作业2 判别式分类器

作业提交截止日期：10月25日

1. 一个三类问题，其判别函数为：

（1）设这些函数是在多类情况(1)条件下确定的，绘出其判别界面和每一个模式类别的区域。

（2）设为多类情况2，并使：, 。绘出其判别界面和多类情况2的区域。

（3）设和是在多类情况(3)的条件下确定的，绘出其判别界面和每类的区域。

答：略

2. 两类模式，每类包括5个3维不同的模式向量。

（1）如果它们是线性可分的，权向量的维度是多少？

答：线性模型的权向量的维度同输入模型的维度为3。还有额外的一个偏差参数。

（2）假如要建立二次的多项式判别函数，权向量的维度是多少？

答：维向的次多项式，权系数的总项数为

3. 用感知器算法求下列模式分类的解向量，并尝试使用2种不同的初始值和2种不同的样本迭代顺序。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 0 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 |
| 0 | 0 | 1 | -1 |
| 0 | 1 | 1 | -1 |
| 0 | 1 | 0 | -1 |
| 1 | 1 | 1 | -1 |

答：略

4. 给定两类样本，其中

类别1有5个样本：

类别2有6个样本：

（1）求这两类的Fisher判别投影向量、判别函数，并在图上画出；

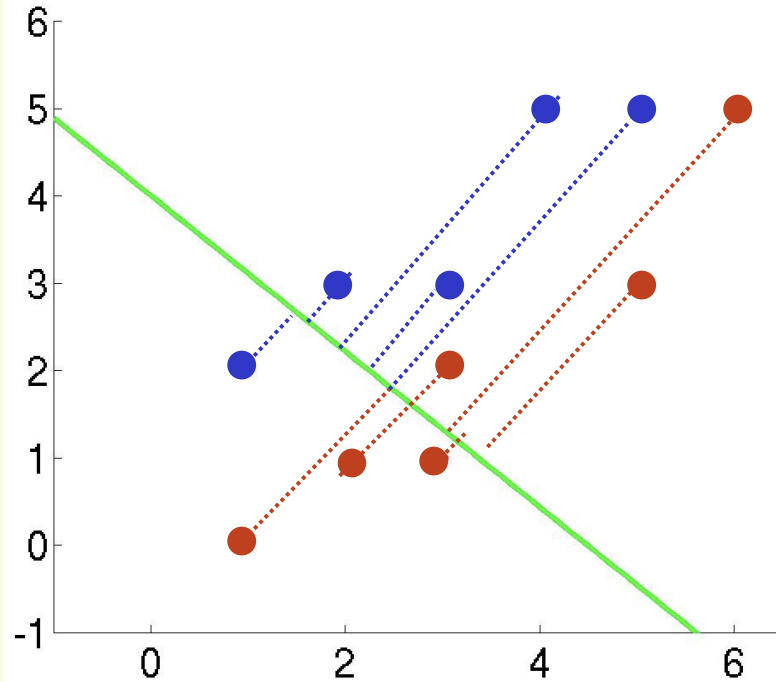
答：

分类阈值为：

判别面方程为：

即。

（2）计算每个样本点投影后的坐标。



5. 对给定的如下样本，根据Gini指数，构造二叉决策树：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 3 | 2 | 5 | 1 |
| 5 | 7 | 7 | 0 |
| 8 | 6 | 9 | 1 |
| 1 | 3 | 2 | 0 |
| 6 | 4 | 8 | 1 |
| 4 | 5 | 6 | 0 |

1. 根结点的Gini指数为多少？

答：（1）

（2）请给出第一个决策规则（选择的特征及阈值）。

答：根据特征 的取值，对所有样本进行排序：得到1，3，4，5，6，8，所以可能的阈值为：，3.5， 4.5，5.5，7。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 1 | 3 | 2 | 0 |
| 3 | 2 | 5 | 1 |
| 4 | 5 | 6 | 0 |
| 5 | 7 | 7 | 0 |
| 6 | 4 | 8 | 1 |
| 8 | 6 | 9 | 1 |

根据每个可能的阈值将数据集划分成左右分支，划分后的阈值分别为：

所以对特征，决策规则为时最佳，划分后的Gini指数为。

对特征类似处理，所以可能的阈值为：，3.5， 4.5，5.5，6.5, 根据每个可能的阈值将数据集划分成左右分支，划分后的阈值分别为：

这里最小的Gini指数大于已有的最佳决策的Gini指数，所以不考虑特征。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 3 | 2 | 5 | 1 |
| 1 | 3 | 2 | 0 |
| 6 | 4 | 8 | 1 |
| 4 | 5 | 6 | 0 |
| 8 | 6 | 9 | 1 |
| 5 | 7 | 7 | 0 |

对特征类似处理，所以可能的阈值为：，5.5， 6.5，7.5，8.5, 根据每个可能的阈值将数据集划分成左右分支，划分后的阈值分别为：

这里最小的Gini指数等于已有的最佳决策的Gini指数，所以第一次的最佳决策规则为或者。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| 1 | 3 | 2 | 0 |
| 3 | 2 | 5 | 1 |
| 4 | 5 | 6 | 0 |
| 5 | 7 | 7 | 0 |
| 6 | 4 | 8 | 1 |
| 8 | 6 | 9 | 1 |