

# 数学建模实验报告——聚类分析

杜明浩 0108170318

资源勘查工程 1703 班 地球科学与信息物理学院

## 1 实验问题

2007 年江苏省的 13 个地市的国民经济分布规律，在众多衡量经济水平的指标中我们将采用下列指标： $x_1$ ：年末户籍人口（万人）； $x_2$ ：城镇化率（%）； $x_3$ ：地区生产总值 GDP（亿元）； $x_4$ ：第三产业占 GDP 的比重（%）； $x_5$ ：城镇固定资产投资额（亿元）； $x_6$ ：社会消费品零售总额（亿元）； $x_7$ ：城市居民人均可支配收入（元）； $x_8$ ：恩格尔系数（城市）（%）； $x_9$ ：农村居民人均纯收入（元）； $x_{10}$ ：恩格尔系数（农村）（%），对 2007 年江苏省的 13 个地市的国民经济进行了聚类

地区	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
苏州	624.43	65.6	5700.85	7.4	1704.27	1250.05	21260	37.9	10475	35.7
无锡	461.74	67.4	3858.54	9.1	1180.74	1134.75	20898	39.8	10026	37.6
常州	357.38	60.9	1881.28	18.6	748.89	610.85	19089	35.0	9033	38.0
南京	617.17	76.8	3283.73	11.0	1443.40	1380.46	20317	35.3	8020	37.4
镇江	268.78	59.6	1206.69	24.5	363.73	331.36	16775	38.7	7668	39.4
南通	766.13	48.6	2111.88	35.1	633.94	736.54	16451	38.5	6905	37.9
扬州	459.25	50.2	1311.89	35.3	438.35	418.90	15057	37.9	6586	38.9
泰州	500.70	47.6	1201.82	33.2	347.73	321.07	14940	43.1	6469	38.1
徐州	940.95	45.8	1679.56	36.0	769.59	543.01	14875	34.9	5534	39.0
连云港	482.23	40.5	618.18	36.2	409.56	249.08	13254	38.9	4828	43.7
淮安	534.00	39.9	765.23	34.8	394.91	269.40	12164	38.9	5010	43.2
盐城	809.79	43.7	1371.26	34.1	470.06	433.74	13857	38.5	6092	41.7
宿迁	531.53	34.1	542.00	32.0	256.18	158.87	9468	42.4	4783	46.0

## 2 实验目的

掌握聚类分析的原理和实现

## 3 实验内容

- (1) 掌握聚类分析的基本内容
- (2) 使用聚类分析对上述实验问题进行解答

## 4 实验平台

R 4.0.0

## 5 实验代码

见附件

## 6 实验结果

使用欧式距离（Euclidean）的聚类分析结果如下图所示（图 1）：

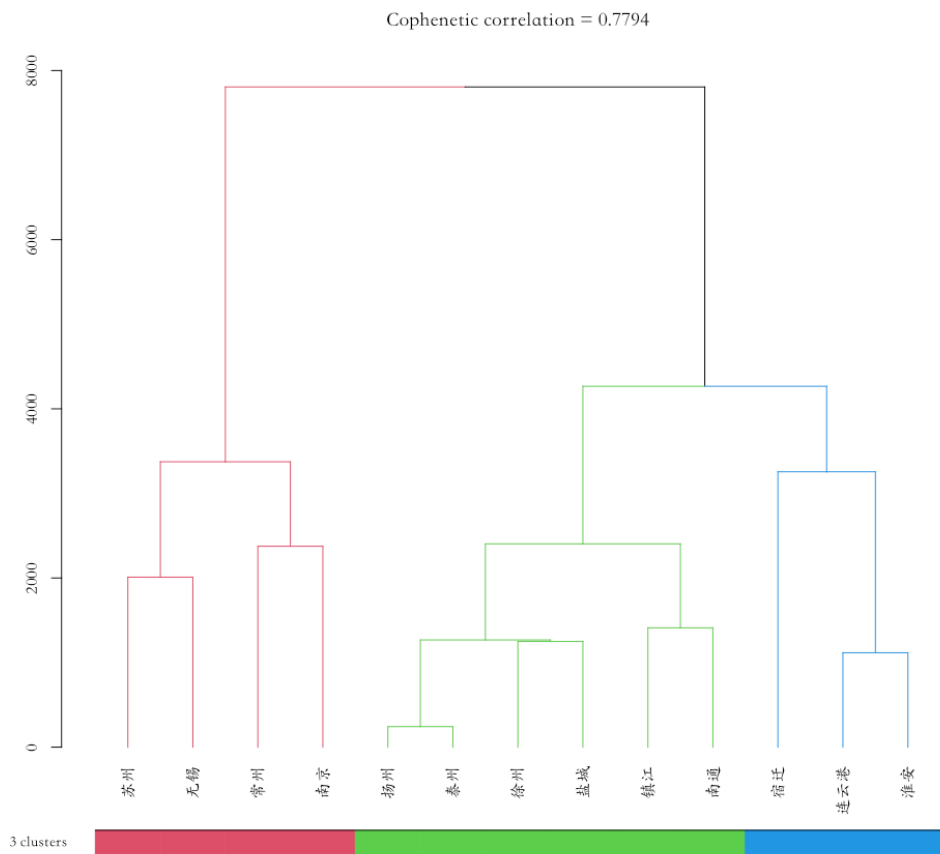


图 1 使用欧式距离的聚类分析结果

使用最长距离（Furthest neighbor）的聚类分析结果如下图所示（图 2）：

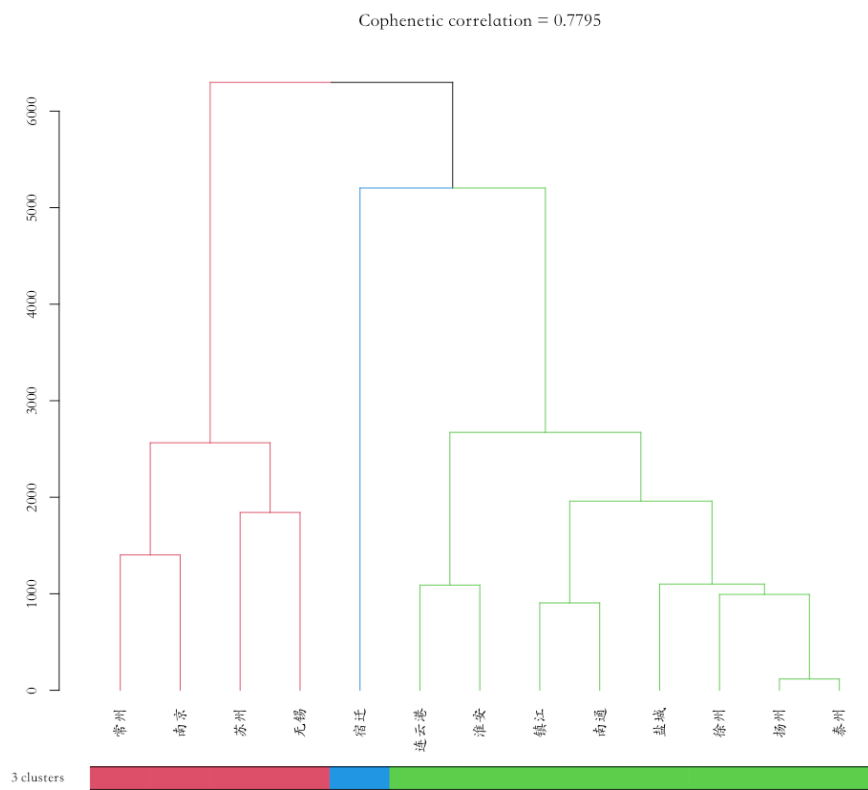


图 2 使用欧式距离的聚类分析结果

欧式距离与最长距离聚类分析结果的比较如下图所示（图 3）：

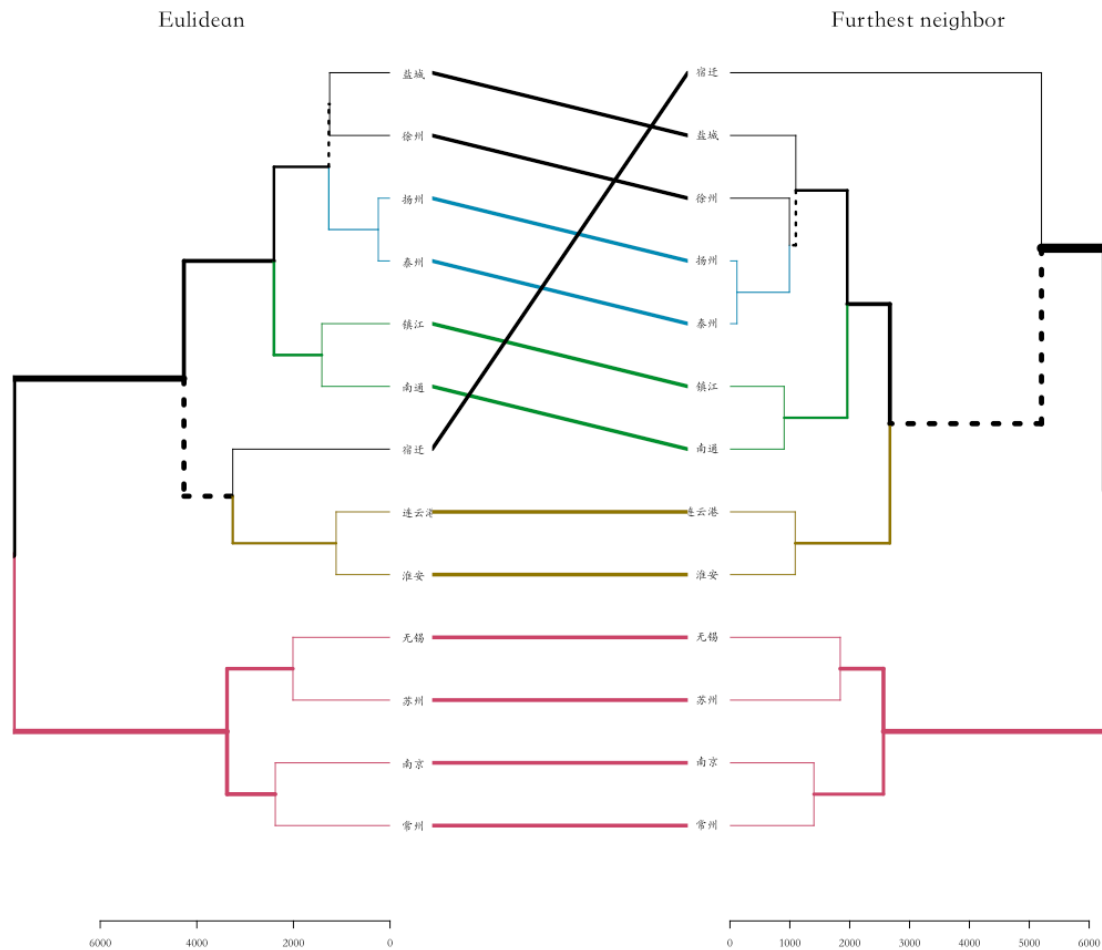


图 3 使用欧式距离和最长距离的聚类分析结果比较

## 7 结果分析

欧式距离将这 13 省分为 3 组，分别为：(1) 苏州、无锡、常州、南京；(2) 扬州、泰州、徐州、盐城、镇江、南通；(3) 宿迁、连云港、淮安。其 cophenetic 相关系数为 0.7794。

最长距离将这 13 省分为 3 组，分别为：(1) 苏州、无锡、常州、南京；(2) 宿迁；(3) 宿迁、连云港、淮安、扬州、泰州、徐州、盐城、镇江、南通。其 cophenetic 相关系数为 0.7795。

相交于最长距离，欧式距离的聚类结果分组比较均匀，因此我们更推荐使用欧式距离。而且通过对这两种方法聚类分析的比较，我们发现，苏州、无锡、常州、南京这四个城市在两种方法中都被分在一组，因此他们四个关系较近。同样

的，连云港和淮安；镇江和南通；扬州和泰州；盐城和徐州在两种方法中都分在一组，两组最大的差别是宿迁。如果最长距离法的连云港、淮安和宿迁关系较近时，两种方法的结果就会较为接近了。

## 8 实验收获

通过本实验我掌握了聚类分析的基本原理，并能通过 R 语言实现聚类分析，收获很大！