# 模式识别实验报告

## 实验三 线性分类器

学院：计算学部

姓名： 杨富祥

学号：21S003071

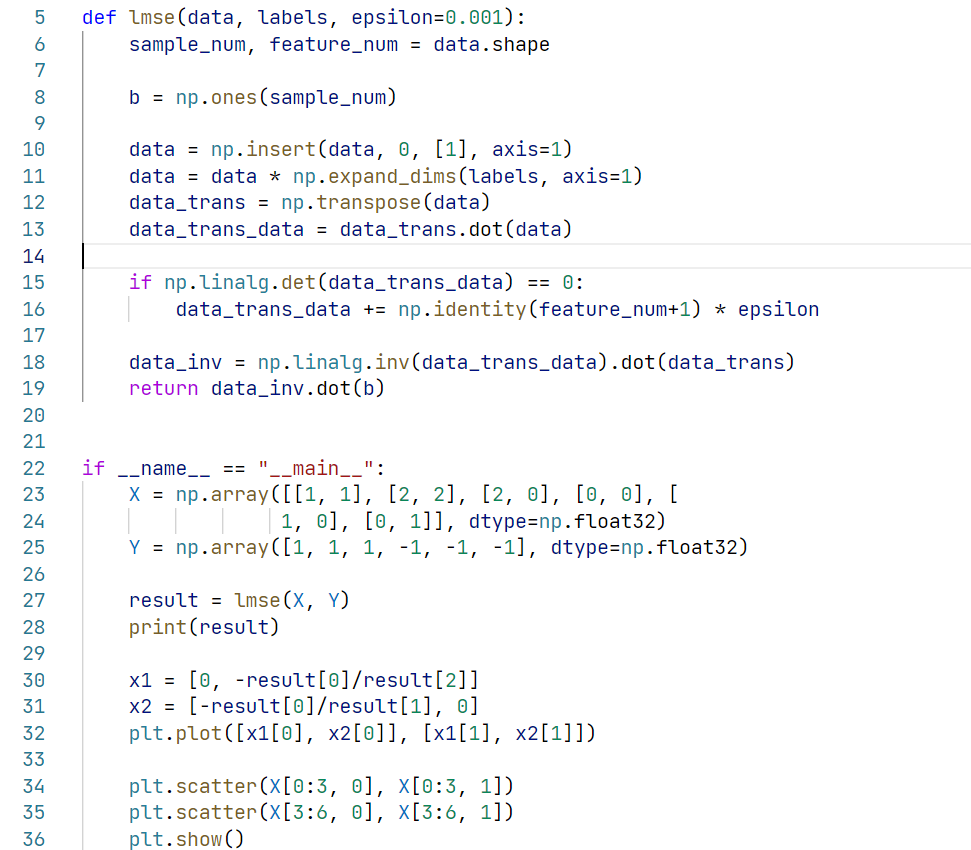
1. **实验内容**
2. 使用Python或Matlab编程实现感知器算法和最小平方误差算法；
3. 分别使用感知器算法和最小平方误差算法学习区分下列两类样本的线性分类器：



1. MNIST-17数据集实验
   1. MNIST-Train-Samples-17.csv中包含30000个17维特征手写数字训练样本，MNIST-Train-Labels-17.csv中包含训练样本的标签；
   2. 使用最小平方误差算法学习区分10个类别的线性分类器；
   3. 测试线性分类器对MNIST-Test-Samples-17.csv中10000个样本的识别正确率。
2. MNIST数据集实验
   1. MNIST-Train-Samples.csv中包含60000个手写数字训练样本，MNIST-Train-Labels.csv中为训练样本的类别标签，MNIST-Test-Samples.csv包含10000个测试样本；
   2. 学习SVM分类器区分10个类别，可以使用函数库中的相应SVM学习函数；
   3. 调节SVM分类器的超参数，包括核函数及核函数参数，调和系数，报告实验找到的最优参数及最优的分类正确率。
3. **程序代码**

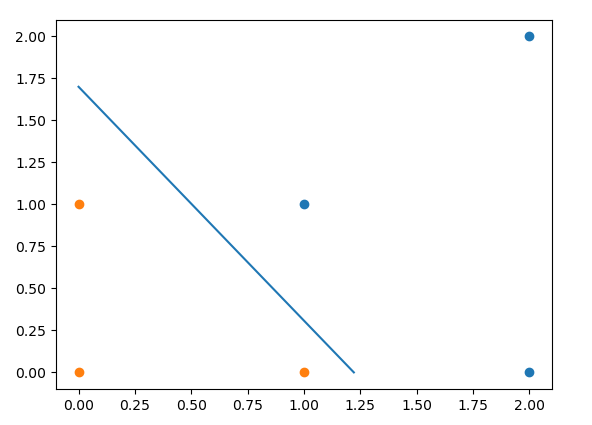
（感知器算法和最小平方误差算法，矩阵乘法和求逆可以调用函数库中的程序）



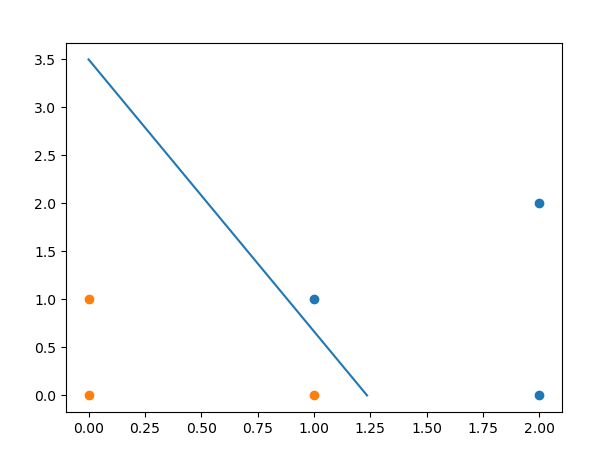


1. **实验结果**
2. 仿真数据实验结果：分别给出使用感知器算法和最小平方误差算法得到的线性判别函数

感知器算法：



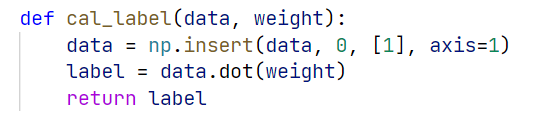
最小平方误差算法：



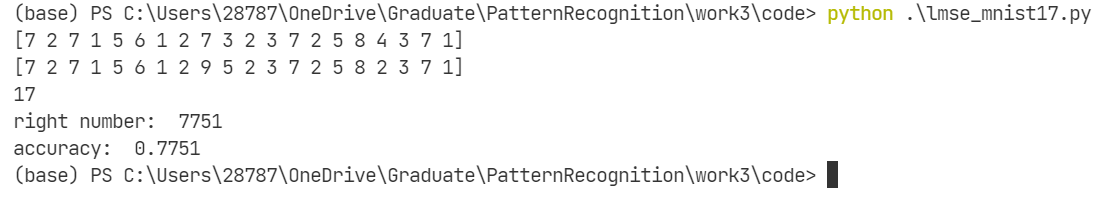
1. MNIST-17数据集实验结果：

如下图所示，在前20个测试用例中，正确的有17个。

在所有的10000个测试用例中，正确的有7751个，正确率为77.51%







1. MNIST数据集实验结果：



实验代码

1）使用kernel='linear'参数

C=1，正确率：无（运行时间过长，未跑出来结果）

2）使用kernel='rbf'参数

C=1，gamma=’scale’，正确率：0.9792

C=100，gamma=’scale’，正确率：0.9833

C=0.01，gamme=’scale’，正确率：0.9257

C=1，gamma=’auto’，正确率：0.1135

3）使用kernel='poly'参数

C=1，gamme=’scale’，degree=3，正确率：0.9771

SVM最优超参数

kernel='rbf'，C=100，gamma=’scale’

最优的分类正确率：0.9833