在给定训练集（train.txt）和测试集(test.txt)的基础上，

（1）实现Logistic回归分类算法。假设模型最高阶数为6，因此，在读入数据后需做特征映射以刻画高阶维度，此外为防止过拟合，需在算法中设计正则化项；

（2）给正则化项系数λ以lamba.txt中列出的不同取值（共50项），进行模型训练，并画出不同λ下模型参数值

的变化情况。为便于观察，可将横坐标取为 log10λ ，画出的图应大致如图1所示；

（3）相应地，绘制出不同λ下模型在训练集本身和测试集上的分类准确率（accuracy），应大致如图2,3所示；

（4）你也可以考虑对训练集进行10折交叉验证（cross validation）来获取更具普遍意义的准确率评估，此时画出的图大致应如图4所示；

（5）请思考并阐述你在此次实验基础上对过拟合和正则化系数的理解。

注1：第（4）小题为选做，如完成此小题，可使用KFold包方便地实现交叉验证；如不做此小题，可直接参考图4；

注2：不得调用除 numpy, pandas, matplotlib, scipy.optimize (用于半自动化参数寻优）和KFold外的其他包。

