

控制与决策

Control and Decision

基于面板数据的灰色曲率关联模型

吴鸿华, 屈忠锋

引用本文:

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.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0441>

您可能感兴趣的其他文章

Articles you may be interested in

基于数据融合的目标检测方法综述

An overview of target detection methods based on data fusion

控制与决策. 2020, 35(1): 1–15 <https://doi.org/10.13195/j.kzyjc.2019.0649>

分层构权灰色主成分评价模型及其应用

Multi-layer weighted grey principal component evaluation model and its application

控制与决策. 2019, 34(6): 1300–1306 <https://doi.org/10.13195/j.kzyjc.2017.1405>

基于实例的离散制造系统能耗知识建模与预测

Case-based energy-consuming knowledge modeling and prediction of discrete manufacturing system

控制与决策. 2019, 34(1): 9–17 <https://doi.org/10.13195/j.kzyjc.2017.1478>

一种无模型自适应积分终端滑模控制方法

A model-free adaptive integral terminal sliding mode control method

控制与决策. 2018, 33(9): 1591–1597 <https://doi.org/10.13195/j.kzyjc.2017.0556>

基于决策粗糙集的多属性灰色关联聚类方法

Decision-theoretic rough set based multi-attribute grey incidence clustering method

控制与决策. 2017, 32(11): 2034–2038 <https://doi.org/10.13195/j.kzyjc.2016.1186>

基于维度最大熵数据流聚类的异常检测方法

Data stream clustering algorithm based on the maximum entropy of data dimension and its applications for anomaly detection

控制与决策. 2016(2): 343–348 <https://doi.org/10.13195/j.kzyjc.2014.1783>

基于泡沫图像特征和物元可拓模型的锑浮选工况识别

Performance recognition of antimony flotation based on froth image features and matter-element extension model

控制与决策. 2015, 30(8): 1485–1490 <https://doi.org/10.13195/j.kzyjc.2014.0971>

基于项权值变化和SCCI框架的加权正负关联规则挖掘

Weighted positive and negative association rules mining based on dynamic item weight and SCCI framework

控制与决策. 2015(10): 1729–1741 <https://doi.org/10.13195/j.kzyjc.2014.1102>

基于面板数据的灰色曲率关联模型

吴鸿华[†], 屈忠锋

(济南大学 数学科学学院, 济南 250022)

摘要: 对于面板数据, 首先将面板数据投射为时间维度离散曲线和指标维度离散曲线; 然后, 基于离散曲率思想分别从时间维度和指标维度构建时间维度和指标维度关联系数公式, 并以此构建面板数据灰色曲率关联模型, 新的模型具有规范性、对称性、相似性、平移不变性等性质; 最后通过与已有模型的对比和实例分析新的模型的合理性, 结果表明, 新的关联度能较好地反应面板数据的相关程度。

关键词: 关联度; 灰色系统; 曲率; 面板数据; 离散曲线

中图分类号: TP941.5 **文献标志码:** A

The grey curvature incidence model based on panel data

WU Hong-hua[†], QU Zhong-feng

(School of Mathematical Sciences, University of Jinan, Jinan 250022, China)

Abstract: The panel data is projected to the time dimension discrete curves and the index dimension discrete curves. Then, based on the thinking of discrete curvature, the incidence coefficient formulas of the time dimension and index dimension are respectively constructed. The grey curvature incidence model is constructed. The properties of the model, e.g., normalization, symmetry, similarity and translation invariance, are also satisfied. Finally, the comparison to some existing models and an example are given to illustrate the rationality of the proposed model, and the results show that the proposed model can better reflect incidence degree of panel data.

Keywords: degree of incidence; grey system; curvature; panel data; discrete curve

0 引言

灰色系统主要研究“少数据”“贫信息”问题, 现已广泛应用于经济、农业、工业、生物等领域. 灰色关联分析是灰色系统理论的重要组成部分, 是灰色系统建模、灰色决策以及灰色控制模型的基础. 自 Deng^[1] 提出灰色关联分析理论以来, 众多学者参与研究并相继构造了大量灰色关联模型, 如: 灰熵关联度^[2]、欧几里德关联度^[3]、T 型关联度^[4]、B 型关联度^[5]、斜率关联度^[6]、绝对关联度、相对关联度和广义关联度^[7]等.

上述关联度公式研究对象为一维时间序列. 目前, 关于面板数据灰色关联分析的研究仍处于起步阶段. 文献[8]基于三维空间距离构造出三维对象的灰色关联分析模型, 弥补了灰色关联理论在面板数据中的空白; 文献[9]将灰关联分析与谱聚类结合, 利用序列的均衡接近度描述数据点的相似程度; 文献[10]将面板数据表征为时空特征的“水平”距离、“增量”距离和“变异”距离, 以此构建了灰色关联模型; 文献

[11]将面板数据投射为空间网格, 利用线段在空间中的斜率构建了面板数据的关联模型; 文献[12]基于空间向量的夹角和投影分别建立了面板数据的相似性和接近性关联度模型; 文献[13]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[14]在二维灰色凸关联度的基础上提出了三维面板数据的凸关联度模型. 针对面板数据, 上述研究从不同角度构建了关联模型, 在一定程度上丰富了面板数据灰色关联分析理论. 文献[14]利用黑塞矩阵定义了凸度, 即曲线的弯曲方向, 但是未能很好刻画曲线的弯曲程度.

鉴于此, 本文引入曲线曲率思想. 曲率刻画了曲线的弯曲程度, 能够很好地反映曲线的几何特征. 另外, 该度量还有其他优良性质, 如关联系数与曲线的空间位置无关等. 为此, 将面板数据投射为时间维度离散曲线和指标维度离散曲线, 分别计算他们各点离散曲率, 以此构建关联系数公式, 最终得到灰色曲率关联度公式. 通过对比分析, 新的关联度公式能够很

收稿日期: 2019-04-10; 修回日期: 2019-09-20.

基金项目: 山东省高等学校科技计划项目(J18KA249); 济南大学科技发展基金项目(XKY1614).

责任编辑: 王光臣.

[†]通讯作者. E-mail: ss_wuhh@ujn.edu.cn.

好地反映面板数据的相关程度. 最后, 将新的关联度公式应用于济南市7区的空气质量评价分析, 结果表明, 新的关联度能较好地反映面板数据的关联程度.

1 相关知识

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. 文献[15-16]采用三维表描述面板数据. 为了描述面板数据的几何特征, 文献[17]将面板数据对应为 $m \times n$ 矩阵.

定义1^[17] 样本 i 关于指标 s 在时间 t 的数值为 $x_i(s, t)$, 称

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

为样本 i 的行为矩阵, 面板数据序列 $X = (X_1, X_2, \dots, X_N)$ 称为样本序列.

样本 i 的行为矩阵可简记为 $X_i = (x_i(s, t))_{m \times n}$. 为了进一步刻画面板数据的几何特征, 将面板数据分别投射为时间维度离散曲线和指标维度离散曲线.

定义2 设样本 i 的行为矩阵 $X_i = (x_i(s, t))_{m \times n}$, 记

$$X_{is} = ((1, x_i(s, 1)), (2, x_i(s, 2)), \dots, (n, x_i(s, n))), \\ s = 1, 2, \dots, m,$$

X_{is} 称为面板数据 X_i 关于指标 s 的时间维度离散曲线.

定义3 设样本 i 的行为矩阵 $X_i = (x_i(s, t))_{m \times n}$, 记

$$\hat{X}_{it} = ((1, x_i(1, t)), (2, x_i(2, t)), \dots, (m, x_i(m, t))), \\ t = 1, 2, \dots, n.$$

\hat{X}_{it} 称为样本 X_i 关于时刻 t 的指标维度离散曲线.

定义4 设样本 i 的行为矩阵 $X_i = (x_i(s, t))_{m \times n}$, X_{is} 、 \hat{X}_{it} 分别为样本 i 的时间维度和指标维度离散曲线, 记

$$(x_i(s, t))'_t = x_i(s, t) - x_i(s, t - 1), \\ (x_i(s, t))'_s = x_i(s, t) - x_i(s - 1, t), \\ (x_i(s, t))''_{tt} = x_i(s, t + 1) + x_i(s, t - 1) - 2x_i(s, t), \\ (x_i(s, t))''_{ss} = x_i(s + 1, t) + x_i(s - 1, t) - 2x_i(s, t).$$

$(x_i(s, t))'_t$ 、 $(x_i(s, t))'_s$ 分别称为样本 X_i 在点 (s, t) 处关于时间和指标的一阶差分. $(x_i(s, t))''_{tt}$ 、 $(x_i(s, t))''_{ss}$ 分别称为样本 X_i 在点 (s, t) 处关于时间和指标的二阶差分.

定义5 设样本 i 的行为矩阵 $X_i = (x_i(s, t))_{m \times n}$,

X_{is} 为样本 i 的时间维度离散曲线, 记

$$K_{is}(t) = \begin{cases} \frac{(x_i(s, t))''_{tt}}{(1 + (x_i(s, t))'_t)^2)^{3/2}}, & x_i(s, t) > \bar{x}_i(s, t); \\ -\frac{(x_i(s, t))''_{tt}}{(1 + (x_i(s, t))'_t)^2)^{3/2}}, & x_i(s, t) < \bar{x}_i(s, t); \\ 0, & x_i(s, t) = \bar{x}_i(s, t). \end{cases}$$

其中: $\bar{x}_i(s, t) = \frac{1}{2}(x_i(s, t + 1) + x_i(s, t - 1))$, $K_{is}(t)$ 称为样本 X_i 在点 (s, t) 关于时间维度的离散曲率.

$|K_{is}(t)|$ 值越大, 表明曲线在时间维度的弯曲程度越大, 反之越小. 若 $K_{is}(t) > 0$, 则曲线的弯曲方向向上; 若 $K_{is}(t) < 0$, 则曲线的弯曲方向向下; 若 $K_{is}(t) = 0$, 则曲线退化为直线.

定义6 设样本 i 的行为矩阵 $X_i = (x_i(s, t))_{m \times n}$, \hat{X}_{it} 为样本 i 的指标维度离散曲线, 记

$$\hat{K}_{it}(s) = \begin{cases} \frac{(x_i(s, t))''_{ss}}{(1 + (x_i(s, t))'_s)^2)^{3/2}}, & x_i(s, t) > \tilde{x}_i(s, t); \\ -\frac{(x_i(s, t))''_{ss}}{(1 + (x_i(s, t))'_s)^2)^{3/2}}, & x_i(s, t) < \tilde{x}_i(s, t); \\ 0, & x_i(s, t) = \tilde{x}_i(s, t). \end{cases}$$

其中: $\tilde{x}_i(s, t) = \frac{1}{2}(x_i(s + 1, t) + x_i(s - 1, t))$, $\hat{K}_{it}(s)$ 称为样本 X_i 在点 (s, t) 关于指标维度的离散曲率.

同理, $|\hat{K}_{it}(s)|$ 值越大, 表明曲线在指标维度的弯曲程度越大, 反之越小. 若 $\hat{K}_{it}(s) > 0$, 则曲线的弯曲方向向上; 若 $\hat{K}_{it}(s) < 0$, 则曲线的弯曲方向向下; 若 $\hat{K}_{it}(s) = 0$, 则曲线退化为直线.

定义7 设面板数据样本序列 $X = (X_1, X_2, \dots, X_N)$, 其中 $X_i = (x_i(s, t))_{m \times n}$. 记

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

1) 如果数据为成本型, 则有

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

2) 如果数据为效益型, 则有

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

3) 如果数据为居中型, 则有

$$x_i(s, t) d = \begin{cases} \frac{2(x_i(s, t) - m_s)}{M_s - m_s}, & m_s \leq x_i(s, t) \leq \frac{M_s + m_s}{2}; \\ \frac{2(M_s - x_i(s, t))}{M_s - m_s}, & \frac{M_s + m_s}{2} < x_i(s, t) \leq M_s. \end{cases}$$

其中 D 称为极差变换算子.

原始数据经过极差变换算子的处理,消除了量纲量级对模型的影响,转换为标准值,且 $x_i(s, t)d \in [0, 1]$.

2 灰色曲率关联度

关联分析通过计算关联度反映两个序列之间的相关程度,为了能够正确度量这种相关性,必须最大程度地反应序列之间的实际关系. 对于面板数据,从相似性角度看,空间曲线曲率的差异反应了他们在不同维度上相关程度的大小,即曲率值越接近,相关性越好,反之越差. 所以可以利用曲率刻画面板数据的相关程度大小.

定义8(时间维度关联系数) 设 $X = (X_1, X_2, \dots, X_N)$ 为样本序列, $K_{is}(t)$ 、 $K_{js}(t)$ 分别为样本 X_i 和 X_j 在点 (s, t) 关于时间维度的离散曲率. 记

$$\rho_{ij}(s, t) = \frac{1}{1 + |K_{is}(t) - K_{js}(t)|},$$

$\rho_{ij}(s, t)$ 称为样本 X_i 和 X_j 在点 (s, t) 关于时间维度关联系数.

定义9(指标维度关联系数) 设 $X = (X_1, X_2, \dots, X_N)$ 为样本序列, $\hat{K}_{it}(s)$ 、 $\hat{K}_{jt}(s)$ 分别为样本 X_i 和 X_j 在点 (s, t) 关于指标维度的离散曲率. 记

$$\sigma_{ij}(s, t) = \frac{1}{1 + |\hat{K}_{it}(s) - \hat{K}_{jt}(s)|},$$

$\sigma_{ij}(s, t)$ 称为样本 X_i 和 X_j 在点 (s, t) 关于指标维度关联系数.

定义10(时间维度关联度) 设 $X = (X_1, X_2, \dots, X_N)$ 为样本序列, $\rho_{ij}(s, t)$ 为样本 X_i 和 X_j 在点 (s, t) 关于时间维度关联系数. 记

$$\rho_{ij} = \frac{1}{m} \frac{1}{n-2} \sum_{s=1}^m \sum_{t=2}^{n-1} \rho_{ij}(s, t),$$

ρ_{ij} 称为样本 X_i 和 X_j 的时间维度关联度.

定义11(指标维度关联度) 设 $X = (X_1, X_2, \dots, X_N)$ 为样本序列, $\sigma_{ij}(s, t)$ 为样本 X_i 和 X_j 在点 (s, t) 关于指标维度关联系数. 记

$$\sigma_{ij} = \frac{1}{m-2} \frac{1}{n} \sum_{s=2}^{m-1} \sum_{t=1}^n \sigma_{ij}(s, t),$$

σ_{ij} 称为样本 X_i 和 X_j 的指标维度关联度.

上述关联度从不同维度刻画了面板数据的相关程度. 为了整体刻画面板数据的相关性的程度,将时间维度关联度和指标维度关联度加权得到灰色曲率关联度.

定义12(灰色曲率关联度) 设 $X = (X_1, X_2, \dots, X_N)$ 为样本序列, ρ_{ij} 和 σ_{ij} 分别为样本 X_i 和 X_j 时间维度关联度和指标维度关联度. 记

$$\varepsilon_{ij} = \alpha\rho_{ij} + \beta\sigma_{ij},$$

其中 $\alpha > 0, \beta > 0, \alpha + \beta = 1, \varepsilon_{ij}(s, t)$ 称为样本 X_i 和 X_j 的灰色曲率关联度.

一般情况下, α, β 取值 $1/2$, 即 $\alpha = \beta = 1/2$.

定理1 基于面板数据的灰色曲率关联度满足灰色关联公理.

证明 1) 规范性. 显然 $0 < \varepsilon_{ij}$, 对于时间维度关联度, 由于 $|K_{is}(t) - K_{js}(t)| \geq 0$, 有

$$\frac{1}{1 + |K_{is}(t) - K_{js}(t)|} \leq 1,$$

即 $\rho_{ij}(s, t) \leq 1$, 则有 $\rho_{ij} \leq 1$. 同理, 对于指标维度关联度, $\sigma_{ij} \leq 1$, 所以 $0 < \alpha\rho_{ij} + \beta\sigma_{ij} \leq 1 (\alpha > 0, \beta > 0, \alpha + \beta = 1)$, 满足规范性公理.

2) 接近性. 由两面板数据在点 (s, t) 处关于时间维度和指标维度关联系数公式可以得到, 当他们关于时间维度和指标维度的离散曲率越接近, 相应的关联系数越大, 从而灰色曲率关联度越大, 满足接近性公理. \square

定理2 基于面板数据的灰色曲率关联模型具有如下性质: 1) $0 < \varepsilon_{ij} \leq 1$; 2) 对称性 $\varepsilon_{ij} = \varepsilon_{ji}$; 3) $\varepsilon_{ii} = 1$; 4) ε_{ij} 与空间位置无关, 仅与离散曲线的离散曲率值有关; 5) 相似性, 两面板数据的形状越相似, ε_{ij} 越接近于 1 ; 6) 平移不变性, 若 $x'_i(s, t) = x_i(s, t) + c, x'_j(s, t) = x_j(s, t) + c$, 其中 c 为常数, 则 $\varepsilon'_{ij} = \varepsilon_{ij}$.

证明 由定理1和关联系数公式易得性质1)~性质5)成立, 下面证明性质6). 对于时间维度, 设 ρ'_{ij} 为 X'_i 与 X'_j 的时间维度关联度, ρ_{ij} 为 X_i 与 X_j 的时间维度关联度, 由于 $x'_i(s, t) = x_i(s, t) + c, x'_j(s, t) = x_j(s, t) + c$, 又 $(x'_i(s, t))'_t = (x_i(s, t) + c) - (x_i(s, t-1) + c) = x_i(s, t) - x_i(s, t-1) = (x_i(s, t))'_t, (x'_i(s, t))'_{tt} = (x_i(s, t+1) + c) + (x_i(s, t-1) + c) - 2(x_i(s, t) + c) = x_i(s, t+1) + x_i(s, t-1) - 2x_i(s, t) = (x_i(s, t))'_{tt}$, 所以 $K'_{is}(t) = K_{is}(t)$, 即 $\rho'_{ij}(s, t) = \rho_{ij}(s, t), \rho'_{ij} = \rho_{ij}$. 同理可得 $\sigma'_{ij} = \sigma_{ij}$, 所以 $\varepsilon'_{ij} = \varepsilon_{ij}$. \square

3 比较分析

已知4组面板数据 X_0, X_1, X_2, X_3 , 行为指标维度, 列为时间维度, 具体如下:

$$X_0 = \begin{bmatrix} 1.2 & 1.5 & 1.8 & 2.1 \\ 1.4 & 2 & 2.2 & 2.3 \\ 1.8 & 2.1 & 2.5 & 2.9 \\ 2 & 2.3 & 2.7 & 3.5 \end{bmatrix}, X_1 = \begin{bmatrix} 1.5 & 1.1 & 1 & 0.7 \\ 1.3 & 1 & 0.6 & 0.5 \\ 0.9 & 0.8 & 0.4 & 0.3 \\ 0.5 & 0.4 & 0.1 & 0.1 \end{bmatrix},$$

$$X_2 = \begin{bmatrix} 1.4 & 1.9 & 2.1 & 2.4 \\ 1.6 & 2.3 & 2.5 & 2.8 \\ 1.9 & 2.5 & 2.8 & 3.2 \\ 2.3 & 2.9 & 3.1 & 3.5 \end{bmatrix}, X_3 = \begin{bmatrix} 1 & 1.8 & 1.6 & 2 \\ 1.2 & 2.3 & 3 & 3.2 \\ 1.9 & 2.4 & 2.2 & 3.3 \\ 1.7 & 2.2 & 2.8 & 4.3 \end{bmatrix}.$$

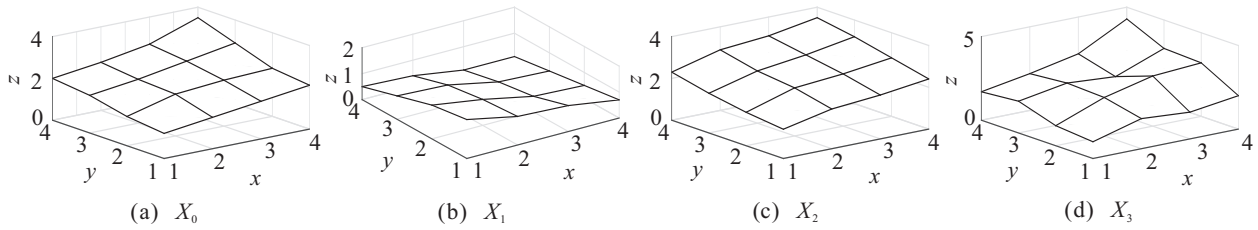


图1 各组面板数据的三维曲线

其中: X_0 为特征样本矩阵, X_1, X_2, X_3 为行为样本矩阵. 利用灰色曲率关联度公式计算得到 $\varepsilon_{01} = 0.666$, $\varepsilon_{02} = 0.764$, $\varepsilon_{03} = 0.727$, 则关联序为 $\varepsilon_{02} > \varepsilon_{03} > \varepsilon_{01}$. 将各组面板数据投射为空间曲线, 得到空间三维曲线如图1所示.

由图1可见, 4组面板数据无论是时间维度还是指标维度, 面板数据 X_0 与 X_1 中的元素变化趋势均相反, 他们之间最不相似. 面板数据 X_0, X_2, X_3 发展趋势大致相似, 但 X_3 数据波动较大, 所以面板数据 X_0 与 X_2 发展趋势比 X_0 与 X_3 的趋势要接近. 利用文献[14]的关联度公式计算得到关联序为 $\varepsilon_{03} > \varepsilon_{02} > \varepsilon_{01}$, 由上述分析可知, 该结果不符合实际情况. 究其原因, 文献[14]定义了凸度, 能够反映空间曲线的弯曲方向, 但未能很好地反映曲线弯曲的程度, 即针对 X_3 数据的波动情况, 文献[14]未能很好地体现. 文献[12-13]结果也是 $\varepsilon_{02} > \varepsilon_{03} > \varepsilon_{01}$, 由此可见, 本文结果能正确表明面板数据间的关系, 说明基于面板数据的曲率关联度模型较为客观、实用.

4 实例分析

为监测济南市大气环境质量状况, 分别在历下区 (X_1)、市中区 (X_2)、高新区 (X_3)、天桥区 (X_4)、槐荫区 (X_5)、历城区 (X_6)、长清区 (X_7) 设置多个环境监测点, 同时设置一个对照子站跑马岭 (X_0). 主要监测可吸入颗粒物 (s_1)、细颗粒物 (s_2)、二氧化硫 (s_3)、二氧化氮 (s_4)、一氧化碳 (s_5) 和臭氧 (s_6), 其中指标 s_6 的单位为 mg/m^3 , 其他指标单位为 $\mu\text{g}/\text{m}^3$. 数据来源于2015~2018年济南市环境质量简报, 具体数据见表1.

由于各指标均为成本型指标, 选择成本型极差变换算子消除量纲量级的影响, 由面板数据灰色曲率关联度公式计算得到 $\varepsilon_{01} = 0.810$, $\varepsilon_{02} = 0.792$, $\varepsilon_{03} = 0.738$, $\varepsilon_{04} = 0.758$, $\varepsilon_{05} = 0.816$, $\varepsilon_{06} = 0.750$, $\varepsilon_{07} = 0.800$.

根据灰色曲率关联度, 得到关联序为 $\varepsilon_{05} > \varepsilon_{01} > \varepsilon_{07} > \varepsilon_{02} > \varepsilon_{04} > \varepsilon_{06} > \varepsilon_{03}$. 具体来看, 近4年, 济南市7区的排序为: 槐荫区、历下区、长清区、市中区、天桥区、历城区、高新区. 事实上, 槐荫区、历下区、长清区之间的区分不大, 他们各有特色. 槐荫区地处

表1 2015~2018年济南市主要污染物浓度情况

	年度	s_1	s_2	s_3	s_4	s_5	s_6
X_0	2015	95.0	60.0	49.0	20.0	1.7	196.0
	2016	88.0	51.0	38.0	17.0	1.5	213.0
	2017	81.0	44.0	22.0	17.0	1.0	172.0
	2018	71.0	35.0	14.0	19.0	1.1	172.0
X_1	2015	147.3	85.3	54.0	53.0	2.6	207.8
	2016	134.2	71.4	40.6	47.2	2.2	196.2
	2017	122.6	61.4	23.8	40.6	2.1	190.2
	2018	104.6	48.2	16.6	42.0	1.7	192.8
X_2	2015	148.5	79.5	43.0	46.5	2.5	181.5
	2016	133.0	67.5	31.0	42.0	2.1	175.0
	2017	121.5	58.0	19.0	37.5	1.8	171.5
	2018	107.0	48.5	15.0	33.5	1.7	202.0
X_3	2015	145.5	77.0	42.5	51.0	2.6	182.0
	2016	149.0	66.5	33.0	49.0	2.1	168.0
	2017	125.0	58.5	28.5	44.5	2.1	184.5
	2018	111.5	51.0	14.0	47.5	1.7	195.5
X_4	2015	167.0	97.0	56.5	61.0	2.7	179.5
	2016	159.0	82.5	44.0	61.0	2.4	186.0
	2017	138.0	68.5	26.5	57.5	2.3	176.5
	2018	123.0	58.0	21.0	55.5	2.0	192.5
X_5	2015	169.5	86.0	39.5	45.5	2.6	199.5
	2016	152.0	77.0	39.0	45.5	2.6	197.0
	2017	145.0	67.3	24.0	48.0	2.3	201.0
	2018	123.7	56.3	18.7	46.3	1.8	208.7
X_6	2015	161.5	87.0	55.3	47.0	2.9	199.0
	2016	148.5	76.0	41.5	46.3	2.2	200.8
	2017	117.7	59.7	28.0	43.3	2.2	202.0
	2018	110.3	49.7	17.7	48.0	1.9	201.3
X_7	2015	142.5	80.5	48.0	55.0	2.2	202.0
	2016	127.0	70.5	41.5	48.5	2.1	204.0
	2017	110.0	59.5	26.0	44.5	1.9	208.0
	2018	103.0	52.5	19.0	37.0	1.7	222.5

济南中西部, 拥有济南湿地公园, 紧靠国家湿地公园, 对空气的改善起到积极作用; 历下区位于市中心, 为重要的文化艺术中心, 污染相对较少, 空气质量较好; 长清区大部分位于郊区, 有“八山一洼一平原”之称, 植被较多, 对空气质量有较好的改善. 所以这3个区排名靠前符合实际情况. 天桥区、历城区、高新区排序相对靠后, 天桥区作为老城区, 拥有火车站、服装城、化工厂等, 不管是生活排放还是工业生产能源消费相对较多; 历城区拥有钢铁冶炼企业、热电厂等, 均对空气质量产生了较大影响; 高新区作为一个新建城区, 近几年基础设施建设较多, 引进工厂、建筑排放以及工业生产能源消费都较高. 利用文献[14]得到市中区和高新区排名靠前, 由上述分析可知, 不

符合实际情况. 从这些来看, 新的关联度得到的关联序能够较好地反映实际情况.

5 结论

本文基于离散曲线的离散曲率思想, 构建了基于面板数据的灰色曲率关联公式, 将面板数据分别投射为时间维度离散曲线和指标维度离散曲线. 新的关联度公式不仅考虑了离散曲线的弯曲方向, 而且考虑了曲线的弯曲程度, 能够更好地反映面板数据的相关程度. 新的模型更符合实际情况, 提高了实用性, 并且为研究面板数据型关联度提供了新的思路.

参考文献(References)

- [1] Deng J L. Spread of grey relational space [J]. *Journal of Grey Systems*, 1995, 7(3): 96-100.
- [2] 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.)
- [3] 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.)
- [4] 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): 33-37.
(Tang W X. On grey incidence degree model of type T with its calculation[J]. *Mathematical Statistics and Management*, 1995, 14(1): 33-37.)
- [5] 王清印. 灰色B型关联分析[J]. *华中理工大学学报*, 1989, 16(6): 77-82.
(Wang Q Y. On grey incidence degree model of type B[J]. *Journal of Huazhong University of Science and Technology*, 1989, 16(6): 77-82.)
- [6] 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.)
- [7] 刘勇, 刘思峰, Jeffrey Forrest. 一种新的灰色绝对关联度模型及其应用[J]. *中国管理科学*, 2012, 20(5): 173-177.
(Liu Y, Liu S F, Jeffrey F. A new grey absolute degree of grey incidence model and application[J]. *Chinese Journal of Management Science*, 2012, 20(5): 173-177.)
- [8] 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 incidence for panel data and its application[J]. *System Engineering — Theory & Practice*, 2010, 30(7): 1253-1259.)
- [9] 郭昆, 张岐山. 基于灰关联分析的谱聚类[J]. *系统工程理论与实践*, 2010, 30(7): 1260-1265.
(Guo K, Zhang Q S. Spectral clustering based on grey relational analysis[J]. *Systems Engineering — Theory & Practice*, 2010, 30(7): 1260-1265.)
- [10] 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.)
- [11] 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]. *System Engineering — Theory & Practice*, 2014, 34(4): 991-996.)
- [12] 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. The similarity and nearness relational degree based on panel data[J]. *Control and Decision*, 2016, 31(3): 555-558.)
- [13] 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*, DOI: 10.13195/j.kzyjc.2018.1054.
(Luo D, Zhang M M. The grey incidence model of B-mode based on panel data and its applicaiton[J]. *Control and Decision*, DOI: 10.13195/j.kzyjc.2018.1054.)
- [14] 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.)
- [15] 郑兵云. 多指标面板数据的聚类分析及其应用[J]. *数理统计与管理*, 2008, 27(2): 265-270.
(Zheng B Y. The clustering analysis of multivariable panel data and its application[J]. *Application of Statistics and Management*, 2008, 27(2): 265-270.)
- [16] Cheng H. *Analysis of panel data*[M]. Cambridge: Cambridge University Press, 1988: 1-13.
- [17] 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.)

作者简介

吴鸿华(1978—), 男, 副教授, 从事灰色系统理论、信息安全等研究, E-mail: ss_wuhh@ujn.edu.cn;

屈忠锋(1979—), 男, 讲师, 从事应用数学等研究, E-mail: ss_quzf@ujn.edu.cn.