作业一

学号：xxxxxxx 姓名：xxxx 学院：xxxxxxx

**题目1答：**

1. 根据题意可知，对于所有随机变量有：



假设对于所有的都有：



那么：



而又根据题意可知，此时：



因此，只要存在，即可得到，等号仅在均匀分布是成立。

1. 错误的概率就是1减去正确的概率，也就是



(3)



**题目2答：**

1. 由贝叶斯公式，有：



条件风险：



贝叶斯风险最小决策：



最小错误率决策：



1. 引入拒识后，条件风险为：



最小风险决策为：



**题目3答：**

1. 构造白化变换，计算分别表示本征向量和本征值的矩阵Φ和Λ。







由



当时，



当时，



当时，





由此，我们可以得到：





1. 将(1)中的白化变换应用于点，求其经过白化变换后的点



1. 通过详细计算，证明原分布中从**x0**到均值*µ*的Mahalanobis距离与变换后的分布中从到**0**的Mahalanobis距离相等。

原分布中从**x0**到均值*µ*的Mahalanobis距离：









变换后的分布中从到**0**的Mahalanobis距离：



原分布中从**x0**到均值*µ*的Mahalanobis距离与变换后的分布中从到**0**的Mahalanobis距离相等，白化变换为线性变化。

1. 概率密度在一个一般的线性变换下是否保持不变？换句话说，对于某线性变换T，是否有?解释原因。

对于线性变换T，我们可以得到，







又



对于线性变换T，成立。

**题目4答：**

对一个c类分类问题, 特征向量x∈Rd，假设各类先验概率相等，每一类条件概率密度为高斯分布。

1. 请写出类条件概率密度函数的数学形式；



1. 请写出在下面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;



其中，







1. 所有类协方差矩阵相等.



其中，





1. 在基于高斯概率密度的二次判别函数中，当协方差矩阵为奇异时，判别函数变得不可计算。请说出两种克服协方差奇异的方法。
   * 1. 降维
2. 求伪逆矩阵Σ†=(Σ⊤Σ)−1Σ⊤来代替逆矩阵，伪逆矩阵也满足Σ†Σ=I。
3. 使用Shrinkage策略(正则判别分析)，将各个类的协方差矩阵向同一矩阵缩并，还可把矩阵再向单位矩阵缩并：





式(4.11)可以认为是对线性判别函数(LCF)做正则，式(4.10)可以认为是对二次判别函数(QDF)做正则。这种策略也可以用来缓解过拟合。

**题目5. 编程题：**

(1) 写一个程序产生*d*维空间的样本点，服从均值为*µ*和协方差矩阵Σ的正态分布。

(2) 考虑正态分布



其中，**I**为单位矩阵，且。说明贝叶斯判决边界。

(3) 产生 n = 100个点(50个ω1类的点，50个ω2类的点)，并计算经验误差。

(4) 对于不断增加的n值重复以上步骤，100 ≤ n ≤1000，步长为100，并绘出所得的经验误差。

**运行环境**：Ubuntu18.04.6 + Python3.10

**环境依赖**：numpy1.23.3，matplotlib3.6.0

**解决思路**：

1. 直接调用numpy中的函数multivariate\_normal (mean, cov, size=None, check\_valid=None, tol=None)生成一个服从多元正态分布的数组
2. 首先，根据(1)中所给信息计算确定判决边界。由于和均服从正态分布，且有，因此，贝叶斯判决边界为



贝叶斯判决边界为：



（3）（4）代码见附件。

实验结果如下：





