Lecture4习题作业

李星毅 U201712072 自实1701

1，在Logistic regression中当标签y={+1,-1}时的误差函数为：

，请推导出该误差函数的梯度表达式。

解：



其中，

2，已知两类样本的数据如下：





试用Fisher判别函数法，求出最佳投影方向W，及分类阈值y0

解：由题意知：



则可计算出类内离差阵：



从而可计算出最佳投影方向：



每一类的中心：



故可以计算出分类阈值为：



3，已知样本为{（2,10）,（2,5）,（8,4）,（5,8）,（7,5）,（6,4）,（1,2）,（4,9）}，请用K均值聚类法将其聚类成3类，并计算各类的类心。

解：假设随机初始化三类的类心为：



可计算出：属于的样本有，属于的样本有，属于的样本有

1. 更新类心，故可计算出：属于的样本有，属于的样本有，属于的样本有
2. 更新类心，故可计算出：属于的样本有，属于的样本有，属于的样本有
3. 更新类心，故可计算出：属于的样本有，属于的样本有，属于的样本有
4. 类心不更新，，故可认为收敛。

综上所述：利用K均值聚类法聚成的三类分别为：

