之间的类别为w1，在x=0.5处w1和w2各占一半，则题目中所述分类规则在θ等于0.5的时候取到最大值，而且符合公式：



1. 第四题

解：



1. 编程题：

语言：python 依赖包：numpy

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | average |
| LDF | 0.963 | 0.976 | 0.785 | 0.881 | 0.889 | 0.702 | 0.914 | 0.855 | 0.761 | 0.797 | 0.855 |
| QDF | 0.935 | 0.736 | 0.935 | 0.894 | 0.900 | 0.717 | 0.890 | 0.875 | 0.890 | 0.822 | 0.859 |

总体上来看，QDF的分类效果略好于LDF，从各个数字的分类效果上来看，QDF的大部分数字的分类准确率高于LDF，但是数字1的分类准确率下降比较严重，可能是因为过拟合造成的，测试数据集当中有较多歪斜的数字，导致误判的发生。