

基于聚类分析预测大坝扬压力的新方法

贾彩虹 顾冲时

(河海大学水利水电工程学院 ,江苏 南京 210098)

摘要 :应用灰色关联度和模糊聚类分析原理 ,对新安江大坝坝基扬压力观测资料进行了分析 ,在分析各测点测值变化规律的基础上 ,利用同一类型的已知测点数据预测了未知测点的数据 ,为进一步分析和研究大坝的工作性态和安全监控提供了依据 .

关键词 :灰色关联度 ;模糊聚类分析 ;观测资料 ;扬压力

中图分类号 :TV314 文献标识码 :A 文章编号 :1000-198X(2002)03-0110-03

大坝的变形性态一般用观测数据来表示 ,但这些数据常受到随机性因素的影响 ,以及人们在认识观测结果、判断观测现象时常使用一些含义不分明的模糊概念 ,因此测值具有精确性和模糊性。“ 当一个系统复杂性增大时 ,使它精确化的能力将减小 ,在达到一定阈值时 ,复杂性和精确性将互相排斥 ” ,这就是大系统的不相容原理 .聚类分析是按某一个标准来鉴别事物之间的接近程度 ,并把彼此接近的事物归为一类 ,为进行分析、决策提供依据 .利用聚类分析对各时段大坝观测数据进行分类后 ,不仅可看出各测点间横向的相互关系 ,而且可以由同类中已知测点的数据来预测未知测点的数据 .

1 原 理

灰色关联度^[1]作为数据序列之间的相似系数 ,形成模糊相似矩阵 ,利用直接聚类法^[2]对测点进行分类 .设聚类论域 $p_i=(p_1,p_2,\dots,p_n)$ 在时间长度 k 内均匀收集反映系统动态变化过程的时间序列数据 .

1.1 关联系数计算

关联系数反映了事物数据系列间的关联程度 ,反映了序列发展趋势和接近程度 ,量化了序列间变化过程的相似度 .设 $\xi_{ik}(k)$ 为参考列和各比较列之间的灰关联系数 ,表示各时间段内相应序列之间的接近程度 .则

$$\xi_{ik}(k)=[\min_i\min_l|x_i(k)-x_l(k)|+\rho\max_i\max_l|x_i(k)-x_l(k)|\wedge|x_i(k)-x_l(k)|+\rho\max_i\max_l|x_i(k)-x_l(k)|]\quad(1)$$

式中 : $x_i(k)$ ——数据系列 ; k ——时间序列 ; ρ ——分辨系数 , $\rho\in(0,1)$.

在一般的灰关联度计算中 ,为简化计算取 $\rho=0.5$,运算中可根据各数据序列间的关联程度对 ρ 值进行调节 ,以增加对比分辨能力 ,其取值可按以下原则进行 .

$$\Delta=(\sum_{i=1}^m\sum_{l=1}^m|x_i(k)-x_l(k)|)\wedge(m\times m)$$

并记 $\epsilon=\Delta/\max_i\max_l|x_i(k)-x_l(k)|$

当 $\max_i\max_l|x_i(k)-x_l(k)|>3\Delta$ 时 , $\epsilon<\rho<1.5\epsilon$;当 $\max_i\max_l|x_i(k)-x_l(k)|<3\Delta$ 时 , $\epsilon<\rho<2.0\epsilon$.实际运算中可采用试算 ρ 值的方法得到较好的对比分辨能力 .

在计算灰关联系数时 ,首先对原始数据进行标准化处理 ,即采用初值法或均值法 ;其次对标准化数据进行一步累加生成来改善数据的拟合光滑性 .关联度分析事实上是动态过程发展态势的量化分析 ,即是各序列间发展态势的量化比较分析 .关联度

$$\gamma_{il} = \frac{1}{t} \sum_{k=1}^t \xi_{il}(k)$$

(2)

1.2 构造关联相似矩阵

由于此相似矩阵元素 γ_{il} 是由不同序列的 $x_i(k)$ 计算所得 ,因参考序列的不同 ,使得对比列与参考列绝对差值的最大值可能不同 ,即 $\max_i \max_l |x_i(k) - x_l(k)|$ 可能不相同 ,故不能保证 $\gamma_{il} = \gamma_{li}$,即关联相似矩阵并不一定是模糊相似矩阵(满足自反性 $\gamma_{ii} = 1$ 、对称性 $\gamma_{il} = \gamma_{li}$) ,此时 ,可对关联相似矩阵进行改造。
可令 $\gamma_{il} = \gamma_{li} = (\gamma_{il} + \gamma_{li})/2$,改造后的关联相似矩阵(见式(3))为模糊相似矩阵 ,表达了聚类论域 p 在具体指标 Y_i 上的动态相似关系 ,反映了系统中时间序列动态变化过程中的局部相似性。

$$R = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ & \ddots & \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{bmatrix}$$

(3)

1.3 直接聚类法

将模糊相似矩阵 $R = (\gamma_{il})_{m \times m}$ 中 γ_{il} 所有不相同的值由大到小依次表示为 $\lambda_k, k = 1, 2, \dots, n$. 对于 $\lambda = \lambda_k$,若 $\gamma_{il} > \lambda_k$,则将 x_i 与 x_l 分为一类 ,若两个类之间存在公共元素 ,则称它们是相连的 . 将所有相连的类合并 ,最后得到的分类 ,即为 λ_k 水平上的等价分类。

2 应用实例

选用新安江大坝坝基扬压力的部分观测资料 ,来验证模糊聚类分析的应用 ,并对个别测点的测值进行预测 . 取河床部分 7[#] , 8[#] , 13[#] , 14[#] 坝段的扬压力观测资料进行分析 . 由于此方法计算简便 ,不需要很长的数据序列 ,因此采用库水位较高的 1983 年和 1999 年的测值进行分析 ,见表 1 2.

表 1 1983 年各坝段扬压力的实测值

Table 1 Observed value of uplift pressure of each dam section for 1983

m

坝段	月 份											
	1	2	3	4	5	6	7	8	9	10	11	12
7 [#]	11.97	11.99	11.94	11.98	11.95	12.01	11.96	11.93	11.85	11.84	11.99	11.92
8 [#]	11.24	11.28	11.27	11.23	11.13	11.16	11.19	11.03	11.07	11.23	11.17	11.19
13 [#]	18.65	18.64	18.63	18.69	18.74	18.83	18.81	18.74	18.67	18.74	18.71	18.72
14 [#]	18.56	18.63	18.62	18.64	18.66	18.64	18.56	18.50	18.63	18.60	18.56	18.60

表 2 1999 年各坝段扬压力的实测值

Table 2 Observed value of uplift pressure of each dam section for 1999

m

坝段	月 份											
	1	2	3	4	5	6	7	8	9	10	11	12
7 [#]	11.99	12.00	11.99	11.97	11.90	11.98	11.95	11.86	11.75	11.67	11.72	11.72
8 [#]	11.42	11.40	11.40	11.37	11.38	11.39	11.42	11.27	11.19	11.06	11.11	11.07
13 [#]	18.68	18.71	18.70	18.70	18.70	18.74	18.79	18.67	18.68	18.61	18.54	18.54
14 [#]	18.62	18.64	18.64	18.62	18.63	18.60	18.63	18.58	18.53	18.46	18.52	18.45

采用初值化对原始数据进行标准化处理 ,

$$x_i'(k) = x_i(k)/x_i(1) \quad k = 1, 2, \dots, 12$$

对标准化数据进行一步累加生成后按式(1)(2)计算各点的相对关联系数 $\xi_{il}(k)$ 和关联度 γ_{il} ,按式(3)构造模糊相似矩阵 : R .

R_1 和 R_2 分别为 1983 年和 1999 年 7[#] , 8[#] , 13[#] , 14[#] 坝段的扬压力实测值计算所得的模糊相似矩阵 :

$$R_1 = \begin{bmatrix} 1 & 0.90 & 0.64 & 0.68 \\ 0.90 & 1 & 0.65 & 0.70 \\ 0.64 & 0.65 & 1 & 0.85 \\ 0.68 & 0.70 & 0.85 & 1 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 1 & 0.95 & 0.69 & 0.72 \\ 0.95 & 1 & 0.69 & 0.72 \\ 0.69 & 0.69 & 1 & 0.89 \\ 0.72 & 0.72 & 0.89 & 1 \end{bmatrix}$$

采用直接聚类法对模糊相似矩阵进行聚类分析,当 $\lambda_k = 0.85$ 时, $\gamma_{12} = 0.90$ (或 0.95) > 0.85 , 即 7# 和 8# 坝段的扬压力测值为同一类;当 $\lambda_k = 0.70$ 时, $\gamma_{12} = 0.90$ (或 0.95) > 0.85 , $\gamma_{34} = 0.85$ (或 0.89) > 0.70 , 即 7# 和 8# 坝段扬压力为同一类, 13# 和 14# 坝段扬压力为同一类. 由此可看出, 在相同环境下, 在同一时间内, 7# 和 8# 坝段的扬压力变化相同, 13# 和 14# 坝段的扬压力变化相同.

用本文介绍的方法对 7# 及 13# 坝段 2000 年的扬压力测值进行了预报, 结果见表 3. 由表 3 可见, 预测效果还是令人满意的.

3 结 论

用模糊聚类分析方法所作的预报精度较高, 且计算简便, 可作为一种快捷的定量分析方法来弥补定性分析方法的不足. 模糊数学方法应用于大坝观测资料分析, 在深度和广度上都有待开拓, 今后随着研究的深入, 方法的改进, 将为大坝安全监测(控)提供更丰富更有效的信息.

参考文献:

[1] 邓聚龙. 灰色系统理论教程[M]. 武汉: 华中理工大学出版社, 1990. 33 ~ 37.
[2] 贺仲雄. 模糊数学及其应用[M]. 天津: 天津科学技术出版社, 1983. 45 ~ 51.

表 3 典型坝段扬压力测值的预报结果

Table 3 Predicted value of uplift pressure of typical dam sections m

坝段	测次	实测值	预报值	残差
7#	1	11.81	11.23	0.58
7#	2	11.81	11.23	0.58
7#	3	11.80	11.17	0.63
7#	4	11.78	11.14	0.64
13#	1	18.63	18.56	0.07
13#	2	18.65	18.57	0.08
13#	3	18.58	18.56	0.02
13#	4	18.54	18.45	0.09

Cluster Analysis-Based New Method for Prediction of Uplift Pressure of Dams

JIA Cai-hong , GU Chong-shi

(College of Water Conservancy and Hydropower Engineering , Hohai Univ. , Nanjing 210098 , China)

Abstract : By use of the principles of gray relative degree and fuzzy cluster analysis , the observed data of uplift pressure at each observation points of the Xin ' anjiang Dam foundation are analyzed. Based on the regularity of variation of observed data , data for some points are predicted , providing a basis for analysis and research of performance and safety monitoring of dams.

Key words : gray relative degree ; fuzzy cluster analysis ; observed data ; uplift pressure