模式识别作业二

姓名：李宁　　　　学号：2120180459　　　　专业：计算机科学与技术

问题一

一、问题描述

证明在 M 类问题中，最优分类器的分类错误边界为$P_e \le \frac{M-1}{M}$。

二、基本思路

在多分类问题的最小错误率决策规则中，可以将错误决策概率表示为正确决策概率的形式：

$$P_{error} = 1 - P_{correct}$$

因此，最小错误率的决策边界即为最大正确率的决策边界。

又因为在观测到模式后的最优决策规则是最大后验决策规则，所以可以通过证明对于每个 $\boldsymbol{x}$ 的后验概率$P(\omega_i|\boldsymbol{x}), i = 1,2, \dots, M$的最大值大于或等于 1/M（且在所有后验概率相等时等式成立）即可。

三、问题求解

证明：在 M 个类别中易知对于每个类别ω_i

$$\sum_{i=1}^{M} P(\omega_i|\boldsymbol{x}) = 1$$

所以 $\max_i\big(P(\omega_i|\boldsymbol{x})\big) \ge \frac{1}{M}$，当且仅当每类的先验概率$P(\omega_i)$相同时，等号成立。

又因为，在观测到模式后的最优决策规则式最大后验决策规则（后验概率最大）

令 P 为贝叶斯分类决策，则

$$P = \max_i\big(P(\omega_i|\boldsymbol{x})\big) \ge \frac{1}{M}$$

由多类问题中最小错误率决策规则可知：

$$\begin{aligned} P_{errror} &= 1 - P_{correct} \\ &= 1 - \max_i\big(P(\omega_i|\boldsymbol{x})\big) \\ &\le 1 - \frac{1}{M} = \frac{M-1}{M} \end{aligned}$$

问题得证。

问题二

一、问题描述

在两类分类问题中，可以约束某类的错误率不变，即让$\varepsilon_1 = \varepsilon$。试证明最小化另一类错误率得到的似然测试规则为：

$$Decide\ \ x \in \omega_1\ \ if\ \ \frac{P(\omega_1|\boldsymbol{x})}{P(\omega_2|\boldsymbol{x})} > \theta$$

其中，选择参数θ来满足约束条件。

二、基本思路

由题意可知，该问题为 *Neyman-Pearson* 测试规则，他和贝叶斯最小风险规则相似。

使用拉格朗日乘子法证明这个问题等价于最小化$q = \theta(\varepsilon_1 - \varepsilon) + \varepsilon_2$的问题。下面对拉格朗日乘子法进行简要讲解。

拉格朗日乘子法试一种经典的求解**条件极值**的解析方法，可将所有约束的优化模型问题转化为无约束极值问题的求解。其计算过程大致如下：

- 假设需要求极值的目标函数为$f(x,y)$，限制条件为$\varphi(x,y) = M$
- 设$g(x,y) = \varphi(x,y) - M$
- 定义一个新函数$F(x,y,\lambda) = f(x,y) + \lambda g(x,y)$
- 用偏导数方法列出方程：

$$\begin{cases} \dfrac{\partial F}{\partial x} = 0 \\ \dfrac{\partial F}{\partial y} = 0 \\ \dfrac{\partial F}{\partial \lambda} = 0 \end{cases}$$

- 求出x, y, λ的值,代入即可得到目标函数的极值。

三、问题求解

证明：由题，利用拉格朗日乘子法可以构造如下公式

$$\begin{aligned} J &= P(error|\omega_2) + \lambda[P(error|\omega_1) - \varepsilon] \\ &= \frac{P(\omega_2)}{p(\boldsymbol{x})}\int_{R_1} p(\boldsymbol{x}|\omega_2)\, d\boldsymbol{x} + \lambda\left[\frac{P(\omega_1)}{p(\boldsymbol{x})}\int_{R_2} p(\boldsymbol{x}|\omega_1)\, d\boldsymbol{x} - \varepsilon\right] \end{aligned}$$

显然，对λ求导可知，若$P(error|\omega_1) = \varepsilon$，则$J$达到最小，即$P(error|\omega_2)$最小。

又因为在上式J中，$P(error|\omega_1) + P(correct|\omega_1) = 1$，即

$$\frac{P(\omega_1)}{p(\boldsymbol{x})}\int_{R_2} p(\boldsymbol{x}|\omega_1)\, d\boldsymbol{x} + \frac{P(\omega_1)}{p(\boldsymbol{x})}\int_{R_1} p(\boldsymbol{x}|\omega_1)\, d\boldsymbol{x} = 1$$

则原式可以写成

$$J=\frac{P(\omega_2)}{p(\boldsymbol{x})}\int_{R_1} p(\boldsymbol{x}|\omega_2)\,d\boldsymbol{x}+\lambda\left[1-\frac{P(\omega_1)}{p(\boldsymbol{x})}\int_{R_1} p(\boldsymbol{x}|\omega_1)\,d\boldsymbol{x}-\varepsilon\right]$$

$$=(1-\varepsilon)\lambda+\int_{R_1}\frac{P(\omega_2)p(\boldsymbol{x}|\omega_2)-\lambda P(\omega_1)p(\boldsymbol{x}|\omega_1)}{p(\boldsymbol{x})}d\boldsymbol{x}$$

最小化上式J，因为第一项非负，所以只需要把被积分项为负的 $\boldsymbol{x}$ 值划分给R_1，即

$$Decide\ \ \omega_1\ \ if\ \ P(\omega_2)p(\boldsymbol{x}|\omega_2)-\lambda P(\omega_1)p(\boldsymbol{x}|\omega_1)<0$$

整理得

$$Decide\ \ \omega_1\ \ if\ \ \frac{P(\omega_1|\boldsymbol{x})}{P(\omega_2|\boldsymbol{x})}>\frac{1}{\lambda}=\theta$$

问题得证。

问题三

一、问题描述

在三类分类问题中，每类模式特征矢量均遵循正态分布。三个类别模式矢量的协方差矩阵相同，$\Sigma_1=\Sigma_2=\Sigma_3=\begin{bmatrix}1.2 & 0.4\\ 0.4 & 1.8\end{bmatrix}$。

每类的均值矢量分类为$\boldsymbol{u}_1=[0.1,0.1]$，$\boldsymbol{u}_2=[2.1,1.9]$，$\boldsymbol{u}_3=[-1.5,2.0]$。先验概率相同。

本次作业要求：

① 根据贝叶斯最小错误率分类器给出未知类别特征矢量[1.6,1.5]的类别。

② 画出与模型特征矢量[2.1,1.9]具有相等马氏距离点的轨迹。

二、基本思路

<u>① 对于多类问题的最小错误率规则</u>

可以将错误率决策概率表示为正确决策概率的形式:

$$P(error)=1-P(correct)$$

其中

$$P(correct)=\sum_{i=1}^{C}P(\omega_i)\int_{R_i}p(\boldsymbol{x}|\omega_i)\,d\boldsymbol{x}$$

易知，最小化错误决策概率等价于最大化正确决策概率

$$P(correct)=\sum_{i=1}^{C}P(\omega_i)\int_{R_i}p(\boldsymbol{x}|\omega_i)\,d\boldsymbol{x}$$

$$= \sum_{i=1}^{C} \int_{R_i} p(\boldsymbol{x}|\omega_i)P(\omega_i)\,d\boldsymbol{x}$$

因此，若要最大化正确决策概率只需要使每一个积分项在相应区域取最大即可，即最小错误率规则可以转化为如下决策：

$$Decide\ \ \omega_i\ \ if\ \ p(\boldsymbol{x}|\omega_i)P(\omega_i)取最大$$

<u>② 马氏距离的求解</u>

由马氏距离的计算公式

$$\|\boldsymbol{x}-\boldsymbol{u}\|_{\Sigma^{-1}}^{2} = (\boldsymbol{x}-\boldsymbol{u})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{u})$$

则可以在一个平面上建立一个坐标系，坐标系中的每个点可表示为$[x_1, x_2]$。通过对每一个点计算其到特定点的马氏距离$D = \|\boldsymbol{x}-\boldsymbol{u}\|_{\Sigma^{-1}}^{2}$，即可画出相等马氏距离的点。

三、原理及算法

<u>① 程序设计流程</u>

- 给定未知类别特征矢量 X
- 计算似然函数值：Likelihood(X, Cov, mean)
- 计算每一类的分类函数：F_i = Likelihood(X,)*Pw[i]
- 获取类别：label = index.(max(F_i))

<u>② 计算马氏距离</u>

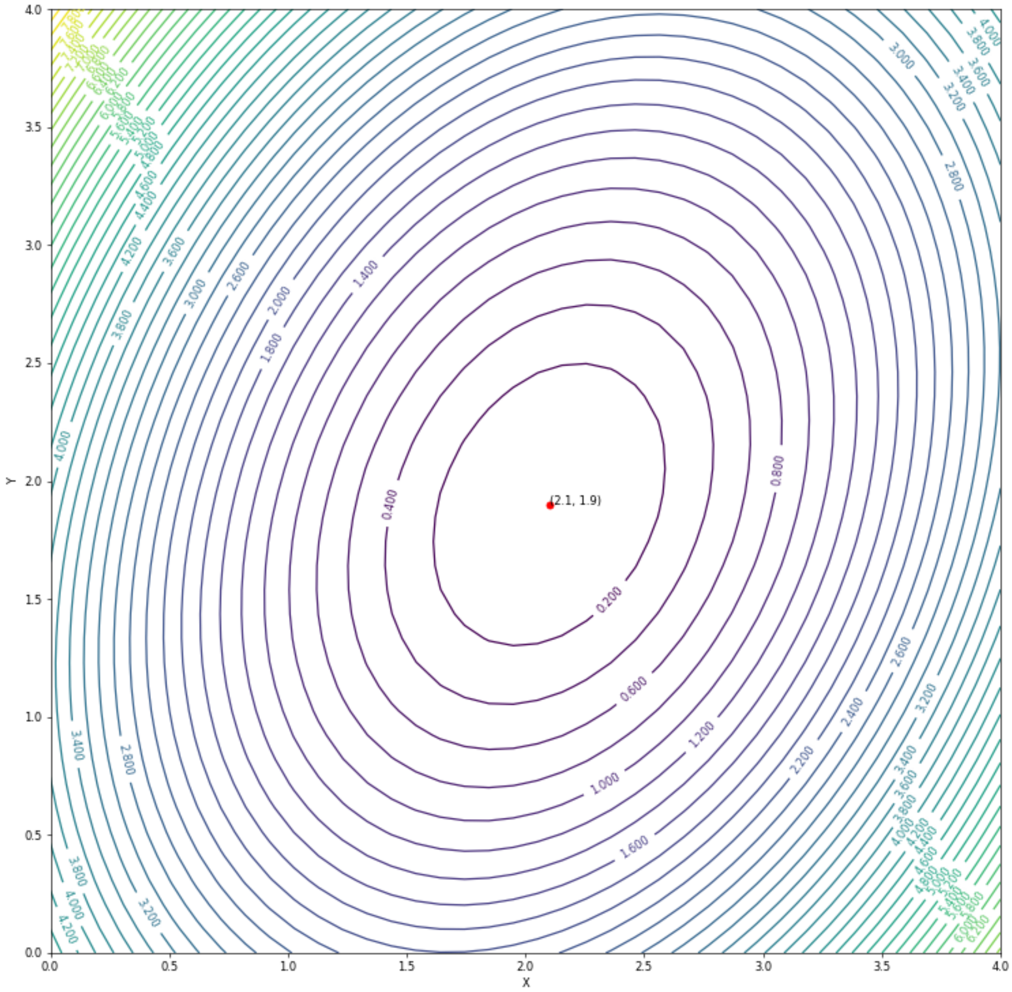
- 生成直角坐标系：y, x = numpy.ogrid[0:4:40j, 0:4:40j] #0 轴为 y 轴，1 轴为 x 轴
- 计算马氏距离函数： Ma_dst = $\|\boldsymbol{x}-\boldsymbol{u}\|_{\Sigma^{-1}}^{2}$
- 画出关于[2.1, 1.9]的马氏距离等值线：pyplt.contour(Ma_dst)

四、实验结果与分析

<u>1、实验结果</u>

```
806ad9-833a-4324-8cd6-18ca4aa74f14' 'RedirectOutput,RedirectOutput' 'd:\Proje
lanobis.py'
---------- classify vector=[1.6, 1.5] ----------
The label of [1.6, 1.5] is: 2
---- Plot the contour lines for [2.1, 1.9] ----
PS D:\Project\VSCodeProject\Python> 
```

由程序运行结果可知，根据贝叶斯最小错误率分类器给出未知类别特征矢量[1.6,1.5]的类别为 2，即平均特征矢量为$\boldsymbol{u}_2$ = [2.1,1.9]的模型。



由上图所示，与特征矢量$\boldsymbol{u}_2=[2.1,1.9]$具有相等马氏距离的点在$x\in[0,4]$，$y\in[0,4]$坐标系上绘制的等值线。

<u>2、结果分析</u>

针对第二问马氏距离所绘制的图形可知。

模型 2 的协方差矩阵为$\Sigma_2=\begin{bmatrix}1.2 & 0.4\\0.4 & 1.8\end{bmatrix}$,

- 由于次对角线上协方差矩阵的值不为 0，因此对于模型 2 中的散点分布，其 x 的分布与 y 的分布是线性相关的。故所绘制的图像呈现椭圆状。
- 由于主对角线上 x 的方差小于 y 的方差，所以散点分布在 x 分支上的分布比在 y 分支上的分布更集中。导致过中心点特征矢量所在直线的斜率大于 1。