



基于 Matlab的水质监测断面优化设置研究

马 飞 蒋 莉

(江苏技术师范学院, 常州 213001)

摘要: 采用 MATLAB统计工具箱中系统聚类分析法编程进行了京杭运河(常州市区段)水质监测断面的优化设置。分类结果直观,优化计算过程体现了 MATLAB监测断面聚类分析程序简便、快速、可视化的优点,此程序可应用于大气、土壤、噪声环境监测站位的优化设置中,也可作为环境管理信息系统的一个子系统,具有良好的推广价值。

关键词: 监测断面; 优化; Matlab

中图分类号: X832 **文献标识码:** B **文章编号:** 1007-0370(2006)03-0048-03

RESEARCH ON OPTIMIZING DESIGN OF WATER QUALITY MONITORING SECTION BASED ON MATLAB

MA Fei JIANG Li

(The Technology and Normal University of Jiang Su Chang Zhou 213001)

Abstract Adopted by systematic accumulation programming under MATLAB statistic tool the optimizing design of water quality monitoring section of Chang Zhou city part of Beijing-Hang Zhou Canal was developed. The result of classification is directly perceived, and the accumulation program of monitoring section by Matlab are simple, rapid and visible through optimizing calculation process. The program can be applied to monitoring sites optimizing design of air, soil and noise environment.

Key words: monitoring section; optimization; Matlab

前言

监测站位布设关系着环境监测工作的成败,优化布点是体现环境监测科学性的重要环节。任何环境要素的监测点常需经历一个由少到多、又由多到少的探索过程。在一定空间范围内布设的测点数量大,虽然能较好地反映环境质量状况,但却需要较高的经济代价。所谓优化就是要合理地组合测点数(含站位数)、代表性(重复性)和监测费用;而优化过程就是最佳测点数的选择过程^[1]。

MATLAB软件中的统计工具箱聚类分析功能强大,并可以把聚类分析结果以图形的形式显示出来,实现了计算的可视化。故可以应用 MATLAB的统计工具箱里的系统聚类软件包编程,对河流水质监测断面进行优化设置^[2,3]。

1 水质监测断面优化设置

系统聚类法是将 n 个元素(样品或变量)看成 n 类,然后将性质最接近(或相似程度最大)的 2 类合并为一个新类,得到 $n-1$ 类,再从中找出最接近的 2 类加以合并变成了 $n-2$ 类,如此下去,最后所有的元素全聚在一类之中。

1.1 建立原始矩阵

下面由京杭运河(常州市区段)水质监测所得的水

质指标建立样本原始矩阵。有 6 个监测断面要进行分
类,每个监测断面有 9 个水质指标,其原始矩阵为:

X=

氟化物	COD	BOD ₅	六价铬	总氮	总磷	氨氮	溶解氧	挥发酚
1.72	80.32	10	0.06	2.62	13.82	2.42	2.22	0.18
1.81	65.60	5	0.06	2.83	14.99	2.54	2.41	0.16
1.93	70.94	5	0.06	2.93	6.97	2.58	3.01	0.14
1.95	65.60	10	0.07	2.92	5.75	2.60	3.34	0.15
2.00	87.01	5	0.06	2.95	2.19	2.62	3.82	0.17
2.02	69.61	5	0.07	2.73	0.88	2.49	2.28	0.17

矩阵中的第一行至第六行分别是中村桥、采菱桥、
龚巷河桥、夏城桥、常武桥、吴家桥的水质指标。

1.2 建立标准化矩阵

因为各水质指标的量纲不同,所以需要由对各监

测断面水质指标值组成的原始矩阵进行标准化,也就
是将各水质指标值除以相应的水质标准值所得到的比
值。水质标准为《地表水环境质量标准基本项目标准
限值》中的Ⅴ级标准。将上矩阵中数据标准化后得到:

X=

氟化物	COD	BOD ₅	六价铬	总氮	总磷	氨氮	溶解氧	挥发酚
1.15	2.11	1	1.2	1.31	34.55	1.21	1.11	1.8
1.21	1.64	1.25	1.2	1.42	37.48	1.27	1.21	1.6
1.29	1.76	1.25	1.2	1.47	17.43	1.29	1.00	1.4
1.30	1.64	1	1.4	1.47	14.38	1.30	1.11	1.5
1.55	1.64	1.25	1.6	1.48	5.48	1.31	1.27	1.7
1.33	2.18	1.25	1.2	1.37	2.2	1.24	1.14	1.7

1.3 Matlab程序设计及运行

1.3.1 寻找变量之间的相似性

用 pdist函数计算计算监测断面两之间的亲疏距
离:

Y= pdist(X), 其中 X为上述标准化矩阵。

1.3.2 定义变量之间的连接

数据中变量之间的亲疏距离计算出来以后,使用
linkage函数就可以决定数据集合中哪一些变量可以归
入某类中。

Z= linkage(Y)

1.3.3 评价聚类信息

计算距离的方法不同,得出的聚类信息是不
同的。反映聚类效果的是相关系数,相关系数越接近 1,
说明两个监测断面的性质越相似,可以归为一类。

(1)修改聚类树

用 cophenet函数来计算相关系数 C:

C= cophenet(Z, Y)

C = 0.8902

将 pdist函数中距离的计算方法分别指定为“Ma-
hal”、“SEuclid”、“Cityblock”重新计算 pdist函数以后,
用 cophenet函数计算,得相关系数 C值分别等于 0.
6016、0.6499、0.8916。其中 0.8916最接近 1,所以使
用将 pdist函数中距离的计算方法分别指定为“City-
block”时的效果最佳,也就是采用布洛克距离计算最为
合理。

(2) cluster函数用不连续性的度量决定在哪里对
数据进行分类。

用 dendrogram函数生成聚类树:

dendrogram(Z)

同时得到一个 Matlab生成的冰柱图,如图 3- 1,纵
轴为变量之间的距离,横轴为聚类变量。此冰柱图即
为监测断面的聚类树,聚类树从左到右反映了聚类的



先后次序, 该图与 linkage函数的计算结果相对应。

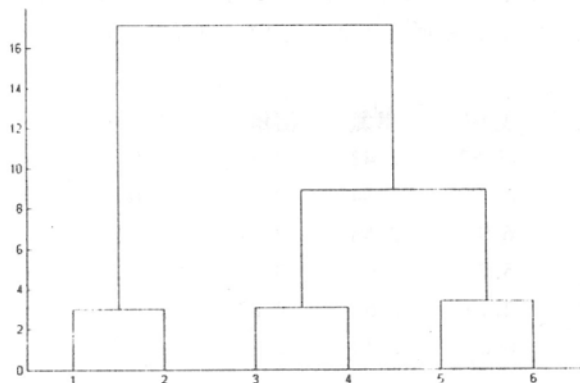


图 3-1 聚类树

1.3.4 创建聚类

有两种方式创建聚类:

(1) 找到原始数据集合的自然分界: 用 cluster函数对数据进行分类, 指定一个不连续系数 0.9 作为 cutoff 参数的值. 则 cluster函数将所有变量分为 3 类。

```
T= cluster(Z, 0.9)
```

```
T= [ 1 1 2 2 3 3]
```

(2) 通过指定类数进行聚类: 如果所有变量归为 3 类, 将 cutoff 参数设置为 3 用 cluster函数进行分类。

```
T= cluster(Z, 3)
```

```
T= [ 3 3 1 1 2 2]
```

可以看出: 无论选择不连续系数为 0.9 或是指定类数进行分类, 通过 MATLAB 系统聚类分析, 都是将监测断面划分为 3 类: 中村桥与采菱桥监测断面归为一类; 龚巷河桥与夏城桥监测断面归为一类; 常武桥与吴家桥监测断面归为一类。

那么, 在监测断面的调整中就可以根据此聚类分析的结果进行优化设置。首先, 要保证每一类中至少

要设一个监测断面, 然后, 可根据实际情况进行取舍。

2 结语

采用 MATLAB 统计工具箱中系统聚类分析法编程进行了运河常州市区河段水质监测断面的聚类分析, 综合利用多个变量的信息对监测断面进行分类, 分类结果直观, 聚类谱系图非常清楚地表现其数值分类结果, 同时聚类分析所得到的结果比传统分类方法更细致、全面、合理。经过运行 MATLAB 监测断面聚类分析程序, 将六个监测断面划分为三类。计算过程体现了 MATLAB 监测断面聚类分析程序简便、快速、可视化的优点, 此程序可应用于大气、土壤、噪声环境监测站位的优化设置中, 也可作为环境管理信息系统的一个子系统, 对大量的环境监测数据进行统计分析, 进而对环境监测站位的布设进行动态管理, 具有良好的推广价值。

参考文献

- [1] 中国环境监测总站. 环境水质监测质量保证手册. 化学工业出版社, 1994 8 10~11
- [2] KEITH LH. Principles of Environmental Sampling [M]. 2nd Edition Washington DC: American Chemical Society 1996
- [3] KEITH LH PATTON GL EDWARDS PG DQO PRO [CP/OL]. distributed by the American Chemical Society Division of Environmental Chemistry 1996
- [4] 苏金明, 阮沈勇, 王永利. MATLAB 与工程数学. 电子工业出版社, 2005 8 110~126