## 实验目的：

实现对数几率回归算法对数据分类，进行线性判别分析

## 实验内容：

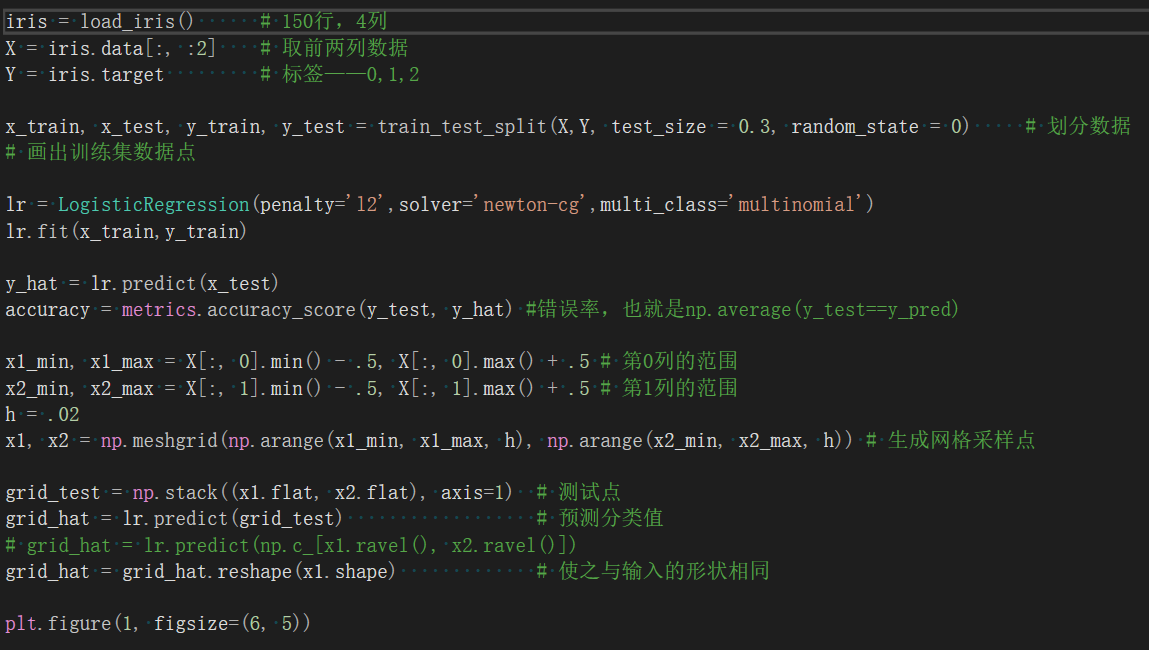
1. **多元线性回归（matlab实现）**

使用iris数据集，将三种花的类别设为1，2，3，将它们打乱并以7：3的比例划分为训练集和测试集。挑选训练集的第一第二条属性进行回归。



1. **对数几率回归算法（Python实现）**

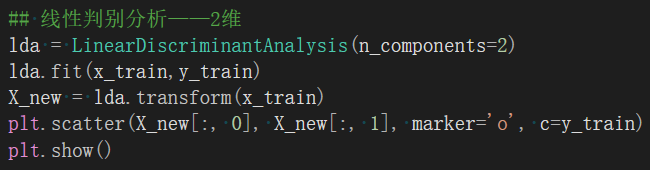
导入鸢尾花数据集，它们的标签为0，1，2，分别代表山鸢尾、变色鸢尾、维吉尼亚鸢尾。用train\_test\_split函数划分数据。这里可以直接用上一步划分好的数据，调用逻辑回归[LogisticRegression()](http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html)函数，用fit(x,y)的方法进行训练。之后用matplotlib包自带的绘图方法进行散点图和区块划分的绘制。

****

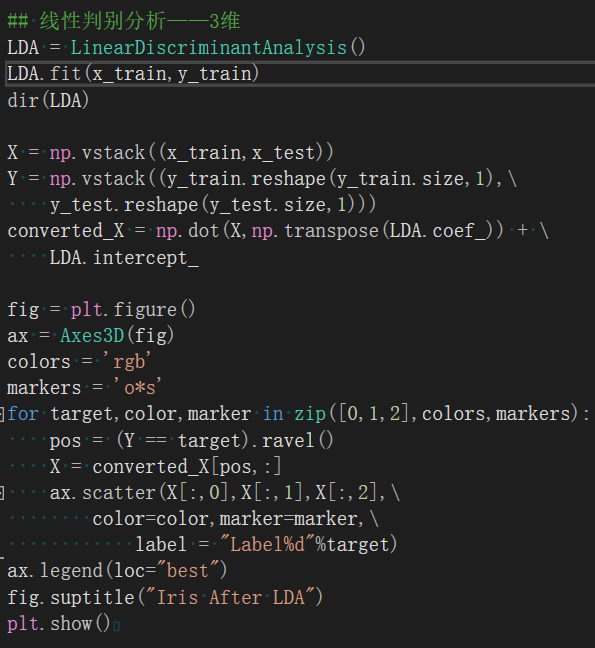
1. **线性判别分析（python实现）**

用刚才划分好的训练集与验证集，通过LinearDiscriminantAnalysis()函数，用fit(x,y)的方法进行拟合，画图对拟合后的结果进行展示。

二维：

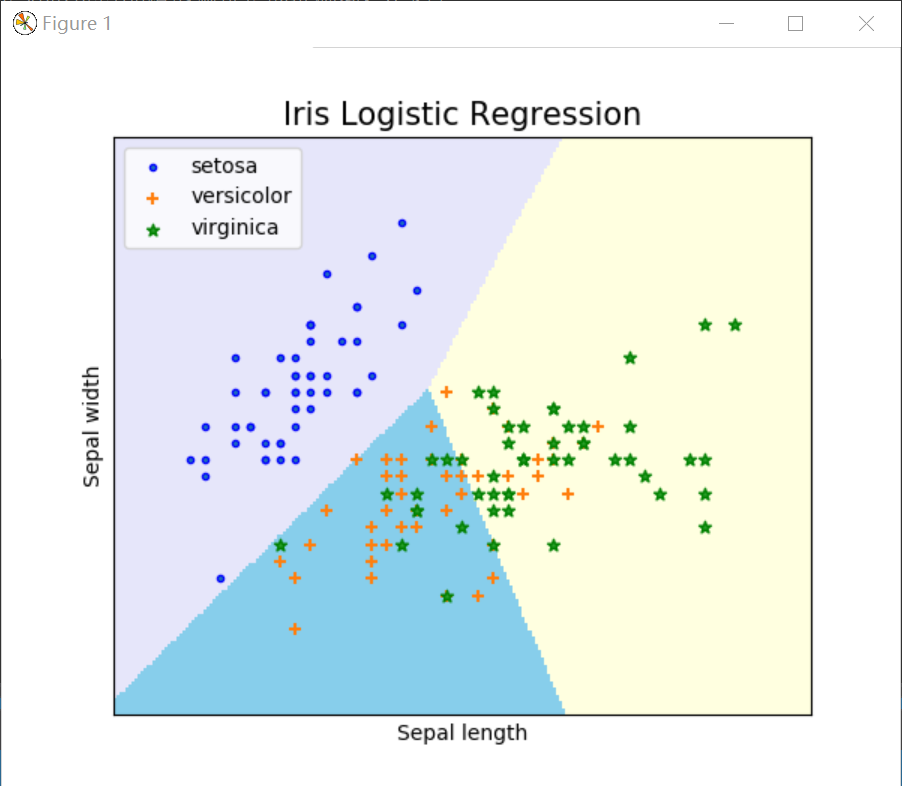


三维：



## 实验结果：

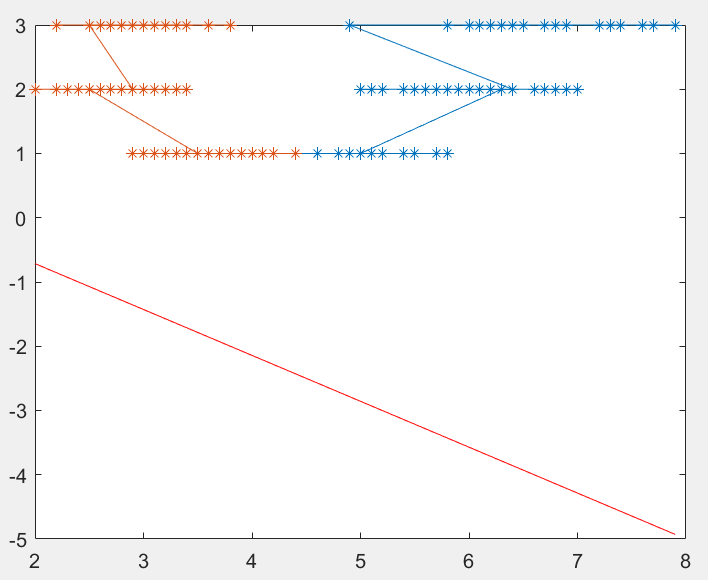
1.对数几率

****

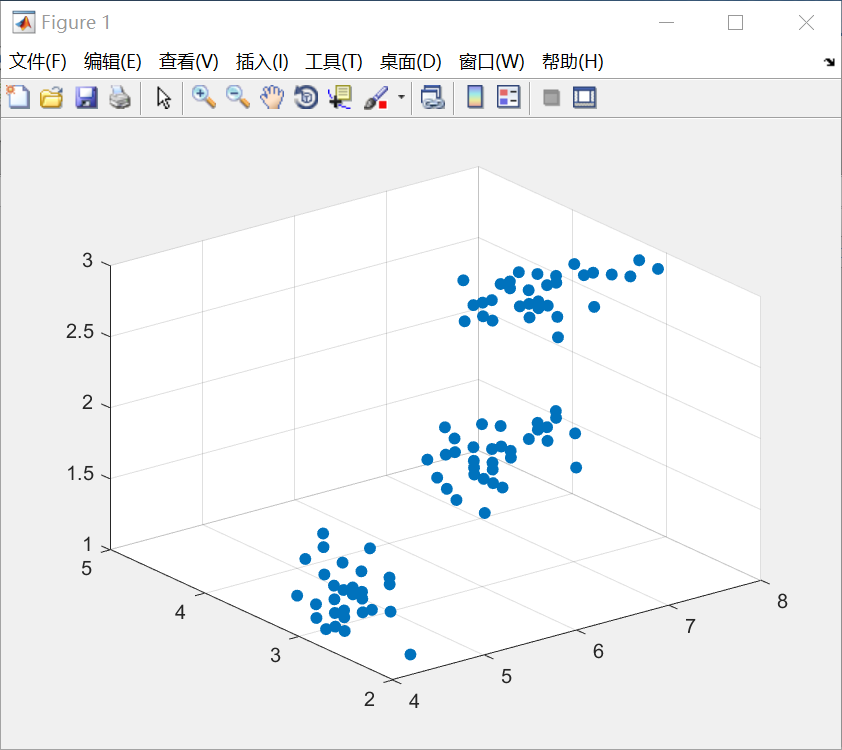
1. 多元线性

二维：

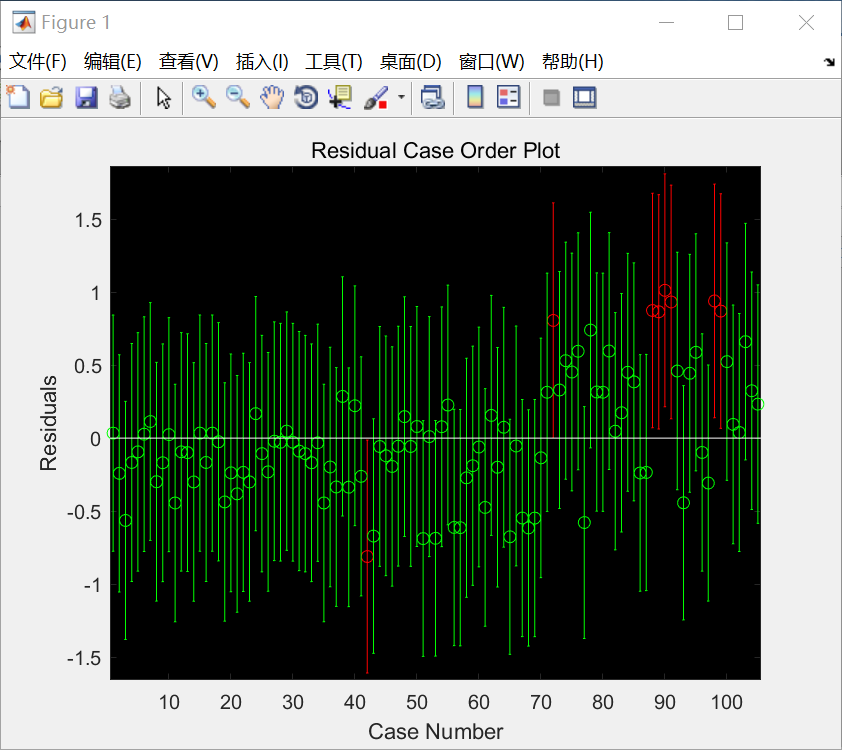
这里应该是数据的问题，导致前两个属性划分不开



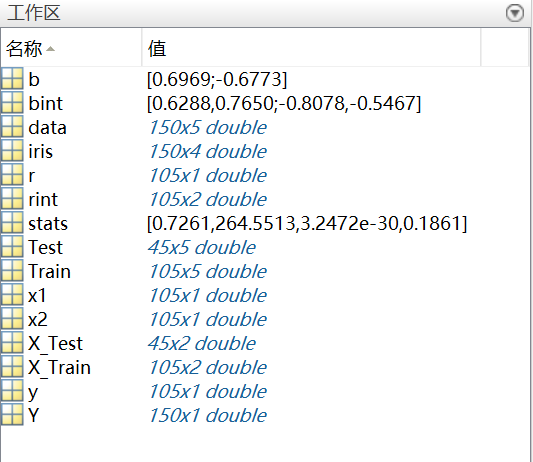
三维视图：



残差分析：

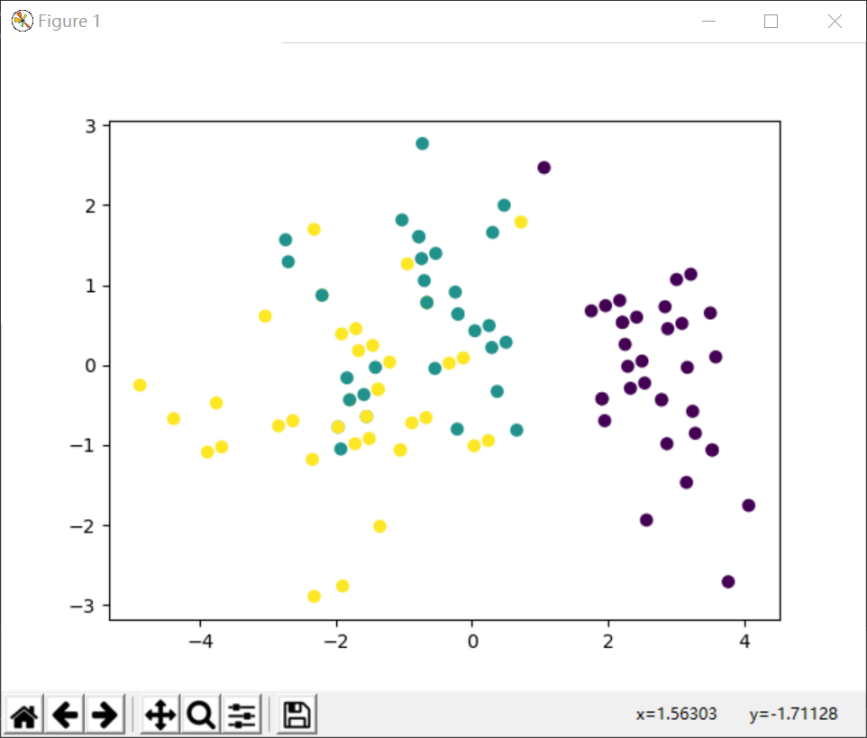


得到的相关参数：

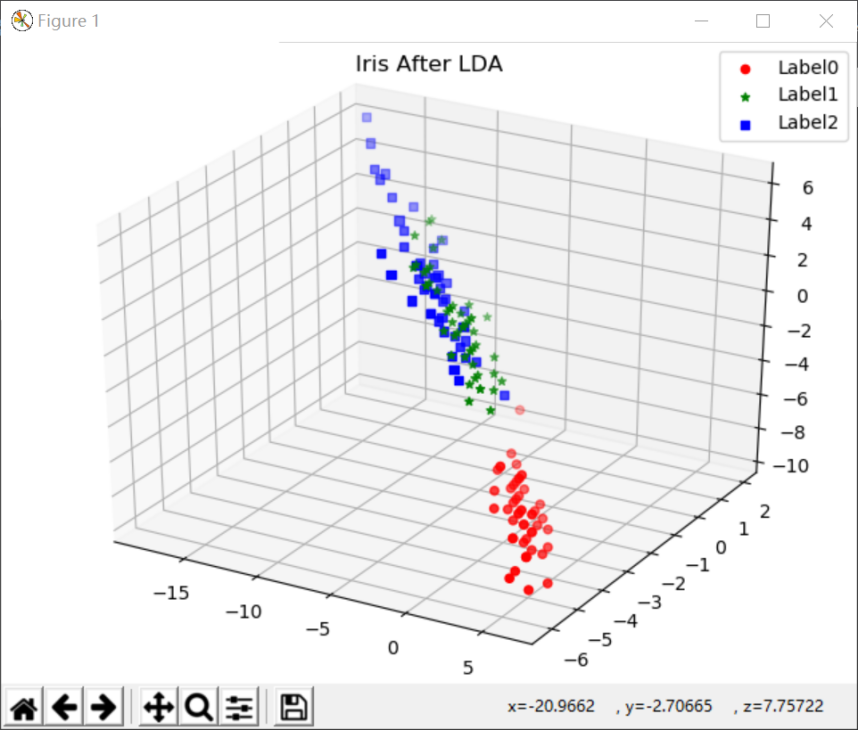


1. 线性分析

二维：



三维：



## 实验结论：

逻辑回归的思想就是通过数据集找到一条决策边界，它是用线性回归模型的结果取逼近真实标记的对数几率。

线性判别分析的目标是使同类别之间的协方差最小，不同类别之间的协方差最大，可以用于数据预处理中的降维步骤。