老师好，我是9191040G0328宋禹润。我完成的作业是分组作业五。截止到目前，我完成了在GMM分类集上编程实现softmax回归和多类感知机，以及在GMM聚类数据集上实现k-均值聚类算法。

首先展示的是softmax回归的分类问题，这是在数据集为GMM6、学习率为0.1，采用梯度下降法，迭代了100次后的结果，从左到右依次为误差、测试集正确率、和分类曲线。

这是基于迭代100次后得到的权重，再迭代100次，得到的结果。我们可以看到在迭代过程中，出现了正确率震荡非常大的现象，针对这个现象，我认为是在迭代进行多次后，过大的学习率导致了权重更新出现了极大的误差，因此在每次更新权重之后，减小权重的值。

这是添加以后的运行结果，可以看到在平稳之后准确率就没有再剧烈变化过。

多类感知机也添加了同样的代码，变得更加平稳

在对GMM4数据集进行分类时出现了分类效果始终不佳的情况，猜测是由于陷入了局部最优解，并且学习率下降，始终无法跳出局部最优解导致了这种情况的产生，我认为理想情况应当为这样，但在实际迭代过程中较好的情况为这样，沿着两类的对称轴划开，依然可以得到总体上较好的正确率，但是对于这两类就存在很大的问题。对于该问题我的解决思路是对一开始的theta设置时，尽可能的随机设置，但用正态分布尝试后，效果并不好。即使是最有解依然不能达到理想的结果。

K-均值聚类算法的结果如图所示，可以看到，存在着错误地将某一类分为了两类甚至三类，针对这个问题，解决方法为，一开始构造聚类中心是，尽可能地使聚类中心分散开，使他们地欧氏距离尽可能地大以此来保证聚类结果不会出现错位。

我的思路是，进行多次随机，得到多组初始聚类中心的值，求出每一组之间最小的距离值，再取这些最小值中最大的那组就是我们要求的初始聚类中心，就尽可能地保证了，初始值之间距离够远。

采用这种方法后，经过多次测试，虽然仍会出现错误分类的情况，但是出现的频率降低，可以认为是有效的。

以上就是我的汇报结果，谢谢老师，