考虑上机题2中的3个类别，设P(wi)=1/3。

(a)以下各测试点与上机练习2中各类别均值间的Mahalanobis距离分别是多少:

(1,2,1)^t,(5,3,2)^t,(0,0,0)^t ,(1,0,0)^t

(b)对以上各点进行分类。

(c)若设P(w1)=0.8,P(w2)= P(w3)=0.1,再对以上测试点进行分类。

**机器学习实验报告**

**（一）**

****

姓名：崔玉峰

学号：201600301079

班级：2016级4班

1. **题目陈述**

考虑上机题2中的3个类别，设P(wi)=1/3。

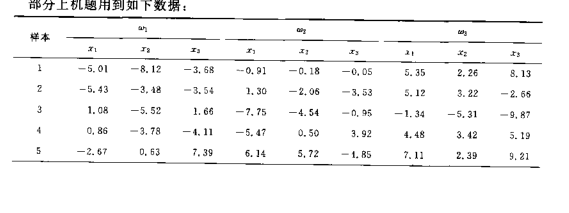
(a)以下各测试点与上机练习2中各类别均值间的Mahalanobis距离分别是多少:

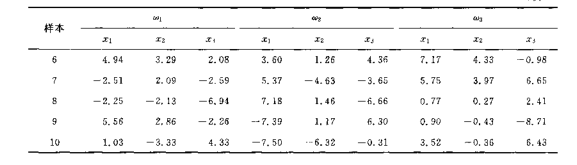
(1,2,1)^t,(5,3,2)^t,(0,0,0)^t ,(1,0,0)^t

(b)对以上各点进行分类。

(c)若设P(w1)=0.8,P(w2)= P(w3)=0.1,再对以上测试点进行分类。

**数据：**





1. **实验思路**

本实验共有三个题目，下面分析每个题目的具体解题思路：

1. **题目a的求解**

题目a的相对简单，主要用到了马氏距离公式，难点在于对与马氏距离公式的掌握情况，以及如何用代码实现。

*马氏距离公式：*



*变量说明：*

 d维均值向量

X d维向量

 d\*d协方差矩阵

r 马氏距离

 样本均值 可以通过样本数据求出，每个类别的样本均值为

X 为给出的各测试点

 为样本协方差矩阵 可以通过样本数据通过协方差公式求出，每

个类别的样本协方差为

所有的自变量都能求出，所以只要将它们带入公式就能求出测试点到各类别样本均值的马氏距离r

1. **题目b的求解**

① 根据题目可以假设 似然函数p(x|) 服从多元正态密度函数即：



***(1)***

② 根据贝叶斯公式可知：



③ 其中=1，所以此公式可以化为：



④ 通过对数分解简化可转化为：



***(2)***

⑤ 将***(1)***式带入***(2)***式可得正态函数判别式：

***(3)***

测试点在该类别算出的越大，则后验概率越大，则该测试点则更有可能属于该类别。只需算出每个测试点在在不同类别样本中，取最大值的那个类别即可.

1. **题目c的求解**

题目c与题目b的区别，就在于先验概率不同，只需将先验概率，更改再代入***(3)***即可。