# 模式识别第1次作业

周强 202128019427002 电子学院

1. 对于一个c类分类问题，假设各类先验概率为条件概率密度为(x表示特征向量);将第j类样本判别为第i类的损失为。
   1. 请写出贝叶斯风险最小决策和最小错误决策的决策规则；

答：第类的后验概率为

其中。

最小错误决策即选择后验概率最大的类别，即

第i类的条件风险是

最小风险决策即选择条件风险最小的类别，即

* 1. 引入拒识（表示为第c+1类），假设决策损失为

请写出最小损失决策的决策规则，(包括分类规则和拒识规则）。

答：当时，

当时，

当第类的风险最小时拒识，即

即

不满足拒识条件时，则输出分类结果，选择条件风险最小的类别，即

综上所述，分类规则如下

1. 对于特征维数为1维的二分类问题，假设，且。
   1. 证明最小错误率为

其中

答：不妨设，则判定规则如下：

则类错判为类的错误率为

设，则

同理，类错判为类的错误率为。则总的错误率为

同理可证，当，。

综上，最小错误率为

其中，得证。

* 1. 利用不等式

证明当趋于无穷大时，趋于0。

答：当时，a趋于无穷大，则

1. 对一个c类分类问题，特征向量，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

答：设，则

其中是d维特征向量，是类高斯分布的均值向量，是协方差矩阵。

* 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;
     2. 所有类协方差矩阵相等.

答：由各类先验概率相等可知

最小错误率的决策判别函数为

当类协方差矩阵不等时，

省略判别函数中与无关的常量有

此时，判别函数是关于的二次型。

当所有类的协方差矩阵相等时，基，判别函数如下

在这种情况下，判定规则可以简化为：计算向量到每一个均值向量的平方马氏距离，并将其归类为获得最小平方马氏距离的类别。

将上述二次型展开，并去掉与无关的项，可得线性判别函数，表达式如下：

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。

答：为了避免协方差矩阵奇异带来的计算问题，可以采用如下两种措施。

1. 给每个类的协方差矩阵加一个小扰动，即
2. 降维。协方差矩阵奇异的原因是的秩小于特征向量的维度。因此只需要做PCA降维将原始的维空间降到维空间中，使得新空间中的协方差矩阵是非奇异的即可。
3. 请说明ZCA白化的作用，推导ZCA中的去相关变换矩阵，并分析与PCA的异同。

答ZCA白化的作用是降低输入数据的冗余性，使得不同特征之间的相关性较低，同时具有相同的方差。PCA白化保证所有的特征分布均值为0，方差为1；ZCA白化则是均值为0，方差相同。

为了推导ZCA白化，先说明PCA白化。设是去中心化的数据矩阵，则协方差矩阵为

其中n是特向向量的维度。设PCA降维的变换矩阵为P，则为变换后的数据矩阵。为了使得在新空间中不同维度的特征向量不相关，需要满足为对角阵，即

因此PCA实际上是将协方差矩阵对角化。为了满足PCA白化中方差为1的要求，要需要对将新数据除以特征值的开方，即

其中。

如此，我们便完成了PCA白化。

ZCA白化则要在PCA白化的基础上将数据映射为原始数据空间，即

1. （编程题）请选择编程题MNIST或CIFAR10中的一个数据集，从中挑选两个类别运用LDF与QDF对其进行两类分类，并对结果分析讨论。

答：实验结果如下

编程语言：Python

数据集：MNIST

第三方库：numpy、sklearn

分别使用线性判别函数（LDF）和二次判别函数（QDF）进行二分类和十分类。实验中使用了PCA降维，原因有三。其一，降维可以大大降低计算复杂度，节约计算资源。其二，原始数据的协方差矩阵可能奇异，无法求逆，使用PCA降维可以降低不同维度数据之见的相关性，使得协方差矩阵可逆。其三，使用QDF时需要计算协方差矩阵的行列式，如果其维数太高，行列式太大，可能超出double型变量的存储范围导致分类器性能下降。

下表是二分类结果，此任务很简单，因此两种方法的准确率都很高。

二分类结果

|  |  |
| --- | --- |
|  | 准确率 |
| LDF | 0.999 |
| QDF | 0.999 |

为了对比两种分类器的性能，进行十分类实验，结果如下表。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 降维数 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 不降维 |
| LDF | 0.7748 | 0.8455 | 0.8592 | 0.8679 | 0.8757 | 0.8759 | 0.8752 | 0.8769 | 0.8713 |
| QDF | 0.8918 | 0.9503 | 0.9607 | 0.9627 | 0.9636 | 0.9616 | 0.9617 |  |  |

纵向对比不同的方法可知，QDF的性能更好。这是因为LDF规定不同的类别使用相同的协方差矩阵的假设可能不合理。横向对比不同的降维数可知，当维数较低时，随着维数的增加，分类器的性能有较大的提升，这是因为降维到较低维度时，丢失的有用信息太多。随着维数的增加，分类器的性能趋于稳定。由此说明，适当的PCA降维可以再性能几乎不下降的情况下，大大降低计算复杂度。