**实验目的：**

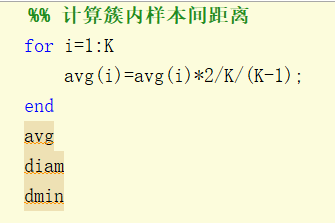
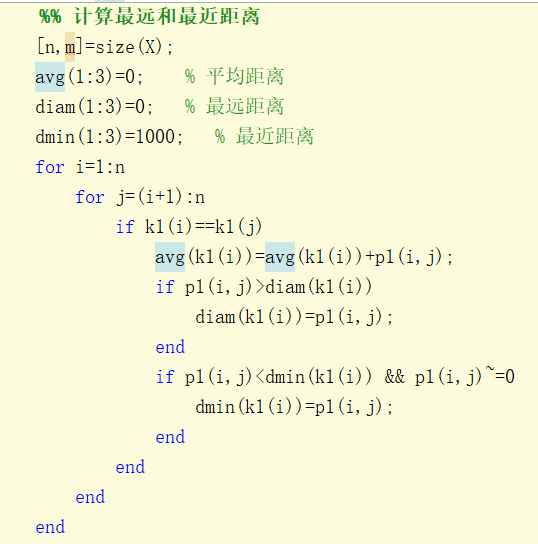
* 计算簇内样本间的平均距离、最近距离、最远距离、簇与簇中心点之间的距离
* 计算DB指数
* 计算VDM
* 计算MinkovDM

**实验内容：**

* 通过二重循环，遍历根据实验6中获得的p=1时的曼哈顿距离，计算每个簇的最远距离diam和最近距离dmin，还有平均距离avg。
* 根据Kmeans函数返回的聚类质心位置C，计算簇与簇中心点之间的距离dcen。
* 根据计算得到的平均距离avg和中心点之间的距离dcen，计算DB指数。
* 统计iris数据集某一属性的取值为a、b的样本数，统计簇i中某一属性的取值为a、b的样本数，计算VDM
* 根据前面计算得到的p1和VDM，相加计算出两两点之间的MinkovDM值。

**实验结果，截图并解释：**

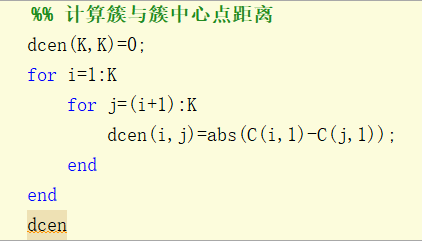
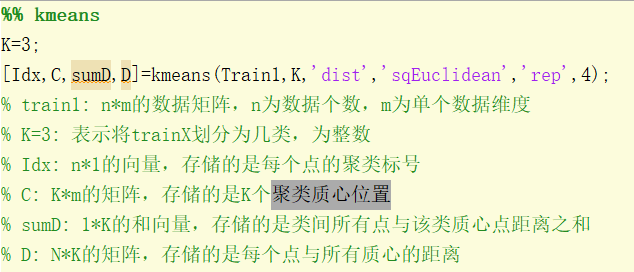
1. 任务4-1：



二重循环遍历曼哈顿距离p1数组；

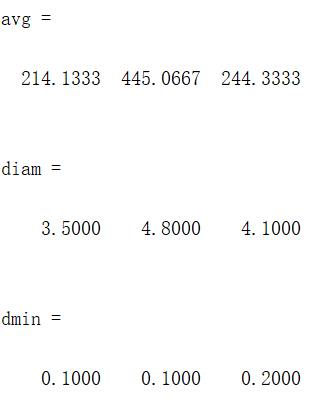
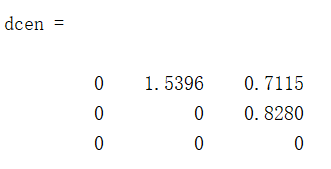
当k1(i)==k1(j)的时候，表示他们在同一个簇中，可以对平均距离先进行相加，然后比较计算最近和最远距离；

循环结束后avg\*2/(K\*(K-1))计算平均距离；

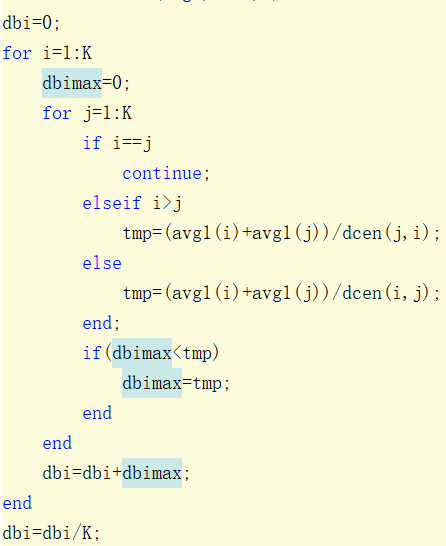


使用kmeans得到的聚类质心位置，计算簇与簇之间距离，dcen(i,j)表示簇i到j的距离。

结果：

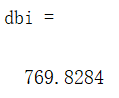
 

1. 任务4-2：



用dbimax暂存每次循环得到的最大值，然后加到dbi中，循环结束后/K。

结果：



1. 任务5-1：

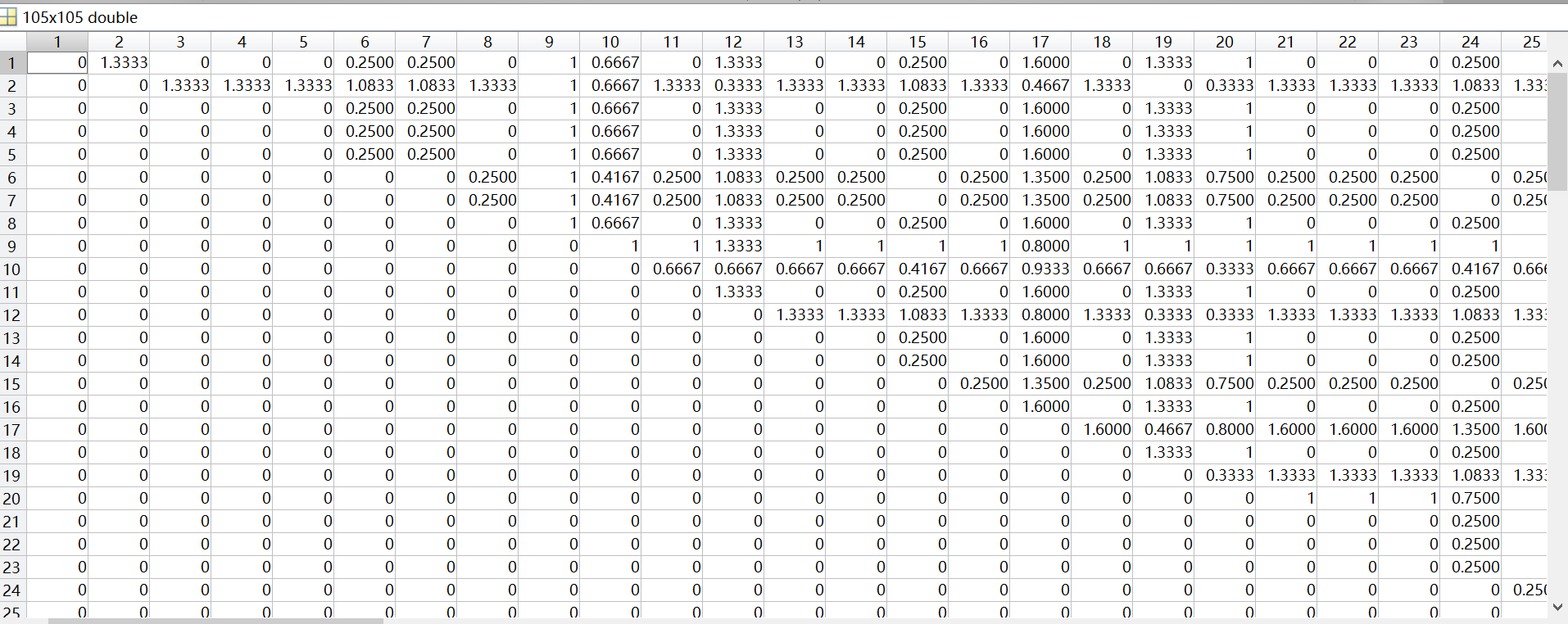


numai、numbi分别表示第i个样本簇，属性j上取值为a和b的样本数；

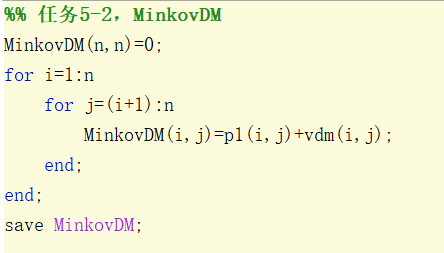
Numa、numb分别表示所有样本中属性j上取值为a和b的样本数；

这里计算iris数据集第一列属性的VDW

结果：

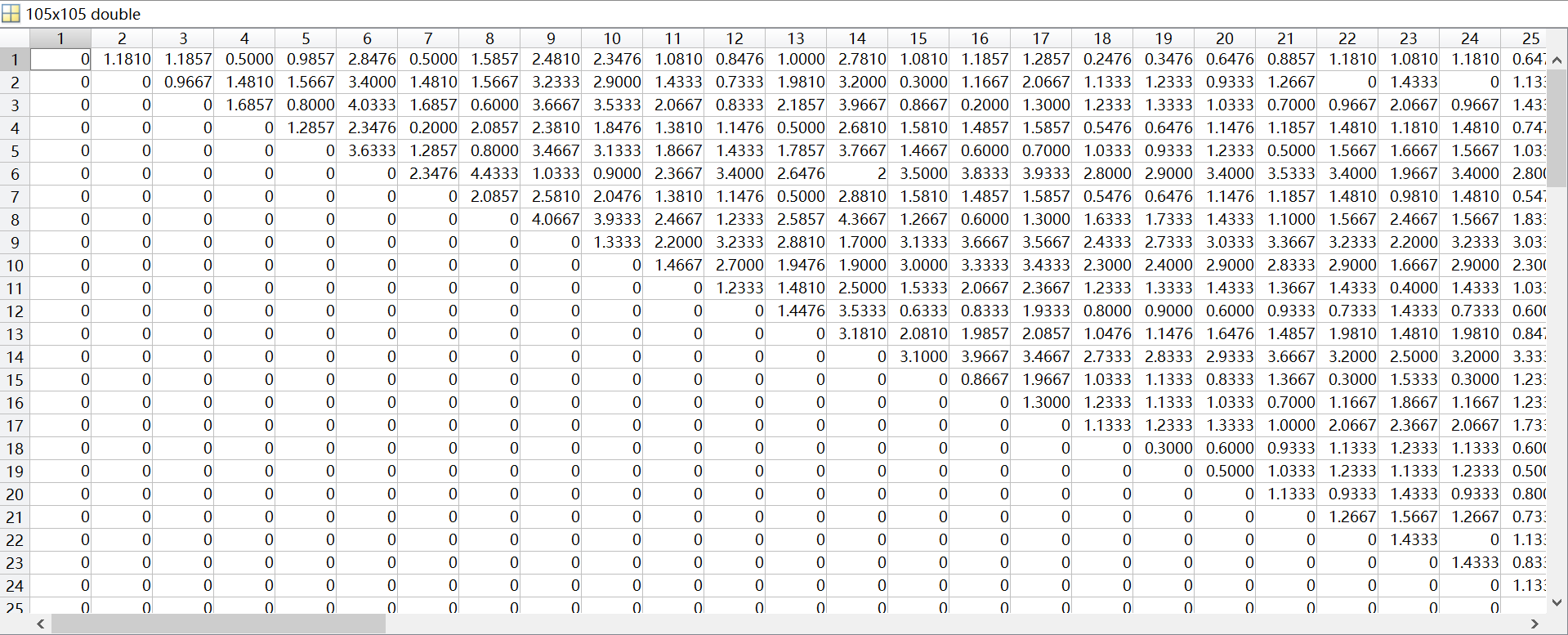


1. 任务5-2：



根据p1和vdm计算出MinkovDM。

结果：



**程序说明：**

导入iris数据集，命令行输入inputdata(iris)即可执行。

1. Inputdata(data)函数进行对iris数据集的划分、kmeans聚类，然后跳转到实现任务的函数clustering;
2. clustering(X,k1,k2,K,C)函数进行任务的总体实现，包含4个函数的跳转Jaccard、d、DBI、VDM。
3. Jaccard(X,k1,k1)函数实现任务1~3，计算Jaccard指数，和闵科夫斯基距离。
4. d(X,K,C,k1,p1)函数实现任务4-1，计算簇内样本平均距离、最远距离、最近距离，簇间中心点距离。
5. DBI(avg1,dcen,K)函数实现任务4-2，计算DB指数
6. VDM(X,K,p1)函数实现任务5，计算VDM和MinkovDM。