第18卷 第5期
Vol 18 No. 5

控 制 与 决 策
Control and Decision

2003年9月
Sept. 2003

文章编号: 1001-0920(2003)05-0593-04

不确定信息条件下的灰色模式识别

赵艳林, 杨绿峰, 吕海波, 韦树英
(广西大学 土木工程学院, 广西 南宁 530004)

摘　要: 为了解决工程中存在的由不确定信息引起的灰色模式识别问题, 首先定义了区间灰数, 建立了区间灰数的距离计算公式; 然后将灰色关联理论与模糊集理论相结合, 提出一种基于区间灰数模糊灰关联分析的灰色模式识别方法, 并给出了这种灰色模式识别结果的可信度。最后通过工程实例说明了该方法的正确性和有效性。

关键词: 区间灰数; 模糊集; 灰关联分析; 模式识别
中图分类号: O235　　**文献标识码**: A

Grey pattern recognition under the condition of uncertain information

ZHAO Yan-lin, YANG Lu-feng, LU Hai-bo, WEI Shu-ying
(College of Civil Engineering, Guangxi University, Nanning 530004, China)

Abstract: The interval grey number is defined and the computational formulation for the distance of the two interval grey numbers is presented in order to tackle the problem of grey pattern recognition which is encountered in dealing with uncertain information in engineering. By combining the grey incidence theory with the fuzzy sets theory, a grey pattern recognition method based on fuzzy grey incidence analysis for interval grey number sequences is proposed and the confidence level of the result of the grey pattern recognition is also evaluated. An example in engineering shows the correctness and effectiveness of the proposed method.

Key words: Interval grey number; Fuzzy sets; Grey incidence analysis; Pattern recognition

1　引　　言

模式识别在工程中有着广泛的应用。模式识别的方法很多, 但目前几乎所有的模式识别方法都以已知系统的全部信息为前提。在信息完全的前提下, 不仅模式中所涉及的指标集是确定的, 而且据此识别的各个指标的特征值也是确定的。然而, 实际工程中的情况并非都是如此, 由于人们认知程度的有限性, 自然界事物本身的复杂性, 信道的噪音干扰以及接收系统能力的局限, 在许多情况下, 人们只能获得非完全信息。非完全信息包含信息不充分与信息不确定两种内涵, 或者说包含"亏损"与"朦胧"[1]两种情况。在不确定信息条件下, 或者说在"朦胧"情况下, 模式识别中的指标集虽然已知, 但各指标的特征值只知大致范围, 而不知其确切的值。这种只知大概范围而不知其确切值的数称为灰数。包含灰数的模式识别问题称为灰色模式识别问题。

工程中最为常见的灰色模式识别是区间灰数型模式识别。本文针对这类问题, 将灰色系统理论与模

收稿日期: 2002-06-28; **修回日期**: 2002-10-08。
基金项目: 广西自然科学基金资助项目(0135051); 广西"十百千人才工程"基金资助项目(99511)。
作者简介: 赵艳林(1958—), 男, 广西全州人, 教授, 博士生导师, 从事工程中的灰色决策理论与应用研究; 杨绿峰(1963—), 男, 河南郑州人, 从事工程中的灰色决策理论与应用研究。

糊集理论相结合, 提出一种基于区间灰数模糊灰关联分析的灰色模式识别方法。

2 区间灰数及其运算

定义 1 若灰数 $\otimes(a)$ 的白化值满足 $\tilde{\otimes}(a) \in [\underline{a}, \overline{a}], \underline{a} \le \overline{a}, \underline{a}, \overline{a} \in R$, 则称 $\otimes(a)$ 为区间灰数。全体区间灰数集合记为 $g(\dot{\otimes})$。

区间灰数表明信息覆盖[2] 为区间, 白化值非唯一, 并满布整个区间, 但真值唯一。当 $\tilde{\otimes}(a) \in [t, t] = t$ 时, $\otimes(a)$ 就是一个普通实数。实数 0 和 1 是两个特殊的灰数, 分别记为 $\otimes(0)$ 和 $\otimes(1)$。

区间灰数的 +, -, ×, / 四则运算规则见文献[3]。

定义 2 $\forall \otimes(a) \in g(\dot{\otimes})$, $\tilde{\otimes}(a) \in [\underline{a}, \overline{a}]$, 则称 $\otimes(a) - \otimes(a) = \otimes(0)$ 为区间灰数的自差; 称 $\otimes(a) / \otimes(a) = \otimes(1)$ 为区间灰数的自商。

定义 2 特别强调“自”的含义, 它表示某个特定时空中的某个区间灰数的本身, 若是两个不同的区间灰数, 尽管两者的白化区间相同, 但由于真值取数不一致, 便不能适用定义 2。

定理 1 设 $\otimes(a), \otimes(b) \in g(\dot{\otimes})$, $\tilde{\otimes}(a) \in [\underline{a}, \overline{a}]$, $\tilde{\otimes}(b) \in [\underline{b}, \overline{b}]$, 则 $\otimes(a)$ 与 $\otimes(b)$ 之间的距离 $\otimes(\rho(\otimes(a), \otimes(b))) \in g(\dot{\otimes})$, 且

$$\tilde{\otimes}(\rho(\otimes(a), \otimes(b))) \in [\underline{\rho}_{ab}, \overline{\rho}_{ab}]$$

其中

$$\underline{\rho}_{ab} = \begin{cases} 0, & [\underline{a}, \overline{a}] \cap [\underline{b}, \overline{b}] \ne \phi \\ \min(|\overline{a} - \underline{b}|, |\overline{b} - \underline{a}|) & \\ & [\underline{a}, \overline{a}] \cap [\underline{b}, \overline{b}] = \phi \end{cases}$$

$$\overline{\rho}_{ab} = \max(|\overline{a} - \underline{b}|, |\overline{b} - \underline{a}|)$$

显然, 实距离 $|a - b|$ 是灰距离的特例。

3 灰色指标特征值的模糊规格化

设共有 s 类模式。s 类模式对于 n 个用于识别的指标所具有的灰色标准特征值矩阵为

$$B(\otimes) = (\otimes(B_{kj}))_{ns} \tag{1}$$

式中: $\otimes(B_{kj}) \in g(\dot{\otimes})$, $\tilde{\otimes}(B_{kj}) \in [\underline{B}_{kj}, \overline{B}_{kj}]$。

设有 m 个待识别对象, 它们对于 n 个指标所具有的灰色特征值矩阵为

$$A(\otimes) = (\otimes(A_{ki}))_{nm} \tag{2}$$

式中: $\otimes(A_{ki}) \in g(\dot{\otimes})$, $\tilde{\otimes}(A_{ki}) \in [\underline{A}_{ki}, \overline{A}_{ki}]$。

由于 n 个指标的量纲不尽相同, 而且数值上往往相差悬殊, 为了有利于灰色关联分析及识别比较, 可采用模糊相对隶属度概念将 $\otimes(B_{kj})$ 和 $\otimes(A_{kj})$ 进行规格化处理, 即

$$\otimes(b_{kj}) = \begin{cases} \dfrac{\overline{B}_{ks}\otimes(1) - \otimes(B_{kj})}{\overline{B}_{ks} - \underline{B}_{k1}}, & \overline{B}_{k1} < \overline{B}_{ks} \\ \dfrac{\otimes(B_{kj}) - \underline{B}_{ks}\otimes(1)}{\overline{B}_{k1} - \underline{B}_{ks}}, & \overline{B}_{k1} > \overline{B}_{ks} \end{cases} \tag{3}$$

$$\otimes(a_{ki}) = \begin{cases} \dfrac{\overline{B}_{ks}\otimes(1) - \otimes(A_{ki})}{\overline{B}_{ks} - \underline{B}_{k1}}, & \overline{B}_{k1} < \overline{B}_{ks} \\ \dfrac{\otimes(A_{ki}) - \underline{B}_{ks}\otimes(1)}{\overline{B}_{k1} - \underline{B}_{ks}}, & \overline{B}_{k1} > \overline{B}_{ks} \end{cases} \tag{4}$$

其中: 当 $\otimes(a_{ki})$ 白化区间中的上限 $\overline{a}_{ki} > 1$ 时, 取 $\overline{a}_{ki} = 1$; 下限 $\underline{a}_{ki} < 0$ 时, 取 $\underline{a}_{ki} = 0$。

经规格化处理后, 则 $\tilde{\otimes}(b_{kj}) \in [\underline{b}_{kj}, \overline{b}_{kj}] \subseteq [0, 1]$, $\tilde{\otimes}(a_{ki}) \in [\underline{a}_{ki}, \overline{a}_{ki}] \subseteq [0, 1]$。

4 灰色关联识别方法

取第 i 个待识别对象规格化后的灰色特征值序列与标准模式规格化后的灰色标准特征值序列进行比较, 则第 i 个待识别对象与第 j 个标准模式在第 k 个指标处的灰关联系数可表示为

$$\otimes(\xi_{ij}(k)) = \frac{\zeta \max_j \max_k \overline{\rho}_{b_{kj}, a_{ki}}}{\otimes(\rho(\otimes(b_{ij}), \otimes(a_{ki}))) + \zeta \max_j \max_k \overline{\rho}_{b_{kj}, a_{ki}}} \tag{5}$$

式中: $\otimes(\xi_{ij}(k)) \in g(\dot{\otimes})$; $\tilde{\otimes}(\xi_{ij}(k)) \in [\underline{\xi}_{ij}(k), \overline{\xi}_{ij}(k)]$; ζ 为分辨系数, 其具体取值可根据“充分体现关联度的整体性和具有抗干扰作用”的原则[4] 选取。一般情况下, 可取 $\zeta = 0.5$ 进行计算。

定义 3 设 $\otimes(b_{kj}), \otimes(a_{ki}) \in g(\dot{\otimes})$, 若 $\tilde{\otimes}(\xi_{ij}(k)) = \tilde{\otimes}(1), \forall k \le n$, 则称第 i 个待识别对象与第 j 类标准模式完全相关。

定义 4 若任给一个 $n \in I$, 至少存在一个 $k \le n$ 使得 $\tilde{\otimes}(\xi_{ij}(k)) \ne \tilde{\otimes}(1)$ 成立, 则称第 i 个待识别对象与第 j 类标准模式非完全相关。

若第 i 个待识别对象与第 j 类标准模式非完全相关, 则表明二者的相关程度与完全相关存在差异, 这种差异程度可用灰色广义权距离表示为

$$\otimes(d_{ij}) = \sqrt[p]{\sum_{k=1}^{n} [w_{kj}(\otimes(1) - \otimes(\xi_{ij}(k)))]^p} \tag{6}$$

式中 w_{kj} 为不同模式对于不同指标的权重。

将灰色关联度 $\otimes(\nu_{ij})$ 视为权重, 引进权广义距离为

$$\otimes(D_{ij})=\otimes(\nu_{ij})\otimes(d_{ij}) \tag{7}$$

设 $\tilde{\otimes}(\nu_{ij}^0)$ 为 $\otimes(\nu_{ij})$ 的白化值真值, $\tilde{\otimes}(d_{ij}^0)$ 为 $\otimes(d_{ij})$ 的白化值真值。为了决定 $\tilde{\otimes}(\nu_{ij}^0)$, 可建立下列目标函数

$$\begin{cases} \min F(\nu_{ij})=\min\sum_{l=1}^{m}\sum_{j=1}^{s}\tilde{\otimes}^2(\nu_{ij}^0)\,\tilde{\otimes}^2(d_{ij}^0) \\ \text{s.t.}\ \sum_{j}^{s}\tilde{\otimes}(\nu_{ij}^0)=1 \end{cases} \tag{8}$$

求解式(8) 可得灰色关联度 $\tilde{\otimes}(\nu_{ij}^0)$ 的最优值为

$$\tilde{\otimes}(\nu_{ij}^0)=\left[\tilde{\otimes}(1)+\sum_{l=1,l\neq j}^{s}\frac{\tilde{\otimes}(T_{ij}^0)}{\tilde{\otimes}(T_{il}^0)}\right]^{-1} \tag{9}$$

式中 $\tilde{\otimes}(T_{ij}^0)$ 为 $\otimes(T_{ij})=\otimes^2(d_{ij})$ 的白化值真值。

由于 $\otimes(T_{ij})$ 的白化值真值未知, 而仅知道 $\tilde{\otimes}(T_{ij})\in[\underline{T}_{ij},\overline{T}_{ij}]$, 故 $\otimes(\nu_{ij})$ 真值也是未知的, 也仅知道 $\tilde{\otimes}(\nu_{ij})\in[\underline{\nu}_{ij},\overline{\nu}_{ij}]$, 即灰色关联度实际上仍为一个区间灰数。因此有

$$\otimes(\nu_{ij})=\left[\otimes(1)+\sum_{l=1,l\neq j}^{s}\frac{\otimes(T_{ij})}{\otimes(T_{il})}\right]^{-1} \tag{10}$$

令

$$\otimes(\nu_{il})=\max_j(\otimes(\nu_{ij})) \tag{11}$$

根据择大原则, 可将第 i 个待识别对象划归第 l 类模式。

5　灰色识别结果的可信度

定义 5　设 $\otimes(a)\in g(\bar{\otimes})$, $\tilde{\otimes}(a)\in[\underline{a},\overline{a}]$, 则称 $(\underline{a}+\overline{a})/2$ 处为 $\otimes(a)$ 的灰心, 记为 $\odot(\otimes(a))$。

设 $\otimes(\nu_{ij})$ 的真值在其白化区间上随机取值, 且服从均匀分布, 则当 $\odot(\otimes(\nu_{il}))=\max_j\odot(\otimes(\nu_{ij}))$ 时, 必有 $\otimes(\nu_{il})=\max_j(\otimes(\nu_{ij}))$ 成立。因此, 只要找出最大的灰心 $\odot(\otimes(\nu_{il}))$, 则可将第 i 个待识别对象划归第 l 类模式。

由于 $\otimes(\nu_{ij})$ $(j=1,2,\cdots,s)$ 的白化值为区间, 故 $\otimes(\nu_{il})$ 的白化区间与 $\otimes(\nu_{ij})$ $(j\neq l)$ 的白化区间可能存在不相交、相交和包含 3 种关系, 故按择大原则来识别对象所属的类别还存在可信度问题。借助于文献[5] 提出的可能度的概念, 可给出关于第 i 个待识别对象划归为第 l 类模式的可信度。

定义 6　设 $\odot(\otimes(\nu_{il}))=\max_j\odot(\otimes(\nu_{ij}))$, 则称

$$\mu(\otimes(\nu_{il}))=\begin{cases} 1.0, & [\underline{\nu}_{il},\overline{\nu}_{il}]\bigcap_{l\neq j}[\underline{\nu}_{ij},\overline{\nu}_{ij}]=\phi \\ \min_j\left[\dfrac{\overline{\nu}_{il}-\overline{\nu}_{ij}}{\overline{\nu}_{il}-\underline{\nu}_{il}}+\dfrac{(\overline{\nu}_{ij}-\underline{\nu}_{il})(\underline{\nu}_{il}-\underline{\nu}_{ij})}{(\overline{\nu}_{il}-\underline{\nu}_{il})(\overline{\nu}_{ij}-\underline{\nu}_{ij})}+\dfrac{(\overline{\nu}_{ij}-\underline{\nu}_{il})^2}{2(\overline{\nu}_{il}-\underline{\nu}_{il})(\overline{\nu}_{ij}-\underline{\nu}_{ij})}\right] & [\underline{\nu}_{il},\overline{\nu}_{il}]\bigcap_{l\neq j}[\underline{\nu}_{ij},\overline{\nu}_{ij}]\neq\phi,\ [\underline{\nu}_{il},\overline{\nu}_{il}]\not\supset_{l\neq j}[\underline{\nu}_{ij},\overline{\nu}_{ij}] \\ \min_j\left(\dfrac{\overline{\nu}_{il}-\overline{\nu}_{ij}}{\overline{\nu}_{il}-\underline{\nu}_{il}}+\dfrac{\overline{\nu}_{ij}-\underline{\nu}_{ij}}{2(\overline{\nu}_{il}-\underline{\nu}_{il})}\right) & [\underline{\nu}_{ij},\overline{\nu}_{ij}]\subset_{l\neq j}[\underline{\nu}_{il},\overline{\nu}_{il}] \end{cases} \tag{12}$$

为第 i 个待识别对象划归第 l 类模式的可信度。

由可信度定义知, 当 $\odot(\otimes\nu_{il_1})=\odot(\otimes\nu_{il_2})=\max_j\odot(\otimes(\nu_{il}))$ 时, 有 $\mu(\otimes(\nu_{il_1}))=\mu(\otimes(\nu_{il_2}))=0.5$, 这时无法将第 i 个待识别对象划归某一确定的模式, 而只有当进一步获取信息后才能断定是 l_1 类还是 l_2 类模式。

6　工程应用实例

参考国内外围岩分类经验, 可将围岩稳定性分为 5 类, 5 类模式对于 5 个指标所具有的灰色标准指标特征值白化矩阵为

$$\begin{array}{cc} & \begin{array}{ccccc} \text{I} & \text{II} & \text{III} & \text{IV} & \text{V} \end{array} \\ \tilde{B}(\otimes)\in & \left[\begin{array}{ccccc} [90,100] & [75,90] & [50,75] & [25,50] & [0,25] \\ [120,200] & [60,120] & [30,60] & [15,30] & [0,15] \\ [0.75,1.0] & [0.45,0.75] & [0.3,0.45] & [0.2,0.3] & [0,0.2] \\ [0.8,1.0] & [0.6,0.8] & [0.4,0.6] & [0.2,0.4] & [0,0.2] \\ [0,5] & [5,10] & [10,25] & [25,125] & [125,300] \end{array}\right] \begin{array}{l} RQD \\ R_w \\ K_v \\ K_f \\ W/L \end{array} \end{array}$$

待识别的某围岩对于 5 个指标所具有的灰色指标特征值白化矩阵为

$$\begin{array}{cc} & \begin{array}{ccccc} RQD & R_w & K_v & K_f & W/L \end{array} \\ \tilde{A}(\otimes)\in & [[77.65,86.35]\ [90.25,99.75]\ [0.665,0.745]\ [0.325\ \ 0.375]\ [19.0,21.0]]^{\mathrm{T}} \end{array}$$

设各指标的权重等权, 由本文提出的灰色模式识别方法, 可求出待识别对象与 5 类标准模式的灰色关联度分别为

$$\tilde{\otimes}(\nu_{11}) \in [0.055\,1,\quad 0.438\,6]$$
$$\tilde{\otimes}(\nu_{12}) \in [0.125\,8,\quad 0.700\,1]$$
$$\tilde{\otimes}(\nu_{13}) \in [0.077\,4,\quad 0.578\,6]$$
$$\tilde{\otimes}(\nu_{14}) \in [0.043\,9,\quad 0.298\,8]$$
$$\tilde{\otimes}(\nu_{15}) \in [0.027\,3,\quad 0.198\,7]$$

通过灰关联度灰心计算后, 有

$$\max_j \odot(\otimes(\nu_{lj})) = \odot(\otimes(\nu_{12})) = 0.413\,2$$

因此, 待识别对象划归为第 2 类模式。由式(12) 可求出划归为第 2 类模式的可信度为

$$\mu(\otimes(\nu_{12})) = 0.644\,2$$

7 结 语

本文将标准模式的指标特征值和待识别对象的指标特征值均视为区间灰数, 能较好地考虑由人们认知能力的有限性、客观事物本身的复杂性以及信息接收系统能力的局限性所带来的不确定性, 从而使模式识别更为科学、更加接近实际。由于普通实数是区间灰数的完全退化, 因此只要将本文所有公式中区间灰数白化值的上、下限取为一致, 则本文模式识别的模型便能方便地应用于标准模式指标值和待识别对象指标值均为普通实数的模式识别。

本文给出了待识别对象划归为某一类模式的可信度, 从而使得人们在进行模式识别时更加心中有数。

参考文献(References):

[1] Deng Julong. Grey hazy sets [J]. *J of Grey System*, 1992, 4(1): 13-30.

[2] Deng Julong. Extent information cover in grey system theory[J]. *J of Grey Systems*, 1995, 7(2): 131-138.

[3] 邓聚龙. 灰色控制系统(第 2 版)[M]. 武汉: 华中理工大学出版社, 1993. 102-110.

[4] 吕峰. 灰色系统关联度之分辨系数的研究[J]. 系统工程理论与实践, 1997, 17(6): 49-54.
(Lu F. Research on the identification coefficient of relational grade for grey system [J]. *Systems Engineering-Theory & Practice*, 1997, 17(6): 49-54.)

[5] 张全, 樊治平, 潘德惠, 等. 不确定性多属性决策中区间数的一种排序方法[J]. 系统工程理论与实践, 1999, 19(5): 129-133.
(Zhang Q, Fan Z P, Pan D H, et al. A ranking approach for interval numbers in uncertain multiple attribute decision making problems [J]. *Systems Engineering-Theory & Practice*, 1999, 19(5): 129-133.)

(上接第 592 页)

参考文献(References):

[1] Brennan M, Schwartz E. Evaluation natural resource investment[J]. *J of Business*, 1985, 58 (2): 135-157.

[2] Pindyck R. Irreversible investment, capacity choice, and the value of the firm [J]. *American Economic Review*, 1988, 79(12): 969-985.

[3] Trigeorgis L. Real options and interactions with financial flexibility [J]. *Financial Management*, 1993, (Autumn): 202-224.

[4] Dangl T. Investment and capacity choice under uncertain demand[J]. *European J of Operational Research*, 1999, 117: 415-428.

[5] Kester W. Today's options for tomorrow's growth[J]. *Harvard Business Review*, 1984, (March-April): 153-160.

[6] Pennings E, Lint O. Market entry, phased rollout or abandonment? A real option approach[J]. *European J of Operational Research*, 2000, 124: 125-138.

[7] Ottoo R. Valuation of internal growth opportunities: The case of a biotechnology company[J]. *The Quarterly Review of Economics and Finance*, 1998, 38 (SI): 537-567.