

矿井构造定量预测中评价指标 权重的确定方法

徐凤银 王桂梁 龙荣生 夏玉成

摘 要

本文以杉木树矿井地质构造预测为例,成功地使用了灰色关联分析法,为模糊综合评判中如何确定权重提供了一种有效方法。首先指出权重的确定意义及应用灰色关联分析法解决该问题的可行性;其次从理论依据到实现过程阐述了确定权重的原理和方法,为深入应用模糊综合评判法解决其它实际问题提供了良好途径。

关键词: 构造预测, 权重, 关联分析, 模糊评判, 评价指标

一、前 言

随着采掘机械化程度的不断提高,开展矿井地质构造预测已成为大中型矿井生产中的一项重要内容。近年来,人们已经从常规法定性预测构造逐步转为用数学方法和电子计算机进行定量、半定量的构造评价。目前广泛采用的有“构造指数法”、“等性块段法”、“模糊综合评判法”等。但无论哪种方法,均涉及到评价指标的权重系数,这是人们常感到为难而又无法逃避的重要问题。这里根据我们的工作体会,以应用模糊综合评判法为例,介绍利用灰色关联分析法获取评价指标权重系数的研究方法,以供参考。

二、模糊综合评判中权重的确定意义

模糊综合评判的数学模型^[1]一般可概括为:

设因素集为 $\underline{U}$ $\underline{U} = \{u_1, u_2, \dots, u_n\}$;

评语集为 $\underline{V}$ $\underline{V} = \{v_1, v_2, \dots, v_m\}$;

权重集为 $\underline{A}$ $\underline{A} = \{a_1, a_2, \dots, a_n\}$;

则以 $\underline{U} \times \underline{V} = \{(u, v) | u \in \underline{U}, v \in \underline{V}\}$ 为论域的一个模糊子集 $\underline{R} \in f(\underline{U} \times \underline{V})$ 可体现 $\underline{U} \times \underline{V}$ 间的模糊关系,对任给 $\underline{R} \in f(\underline{U} \times \underline{V})$ 都可唯一确定一个从 $\underline{U}$ 到 $\underline{V}$ 的模糊映射,反之亦然。由此可构成单因素评价集 $\underline{R}$: $\underline{R} = (r_{ij})_{n \times m}$, 再通过模糊向量 $\underline{A}$ 和模糊关系矩阵 $\underline{R}$ 的合成即可得出模糊综合评判的结果向量。即:

$$\underline{B} = \underline{A} \circ \underline{R} = \{b_1, b_2, \dots, b_n\}$$

由此可见综合评判集 $\underline{B}$ 直接与权重集 $\underline{A}$ 的取值有关。从权重本身的意义看,因素集

中各因素(u_i)权重的大小直接体现出它在综合评判中的重要性。在评价矿井构造复杂程度时,各指标的权重大小,一方面可反映用该指标单独评价某一块段构造复杂程度时的可靠性,并能较好地反映出所选指标的实用性及其合理性^[2]。另一方面,因素集内各指标权重的相对大小直接影响综合评价的结果。因此,合理确定评价指标的权重系数,客观反映它在综合评价中的重要性,就会直接提高最终评价结果的准确性。

三、灰色关联分析确定权重的可行性

用模糊综合评判法评价矿井地质构造复杂程度时,首先应确定一些能够真正反映构造复杂程度的评价指标,构成评价因素集。而要使最终评价结果真正量化,除了寻找能使地质构造复杂程度“数量化”的特殊指标外,还必须确定各评价指标在决定这些特殊指标大小时的权重系数。以便进行综合评价时掌握哪些指标是主要的,哪些指标是次要的;哪些影响大,哪些影响小;哪些是潜在的,哪些是明显的。而这时“权重”如果象以往那样由人为假定,势必会增加主观因素。实际上很难直接识别出来,只有通过因素分析才能取得较为客观的结果。而灰色关联分析正是解决这种因素与因素之间相关程度的一种简便而科学的方法。

四、灰色关联分析的实现原理

关联分析又称灰因素间的关联分析,实质上是动态发展态势的量化比较分析^[3]。除主要用于分析各因素之间随时间变化的动态关系及其特征外,常用于分析因素本身之间的密切程度,它是灰色系统分析、预测、决策的基础。包括关联系数、关联度、关联序及关联矩阵的计算系列。

设有一参考数据列(也叫母因素数列),记作:

$$x_0 = (x_0(1), x_0(2), \dots, x_0(n))$$

同时获得一系列比较数据列(也叫子因素数列)记作:

$$\begin{aligned} x_1 &= (x_1(1), x_1(2), \dots, x_1(n)) \\ x_2 &= (x_2(1), x_2(2), \dots, x_2(n)) \\ &\vdots \\ x_m &= (x_m(1), x_m(2), \dots, x_m(n)) \end{aligned}$$

记数据列 x_i ($i = 1, 2, \dots, m$) 对数据列 x_0 的关联系数为 $\xi_i(K)$, K 为表示两数列比较其关联性的采样点,即:

$$\xi_i(K) = \frac{\Delta_{\min} + k \cdot \Delta_{\max}}{\Delta_i(K) + k \cdot \Delta_{\max}}$$

式中: k 取: $0.1 \leq k \leq 1$, 为分辨系数;

$\Delta_{\min} = \min_i \min_K |x_0(K) - x_i(K)|$ 为因素间最小绝对差,一般可选取零;

$\Delta_{\max} = \max_i \max_K |x_0(K) - x_i(K)|$ 为因素间最大绝对差;

$\Delta_i(K) = |x_0(K) - x_i(K)|$ 为比较点上因素间的绝对差;

为了使信息得以集中, 也便于比较, 一般用各比较数列对参考数列 (x_0) 的关联度 (r_i) 来表示:

$$r_i = \frac{1}{n} \sum_{i=1}^n \xi_i(K)$$

将关联度 r_i ($i = 1, 2, \dots, m$) 按大小排成一行, 即为各比较数列 (x_i) 对参考数列 (x_0) 的关联序; 若依次分别取各比较数列作参考数列, 计算与其它数列的关联度, 并排列成矩阵, 便形成关联阵。

五、权重的确定方法及结果分析

在杉木树矿井地质构造定量预测中, 我们共确定和引用了十个评价指标^[4]。要对矿井不同区段的构造复杂程度作出综合评价, 关键就在于合理确定各评价指标的权重大小。结合煤矿生产的特点和要求, 我们选取构造面积损失系数(S)和构造面积掘进指数(J)作为两个经济技术指标分别代表地质构造的复杂程度, 使其得以数量化; 而取断层密度(M)、断层强度(F)、煤层底板倾角变异系数(R)、地层纵向倾角变异系数(T)、平面褶皱强度指数(D)、剖面褶皱强度指数(Z)、断层与煤层夹角指数(Q)及构造组数(N)等八个指标作为评价不同区段构造复杂程度的地质指标。这样, 衡量各评价指标在决定某一区段构造复杂程度时的重要性, 实际上就是计算其相对于两个经济技术指标的权重值。因此, 通过统计已采区各单元 ($200 \times 200\text{m}^2$) 内每个指标的对应取值 (表 1), 依据关联分析的实现原理, 将两个经济技术指标分别作为母因素 (为互相对比、验证), 其它八个评价指标作为子因素, 作子因素与母因素间的关联分析。通过下列计算过程^[5] (下图), 求出关联度。然后按归一化处理, 所得结果即为各指标相对于评价构造复杂程度时的权重系数。归一化表达式为:

$$a_i = \frac{r_i}{\sum_{i=1}^n r_i}$$

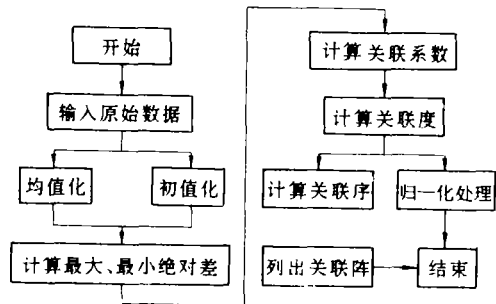
式中: a_i 为归一化后的权重系数;

n 为评价指标个数 (这里 $n = 8$)。

由于矿井内不同采区构造特征不同, 为便于构造预测, 我们将矿井已采区划分为 $N_1 \sim N_4$ 区、 N_8 区、 $S_1 \sim S_6$ 区三个块段, 分别得出每个块段内各指标对应权重系数。这里将构造面积损失系数(S)作为母因素, 以 N_{1-4} 采区为例, 说明权重的计算过程。

根据指标的物理意义, 统计得出 $N_1 \sim N_4$ 区内各单元构造面积损失系数(S)的取值为 (详见表 1):

$$S = (45.5, 36.3, 80.7, 94.3, \dots, 63.0, 54.3)$$



关联分析计算过程

表1 $N_1 \sim N_4$ 区各单元指标统计值

统计值 单元号 \ 指标	S	J	F	M	Q	Z	R	D	T	N
1	45.50	370.00	3221.90	2.35	99.00	22.50	28.00	7.75	30.60	4.00
2	56.30	315.00	4360.30	1.06	6.60	37.50	17.60	9.10	35.90	2.00
3	80.70	383.00	5344.40	1.32	37.00	32.00	23.70	2.00	34.70	3.00
4	94.30	451.00	5925.20	1.59	53.80	32.00	29.70	2.00	32.10	3.00
5	66.20	440.00	4386.30	2.90	164.10	19.00	34.80	1.25	37.10	4.00
6	25.40	300.00	1583.10	1.27	35.80	33.30	14.80	18.37	26.00	4.00
7	20.00	280.00	978.30	1.80	28.00	28.6	35.00	10.13	16.80	2.00
8	36.50	357.00	1882.60	3.43	72.90	17.78	30.90	9.96	27.50	6.00
9	58.70	433.00	2775.80	3.47	133.50	9.18	35.80	9.76	38.20	5.00
10	80.50	434.00	2141.90	2.19	79.70	7.30	40.20	12.55	42.80	5.00
11	58.20	361.00	1693.00	2.62	49.40	16.25	40.20	8.96	39.50	6.00
12	37.70	289.00	1195.70	1.92	34.80	23.89	43.00	4.67	17.20	4.00
13	41.50	301.00	1112.60	3.22	85.60	16.18	30.30	2.15	7.40	4.00
14	41.20	299.00	1491.80	2.24	49.60	23.10	36.90	2.85	21.00	4.00
15	63.80	342.00	2849.80	0.87	50.70	16.92	44.80	9.35	46.10	5.00
16	87.50	404.00	1986.50	1.50	73.90	7.25	44.70	19.35	70.20	4.00
17	69.60	322.00	4404.80	1.30	64.90	15.91	49.90	13.04	40.30	4.00
18	33.80	316.00	632.10	5.15	86.60	15.26	27.30	3.40	9.10	5.00
19	45.70	327.00	305.40	2.93	41.20	17.08	34.80	7.60	120.00	4.00
20	40.60	323.00	418.60	3.13	66.10	19.33	51.10	14.28	21.00	5.00
21	73.50	324	2664.30	1.60	81.30	12.92	63.20	10.30	32.60	5.00
22	74.50	364.00	7242.40	1.64	51.10	9.64	81.10	3.33	26.10	4.00
23	60.10	354.00	748.70	1.81	60.80	15.48	65.10	6.38	19.60	5.00
24	67.70	352.00	1112.10	1.49	64.50	18.00	33.90	7.78	13.50	5.00
25	76.90	386.00	1719.90	1.30	51.70	3.00	91.60	3.97	21.00	4.00
26	66.50	367.00	745.80	1.85	69.30	11.67	78.60	5.32	19.10	4.00
27	66.40	365.00	811.40	1.40	125.70	17.40	43.60	7.77	14.30	5.00
28	54.10	353.00	674.80	1.37	137.70	14.12	21.10	6.88	9.30	3.00
29	50.00	346.00	521.30	0.95	203.80	15.00	19.10	8.15	12.70	3.00
30	56.10	333.00	1017.70	3.02	126.30	15.00	40.90	12.12	17.60	4.00
31	52.80	339.00	744.90	1.80	46.60	7.25	70.90	10.08	19.30	5.00
32	67.20	356.00	954.20	2.59	59.60	7.50	66.20	4.73	14.20	3.00
33	72.30	343.00	122.50	2.12	40.40	5.00	51.50	7.37	12.60	3.00
34	63.00	345.00	799.00	1.77	71.40	7.17	79.00	15.89	19.50	6.00
35	54.30	354.00	1397.40	0.92	85.10	12.14	57.40	29.12	19.30	4.00

同时获得其它指标的对应值分别为:

$$F = (3221.9, 4360.3, 5344.4, 5925.2, \dots, 799.0, 1397.4)$$

$$M = (2.35, 1.06, 1.32, 1.59, \dots, 1.77, 0.92)$$

$$Q = (99.0, 6.6, 37.0, 53.8, \dots, 71.4, 85.1)$$

$$Z = (22.5, 37.5, 32.0, 32.0, \dots, 7.17, 12.14)$$

$$R = (28.0, 17.6, 23.7, 29.7, \dots, 79.0, 57.4)$$

$$D=(7.75, 9.1, 2.0, 2.0, \dots, 15.89, 29.12)$$

$$T=(30.6, 35.9, 34.7, 32.1, \dots, 19.5, 19.3)$$

$$N=(4, 2, 3, 3, \dots, 6, 4)$$

这时, 将各数据分别输入计算机内, 利用关联分析的数学模型, 选择构造面积损失系数(S)为母因素, 计算出每个指标对它在任一单元内的关联系数 $\xi_i(K)$ (数据量太大, 略)。然后分别计算出每个指标在所有单元内关联系数的均值, 即为各指标对母因素的关联度 (r); 即:

$$r=(0.4401, 0.4599, 0.4920, 0.5068, 0.5765, 0.5214, 0.5136, 0.5858)$$

经归一化处理, 就得出各指标在决定构造面积损失系数(S)大小 (也即构造复杂程度) 时的权重系数(a_i):

$$A=(0.1074, 0.1123, 0.1201, 0.1237, 0.1407, 0.1273, 0.1254, 0.1430)$$

如果按关联度大小将每个评价指标排队, 即得出它们的相关序, 对本区来说相关序为:

$$N > R > D > T > Z > Q > M > F$$

如果只计算模糊综合评判或其它综合评价时的权重系数, 上述结论即可满足; 但实际中往往还应对其作一定的可靠性评价。比如, 对本区来说, 反映褶皱复杂程度的指标(如 R 、 D), 关联度、权重均较大, 关联序中排在前; 反映断层复杂程度的指标(如 F 、 M), 关联度、权重均较小。说明次级褶皱为决定本区构造复杂程度的主体构造。从生产实际揭露情况看, 本区是全矿井次级褶皱最发育的地区, 投产近二十年来, 共揭露十二个对生产有影响的次级褶皱, 本区就占十个(83%)。所以, 我们在综合评价和预测本区及相邻区未揭露区的地质构造复杂程度时, 以上述权重集为依据, 作模糊合成, 取得了较为准确的评价结果^[6]。

同理可得出 N_8 区和 $S_1 \sim S_6$ 区各指标权重的对应值。从 N_8 采区的计算结果 (见表 2) 看, 本区反映断层复杂程度的指标 (如 Q 、 M 等), 关联度及权重均较大, 关联序中排在前; 反映褶皱复杂程度的指标(如 D 、 Z 等), 关联度及权重均较小。说明断层为决定本区构造复杂程度的主体构造。实际情况证明, 本区未发育次级褶皱, 只发育不少断层。在评价和预测本区及邻区构造复杂程度时, 利用本区的权重集进行模糊合成, 也得出满意的预测效果。

表 2

 N_8 区 关 联 分 析 结 果

评价指标	F	M	Q	Z	R	D	T	N
关 联 度	0.4958	0.5159	0.6278	0.4048	0.5061	0.3784	0.4584	0.5535
权 重	0.1258	0.1309	0.1593	0.1027	0.1284	0.0961	0.1163	0.1404
关联序	$Q > N > M > R > F > T > Z > D$							

从 $S_1 \sim S_6$ 区的计算结果 (见表 3) 看, 本区很难明确得出以褶皱还是断层指标为主, 相关序不明显。实际揭露发现, 本区断层、褶皱普遍减少, 构造对生产的影响明显受几条大断层的控制, 规律性极为明显。因此, 在评价本区及邻区构造复杂程度时, 我们选择了该结果作为模糊综合评判中的权重集, 取得较为理想的效果。

表3 $S_1 \sim S_6$ 区 关 联 分 析 结 果

评价指标	F	M	Q	Z	R	D	T	N
关 联 度	0.5538	0.4909	0.5031	0.34	0.5089	0.5015	0.4058	0.439
权 重	0.1479	0.1311	0.1344	0.0908	0.1360	0.1339	0.1084	0.1173
关联序	$F > R > Q > D > M > N > T > Z$							

六、方法评述

矿井地质构造定量评价是煤矿生产的一个重要环节,模糊综合评判法的引入使其在方法上真正向前迈进一步,而权重确定的准确性则直接关系到模糊综合评价结果的可靠性。近年来,虽然模糊综合评判法已在不同领域得到广泛应用,但从所见研究成果看^[7],权重问题仍未得到客观的解决。一般只凭专家经验或主观意断,这样势必会增加人为因素,使评价结果的可靠性受到限制。本项课题开始时,我们就把它作为研究重点,在解决评价指标问题的基础上,一直寻求一种确定权重的客观方法。从理论到实践试验,起初选择了因素分析中经常采用的多元回归分析法,求出相关系数,在综合评判过程中,发现该方法缺陷很多,比如要求数据量太大,计算也繁杂,而且所选指标之间应有典型分布,也易出现反常情况。而确定权重一般在解决问题之前,资料常很少,特别是遇到除了要掌握子因素与母因素间的相关性外,还要了解子因素间的相关关系时,回归分析就显得极不方便。关联分析法正好能够克服和弥补上述不足^[8]。从应用效果看,其不仅准确性高,而且灵活方便。比如,可以在资料很少的情况下得到准确的结果,也可以随时按构造特征把矿井划为几个块段,分别确定各块段内不同指标的权重,这正好与实际中不同块段构造发育特征及复杂程度不一致相吻合。从方法本身来说,关联分析法原理简单,易于实现,用它不仅能够解决矿井构造模糊综合评判中的权重问题,而且也可广泛用到其它方面。特别当因素集中各因素的权重由人的主观思维根本无法估计,或者因素较多时,该方法更为理想。比如在本项研究的因素集中,平面褶皱强度指数(D)与煤层底板倾角变异系数(R)两个指标,到底哪个与构造面积损失系数(S)相关性更大,人为很难得出结论。而利用关联分析法只要有几个统计单元的数据,就会很快得出定量的结果。因此,在要求煤矿地质构造或其它预测精度越来越高的今天,权重的科学确定必将给模糊综合评判法的深入应用开辟了广阔的前景。

参加课题实际工作的还有:芙蓉矿务局谢仕康、孙玉德、李汝明、李永云、雷振德、中国煤田遥感公司张光超等。

工作过程中,芙蓉矿务局来和章总工程师、地测处谢高文处长、科研所魏明康所长、童代金高工给予大力支持;煤科院西安分院门桂珍、赵宗沛教授级高工提出许多宝贵意见,在此一并表示衷心感谢。

参 考 文 献

- [1] 汪培庄:《模糊集合论及其应用》,上海科技出版社,1983年。

- [2] 徐凤银、龙荣生: 多层次模糊综合评判法在矿井断层预测中的应用,《西安矿业学院学报》, 1988 年, 第 4 期。
- [3] 邓聚龙:《灰色系统基本方法》, 华中理工大学出版社, 1988 年。
- [4] 徐凤银: 杉木树矿井地质构造的定量评价及其预测, 西安矿业学院硕士论文, 1990 年。
- [5] 王学明、罗建军编著:《灰色系统预测决策建模程序集》, 科学普及出版社, 1986 年。
- [6] 徐凤银等: 矿井地质构造定量评价及预测,《煤炭学报》, 1991 年, 第 16 卷, 第 4 期。
- [7] 煤科总院西安分院: 矿井地质条件分类软件 (报告之一), 1989 年。
- [8] Xu Fengyin, Long Rongshen: An Application of Grey Relational Analysis in Seeking Prediction Indices of Geology Structure in Coal Mine, 14th World Mining Congress, Beijing, 1990.

A Method of Weight Determination of the Assessment Indices in Quantitative Forecast of the Geological Structure in Coal Mine

Xu Fengyin Wang Guiliang et al

Abstract

Taking the prediction of the geological structure in Shanmushu Coal Mine as an example, and using the method of gray correlation analysis, the paper has successfully solved the problem of weight determination in fuzzy comprehensive evaluation. Firstly, it indicates the significance of weight determination and the feasibility to solve the problem by applying the method of grey correlation analysis. Secondly, it discusses in detail the principle and method of weight determination from theoretical foundation to practical process, and it provides a good way to apply the method of fuzzy comprehensive evaluation in practice.

Key words: structure forecast, weight, correlation analysis, fuzzy evaluation, assessment indices