

模式识别导论第四次作业——编程题

杨登天-202028015926089-微电子研究所

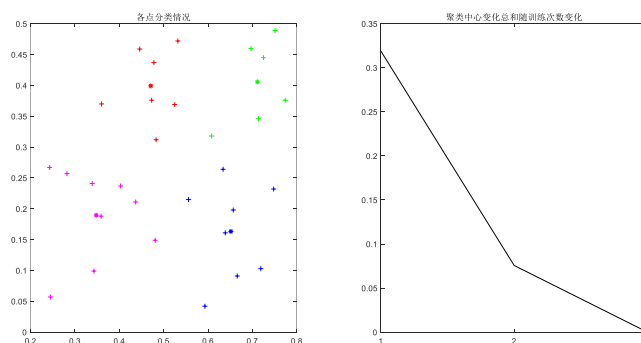
编号	密度	含糖率	编号	密度	含糖率	编号	密度	含糖率
1	0.697	0.460	11	0.245	0.057	21	0.748	0.232
2	0.774	0.376	12	0.343	0.099	22	0.714	0.346
3	0.634	0.264	13	0.639	0.161	23	0.483	0.312
4	0.608	0.318	14	0.657	0.198	24	0.478	0.437
5	0.556	0.215	15	0.360	0.370	25	0.525	0.369
6	0.403	0.237	16	0.593	0.042	26	0.751	0.489
7	0.481	0.149	17	0.719	0.103	27	0.532	0.472
8	0.437	0.211	18	0.359	0.188	28	0.473	0.376
9	0.666	0.091	19	0.339	0.241	29	0.725	0.445
10	0.243	0.267	20	0.282	0.257	30	0.446	0.459

- 对如上的 30 个数据进行 K -均值聚类，聚类个数设置为 $K = 4$ 。
 - 指出所使用的初始聚类中心，并报告在此条件下得到的最终聚类结果以及需要的迭代次数，对应的误差平方和。
 - 重新选择 3 组不同的初始聚类中心，给出对应的聚类结果和误差平方和。
- 对上述数据集进行模糊 K -均值聚类，聚类个数设置为 $K = 4$ 。指出使用的初始聚类中心、初始隶属度，报告在此初始化条件下的聚类结果（即：样本属于不同聚类的隶属度）以及需要的迭代次数。

解答：

任务一：对如上的 30 个数据进行 K -均值聚类，聚类个数设置为 $K = 4$

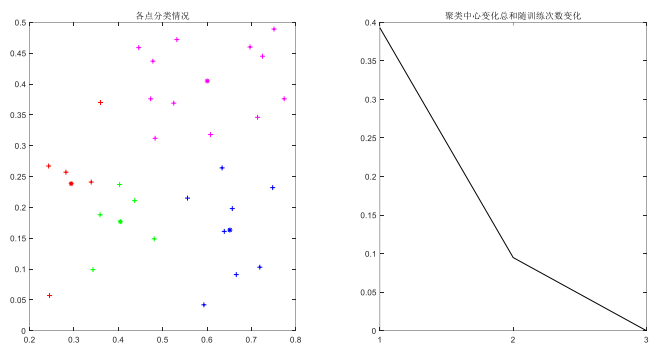
- 使用初始聚类中心为第 2、20、24、25 组数据。（初始聚类中心由随机组序号生成）其最终聚类结果如下所示



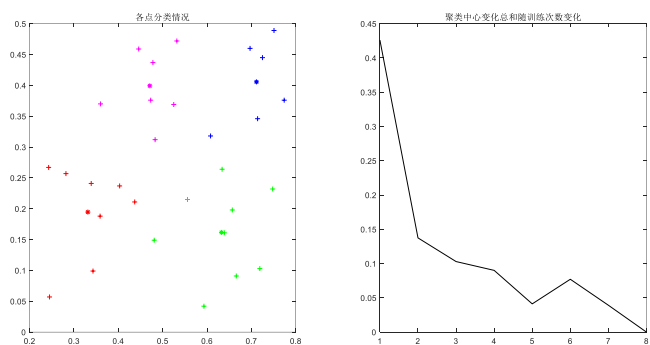
共迭代 3 次，最小误差平方和从 0.318→0.075→0

- 重新选择 3 组不同的初始聚类中心依次得到如下结果

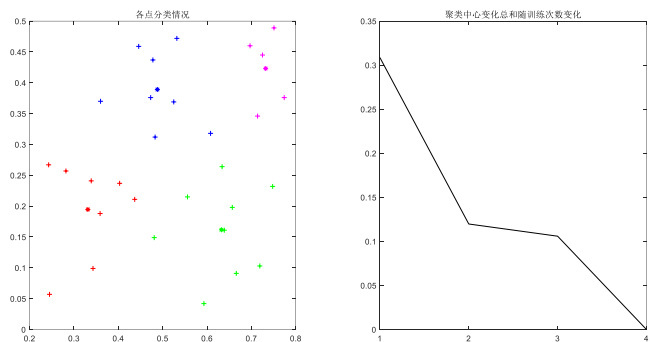
第 1 组：初始聚类中心为第 8、10、7、28 组数据，共迭代 3 次，最小误差平方和从 0.392→0.094→0



第 2 组：初始聚类中心为第 9、11、21、18 组数据，共迭代 8 次，最小误差平方和从 $0.426 \rightarrow 0.137 \rightarrow 0.102 \rightarrow 0.090 \rightarrow 0.041 \rightarrow 0.077 \rightarrow 0.039 \rightarrow 0$ 。



第 3 组：初始聚类中心为第 3、15、22、23 组数据，共迭代 4 次，最小误差平方和从 $0.308 \rightarrow 0.119 \rightarrow 0.105 \rightarrow 0$ 。

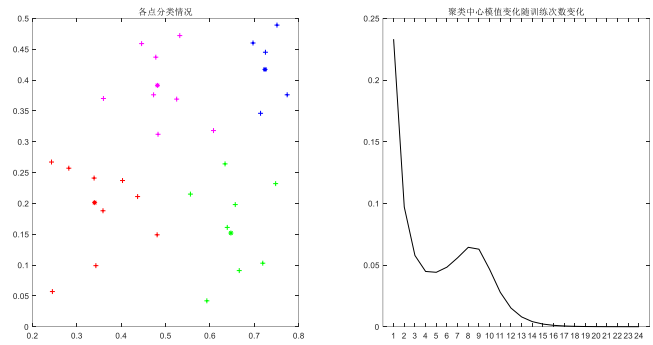


从上面 4 幅图能观察到，最后的聚类结果差距比较明显。

任务二：对上述数据集进行模糊 K -均值聚类，聚类个数设置为 $K = 4$ 。指出使用的初始聚类中心、初始隶属度，报告在此初始化条件下的聚类结果（即：样本属于不同聚类的隶属度）以及需要的迭代次数。

初始聚类中心随机生成，而初始隶属度则由初始聚类中心计算得到。其中参数 b 被设置为 2。初始聚类中心为第 4、11、23、30 个样本。

根据原理可以得到下图。聚类结果依据最大隶属度



最终迭代 24 次。其初始隶属度如下表所示（均采用 3 位有效位截断，0、1 除外）

编号	隶属度 1	隶属度 2	隶属度 3	隶属度 4
1	0.214	0.039	0.516	0.230
2	0.204	0.047	0.588	0.158
3	0.118	0.015	0.826	0.040
4	0	0	1	0
5	0.410	0.049	0.454	0.084
6	0.590	0.123	0.146	0.138
7	0.438	0.181	0.260	0.119
8	0.586	0.119	0.177	0.117
9	0.293	0.135	0.439	0.131
10	0.281	0.380	0.123	0.214
11	0	1	0	0
12	0.129	0.739	0.071	0.059
13	0.285	0.081	0.526	0.106
14	0.238	0.054	0.614	0.091
15	0.375	0.062	0.108	0.453
16	0.307	0.215	0.342	0.133
17	0.275	0.120	0.467	0.136
18	0.358	0.365	0.139	0.136
19	0.421	0.254	0.138	0.184
20	0.324	0.340	0.127	0.207
21	0.215	0.058	0.611	0.115
22	0.157	0.028	0.713	0.101
23	1	0	0	0
24	0.083	0.006	0.042	0.867
25	0.524	0.014	0.277	0.183
26	0.226	0.052	0.471	0.249
27	0.173	0.019	0.164	0.642
28	0.564	0.015	0.109	0.310
29	0.211	0.042	0.540	0.206
30	0	0	0	1