

控制与决策

Control and Decision

基于面板数据的灰色四面体网络关联分析模型及应用

吴鸿华, 刘思峰, 方志耕

引用本文:

吴鸿华, 刘思峰, 方志耕. 基于面板数据的灰色四面体网络关联分析模型及应用[J]. *控制与决策*, 2022, 37(11): 3033–3041.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0753>

您可能感兴趣的其他文章

Articles you may be interested in

[融合粗糙集与GRA的异构信息多准则三支推荐及其在医疗推荐中的应用](#)

Multi-criteria three-way recommendation of heterogeneous information based on rough set and GRA and its application in medical recommendation

控制与决策. 2022, 37(7): 1883–1893 <https://doi.org/10.13195/j.kzyjc.2020.1631>

[考虑多重决策心理的海岛孤网调度综合评价](#)

Comprehensive evaluation of island isolated microgrid dispatch considering multiple decision psychology

控制与决策. 2022, 37(6): 1591–1600 <https://doi.org/10.13195/j.kzyjc.2020.1792>

[区间条件下基于GRA和TOPSIS的辐射源威胁评估](#)

Combining TOPSIS and GRA for emitter threat evaluation with interval number

控制与决策. 2021, 36(6): 1516–1522 <https://doi.org/10.13195/j.kzyjc.2019.1559>

[基于资源特征融合的复杂型面配合决策方法](#)

Decision-making method for complex surface coordination based on resource feature fusion

控制与决策. 2021, 36(6): 1368–1376 <https://doi.org/10.13195/j.kzyjc.2019.1311>

[大群体应急决策中考虑属性关联的偏好信息融合方法](#)

Preference information fusion method of large groups emergency decision-making based on attributes association

控制与决策. 2021, 36(10): 2537–2546 <https://doi.org/10.13195/j.kzyjc.2020.0117>

基于面板数据的灰色四面体网格关联分析模型及应用

吴鸿华^{1,2†}, 刘思峰¹, 方志耕¹

(1. 南京航空航天大学 经济与管理学院, 南京 211106; 2. 济南大学 数学科学学院, 济南 250022)

摘要: 针对现有面板数据关联分析模型中指标的排列顺序变换引起灰色关联序不一致以及面板数据三角曲面构造的不足, 提出一种基于面板数据的灰色四面体网格关联分析模型. 首先, 基于数列的组合原理将样本矩阵分解为二元指标子矩阵; 其次, 将二元指标子矩阵投射为空间四面体网格, 并给出四面体的有向体积; 然后, 基于有向体积构造二元指标的关联系数, 最终得到灰色四面体网格关联分析模型, 并对模型的对称性、规范性、平移不变性等性质进行证明; 最后, 将模型应用到山东省沿海城市的空气质量评价, 验证模型的合理性和有效性, 并分析了灰色四面体网格关联分析模型不受指标排列顺序变化的影响.

关键词: 面板数据; 关联度; 灰色关联分析; 灰色系统理论; 四面体

中图分类号: N941.5

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0753

引用格式: 吴鸿华, 刘思峰, 方志耕. 基于面板数据的灰色四面体网格关联分析模型及应用[J]. 控制与决策, 2022, 37(11): 3033-3041.

Grey tetrahedral grid incidence analysis model based on panel data and its application

WU Hong-hua^{1,2†}, LIU Si-feng¹, FANG Zhi-geng¹

(1. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; 2. School of Mathematical Sciences, University of Jinan, Jinan 250022, China)

Abstract: A grey tetrahedral grid incidence analysis model based on panel data is proposed for the inconsistency of grey incidence order resulting from an order changing of indicators in the grey incidence model and the deficiency of the construction of trigonometric surface for panel data. Firstly, the sample matrix is decomposed into binary indicator sub-matrixes based on the combination principle of sequence. Then, the binary indicator sub-matrix is projected into a spatial tetrahedral grid and the directed volume of the tetrahedroid is given. Based on the directed volume, the incidence coefficient of the binary indicators is constructed. And then, the grey tetrahedral grid incidence analysis model is obtained. Furthermore, the properties of the model including symmetry, normalization, translation-invariance are verified. As a result, the model is used to assess the air quality of coastal cities in Shandong province verifying the rationality and effectiveness of the model, showing that the grey tetrahedron grid incidence analysis model is not related to the order changing of indicators.

Keywords: panel data; degree of incidence; grey incidence analysis; grey system theory; tetrahedroid

0 引言

灰色关联分析是灰色系统理论的一个重要组成部分, 是灰色控制、灰色决策、灰色聚类的基础, 基本思想是根据数据序列之间的几何特征来判别两者的关联程度^[1]. 自邓聚龙教授提出邓氏关联模型以来^[2], 众多学者进行了富有成效的研究, 从不同的角

度构建了一系列的关联模型, 例如灰熵关联模型^[3]、灰色 B 型关联模型^[4]、灰色 C 型关联模型^[5]、灰色 T 型关联模型^[6]、灰色斜率关联模型^[7]、灰色欧几里得关联模型^[8]、灰色生成速率关联模型^[9]、接近性和相似性视角的灰色关联模型^[10] 等等, 这些研究有效推动了灰色关联分析理论的发展. 在众多模型中, 其中

收稿日期: 2021-04-29; 录用日期: 2021-08-18.

基金项目: 国家自然科学基金项目(72071111, 71801127, 71671091); 国家自然科学基金与英国皇家学会国际合作交流项目(71811530338); 国家科技部科技创新引智基地项目(G20190010178); 中国博士后科学基金特别资助项目(2019TQ0150); 中央高校基本科研业务费专项基金项目(NC2019003, XBA19003).

责任编辑: 樊治平.

†通讯作者. E-mail: ss_wuhh@ujn.edu.cn.

以邓氏关联模型最为经典,应用最为广泛.

近年来,随着数据维度的增加,关于面板数据的灰色关联分析已渐渐成为研究热点.文献[11]基于灰色绝对关联模型和二重积分构造出三维对象的灰色关联分析模型,首次将灰色关联模型的适用范围拓展至面板数据;文献[12]将邓氏关联模型拓展至三维空间,构建了包含时间、指标、方案3个维度的灰色关联模型;文献[13]整合“水平”“增量”“变异”3个维度信息,构建了灰色矩阵关联分析模型;文献[14-15]提出了空间网格的概念,分别构建了相似性和接近性视角的面板数据灰色关联模型;文献[16]从个体和时间维度定义了系统特征行为矩阵之间的发展速度指数和增长速度指数,构建了面板数据灰色矩阵相似关联模型;文献[17]将面板数据投射为 n 维空间向量,然后基于空间向量的夹角和向量的模分别建立了面板数据的相似性和接近性关联模型;文献[18]基于B型关联度构建了面板数据的灰色B型关联度公式;文献[19]在二维灰色凸关联度的基础上提出了面板数据的凸关联模型;文献[20]基于空间曲线的曲率分别构建了指标维度和时间维度的关联度,最终通过加权平均得到面板数据的灰色曲率关联模型;文献[21]基于动态权函数建立了面板数据的动态加权关联模型.针对面板数据,上述研究从不同的角度构建了灰色关联分析模型,丰富了面板数据灰色关联分析理论.但是现有的面板数据关联分析模型仍存在以下不足:

1) 在描述面板数据的空间特征时,往往投射为三角曲面^[11],但顶点不同的连接方法得到的离散三角曲面有所不同,导致由三角曲面所构成的几何量(如体积)不同,从而影响由其构建的关联模型的结果.

2) 对于现有模型,不同指标排列顺序所构成的面板数据,得到的关联度往往不同,即关联分析模型受指标顺序改变的影响.文献[22-23]提出了不受指标顺序影响的关联分析模型,但未考虑指标之间的耦合作用.事实上,在空气质量评价中,大气环境指标因素相互影响,如 PM_{10} 与 $PM_{2.5}$ 有较强的关联特征, $PM_{2.5}$ 对 NH_3 较为敏感^[24],不同的季节 $PM_{2.5}$ 对空气质量的6个因素有不同的影响^[25].

鉴于此,本文给出新的面板数据投射方法,即将面板数据投射为空间四面体网格.四面体网格是由体、面、线和点组成的集合,不仅含有实体结构具有的优势,而且还包含一些边界表示的优点^[26].为了解决指标排列顺序的影响,结合数列的组合原理,构建面板数据的灰色四面体关联系数,并最终得到面

板数据的灰色四面体网格关联分析模型.为此,首先给出样本行为矩阵的二元指标行为子矩阵;其次,二元指标行为子矩阵的指标值投射为三维空间中的点,将相邻的四点两两相连,构成面板数据的空间四面体网格;然后,基于四面体的有向体积构建面板数据的关联系数,并加权平均得到灰色四面体关联分析模型;最后,将新的模型应用于山东省沿海城市的空气质量评价,结果表明所提出的模型能够很好反映面板数据的相关程度,通过与已有模型进行对比,进一步说明了模型的合理性和有效性.

1 面板数据的空间四面体网格构建

面板数据包括指标截面数据和时间信息,同时具有时间维度和指标维度.为了描述面板数据的几何特征,文献[11]将面板数据对应为 $m \times n$ 矩阵.

定义1 设 $x_i(s, t)$ 为样本 i 关于指标 s 在时间 t 的数值,则称

$$X_i = \begin{bmatrix} x_i(1, 1) & x_i(1, 2) & \dots & x_i(1, n) \\ x_i(2, 1) & x_i(2, 2) & \dots & x_i(2, n) \\ \vdots & \vdots & \ddots & \vdots \\ x_i(m, 1) & x_i(m, 2) & \dots & x_i(m, n) \end{bmatrix} \quad (1)$$

为样本 i 的行为矩阵,简记为 $X_i = (x_i(s, t))_{m \times n}$,称面板数据序列 $X = (X_1, X_2, \dots, X_N)$ 为面板数据样本序列.

定义1给出了样本 i 的一种描述形式,即样本的指标顺序固定,现有的面板数据关联模型多以此进行建模,关联分析模型的结果往往与指标排列顺序有关.

为此,结合数列的组合原理,下面给出样本矩阵的二元指标行为子矩阵的概念.

定义2 设 $X_i = (x_i(s, t))_{m \times n}$ 为样本 i 的行为矩阵,样本有 m 个指标,指标集为 $S = \{s_1, s_2, \dots, s_m\}$,任取 S 中两个指标记为 s_{j_1} 和 s_{j_2} ,其中 j_1 和 j_2 为 $\{1, 2, \dots, m\}$ 中的任意两个不同的数,则称 $\{s_{j_1}, s_{j_2}\}$ 为指标集的二元组合.所有可能的二元指标组合的集合记为

$$M_S^2 = \{\{s_{j_1}, s_{j_2}\} | j_1 \neq j_2 \in \{1, 2, \dots, m\}\}.$$

由指标 s_{j_1} 和 s_{j_2} 所对应的数据构成 $2 \times n$ 矩阵,记为 $X_{i(j_1, j_2)}$,则称 $X_{i(j_1, j_2)}$ 为矩阵 X_i 的二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵.

$$X_{i(j_1, j_2)} = \begin{bmatrix} x_i(j_1, 1) & x_i(j_1, 2) & \dots & x_i(j_1, n) \\ x_i(j_2, 1) & x_i(j_2, 2) & \dots & x_i(j_2, n) \end{bmatrix}.$$

显然, m 个不同指标的二元组合共有 $m(m-1)/2$ 种可能,所以 X_i 的二元指标的行为子矩阵共有 $m(m-1)/2$ 个.

为了描述面板数据的空间特征, 将 X_i 的二元指标行为子矩阵 $X_{i(j_1, j_2)}$ 的数值投射为三维空间的点, 即 $(1, t, x_i(j_1, t)), (2, t, x_i(j_2, t)), t = 1, 2, \dots, n$.

定义 3 设 $X_i = (x_i(s, t))_{m \times n}$ 为样本 i 的行为矩阵, $X_{i(j_1, j_2)}$ 为矩阵 X_i 的二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵, 将 $X_{i(j_1, j_2)}$ 所对应的相邻 4 个点 $(1, t, x_i(j_1, t)), (2, t, x_i(j_2, t)), (1, t+1, x_i(j_1, t+1)), (2, t+1, x_i(j_2, t+1)) (t = 1, 2, \dots, n-1)$ 两两相连, 所构成的几何体记为 $T_i(j_1, j_2, t)$, 则称 $T_i(j_1, j_2, t)$ 为 $X_{i(j_1, j_2)}$ 在时刻 t 的四面体. $X_{i(j_1, j_2)}$ 所有的四面体 $T_i(j_1, j_2, t)$ 所构成的几何体记为 $T_i(j_1, j_2)$, 则称 $T_i(j_1, j_2)$ 为 $X_{i(j_1, j_2)}$ 的四面体网格.

体积是四面体的一个重要的空间几何量. 但需要注意的是, 如果底面面积相同高度也相同但方向不同的两个四面体的体积相等, 如四面体 $ABCD$, 顶点 A 关于平面 BCD 的对称点为 A' , 则四面体 $ABCD$ 与四面体 $A'BCD$ 的体积相等. 为了刻画四面体的这种差异, 下面给出有向体积的概念.

定义 4 设 $X_i = (x_i(s, t))_{m \times n}$ 为样本 i 的行为矩阵, $X_{i(j_1, j_2)}$ 为矩阵 X_i 的二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵, $T_i(j_1, j_2, t)$ 为 $X_{i(j_1, j_2)}$ 在 t 时刻的四面体. 设 $T_i(j_1, j_2, t)$ 的体积为 $\bar{V}_i(j_1, j_2, t)$, 则 $T_i(j_1, j_2, t)$ 的有向体积为

$$V_i(j_1, j_2, t) = \begin{cases} \bar{V}_i(j_1, j_2, t), & \tilde{x}_i(j_1, j_2, t) > \tilde{x}_i(j_2, j_1, t); \\ 0, & \tilde{x}_i(j_1, j_2, t) = \tilde{x}_i(j_2, j_1, t); \\ -\bar{V}_i(j_1, j_2, t), & \tilde{x}_i(j_1, j_2, t) < \tilde{x}_i(j_2, j_1, t). \end{cases} \quad (2)$$

其中

$$\begin{aligned} \tilde{x}_i(j_1, j_2, t) &= x_i(j_1, t) + x_i(j_2, t + 1), \\ \tilde{x}_i(j_2, j_1, t) &= x_i(j_2, t) + x_i(j_1, t + 1). \end{aligned}$$

当 $x_i(j_1, t) + x_i(j_2, t + 1) > x_i(j_2, t) + x_i(j_1, t + 1)$ 时, 有向体积取正值; 当 $x_i(j_1, t) + x_i(j_2, t + 1) < x_i(j_2, t) + x_i(j_1, t + 1)$ 时, 有向体积取负值; 当 $x_i(j_1, t) + x_i(j_2, t + 1) = x_i(j_2, t) + x_i(j_1, t + 1)$ 时, 空间四面体退化为空间平面, 此时体积为 0.

定理 1 设 $X_i = (x_i(s, t))_{m \times n}$ 为样本 i 的行为矩阵, $X_{i(j_1, j_2)}$ 为矩阵 X_i 的二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵, $T_i(j_1, j_2, t)$ 为 $X_{i(j_1, j_2)}$ 在 t 时刻的四面体, 则 $T_i(j_1, j_2, t)$ 的体积为

$$\bar{V}_i(j_1, j_2, t) = \frac{1}{6} |x_i(j_1, t) + x_i(j_2, t + 1) - x_i(j_2, t) - x_i(j_1, t + 1)|.$$

证明 将点 $(1, t, x_i(j_1, t)), (2, t, x_i(j_2, t)), (1, t + 1, x_i(j_1, t + 1)), (2, t + 1, x_i(j_2, t + 1))$ 分别记为 $A、B、$

$C、D$, 则

$$\begin{cases} \vec{AB} = (1, 0, x_i(j_2, t) - x_i(j_1, t)), \\ \vec{AC} = (0, 1, x_i(j_1, t + 1) - x_i(j_1, t)), \\ \vec{AD} = (1, 1, x_i(j_2, t + 1) - x_i(j_1, t)). \end{cases}$$

由立体几何知识可知, 四面体的体积等于以向量 $\vec{AB}、\vec{AC}、\vec{AD}$ 为棱的平行六面体体积的 1/6, 而 $|\vec{AB}, \vec{AC}, \vec{AD}|$ 表示向量 $\vec{AB}、\vec{AC}、\vec{AD}$ 为棱的平行六面体体积.

又因为

$$[\vec{AB}, \vec{AC}, \vec{AD}] = \begin{vmatrix} 1 & 0 & x_i(j_2, t) - x_i(j_1, t) \\ 0 & 1 & x_i(j_1, t + 1) - x_i(j_1, t) \\ 1 & 1 & x_i(j_2, t + 1) - x_i(j_1, t) \end{vmatrix},$$

所以

$$\bar{V}_i(j_1, j_2, t) = \frac{1}{6} |x_i(j_1, t) + x_i(j_2, t + 1) - x_i(j_2, t) - x_i(j_1, t + 1)|. \quad \square$$

由于面板数据所形成的空间四面体网格往往是一些四面体薄片, 计算体积时去掉前面的系数 1/6, 可适当扩大关联度的区分度, 后面也将证明数乘不影响关联序. 乘以 6 后的体积与四面体的体积一一对应, 这里仍称为四面体的体积, 后面均用去掉 1/6 后的体积做运算.

在给定四面体的体积以及去掉系数后, 定义 4 中的有向体积为

$$V_i(j_1, j_2, t) = x_i(j_1, t) + x_i(j_2, t + 1) - x_i(j_2, t) - x_i(j_1, t + 1). \quad (3)$$

$T_i(j_1, j_2, t)$ 所对应的四面体体积与指标 $s_{j_1}、s_{j_2}$ 顺序无关, 即 $\bar{V}_i(j_1, j_2, t) = \bar{V}_i(j_2, j_1, t)$, 也就是说, 交换二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵的两行四面体体积不变. 再由定义 2 可知, 样本 i 的所有二元指标组合包含了所有可能排列, 即整体包含了所有指标排列顺序. 需要注意的是, 交换指标 s_{j_1} 和 s_{j_2} 的顺序后有向体积的符号会发生改变, 但不同样本有向体积差的绝对值是不变的.

定义 5 设面板数据样本序列为 $X = (X_1, X_2, \dots, X_N)$, 其中 $X_i = (x_i(s, t))_{m \times n}$. 记

$$\begin{aligned} M_s &= \max_{1 \leq i \leq N; 1 \leq t \leq n} \{x_i(s, t)\}, \\ m_s &= \min_{1 \leq i \leq N; 1 \leq t \leq n} \{x_i(s, t)\}, \\ X_i D &= x_i(s, t)d. \end{aligned}$$

1) 如果数据为成本型, 则有

$$x_i(s, t)d = \frac{M_s - x_i(s, t)}{M_s - m_s}; \quad (4)$$

2) 如果数据为效益型,则有

$$x_i(s, t)d = \frac{x_i(s, t) - m_s}{M_s - m_s}. \quad (5)$$

称 D 为极差变换算子, $X_i D$ 为 X_i 的极差变换像.

原始数据通过极差变换处理,消除了原始数据量纲量级的影响,且 $x_i(s, t)d \in [0, 1], s = 1, 2, \dots, m, t = 1, 2, \dots, n$.

2 灰色四面体网格关联分析模型的构建及性质

对于面板数据,不同的四面体反映了面板数据的差异性. 也就是说,四面体的有向体积越接近,他们之间的相关性越好,反之越差. 所以,可以利用面板数据的有向体积来刻画不同面板数据的相关程度.

定义6 设 $X_0 = (x_0(s, t))_{m \times n}$ 为系统特征行为矩阵, $X_i = (x_i(s, t))_{m \times n}$ 为系统相关行为矩阵, $X_{0(j_1, j_2)}$ 和 $X_{i(j_1, j_2)}$ 分别是 X_0 及 X_i 的二元指标 s_{j_1} 和 s_{j_2} 的行为子矩阵, $V_0(j_1, j_2, t)$ 和 $V_i(j_1, j_2, t) (t = 1, 2, \dots, n - 1)$ 分别为 t 时刻的四面体 $T_0(j_1, j_2, t)$ 和 $T_i(j_1, j_2, t)$ 的有向体积. 记

$$\varepsilon_{0i}(j_1, j_2, t) = \frac{1}{1 + |V_0(j_1, j_2, t) - V_i(j_1, j_2, t)|}, \quad (6)$$

则称 $\varepsilon_{0i}(j_1, j_2, t)$ 为 X_0 与 X_i 在 t 时刻关于指标 s_{j_1} 和 s_{j_2} 的灰色四面体网格关联系数.

定义7 设 $X_0 = (x_0(s, t))_{m \times n}$ 为系统特征行为矩阵, $X_i = (x_i(s, t))_{m \times n}$ 为系统相关行为矩阵, $\varepsilon_{0i}(j_1, j_2, t)$ 为 X_0 与 X_i 在 t 时刻关于指标 s_{j_1} 和 s_{j_2} 的灰色四面体网格关联系数. 记

$$\varepsilon_{0i} = \frac{2}{m(m-1)(n-1)} \sum_{s_{j_1}, s_{j_2} \in M_s^2} \sum_{t=1}^{n-1} \varepsilon_{0i}(j_1, j_2, t), \quad (7)$$

则称 ε_{0i} 为 X_0 与 X_i 的灰色四面体网格关联度.

性质1 基于面板数据的灰色四面体网格关联度 ε_{0i} 满足 $0 < \varepsilon_{0i} \leq 1$.

证明 对于 X_0 和 $X_i, \forall j_1, j_2 \in \{1, 2, \dots, m\}$, 显然 $|V_0(j_1, j_2, t) - V_i(j_1, j_2, t)| \geq 0$, 则 $0 < 1/(1 + |V_0(j_1, j_2, t) - V_i(j_1, j_2, t)|) \leq 1$, 即 $0 < \varepsilon_{0i}(j_1, j_2, t) \leq 1$, 所以 $0 < \varepsilon_{0i} \leq 1$. \square

性质2 设 X_0 和 X_i 为两同型样本行为矩阵, 则 $\varepsilon_{0i} = \varepsilon_{i0}$.

证明 对于 X_0 和 $X_i, \forall j_1, j_2 \in \{1, 2, \dots, m\}$, 显然 $|V_0(j_1, j_2, t) - V_i(j_1, j_2, t)| = |V_i(j_1, j_2, t) - V_0(j_1, j_2, t)|$, 得 $\varepsilon_{0i}(j_1, j_2, t) = \varepsilon_{i0}(j_1, j_2, t)$, 所以 $\varepsilon_{0i} = \varepsilon_{i0}$. \square

性质3 设 X_0 和 X_i 为两同型样本行为矩阵, 如果 $X_0 = X_i$, 则 $\varepsilon_{0i} = 1$.

证明 对于 X_0 和 $X_i, \forall j_1, j_2 \in \{1, 2, \dots, m\}$, 由于 $X_0 = X_i$, 显然 $|V_0(j_1, j_2, t) - V_i(j_1, j_2, t)| = 0$, 即 $\varepsilon_{0i}(j_1, j_2, t) = 1$, 所以 $\varepsilon_{0i} = 1$. \square

性质4 设 X_0, X'_0, X_i, X'_i 为4个同型样本行为矩阵, 如果 $X'_0 = X_0 + c_1, X'_i = X_i + c_2$, 其中 c_1, c_2 为任意常数, 则 $\varepsilon'_{0i} = \varepsilon_{0i}$.

证明 对于 X_0 和 $X'_0, \forall j_1, j_2 \in \{1, 2, \dots, m\}$, 由于 $X'_0 = X_0 + c_1$, 有

$$\begin{aligned} V'_0(j_1, j_2, t) &= \\ x'_0(j_1, t) + x'_0(j_2, t+1) - x'_0(j_2, t) - x'_0(j_1, t+1) &= \\ (x_0(j_1, t) + c) + (x_0(j_2, t+1) + c) - & \\ (x_0(j_2, t) + c) - (x_0(j_1, t+1) + c) &= \\ x_0(j_1, t) + x_0(j_2, t+1) - x_0(j_2, t) - x_0(j_1, t+1) &= \\ V_0(j_1, j_2, t), \end{aligned}$$

即 $V'_0(j_1, j_2, t) = V_0(j_1, j_2, t)$.

同理可得 $V'_i(j_1, j_2, t) = V_i(j_1, j_2, t)$, 则

$$\begin{aligned} |V'_0(j_1, j_2, t) - V'_i(j_1, j_2, t)| &= \\ |V_0(j_1, j_2, t) - V_i(j_1, j_2, t)|, \end{aligned}$$

故 $\varepsilon_{0i}(j_1, j_2, t) = \varepsilon'_{0i}(j_1, j_2, t)$, 所以 $\varepsilon'_{0i} = \varepsilon_{0i}$. \square

性质5 设 $X_0, X'_0, X_i, X'_i, X_k, X'_k$ 为6个同型样本行为矩阵, 满足 $X'_0 = aX_0, X'_i = aX_i, X'_k = aX_k$, 其中 a 为非零常数, 如果 $\varepsilon_{0i}(j_1, j_2, t) > \varepsilon_{0k}(j_1, j_2, t)$, 则 $\varepsilon'_{0i}(j_1, j_2, t) > \varepsilon'_{0k}(j_1, j_2, t)$.

证明 由于 $\varepsilon_{0i}(j_1, j_2, t) > \varepsilon_{0k}(j_1, j_2, t)$, 有

$$\frac{1}{1 + |V_0(j_1, j_2, t) - V_i(j_1, j_2, t)|} > \frac{1}{1 + |V_0(j_1, j_2, t) - V_k(j_1, j_2, t)|},$$

可得

$$\begin{aligned} |V_0(j_1, j_2, t) - V_i(j_1, j_2, t)| &< \\ |V_0(j_1, j_2, t) - V_k(j_1, j_2, t)|. \end{aligned}$$

对于 $\forall a \neq 0$, 有

$$\begin{aligned} |a||V_0(j_1, j_2, t) - V_i(j_1, j_2, t)| &< \\ |a||V_0(j_1, j_2, t) - V_k(j_1, j_2, t)|, \end{aligned}$$

即

$$\begin{aligned} |V'_0(j_1, j_2, t) - V'_i(j_1, j_2, t)| &< \\ |V'_0(j_1, j_2, t) - V'_k(j_1, j_2, t)|, \end{aligned}$$

则

$$\frac{1}{1 + |V'_0(j_1, j_2, t) - V'_i(j_1, j_2, t)|} > \frac{1}{1 + |V'_0(j_1, j_2, t) - V'_k(j_1, j_2, t)|},$$

即 $\varepsilon'_{0i}(j_1, j_2, t) > \varepsilon'_{0k}(j_1, j_2, t)$. \square

由性质 1~性质 5 及第 1 节分析, 可得如下定理.

定理 2 基于面板数据的灰色四面体网格关联度具有如下性质:

- 1) 规范性, $0 < \varepsilon_{0i} \leq 1$;
- 2) 对称性, $\varepsilon_{0i} = \varepsilon_{i0}$;
- 3) $\varepsilon_{00} = \varepsilon_{ii} = 1$;
- 4) 数据平移变换关联度不变性;

5) 数据数乘变换关联序不变性;

6) 关联度与指标排列顺序无关.

3 实例分析

山东省为北方大省, 地理位置优越, 在渤海和黄海两处海域都占据了重要位置. 山东半岛陆地海岸线全长 3 290 km, 海岸线有威海市、青岛市、烟台市、日照市、东营市、滨州市、潍坊市 7 个地级市. 2012 年

表 1 山东沿海 7 城市的空气质量数据表

城市	年度	AQI	PM _{2.5} /($\mu\text{g}/\text{m}^3$)	PM ₁₀ /($\mu\text{g}/\text{m}^3$)	SO ₂ /($\mu\text{g}/\text{m}^3$)	CO/(mg/m^3)	NO ₂ /($\mu\text{g}/\text{m}^3$)	O ₃ /($\mu\text{g}/\text{m}^3$)
潍坊	2014	124.17	78.08	145.58	56.25	1.07	39.08	125.00
	2015	114.92	71.83	134.58	43.75	1.00	35.58	118.33
	2016	104.50	64.50	123.83	37.42	0.95	35.42	109.08
	2017	105.58	58.33	114.08	24.92	0.96	35.42	114.42
	2018	92.92	50.75	100.08	16.25	0.75	32.83	105.50
	2019	100.83	56.83	106.67	12.58	0.76	37.75	106.00
	2020	88.33	49.75	88.75	10.67	0.72	32.58	102.33
东营	2014	124.83	77.42	147.42	71.92	1.30	46.75	123.42
	2015	120.25	72.18	128.73	51.45	1.13	37.36	119.27
	2016	106.75	64.17	120.33	47.75	1.17	39.67	112.58
	2017	103.75	56.42	107.67	38.08	1.14	37.08	114.08
	2018	90.92	45.25	90.83	19.00	0.72	33.08	112.00
	2019	100.17	48.67	92.67	15.42	0.73	35.50	116.25
	2020	90.17	45.67	80.50	15.08	0.68	31.33	110.00
滨州	2014	117.25	84.92	134.50	69.00	2.03	45.75	69.25
	2015	110.67	73.45	116.64	56.18	2.02	41.09	87.27
	2016	105.50	74.00	126.00	36.75	1.73	36.50	79.17
	2017	109.25	66.00	101.17	29.75	1.53	40.83	105.17
	2018	98.75	52.83	86.50	21.25	1.02	36.83	113.17
	2019	103.50	53.33	88.17	19.33	1.14	40.25	115.33
	2020	93.67	48.83	77.42	16.67	0.95	38.42	110.17
青岛	2014	91.42	55.00	107.08	33.00	0.94	40.08	105.17
	2015	86.00	46.27	92.64	25.55	0.86	32.27	101.91
	2016	79.25	44.58	86.83	21.58	0.78	33.33	95.42
	2017	82.00	38.00	77.42	15.25	0.78	37.83	104.50
	2018	68.92	33.08	69.83	8.83	0.59	32.17	89.17
	2019	80.17	39.67	79.00	8.17	0.70	35.58	96.17
	2020	70.08	31.75	61.17	8.00	0.64	33.42	94.58
烟台	2014	82.33	51.42	81.17	27.25	0.82	39.75	96.75
	2015	78.33	41.91	73.64	18.73	0.80	31.82	102.36
	2016	72.58	39.75	76.00	22.25	0.79	34.67	91.83
	2017	76.08	34.25	67.75	18.08	0.85	32.67	105.75
	2018	65.75	28.17	62.75	10.58	0.55	25.25	96.17
	2019	80.83	37.92	77.17	9.17	0.56	28.17	103.75
	2020	70.92	31.75	61.67	10.42	0.56	28.00	102.00
威海	2014	81.58	40.83	73.92	21.33	0.62	25.25	116.50
	2015	72.92	35.36	67.00	15.18	0.55	21.27	105.91
	2016	66.42	33.83	62.92	14.50	0.61	21.92	95.08
	2017	67.58	28.33	54.08	11.00	0.65	17.67	102.00
	2018	62.50	25.33	47.92	7.25	0.48	17.33	100.58
	2019	74.50	31.25	57.42	5.75	0.54	18.58	109.17
	2020	63.75	25.83	45.08	4.58	0.51	15.50	102.33
日照	2014	95.83	62.50	111.67	29.75	0.94	36.67	97.25
	2015	96.17	56.18	99.09	24.18	0.98	35.73	108.36
	2016	94.08	58.33	100.92	21.00	0.93	38.33	97.92
	2017	85.67	46.50	83.25	14.67	0.86	37.58	102.00
	2018	73.75	38.58	73.58	10.25	0.61	32.25	95.50
	2019	86.92	46.50	88.25	8.33	0.58	34.67	102.83
	2020	77.25	37.50	68.42	8.00	0.61	33.00	103.00

中国颁布了空气质量标准(GB3095-2012),第2年,山东省发起了“保卫蓝天”计划.现评估山东省沿海城市的“保卫蓝天”的效果,即各城市的空气质量情况.需要注意的是,由于新的标准的颁布以及地方政府的干预,所得的空气质量数据具有“贫信息”特征,这恰恰是灰色系统理论所解决的问题.

根据环境空气质量标准以及数据的可获得性,选择AQI、PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃作为评价指标,其中指标CO的单位为mg/m³,其他指标单位为μg/m³.山东沿海7城市2014~2020年的数据来源于中国空气质量在线监测分析平台历史数据,如表1所示.

3.1 基于灰色四面体网格关联分析模型的空气质量评价

各指标原始数值越小,空气质量越好,所以选择成本型极差变换公式(4)处理原始数据.标准化后的数据记为{X_i, i = 1, 2, ..., 7},其中X_i(i = 1, 2, ..., 7)分别表示潍坊、东营、滨州、青岛、烟台、威海、日照的标准化数据.

标准化数据越大越好,为此选择各城市在各时刻各指标的最大值作为系统特征行为矩阵,记为

$$X_0 = \begin{bmatrix} 0.69 & 0.74 & 0.72 & 0.75 & 0.91 & 0.69 & 1.00 \\ 0.83 & 0.83 & 0.79 & 0.84 & 0.96 & 0.82 & 0.68 \\ 0.94 & 0.86 & 0.83 & 0.85 & 0.92 & 0.80 & 0.82 \\ 0.92 & 0.95 & 0.91 & 0.91 & 0.89 & 0.93 & 0.41 \\ 1.00 & 1.00 & 0.97 & 0.96 & 1.00 & 0.94 & 0.64 \\ 0.81 & 0.90 & 0.88 & 0.98 & 0.96 & 0.90 & 0.52 \\ 0.98 & 0.99 & 1.00 & 1.00 & 0.98 & 1.00 & 0.55 \end{bmatrix}.$$

由灰色四面体网格关联度公式(7),可得山东沿海7城市与特征行为矩阵X₀的关联度,如表2所示.

表2 山东7城市的关联度结果

城市	关联度	关联序
潍坊	0.889	6
东营	0.903	5
滨州	0.863	7
青岛	0.914	3
烟台	0.922	2
威海	0.967	1
日照	0.906	4

由表2可知,关联度介于0.863与0.967之间,说明山东沿海城市空气质量具有区域一致性的特征.当然也存在一定的差异,由关联度得到关联序为ε₀₆ > ε₀₅ > ε₀₄ > ε₀₇ > ε₀₂ > ε₀₁ > ε₀₃,具体结果为:威海、烟台、青岛排名前三,后面依次为日照、东营、潍坊、滨州.

山东沿海7城市年度空气质量指标折线图如图1

所示,其中图1(a)~图1(g)分别表示2014~2020年的空气质量年度数据,数据越大说明空气中相应指标的污染越严重.由折线图可以看出,折线越靠上,空气污染越严重.

由图1不难看出,威海年度折线图均位于最下方,也就是污染物指标数值最小,说明空气质量最优.另外,2015~2020年全国十佳宜居城市榜单中,威海较其他城市入选次数最多,且于2016~2019连续4年排名第3,而宜居城市榜中一个重要指标为环境质量.

紧挨着威海折线的是烟台、青岛,可见威海、烟台、青岛排名前三符合实际情况.不仅仅如此,烟台也于2014年和2015年入选了全国十佳宜居城市.

而东营、潍坊、滨州的折线图大多位于上方,说明空气污染指标数值较大,空气质量与其他4个城市相比较差.由表1不难看出,潍坊、滨州PM_{2.5}的数值整体相对较大,特别是滨州在2014年达到最大值84.92 μg/m³.沿海城市的海洋气候对PM_{2.5}产生积极的影响^[27],但潍坊和滨州是只有部分地区靠海,受海洋气候的影响较小.虽然东营市为临海城市,但东营市的经济支柱主要是石油,石油的开采易对空气质量产生影响.从这些来看,东营、潍坊、滨州在沿海城市中排名靠后符合实际情况.

综上所述,灰色四面体网格关联分析模型能够很好地反映面板数据的相关程度,并对山东沿海城市的空气质量进行了有效评价.

3.2 与已有模型的对比分析

如引言所述,基于面板数据经典的关联分析模型包括拓展灰色绝对关联模型^[11]、灰色网格关联模型^[14]、灰色凸关联模型^[19],为此选择这3个模型进行对比分析.基于此3个关联分析模型,分别计算X₀与X_i(i = 1, 2, ..., 7)的关联度,具体结果如表3所示.

由表3可看出,文献[19]的模型得到的关联序与本文基本一致.虽然文献[14]能够识别出前3名的城市为威海、烟台、青岛,但东营优于日照,由上述分析可知此结果不符合实际情况.而文献[11]将威海排名第3,可见也不符合实际情况.究其原因,文献[11]的模型基于离散三角曲面所围的体积建立,此体积易受三角曲面构造方法的影响.

再者,3种模型的结果是基于X_i(i = 0, 1, ..., 7)计算,而矩阵X_i(i = 0, 1, ..., 7)所对应的指标顺序是固定的,即对应于AQI、PM_{2.5}、PM₁₀、SO₂、CO、NO₂、O₃.显然这些指标排列顺序与空气质量应该是无关的,也就是说,交换指标的排列顺序不应该影响空气质量评价.但对于上述3种模型,如果交换指标的顺序,以上结果很可能会发生变化.为此,交换X_i(i = 0,

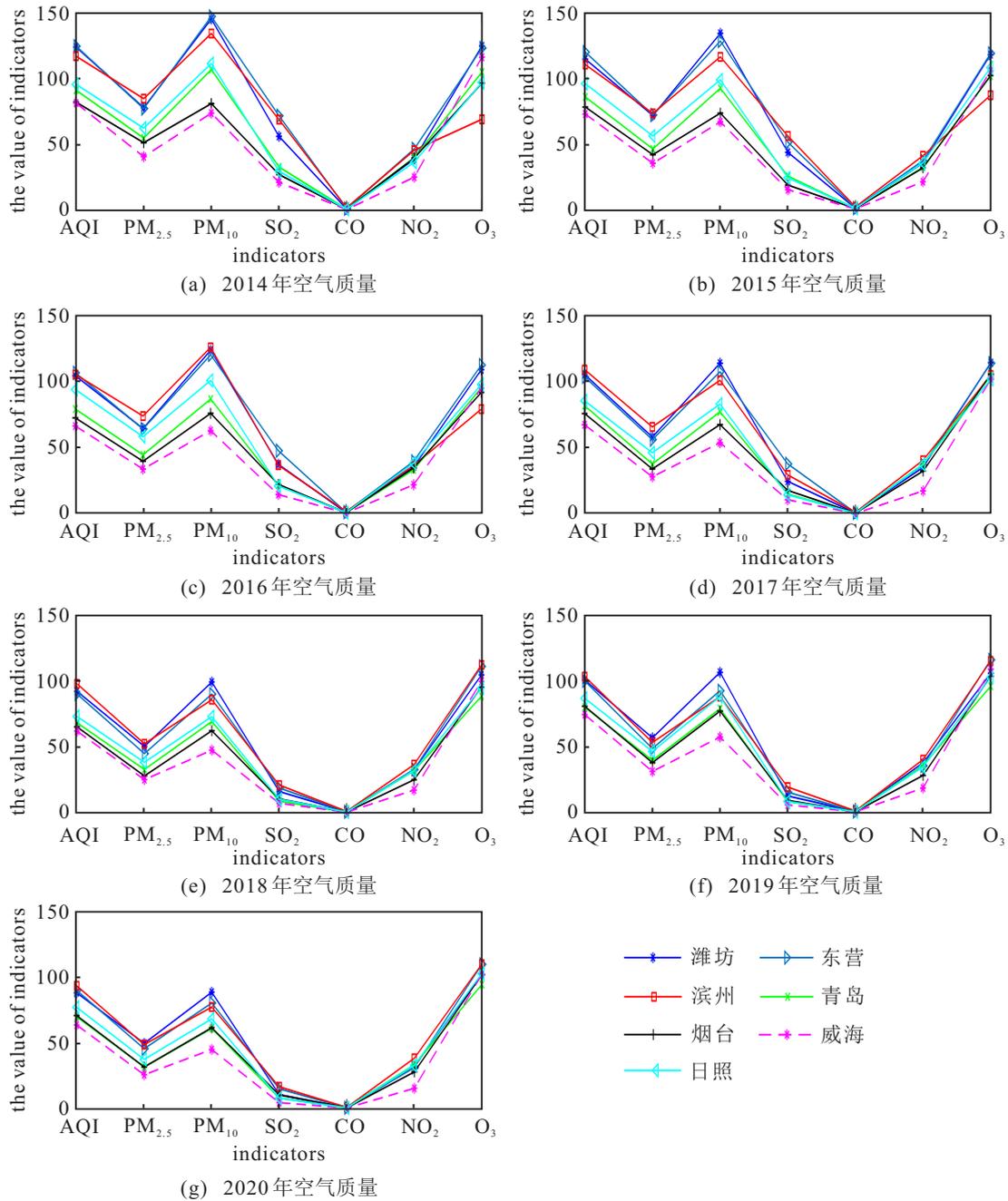


图1 山东沿海7城市的空气质量指标年度折线图

表3 3种模型的关联度结果

模型	潍坊		东营		滨州		青岛		烟台		威海		日照	
	关联度	排序												
文献[11]	0.718	5	0.668	7	0.713	6	0.774	4	0.861	2	0.855	3	0.888	1
文献[14]	0.628	5	0.671	4	0.213	7	0.638	3	0.658	2	0.929	1	0.406	6
文献[19]	0.481	5	0.464	6	0.450	7	0.655	3	0.716	2	0.950	1	0.600	4

表4 交换第1和第3行后的关联度结果

模型	潍坊		东营		滨州		青岛		烟台		威海		日照	
	关联度	排序												
文献[11]	0.819	3	0.660	6	0.813	5	0.885	2	0.814	4	0.967	1	0.647	7
文献[14]	0.611	5	0.695	2	0.239	7	0.649	4	0.657	3	0.929	1	0.408	6
文献[19]	0.478	5	0.460	6	0.447	7	0.651	3	0.715	2	0.948	1	0.598	4
本文	0.889	6	0.903	5	0.863	7	0.914	3	0.922	2	0.967	1	0.906	4

1, ..., 7) 的第1行和第3行, 即交换 AQI 和 PM₁₀ 的数据顺序得到样本新的矩阵形式 $X'_i(i = 0, 1, \dots, 7)$. 基于3种模型以及本文模型再次计算 X'_0 与 $X'_i(i = 1, 2, \dots, 7)$ 的关联度, 具体结果如表4所示.

比较表3和表4可看出, 文献[11]和文献[14]所得到的关联序均发生改变. 虽然文献[19]所确定的关联序没有变化, 但是关联度的大小发生了改变. 可见, 3种模型的结果易受指标排列顺序的影响. 本文的模型不管是关联度的大小还是关联序均未发生变化. 因此, 3种模型适合于固定指标顺序的数据分析, 而对于非固定顺序的数据分析易出现偏差.

总之, 灰色四面体网格关联分析模型的结果与指标顺序无关, 适用于固定或者非固定指标顺序的数据分析, 与上述3种模型相比, 应用范围更广.

4 结论

本文给出了面板数据的新的投射方法, 解决了面板数据的三角网格不同连接方法影响关联度的问题, 并基于数列的组合原理, 构造了面板数据灰色四面体网格关联分析模型. 新的模型不仅不受指标顺序的影响, 而且考虑了成对指标的影响. 最后, 运用该模型分析了山东沿海7城市的空气质量, 验证了模型的合理性和有效性. 如果面板数据从时间维度或者对象维度构建, 则该模型还可以分析山东沿海7城市的时间特征以及识别主要污染物. 尽管灰色四面体网格关联模型克服了其他方法的一些不足, 但该模型的机理还需进一步研究.

参考文献(References)

- [1] Deng J L. Figure on difference information space in Grey relational analysis[J]. Journal of Grey System, 2004, 16(2): 96-100.
- [2] Deng J L. Spread of grey relational space[J]. Journal of Grey Systems, 1995, 7(3): 96-100.
- [3] 张岐山, 郭喜江, 邓聚龙. 灰关联熵分析方法[J]. 系统工程理论与实践, 1996, 16(8): 7-11.
(Zhang Q S, Guo X J, Deng J L. Grey relation entropy method of relation analysis[J]. Systems Engineering — Theory & Practice, 1996, 16(8): 7-11.)
- [4] 王清印. 灰色B型关联分析[J]. 华中理工大学学报, 1989, 17(6): 77-82.
(Wang Q Y. Grey relational analysis of B-mode[J]. Journal of Huazhong University of Science and Technology, 1989, 17(6): 77-82.)
- [5] 王清印, 赵秀恒. C型关联分析[J]. 华中理工大学学报, 1999, 27(3): 75-77.
(Wang Q Y, Zhao X H. The relational analysis of C mode[J]. Journal of Huazhong University of Science and Technology, 1999, 27(3): 75-77.)
- [6] 唐五湘. T型关联度及其计算方法[J]. 数理统计与管理, 1995, 14(1): 34-37.
(Tang W X. The concept and the computation method of t's correlation degree[J]. Application of Statistics and Management, 1995, 14(1): 34-37.)
- [7] 党耀国, 刘思峰, 刘斌, 等. 灰色斜率关联度的改进[J]. 中国工程科学, 2004, 6(3): 41-44.
(Dang Y G, Liu S F, Liu B, et al. Improvement on degree of grey slope incidence[J]. Engineering Science, 2004, 6(3): 41-44.)
- [8] 赵艳林, 韦树英, 梅占馨. 灰色欧几里德关联度[J]. 广西大学学报: 自然科学版, 1998, 23(1): 10-13.
(Zhao Y L, Wei S Y, Mei Z X. Grey Euclid relation grade[J]. Journal of Guangxi University: Natural Science Edition, 1998, 23(1): 10-13.)
- [9] 李雪梅, 党耀国, 王俊杰. 基于灰色准指数律的灰色生成速率关联模型的构建及应用[J]. 控制与决策, 2015, 30(7): 1245-1250.
(Li X M, Dang Y G, Wang J J. Grey generation rate relational analysis model based on grey exponential law and its application[J]. Control and Decision, 2015, 30(7): 1245-1250.)
- [10] 刘思峰, 谢乃明, Forrest Jeffery. 基于相似性和接近性视角的新型灰色关联分析模型[J]. 系统工程理论与实践, 2010, 30(5): 881-887.
(Liu S F, Xie N M, Forrest J. On new models of grey incidence analysis based on visual angle of similarity and nearness[J]. Systems Engineering — Theory & Practice, 2010, 30(5): 881-887.)
- [11] 张可, 刘思峰. 灰色关联聚类在面板数据中的扩展及应用[J]. 系统工程理论与实践, 2010, 30(7): 1253-1259.
(Zhang K, Liu S F. Extended clusters of grey incidences for panel data and its application[J]. Systems Engineering — Theory & Practice, 2010, 30(7): 1253-1259.)
- [12] 王正新, 党耀国, 沈春光. 三维灰色关联模型及其应用[J]. 统计与决策, 2011(15): 174-176.
(Wang Z X, Dang Y G, Shen C G. Three-dimensional grey relational model and its application[J]. Statistics & Decision, 2011(15): 174-176.)
- [13] 钱吴永, 王育红, 党耀国, 等. 基于多指标面板数据的灰色矩阵关联模型及其应用[J]. 系统工程, 2013, 31(10): 70-74.
(Qian W Y, Wang Y H, Dang Y G, et al. Grey matrix relational modeling and its application based on multivariate panel data[J]. Systems Engineering, 2013,

- 31(10): 70-74.)
- [14] 刘震, 党耀国, 钱吴永, 等. 基于面板数据的灰色网格关联度模型[J]. 系统工程理论与实践, 2014, 34(4): 991-996.
(Liu Z, Dang Y G, Qian W Y, et al. Grey grid incidence model based on panel data[J]. Systems Engineering—Theory & Practice, 2014, 34(4): 991-996.)
- [15] 刘震, 党耀国, 周伟杰, 等. 新型灰色接近关联模型及其拓展[J]. 控制与决策, 2014, 29(6): 1071-1075.
(Liu Z, Dang Y G, Zhou W J, et al. New grey nearness incidence model and its extension[J]. Control and Decision, 2014, 29(6): 1071-1075.)
- [16] 崔立志, 刘思峰. 面板数据的灰色矩阵相似关联模型及其应用[J]. 中国管理科学, 2015, 23(11): 171-176.
(Cui L Z, Liu S F. Grey matrix similar incidence model for panel data and its application[J]. Chinese Journal of Management Science, 2015, 23(11): 171-176.)
- [17] 吴鸿华, 穆勇, 屈忠锋, 等. 基于面板数据的接近性和相似性关联度模型[J]. 控制与决策, 2016, 31(3): 555-558.
(Wu H H, Mu Y, Qu Z F, et al. Similarity and nearness relational degree based on panel data[J]. Control and Decision, 2016, 31(3): 555-558.)
- [18] 罗党, 张曼曼. 基于面板数据的灰色B型关联模型及其应用[J]. 控制与决策, 2020, 35(6): 1476-1482.
(Luo D, Zhang M M. Grey incidence model of B-mode based on panel data and its application[J]. Control and Decision, 2020, 35(6): 1476-1482.)
- [19] 吴利丰, 刘思峰. 基于灰色凸关联度的面板数据聚类方法及应用[J]. 控制与决策, 2013, 28(7): 1033-1036.
(Wu L F, Liu S F. Panel data clustering method based on grey convex relation and its application[J]. Control and Decision, 2013, 28(7): 1033-1036.)
- [20] 吴鸿华, 屈忠锋. 基于面板数据的灰色曲率关联模型[J]. 控制与决策, 2020, 35(5): 1072-1076.
(Wu H H, Qu Z F. The grey curvature incidence model based on panel data[J]. Control and Decision, 2020, 35(5): 1072-1076.)
- [21] Wu H H, Qu Z F. Gray clustering model based on the degree of dynamic weighted incidence for panel data and its application[J]. Grey Systems: Theory and Application, 2020, 10(4): 413-423.
- [22] Sun J, Dang Y G, Zhu X Y, et al. A grey spatiotemporal incidence model with application to factors causing air pollution[J]. Science of The Total Environment, 2020, 759(12): 143576.
- [23] 党耀国, 尚中举, 王俊杰, 等. 基于面板数据的灰色指标关联模型构建及其应用[J]. 控制与决策, 2019, 34(5): 1077-1084.
(Dang Y G, Shang Z J, Wang J J, et al. Grey incidence model for relationship between indicators of panel data and its application[J]. Control and Decision, 2019, 34(5): 1077-1084.)
- [24] Xu G Y, Zhang Q Q, Yao Y, et al. Changes in PM_{2.5} sensitivity to NO_x and NH₃ emissions due to a large decrease in SO₂ emissions from 2013 to 2018[J]. Atmospheric and Oceanic Science Letters, 2020, 13(3): 210-215.
- [25] 王鸥, 何秉宇. 基于灰色关联度的乌鲁木齐市PM_{2.5}与环境空气质量六要素关系的探讨[J]. 干旱区资源与环境, 2018, 32(6): 176-181.
(Wang O, He B Y. Relationships between PM_{2.5} and six elements for air quality in Urumqi based on Grey correlation[J]. Journal of Arid Land Resources and Environment, 2018, 32(6): 176-181.)
- [26] 王涛生. 制度创新影响国际贸易竞争优势的机理、模型与实证研究[D]. 长沙: 湖南大学, 2013.
(Wang T S. Study on influence mechanism of competitive advantage in the international trade due to system innovation and its empirical analysis[D]. Changsha: Hunan University, 2013.)
- [27] 陈媛, 岑况, Norra S, 等. 北京市区大气气溶胶PM_{2.5}污染特征及颗粒物溯源与追踪分析[J]. 现代地质, 2010, 24(2): 345-354.
(Chen Y, Cen K, Norra S, et al. Study on pollution characteristics of PM_{2.5} in the aerosol, tracing and tracking atmospheric particulates in Beijing city[J]. Geoscience, 2010, 24(2): 345-354.)

作者简介

吴鸿华(1978—), 男, 副教授, 博士生, 从事灰色系统理论等研究, E-mail: ss_wuhh@ujn.edu.cn;

刘思峰(1955—), 男, 教授, 博士生导师, 从事灰色系统理论、系统方法分析与模型等研究, E-mail: sfliu@nuaa.edu.cn;

方志耕(1962—), 男, 教授, 博士生导师, 从事可靠性复杂装备研制管理、可靠性工程及现代控制理论等研究, E-mail: zhigengfang@163.com.

(责任编辑: 孙艺红)