

# Homework 4

钟赞

2016K8009915009

## 逻辑回归与线性回归的对比：

### 1. 面向的变量不同

线性回归要求变量服从正态分布，逻辑回归对变量分布没有要求；

线性回归要求因变量是连续性的，逻辑回归要求变量是分类型变量；

### 2. 拟合过程不同

线性回归的拟合函数为 $h_{\theta} = \theta^T$ ，拟合损失为基于最小二乘法的均方误差。

逻辑回归的拟合函数为 $h_{\theta} = \frac{1}{1+e^{-\theta^T}}$ ，拟合损失为基于最大似然函数的对数似然函数。

### 3. 功能不同

线性回归的目的是预测。对于输入  $x$ ，利用求得的拟合函数预测输出  $y$  值。

逻辑回归的目的是分类。对于输入  $x$ ，利用求得的拟合函数判断  $x$  所属的类别。

线性回归分析因变量与自变量的关系，逻辑回归分析分析因变量取某个值的概率与自变量的关系。

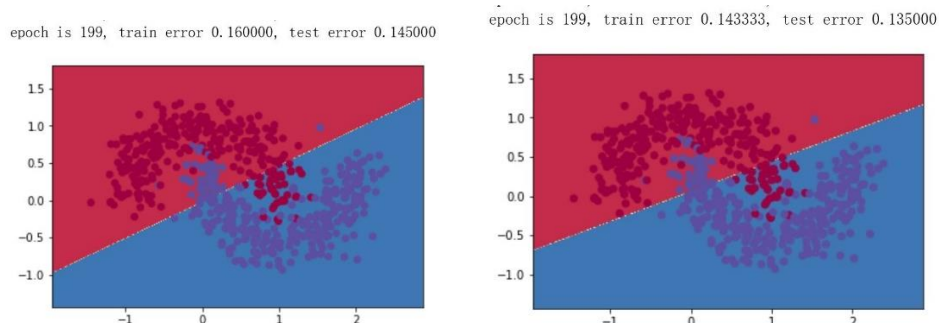


图 1-Logistic Regression for data “moons”

图 2-Linear Regression for data “moons”

由以上两个测试结果可以看出，逻辑回归的曲线比线性回归高，更倾向于把两种样本点分类，而线性回归倾向于综合两种自变量带来的共同影响。

#### 4. 逻辑回归对异常点更加鲁棒

线性回归考虑所有样本点到分类决策面的距离,因此对于两类数据分布不均匀或者有异常点的情况下会降低预测精度;逻辑回归的所有数据都被 sigmoid 函数进行了非线性映射,因此离分类决策面较远的样本点的作用削弱,较少受到异常点的影响。

#### 5. 线性回归对于非线性的关系预测较差

由图 3 和图 4 可以看出,对于非线性的因果关系,逻辑回归和线性回归的预测结果都不够好,相比之下,逻辑回归的测试结果更优一些。

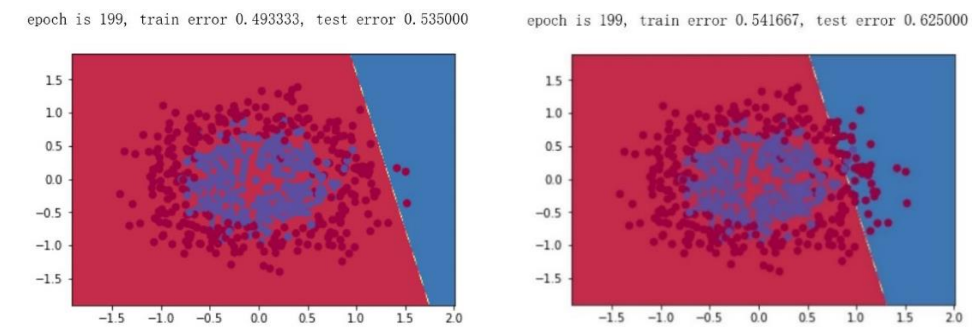


图 3-Logistic Regression for data "circles"

图 4-Linear Regression for data "circles"