

模糊数学在激电异常与富水评判中的应用

丁 华

(浙江有色工程公司 地球物理院,浙江 绍兴 312000)

摘 要: 针对激电异常与富水性间存在着一定的相关性,运用动态聚类方法从量的基础上去分析激电异常与富水性,并对一组实例数据进行分析解释,证明了该方法简单、有效。

关键词: 激电异常;富水性;动态聚类

中图分类号: P631.3

文献标识码: A

文章编号: 1000-8918(2002)05-0381-03

激电法由于不受纯地形起伏及周围岩体电阻率不均的影响,并可充分利用其时间(或频率)特性,从而给激电找水提供了应用空间,尤其是山区找水。但激电异常与测点富水性间的确切关系,仍值得人们深入地研究。

1 问题引述

由于地质体本身的复杂性及测量精度上的有限性,以及一些不可避免的测量误差,使得野外实测数据本身就不够精确,含一定的模糊性。同时含水性的划分也具有一定的模糊性,即不能用一个确切的 η 值或 J 值等,来对测点的含水性进行准确定类,故经典数学中的反演解释很难对此进行客观评价。

激电测量中的测量参数较多。有表征岩石激发极化强弱的极化率 η 、充电率 M ;表征岩石激发极化放电快慢的半衰时 S_t 、衰减度 D ;综合参数激发比($J = \eta \cdot D$)、 $Z = \eta \cdot S_t$ 、相对衰减时($S_R = S_t/\rho$),以及偏离度 r 、电阻率 ρ 等。且每一测量参数与测点富水性都应有密切关系,从而要求我们应运用多参数的方法对测量资料进行解释和处理。

任一客观事物都是由它们的本质属性来确定的,即由它们属性间的异同可划分成不同的类别。对于激电测量点的含水性划分也应如此,即利用各测点的每一个测量参数来作为其本质属性,从而进行分类。

2 方法原理

2.1 数学原理^[1,2]

笔者所采用的动态聚类法为 ISODATA,其原名

是 iterative self-organizing data analysis technique A,即“迭代自组织数据分析技术”。

设样品集 $X = \{x_1, x_2, \dots, x_n\}$, n 为样品数,其中 $x_i = \{x_{i1}, x_{i2}, \dots, x_{is}\}$, 表示第 i 个样品由 s 个因子组成的向量,其中 x_{ij} 为标准化后的数据,即满足 x_{ij} 在 $[0, 1]$ 上取值。对样品集首先进行初始分类。假设分成 c 类,则分类矩阵为 $U = [u_{ij}]$ ($i = 1, 2, \dots, c; j = 1, 2, \dots, n$), 且对 U 需满足以下 3 个条件:①每个元素 u_{ij} 在 $[0, 1]$ 上取值;②每列元素之和为 1,即同列中分属于各类的隶属函数之和为 1;③每行元素之和大于 0,即每类中至少有一个样品。

构造目标函数

$$J_m(U, V) = \sum_{j=1}^n \sum_{i=1}^c (u_{ij})^m \|x_j - v_i\|^2, \quad (1)$$

其中, $m \geq 1$, 反映最终分类的模糊性; v_i 表示第 i 类的聚类中心, $i = 1, 2, \dots, c$; $\|x_j - v_i\|^2$ 表示 x_j 与 v_i 的距离平方, 距离公式取

$$\|x_j - v_i\|_p = \left(\sum_{k=1}^s |x_{jk} - v_{ik}|^p \right)^{1/p} \quad (1 \leq p < \infty). \quad (2)$$

对于上述目标函数, Dunn 证明了极小的可能, Bezdek 给出了 $m \geq 1$ 且 $x_j \neq v_i$ 时的迭代算法:

1. 取定 c, m, ε , 其中 $2 \leq c \leq n, 1 \leq m \leq \infty, \varepsilon > 0$, 置初始分类 $U^{(0)}$ 和 $l = 0$ 。

2. 计算聚类中心 $V = \{v_i^{(l)}\}$

$$v_i^{(l)} = \sum_{k=1}^n (u_{ik}^{(l)})^m x_k - \sum_{k=1}^n (u_{ik}^{(l)})^m, \quad (3)$$

3. 利用聚类中心 $\{v_i^{(l)}\}$ 去修正 $U(l)$, 得

$$u_{ik}^{(l+1)} = 1 / \sum_{j=1}^c \left(\frac{\|x_k - v_i^{(l)}\|}{\|x_k - v_j^{(l)}\|} \right)^{1/(m-1)}, \forall i \forall k; \quad (4)$$

4. 计算 $\|U^{(l+1)} - U^{(l)}\|$, 若 $\|U^{(l+1)} - U^{(l)}\| \leq \varepsilon$, 则停止迭代, 否则转向 2 步, 再进行循环计算, 直至 $\|U^{(l+1)} - U^{(l)}\| \leq \varepsilon$, 停止迭代, 输出 $U^{(l+1)}$ 和 $V^{(l+1)}$ 。

2.2 应用原理

对激电测量来说, 其实际的测量点数即为样品数 n , 而所测到的参数量即为样品的特征因子数 s 。

1. 数据标准化。即将实测的激电数值化为 $[0, 1]$ 上取值, 其公式为:

$$x_{ij} = (x_{ij} - x_{\min}) / (x_{\max} - x_{\min})$$

或 $x_{ij} = x_{ij} (x_{\max} - x_{\min}) + x_{\min}$, (5)

其中 x_{ij} 为实际测量的激电数值, 而 $x_{\min}$ 为同一观测分量中的极小值, $x_{\max}$ 为同一观测分量中的极大值。

2. 参数的选择。 m 的大小表示了聚类的模糊性, m 越大其分类越模糊, 据研究表明, m 的最佳取值范围为 $1.5 \sim 2$, 作者取 $m = 1.5$ 。对于实测数据来说, 我们将测点的含水性划分成四类: 无水、中等富水、富水和富水且可能有矿, 即 $c = 4$ 。文中取迭代精度 $\varepsilon = 0.1$ 。另外对 (1) 式所表述的距离公式中的 p 取为 2, 得欧氏距离公式:

$$\|x_k - v_i\| = \left[\sum_{k=1}^s (x_{jk} - v_{ik})^2 \right]^{1/2}, \quad (6)$$

3. $U^{(0)}$ 的选取。通过上述的数学原理分析可知, $U^{(0)}$ 的选取不仅影响到迭代计算的次数, 而且直接决定了最终的聚类结果。

对于通常的含水层上, η 值和 D 值较高, 为了突出异常, 人们又提出了综合参数 $J = \eta \cdot D$, 则可综合 η 、 D 、 J 三参数, 进行分类得 $U^{(0)}$ 。根据参考文献 [3] 中的总结资料, 可得表 1。假设实测数值属于无

水, 则取 (以 J 为例) $J = 0.1$, 中等富水 $J = 0.4$, 富水 $J = 0.7$, 富水且可能有矿 $J = 1$ 。 η 和 D 也同理进行取值。

表 1 初始判断结果

划分	$J/\%$	$\eta/\%$	$D/\%$
无水	<0.3	<0.8	<30
中等富水	0.3~0.6	0.8~1.5	30~50
富水	0.6~3	1.5~5	50~80
富水且可能有矿	>3	>5	>80

采用多元主向量排序, 定义 $A = \alpha J + \beta \eta + \gamma D$, 其中 α, β, γ 分别为各观测分量的加权系数, 且 $\alpha + \beta + \gamma = 1$ 。为了突出 J 反映富水性的可靠度, 可取 $\alpha = 0.4, \beta = \gamma = 0.3$ 。

为严格反映测点的富水性, 对任一测点 i 有: $A > 0.75, U_i^{(0)} = [0.9; 0.1; 0; 0]$, 即表明该点有 90% 的可能是“富水且可能有矿”, 同时还有 10% 的可能是“富水”。同理分别有: $0.75 \geq A > 0.5, U_i^{(0)} = [0; 0.9; 0.1; 0]; 0.5 \geq A > 0.25, U_i^{(0)} = [0; 0; 0.9; 0.1]; A \leq 0.25, U_i^{(0)} = [0; 0; 0; 1]$, 则可得初始分类 $U^{(0)}$ 。

3 实例分析

表 2 为笔者收集的不同区域 16 个测点实际测量出的激电数据, 首先利用 (5) 式将其标准化, 再由计算出的 A 值得到初始聚类结果 $U^{(0)}$, 通过计算得最终的聚类结果 $U^{(l+1)}$ 和聚类中心 $V^{(l+1)}$, 并利用 (5) 式反计算出实测数值的聚类中心 V 。

由 $U^{(0)}$ 和 $U^{(3)}$, 我们可以得到表 3。

由表 3, 我们可看出最终聚类结果与初始分类相差不大, 且有一个 8 号点, 同时在计算过程中只需

表 2 实测数据

点号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	min	max
$J/\%$	0.86	4.31	0.90	0.48	0.28	0.06	1.20	0.24	0.14	0.38	4.11	0.09	3.23	0.99	0.52	1.00	0.006	4.31
$\eta/\%$	1.95	5.01	2.00	1.30	0.91	0.33	2.50	0.69	0.51	1.11	4.89	0.41	4.36	1.73	1.34	2.10	0.33	5.01
$D/\%$	44.0	86.0	45.0	37.1	31	19.2	50.0	34.1	27.6	34.2	84.0	22.1	74.0	57.0	38.8	46.0	19.2	86.0
A	0.61	1	0.61	0.4	0.28	0.1	0.7	0.19	0.1	0.4	0.91	0.1	0.82	0.7	0.4	0.61		

$$U^{(0)} = \begin{bmatrix} 0 & 0.9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0 & 0.9 & 0 & 0 & 0 \\ 0.9 & 0.1 & 0.9 & 0 & 0 & 0 & 0.9 & 0 & 0 & 0 & 0 & 0.1 & 0 & 0.1 & 0.9 & 0 & 0.9 \\ 0.1 & 0 & 0.1 & 0.9 & 0.9 & 0 & 0.1 & 0 & 0 & 0.9 & 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0.1 & 0.1 & 1 & 0 & 1 & 1 & 0.1 & 0 & 1 & 0 & 0 & 0.1 & 0 & 0 \end{bmatrix},$$

$$V^{(3)} = \begin{bmatrix} 0.9064 & 0.9484 & 0.9351 \\ 0.2193 & 0.3711 & 0.4305 \\ 0.0860 & 0.1761 & 0.2506 \\ 0.0164 & 0.0285 & 0.0805 \end{bmatrix}, \quad V = \begin{bmatrix} 3.9122 & 4.7685 & 81.6647 \\ 0.9920 & 2.0667 & 47.9574 \\ 0.4255 & 1.1541 & 35.9401 \\ 0.1297 & 0.4634 & 24.5774 \end{bmatrix},$$

$$U^{(3)} = \begin{bmatrix} 0.0039 & 0.9756 & 0.0022 & 0.0007 & 0.0031 & 0.0029 & 0.0121 & 0.0039 \\ 0.8868 & 0.0115 & 0.9452 & 0.0172 & 0.0399 & 0.0200 & 0.8945 & 0.0483 \\ 0.0876 & 0.0074 & 0.0414 & 0.9656 & 0.6259 & 0.0723 & 0.0666 & 0.5975 \\ 0.0218 & 0.0056 & 0.0112 & 0.0165 & 0.3310 & 0.9049 & 0.0268 & 0.3503 \\ 0.0008 & 0.0005 & 0.9929 & 0.0010 & 0.8819 & 0.0179 & 0.0020 & 0.0009 \\ 0.0071 & 0.0093 & 0.0034 & 0.0073 & 0.0614 & 0.7878 & 0.0583 & 0.9840 \\ 0.0436 & 0.9668 & 0.0021 & 0.0301 & 0.0335 & 0.1410 & 0.9022 & 0.0116 \\ 0.9485 & 0.0233 & 0.0016 & 0.9616 & 0.0232 & 0.0533 & 0.0374 & 0.0035 \end{bmatrix}。$$

表 3 初始分类与最终聚类结果对比

类型	初始分类点号	最终聚类点号
无水	6,8,9,12	6,9,12
中等富水	4,5,10,15	4,5,8,10,15
富水	1,3,7,14,16	1,3,7,14,16
富水且可能有矿	2,11,13	2,11,13

循环 3 次即可满足 $\|U^{(l+1)} - U^{(l)}\| \leq 0.1$, 说明我们利用表 1 采用多元主向量排序法来确定 $U^{(0)}$ 是合理有效的。而对 8 号点来说, 初始分类为“无水”, 最终为“中等富水”, 其实际井位有水, 出水量约 $40 \text{ m}^3/\text{h}$, 说明最终的聚类结果是合理可靠的。

另外还可利用最终计算出的聚类中心 V , 来对其它实测数据进行分类, 参考文献[4]中的 2 组数据:

A 点: $J=0.75, \eta=1.80, D=41.0, q_{250}=4.63$,

B 点: $J=0.80, \eta=1.90, D=49.0, q_{250}=5.56$,

其中 q_{250} 表示实际井径为 250 mm 的单位涌水量, 单位为 $\text{L}/(\text{s} \cdot \text{m})$ 。将 A、B 实测数值标准化后, 再利用 (6) 式分别计算出其与 $V^{(3)}$ 的距离, 得 (归一化后):

$$d_{AV} = [0.6184; 0.0707; 0.0940; 0.2170]$$

$$d_{BV} = [0.5639; 0.0313; 0.1404; 0.2643]$$

根据距离最近原则, 则知 A 点最有可能属于“富水”, 其次为“中等富水”, 且 A 与“中等富水”和“富水”间的距离较接近; B 点也同样如此, 另外其与“富水”间的距离比 A 要更为接近, 与实际测量结果较为吻合,

说明最终的聚类结果可靠性、适用性较好。

4 结论

本方法是对整个测量区域内测点数据进行综合定量对比分析, 可以整体把握测区的含水性情况, 从而有助于对地下水资源的合理开发和应用。文中采用多参数综合解释, 解释精度相对较高, 且从实例分析来看, 最终的解释结果有一定的合理性。数据的再用率高, 由最终得到的聚类中心可对其它测点数据进行分类。

由于是首次将动态聚类与激电异常和富水性结合起来, 应用过程难免出现诸多不足, 望大家不吝指正。在成文过程中得到河北建筑科技学院李树文老师的诸多鼓励和关怀, 在此深表感谢!

参考文献:

- [1] 袁嘉祖. Fuzzy 数学在苗木质量标化上的应用技术[J]. 模糊数学, 1984, (1).
- [2] 冯德益. 模糊数学方法与应用[M]. 北京: 地震出版社, 1983.
- [3] 李树文, 屈均利, 武强, 等. 激电异常与富水性关系的模糊综合评判[J]. 工程勘察, 1999, (6).
- [4] 杨进, 曲维娟. 激电法预报地下涌水量的回归分析方法研究[J]. 物探与化探, 1997, (5).

THE APPLICATION OF FUZZY MATHS TO THE APPRAISAL OF IP ANOMALY AND ENRICHED GROUNDWATER

DING Hua

(Institute of Geophysics, ZNEC, Shaoxing 312000, China)

Abstract: This paper deals with the relationship between IP anomaly and enriched groundwater. Using ISODATA, the author has analyzed quantitatively the IP anomaly and the enriched groundwater property, and explained a set of sample data. The results show that the method is easy and effective in the application work.

Key words: IP anomaly; enrich groundwater property; ISODATA

作者简介: 丁华(1979-), 安徽滁州人, 2001 年毕业于华东地质学院应用地球物理专业。