

# 控制与决策

Control and Decision

## 基于灰色可能度函数的面板数据聚类方法

耿率帅, 党耀国, 丁松, 尚中举

引用本文:

耿率帅, 党耀国, 丁松, 等. 基于灰色可能度函数的面板数据聚类方法[J]. *控制与决策*, 2020, 35(6): 1483–1489.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2018.0985>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### [基于面板数据的灰色曲率关联模型](#)

The grey curvature incidence model based on panel data

*控制与决策*. 2020, 35(5): 1072–1076 <https://doi.org/10.13195/j.kzyjc.2019.0441>

### [基于面板数据的灰色指标关联模型构建及其应用](#)

Grey incidence model for relationship between indicators of panel data and its application

*控制与决策*. 2019, 34(5): 1077–1084 <https://doi.org/10.13195/j.kzyjc.2017.1538>

### [面板数据的灰色矩阵关联模型及在旱灾脆弱性风险中的应用](#)

Grey matrix incidence model for panel data and its application in drought vulnerability risk

*控制与决策*. 2018, 33(11): 2051–2056 <https://doi.org/10.13195/j.kzyjc.2017.0881>

### [基于灰关联度的面板数据聚类方法及在空气污染分析中的应用](#)

Panel data clustering method based on grey correlation degree and its application in air pollution analysis

*控制与决策*. 2017, 32(12): 2227–2232 <https://doi.org/10.13195/j.kzyjc.2016.1405>

### [面板数据下的灰色指标关联聚类模型与应用](#)

Grey relational clustering model for panel data clustering on indicators and its application

*控制与决策*. 2015, 30(8): 1447–1452 <https://doi.org/10.13195/j.kzyjc.2014.0814>

# 基于灰色可能度函数的面板数据聚类方法

耿率帅<sup>1†</sup>, 党耀国<sup>1</sup>, 丁 松<sup>2</sup>, 尚中举<sup>1</sup>

(1. 南京航空航天大学 经济与管理学院, 南京 211106; 2. 浙江财经大学 经济学院, 杭州 310018)

**摘 要:** 针对静态灰色可能度函数聚类方法的局限性, 综合考虑评价指标发展趋势、指标权重和时间权重的影响, 构建一种体现发展趋势的灰色可能度函数聚类模型用于解决面板数据问题. 该方法引入发展因子概念, 用于表示观测值的发展趋势; 对于观测对象在不同时刻的观测值, 运用发展因子将观测值集结为发展作用值, 并利用时间权重对包含发展作用值的面板数据进行降维, 从而获得信息集结值. 运用所提出的权重求解法确定每个指标各时刻指标权重, 再根据离差平方和最小化方法优化求解总时段综合指标权重, 并对降维后的信息集结值进行灰色可能度函数聚类分析. 最后, 通过对中国 10 个城市 5 年宏观经济状况的聚类分析, 验证所提出模型的有效性和可行性, 实现了灰色可能度函数聚类方法的面板数据处理.

**关键词:** 面板数据; 灰色系统; 可能度; 聚类; 发展趋势

中图分类号: N945.1

文献标志码: A

## Clustering method based on the grey possibility degree function for panel data

GENG Shuai-shuai<sup>1†</sup>, DANG Yao-guo<sup>1</sup>, DING Song<sup>2</sup>, SHANG Zhong-ju<sup>1</sup>

(1. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; 2. School of Economics, Zhejiang University of Finance & Economics, Hangzhou 310018, China)

**Abstract:** Aiming at the limitations of the static grey clustering method, a grey possibility clustering method is proposed to solve the problems of panel data, synthetically considering the growing trend, index weight, and time weight. In this method, the development factor is defined to express development tendency for observed values. Subsequently, according to the time weight in different time point, each index value of the evaluation object is aggregated into development action value by growing factor, and development action values are aggregated to obtain information assemble values for reducing the dimension of the panel data. Meanwhile, the index weights in different time points are determined by using the proposed method. Apart from these above index weights, the comprehensive weight for the whole time period is measured by minimizing the sum of squares of deviations. Furthermore, the information assemble values after information aggregation are clustering analyzed by utilizing the grey possibility function. Finally, the experimental results generated by using the economic and social data from China's 10 cities in 5 years verify the practicality and effectiveness of this proposed model, and the proposed model realizes the grey possibility function clustering for panel data.

**Keywords:** panel data; grey system; possibility degree; clustering; development tendency

## 0 引 言

灰色聚类是灰色系统理论中对于不确定性研究的重要组成部分, 其中包括灰色关联聚类和灰色可能度函数聚类<sup>[1]</sup>. 按研究数据类型不同, 灰色聚类可分为截面数据的灰色聚类和面板数据的灰色聚类. 对于截面数据的灰色聚类研究可归纳为静态灰色聚类研究, 而面板数据的灰色聚类研究则为动态灰色聚类研究. 对于静态灰色聚类研究, 有静态灰色关联聚类和静态灰色可能度聚类. 在静态灰色可能度函数聚

类方面, 刘思峰等<sup>[2]</sup>利用定性及定量分析方法确定各评估指标权重, 分别构建了基于端点和中心点的可能度函数灰色聚类模型; 党耀国等<sup>[3]</sup>探讨了在聚类系数无显著性差异下的综合聚类评估问题, 提出了两阶段灰色聚类方法; 董一哲等<sup>[4]</sup>基于离差最大化的思想确定指标权重, 构建了一种新型灰色可能度聚类模型; 钱丽平等<sup>[5]</sup>运用灰色聚类模型针对区间灰数问题, 基于灰色熵权构造了基于熵权和区间灰数信息的聚类评估方法. 关于灰色静态关联聚类方面, 刘勇等<sup>[6]</sup>针

收稿日期: 2018-07-17; 修回日期: 2018-12-12.

基金项目: 国家自然科学基金项目(71771119); 江苏省社会科学基金项目(NP2017301); 南京航空航天大学短访基金项目(180907DF09).

†通讯作者. E-mail: GENGshuaishuai1987@163.com.

对多属性灰色关联聚类的阈值确定问题,利用决策粗糙集方法构建了基于决策粗糙集的多属性灰色关联聚类方法;韩敏等<sup>[7]</sup>针对灰色绝对关联度模型和灰色相似关联度模型存在的问题,提出了一种基于相对变化面积的改进灰色关联度模型;宋捷等<sup>[8]</sup>基于决策者间差异度最小的思想构建了多目标优化模型求解指标权重,给出了新型加权灰色关联度聚类方法。

随着经济社会的发展,上述静态灰色聚类方法已经不能满足现实中面板数据问题的解决。而针对面板聚类研究,关于灰色关联聚类的研究已逐步涌现。钱吴永等<sup>[9]</sup>提出了具有时空属性的灰色矩阵关联模型;罗党等<sup>[10]</sup>针对现有灰色关联模型中对面板数据存在指标的排列顺序影响关联序问题,构建了面板数据下新型灰色关联模型;李雪梅等<sup>[11]</sup>利用指标内所有对象的平均生成速率作为该指标的序列,并结合偏离、差离和分离的三重差异信息构建了面板数据指标关联聚类模型;崔立志等<sup>[12]</sup>提出了存在发展速度和增长速度的灰色矩阵相似关联模型;党耀国等<sup>[13]</sup>提出了基于增量和离差构成的面板数据灰关联度模型;Wang等<sup>[14]</sup>提出了基于序列连线间角度的灰色动态趋势关联模型。以上这些研究都对实现面板数据问题的解决做出了重要贡献。

目前,运用灰色可能度函数聚类方法解决面板数据的研究还较少有相关学者涉猎。因此,本文针对面板数据,引入相应数据中时间序列的发展趋势以及时间维度权重的概念,构建一种新型灰色可能度函数面板数据聚类模型。基于灰色可能度函数的面板数据聚类模型,将面板数据通过降维方式压缩成截面数据再进行灰色可能度函数聚类分析。本文提出的灰色聚类评价方法将传统的灰色聚类方法从原有的二维截面静态情景问题延伸至三维面板数据问题,拓展了灰色聚类方法的研究范围,而且此方法从发展趋势的变化可以体现评价对象之间的隐性关系,从而增加评价主体的可靠性。最后,通过案例表明了该方法的有效性和可行性。

## 1 面板灰色可能度聚类模型构建

本文将运用截面数据间发展因子、时间权重和不同时刻指标值对观测值集结,实现对面板数据进行降维处理,从而得到观测对象在整个观测时段上的信息集结值,然后再进行灰色可能度聚类分析。

### 1.1 面板数据表示

面板数据中常见的三维观测值是集观测对象、观测指标以及时间序列为一体的数据集合,每一个数值实际是其所对应的对象和指标在相应时刻的观测

值,针对这种三维观测值集合可用矩阵进行表示。下面给出面板数据的矩阵 $Y_{ijt}^*$ 以及经过标准化后的矩阵 $Y_{ijt}$ 的表示方法。

**定义1** 设在面板数据中第 $i(i = 1, 2, \dots, n)$ 个观测对象、第 $j(j = 1, 2, \dots, m)$ 个指标在 $t(t = 1, 2, \dots, p)$ 时刻的观测值为 $y_{ijt}^*$ ,则称面板数据矩阵集 $Y_{ijt}^*$ 为

$$Y_{ijt}^* = \begin{bmatrix} y_{11t}^* & y_{12t}^* & \cdots & y_{1mt}^* \\ y_{21t}^* & y_{22t}^* & \cdots & y_{2mt}^* \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1t}^* & y_{n2t}^* & \cdots & y_{nmt}^* \end{bmatrix}.$$

对观测值 $y_{ijt}^*$ 进行标准化处理可得到 $y_{ijt}$ ,于是称 $Y_{ijt}$ 为面板数据标准化矩阵集,记为

$$Y_{ijt} = \begin{bmatrix} y_{11t} & y_{12t} & \cdots & y_{1mt} \\ y_{21t} & y_{22t} & \cdots & y_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1t} & y_{n2t} & \cdots & y_{nmt} \end{bmatrix},$$

其中标准化方程为

$$y_{ijt} = \frac{y_{ijt}^*}{\sqrt{\sum_{j=1}^m y_{ijt}^{*2}}}. \quad (1)$$

### 1.2 发展因子的构造

**定义2**<sup>[15]</sup> 设对象 $i(i = 1, 2, \dots, n)$ 关于指标 $j(j = 1, 2, \dots, m)$ 在 $t-1$ 和 $t$ 时刻的观测值分别为 $y_{ijt-1}$ 和 $y_{ijt}$ ,称

$$v_{ijt} = \begin{cases} 0, & t = 1; \\ \frac{y_{ijt} - y_{ijt-1}}{t - (t-1)}, & t = 2, 3, \dots, T \end{cases} \quad (2)$$

为对象 $i(i = 1, 2, \dots, n)$ 关于指标 $j(j = 1, 2, \dots, m)$ 在时段 $[t-1, t]$ 的变化速度。称

$$\alpha_{ijt} = \begin{cases} 0, & t = 1; \\ \frac{v_{ijt} - v_{ijt-1}}{t - (t-1)}, & t = 2, 3, \dots, T \end{cases} \quad (3)$$

为对象 $i(i = 1, 2, \dots, n)$ 关于指标观测值 $j(j = 1, 2, \dots, m)$ 在时段 $[t-1, t]$ 的加速度。则

$$\rho_{ijt} = \frac{2}{1 + \exp(-\alpha_{ijt})} \quad (4)$$

为对象 $i(i = 1, 2, \dots, n)$ 关于指标观测值 $j(j = 1, 2, \dots, m)$ 在时段 $[t-1, t]$ 的发展因子。

式(4)中 $\rho_{ijt}$ 表示观测对象 $i$ 关于指标 $j$ 在时段 $[t-1, t]$ 的发展程度,且 $\rho_{ijt}$ 是关于 $\alpha_{ijt}$ 的单调递增函数。若 $\alpha_{ijt} > 0$ ,则 $\rho_{ijt} = \frac{2}{1 + \exp(-\alpha_{ijt})} > 1$ ,即对于评价对象在 $[t-1, t]$ 时段观测值变化速度有上升趋势,发展因子对观测值进行相应增进;若 $\alpha_{ijt} < 0$ ,则 $\rho_{ijt} = \frac{2}{1 + \exp(-\alpha_{ijt})} < 1$ ,即对于评价对象

在  $[t - 1, t]$  时段观测值变化速度有下降趋势, 发展因子对观测值进行相应弱化; 若  $\alpha_{ijt} = 0$ , 则  $\rho_{ijt} = \frac{2}{1 + \exp(-\alpha_{ijt})} = 1$ , 即对于评价对象在  $[t - 1, t]$  时段观测值没有任何加速变化的情况, 没有趋势显示增进或弱化作用, 因此  $y_{ijt}^p$  不发生变化。

### 1.3 时间权重的确定

时间权重反映了在面板聚类中, 不同时刻的数据对于总体评价研究的重要程度. 根据灰色系统理论中“重视新信息, 弱化旧信息”的新信息优先原理, 以下将给出相应的赋权方法. 根据文献 [16], 运用熵权法确定不同时间点的权重, 同时引入时间度的概念, 用于调节权重的分布规律.

**定义 3** 设  $\varepsilon_t = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)^T$  为不同时刻  $t(t = 1, 2, \dots, p)$  的时间权重向量, 称

$$I = - \sum_{t=1}^p \varepsilon_t \ln \varepsilon_t \quad (5)$$

为时间权向量的熵; 称

$$\lambda = \sum_{t=1}^p \frac{p-t}{p-1} \times \varepsilon_t \quad (6)$$

为时间度.

式(5)中  $I$  反映了对样本的集结过程中权重包含信息的平均程度, 它表示对时序内时点重要程度的判断. 当  $\lambda = 1$  时,  $\varepsilon_t = (1, 0, \dots, 0)$ , 表示评价者极度重视起始时刻信息; 当  $\lambda = 0$  时,  $\varepsilon_t = (0, 0, \dots, 1)$ , 表示评价者极度重视临近时刻信息; 当  $\lambda = 0.5$  时,  $\varepsilon_t = (1/p, 1/p, \dots, 1/p)$ , 表示评价者对于各时刻信息重视程度相同. 在本文中, 基于“重视新信息, 弱化旧信息”的新信息优先原理,  $\lambda$  取值规定在  $(0, 0.5)$  之内.

根据定义 3,  $\varepsilon_t$  的求解按照  $\lambda$  确定后, 以挖掘数据熵值最大为标准构建如下非线性规划求解各时刻  $t(t = 1, 2, \dots, p)$  的时间权重:

$$Z = \max \left( - \sum_{t=1}^p \varepsilon_t \ln \varepsilon_t \right).$$

$$\varepsilon_t : \begin{cases} \lambda = \sum_{t=1}^p \frac{p-t}{p-1} \times \varepsilon_t; \\ \sum_{t=1}^p \varepsilon_t = 1, \varepsilon_t \in (0, 1); \\ t = 1, 2, \dots, p; \\ \varepsilon_{t+1} > \varepsilon_t. \end{cases} \quad (7)$$

### 1.4 指标权重的确定

介于指标权重在不同时刻的差异, 首先通过本文提出的改进 criteria importance through intercriteria correlation(CRITIC)法, 求解各指标在不同时刻的权

重; 然后, 通过最小化总离差平方和方法求解指标的综合权重.

**定义 4** 设在时刻  $t(t = 1, 2, \dots, p)$ , 面板数据处理矩阵集中任何两个指标  $j(j = 1, 2, \dots, m)$  和  $b(b = 1, 2, \dots, m)$  且  $j \neq b$  关于对象  $i(i = 1, 2, \dots, n)$  的序列为  $Y_j^p(y_1(j), y_2(j), \dots, y_n(j))$  和  $Y_b^p(x_1(b), x_2(b), \dots, x_n(b))$ , 则

$$\eta_{jb} = 1 - \frac{1}{n} \sum_{d=1}^n |y_d^p(j) - y_d^p(b)| \quad (8)$$

称为序列  $Y_j^p$  与  $Y_b^p$  的点距离关联度; 称

$$C_{jt} = (\delta_j^p / \bar{y}_j^p) \sum_{j \neq b}^m (1 - \eta_{jb}) \quad (9)$$

为指标在  $t$  时刻的信息量; 称

$$w_{jt} = \frac{C_{jt}}{\sum_{j=1}^m C_{jt}} \quad (10)$$

为评价指标  $j$  在时刻  $t$  的权重. 其中:  $\delta_j^p$  为指标  $j$  的标准差,  $\bar{y}_j^p$  表示指标中观测值的平均数,  $\delta_j^p / \bar{y}_j^p$  为  $j$  指标的变异系数,  $|y_d^p(j) - y_d^p(b)|$  为  $j$  指标序列与  $b$  指标序列在  $d$  点的距离.

定义 4 中求解方法称为改进的 CRITIC 法.

由定义 4 可以看出: 对于式(8), 它表示  $j$  指标与其他指标的综合点距离差值; 式(9)反映  $j$  指标内部变异波动程度以及  $j$  指标与外部指标信息差异的综合测度. 式(9)中  $\sum_{j \neq b}^m (1 - \eta_{jb})$  反映了第  $j$  个指标与其他指标的差异程度, 评价指标  $j$  内部的变异程度是用变异系数  $\delta_j^p / \bar{y}_j^p$  衡量的, 两者的乘积  $C_{jt}$  反映了指标  $j$  对于整个指标系统的信息量.  $C_{jt}$  值的大小, 反映了评价指标  $j$  对整个体系的影响程度, 评价指标  $j$  的信息量越大, 其对应权重  $w_{jt}$  的重要性就越大.

根据定义 4, 求出指标各时刻  $t(t = 1, 2, \dots, p)$  的指标权重  $w_{jt}$ , 然后构建关于每一指标在时刻  $t(t = 1, 2, \dots, p)$  综合权重  $w_j$  的最小化总离差平方和求解模型. 规划模型求解如下:

$$\min \sigma^2 = \sum_{t=1}^p \sum_{i=1}^n \sum_{j=1}^m [y_{ijt}^p (w_{jt} - w_j)]^2.$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^m w_j = 1; \\ w_j \geq 0, j = 1, 2, \dots, m. \end{cases} \quad (11)$$

### 1.5 信息集结值的确定

**定义 5** 设如定义 1 中多维面板数据处理矩阵集  $Y_{ijt}^*$  中每一元素  $y_{ijt}$  经过式(4)相应的发展因子  $\rho_{ijt}$  作用后, 称

$$y_{ijt}^{\rho} = y_{ijt} \times \rho_{ijt} \quad (12)$$

为观测对象  $i(i = 1, 2, \dots, n)$  的观测指标  $j(j = 1, 2, \dots, m)$  在  $t(t = 1, 2, \dots, p)$  时刻的发展作用值。根据式(12)与定义3时间权重  $\varepsilon_t$  所得计算结果,则

$$\theta_{ij} = \sum_{t=1}^p \varepsilon_t \times y_{ijt}^{\rho} \quad (13)$$

称为发展作用值  $y_{ijt}^{\rho}$  关于时刻  $t(t = 1, 2, \dots, p)$  的信息集结值。

### 1.6 可能度函数确定及聚类划分

面板数据可能度函数的确定是通过将面板数据经过降维处理后获得的信息集结值进行相应的聚类划分,实现对面板数据进行静态化处理。

**定义6**<sup>[1]</sup> 设有  $n$  个聚类对象,  $m$  个聚类指标,  $s$  个不同灰类。令  $\lambda_k(k = 1, 2, \dots, s)$  为指标  $j(j = 1, 2, \dots, m)$  所划分成所有灰类的灰类代表  $k(k = 1, 2, \dots, s)$ ;  $\lambda_j^0$  和  $\lambda_j^{k+1}$  表示延拓灰类0和  $k+1$  的灰类中心;  $f_j^k(\cdot)(j = 1, 2, \dots, m, k = 1, 2, \dots, s)$  是第  $i$  个对象对第  $j$  个指标关于灰类  $k$  的可能度函数。则称

$$f_j^k = f_j^k(\theta_{ij})$$

为第  $i$  个对象的信息集结值对第  $j$  个指标关于灰类  $k$  的可能度函数值。称

$$f_j^k(\theta_{ij}) = \begin{cases} 0, & \theta_{ij} \notin (\lambda_j^{k-1}, \lambda_j^k); \\ \frac{\theta_{ij} - \lambda_j^{k-1}}{\lambda_j^k - \lambda_j^{k-1}}, & \theta_{ij} \in (\lambda_j^{k-1}, \lambda_j^k); \\ \frac{\lambda_j^{k+1} - \theta_{ij}}{\lambda_j^{k+1} - \lambda_j^k}, & \theta_{ij} \in (\lambda_j^k, \lambda_j^{k+1}) \end{cases} \quad (14)$$

为第  $i$  个对象对第  $j$  个指标关于灰类  $k$  的可能度函数值表达式。

**定义7**<sup>[1]</sup> 若  $j(j = 1, 2, \dots, m)$  指标的  $k$  灰类与定义4中的权重  $w_j$  没有关系,则称

$$\delta_i^k = \sum_{j=1}^m f_j^k(\theta_{ij}) w_j \quad (15)$$

为对象  $i(i = 1, 2, \dots, n)$  属于灰类  $k$  的灰色聚类系数。若

$$\max_{1 \leq k \leq s} \delta_i^k = \delta_i^k, \quad (16)$$

则称对象  $i$  属于灰类  $k$ 。根据信息集结值  $\theta_{ij}$ , 将第  $i(i = 1, 2, \dots, n)$  个对象归入第  $k(k = 1, 2, \dots, s)$  类,称为灰色可能度聚类。

综上,通过式(13)确定信息集结值,然后用求得的指标权重通过式(15)和(16)实现灰色面板数据可能度聚类。

### 1.7 面板数据的灰色可能度聚类方法步骤

面板数据灰色聚类方法应用步骤归纳如下。

step 1: 对多指标面板数据的观测值  $y_{ijt}$  通过式(1)进行一致化和无量纲化处理,得到多维面板数据处理值;

step 2: 依次利用式(2)、(3)、(4)计算出发展因子,利用式(5)、(6)、(7)计算出时间权重,再根据式(12)和(13)对多维面板数据处理值进行信息集结;

step 3: 对评价指标进行灰类划分,并确定划分灰类的可能度函数  $f_j^k$ ,然后通过式(14)计算出各对象的信息集结值的可能度函数值  $f_j^k(\theta_{ij})$ ;

step 4: 依据式(8)~(11)求得各时刻观测对象的指标权重  $w_j(t)$  以及各指标最终的指标权重  $w_j$ ;

step 5: 通过式(15)和(16)计算得出观测对象  $i(i = 1, 2, \dots, n)$  属于已设定灰类  $k$  的灰色聚类系数  $\delta_i^k$ ,根据对象在每一类中的综合聚类系数  $\delta_i^k$ ,判定哪一个聚类系数最大,则该评价对象就归属于哪一类。

## 2 实例分析

本文以可操作性及统计数据支持程度为出发点,选取城市宏观经济状况的8个经济指标(包括:人均国内生产总值(元/人)、第三产业增加值比重(%),人均全社会固定资产投资额(元/人)、人均社会消费品零售总额(元/人)、居民消费水平(元/人)、每10万人拥有医疗床位数(个)、每万人拥有在校(高校)学生人数(人)和人均邮电业务量(万元/人))作为灰色可能度函数聚类评价与分析内容。这些指标的选取能够有效、综合地体现出在某一阶段内城市发展水平状况<sup>[17]</sup>。

本文选取的评价对象主要考虑所选城市的行政级别以及时空特征两个方面的表现。最终所选的10个城市为天津、重庆、成都、武汉、杭州、南京、苏州、青岛、无锡和常州。这些城市的选择是因为他们是除北京、上海、广州、深圳以外涵盖我国其他地域较为发达的城市代表,不管从现有发展情况还是未来发展趋势上都是需要重点关注和研究的城市代表,所以对这些城市的经济社会发展进行评估具有重要意义,可以为地方以及区域经济发展政策的制定提供参考。同时,由于每一次我国大的规划政策出台都会对指标数据造成一定影响,而我国重大经济规划是依据五年规划制定实施,本文选取的时段充分考虑了以上政策冲击的影响,故选取我国2011~2015年为评估时段,数据来源于2012~2016年各地《统计年鉴》。

上述方法的具体计算步骤如下。

step 1: 通过式(1)进行一致化和无量纲化处理,得到面板数据处理值;

step 2: 通过式 (12) 和 (13) 计算出时间维的集结信息值;

step 3: 根据式 (14) 确定各指标的灰类以及划分各灰类的可能度函数;

step 4: 根据式 (7) 计算各时刻权重 (见表 1), 并根据式 (11) 计算评价时间段内各时刻和综合指标权重 (见表 2);

step 5: 根据式 (15) 计算评价对象的各灰类系数 (见表 3).

表 1 2011~2015 年各年度时间权重赋值

	年份				
	2011	2012	2013	2014	2015
权重 ( $\varepsilon_k$ )	0.002 0	0.006 3	0.077 2	0.345 9	0.570 4

表 2 2011~2015 年各指标权重及综合权重赋值

指标权重	2011 年	2012 年	2013 年	2014 年	2015 年	2011~2015 年
$w_1$	0.08	0.07	0.07	0.07	0.07	0.07
$w_2$	0.06	0.06	0.06	0.05	0.05	0.06
$w_3$	0.05	0.05	0.05	0.06	0.06	0.05
$w_4$	0.13	0.13	0.13	0.13	0.13	0.13
$w_5$	0.08	0.08	0.08	0.07	0.07	0.08
$w_6$	0.41	0.38	0.39	0.39	0.35	0.38
$w_7$	0.06	0.06	0.07	0.07	0.07	0.07
$w_8$	0.13	0.17	0.15	0.16	0.20	0.16

表 3 中国 10 个城市 2011~2015 年的宏观经济状况综合灰类系数

灰类	天津	重庆	成都	武汉	杭州	青岛	南京	苏州	无锡	常州
高	0.130	0.000	0.065	0.450	0.040	0.000	0.464	0.076	0.273	0.001
中	0.356	0.301	0.460	0.270	0.617	0.634	0.318	0.401	0.253	0.554
低	0.513	0.690	0.362	0.108	0.037	0.275	0.095	0.460	0.381	0.283
所属灰类	低	低	高	高	中	中	高	低	低	中

本文方法所给出的评价结果及分析如下: 属于高灰类的评价对象有南京、武汉、成都; 属于中灰类的评价对象有杭州、青岛、常州; 属于低灰类的评价对象有无锡、苏州、天津、重庆. 从最终结果分类可以看出: 成都、武汉和南京综合表现较好; 杭州、青岛和常州次之, 天津、重庆、苏州和无锡相对较弱. 总体来看, 这些被评价城市在大部分经济类别指标上差异区分度并不大, 而在代表医疗、教育、邮电业务量这些最关乎民生的问题上则有着不同的表现, 因此, 在评价结果上这些指标起到了较为重要的作用.

在高灰类的城市中, 成都、武汉和南京都有着较好的教育资源, 高校众多, 而且这些城市的居民医疗条件都表现良好, 人均邮电业务量也都表现良好. 在中灰类的城市中, 杭州、青岛和常州次之, 这主要是因

为这些城市在社会宏观经济发展中总体上都表现得中规中矩, 动态发展水平稳定. 天津、重庆、苏州、无锡这些城市归属于低灰类城市主要有两个原因: 一是天津、重庆和苏州这些城市人口基数较大, 在人均水平上并没有较好的优势; 二是自身的民生相关的考量指标各年度并不抢眼, 且动态发展水平也不突出, 所以综合表现被归为低灰类.

为了进一步体现本文方法的有效性, 将运用本文方法得到的 2011~2015 年期间中国 10 城市宏观经济发展状态最终评价结果与文献 [17] 的结果 (见表 4) 进行比较, 对比结果如表 5 所示. 由表 4 可知, 运用文献 [17] 的动态综合评价法得出的中国 10 城市的综合值由高到低依次为南京、武汉、杭州、成都、无锡、常州、青岛、天津、苏州和重庆.

表 4 中国 10 个城市 2011~2015 年的宏观经济状况综合值及排名

城市名称	天津	重庆	成都	武汉	杭州	青岛	南京	苏州	无锡	常州
综合值	0.196	0.149	0.243	0.353	0.276	0.218	0.358	0.193	0.236	0.221
排序	8	10	4	2	3	7	1	9	5	6

表5 中国10个城市2011~2015年的宏观经济状况排名对比

城市名称	天津	重庆	成都	武汉	杭州	青岛	南京	苏州	无锡	常州
本文排序	9	10	3	2	4	5	1	8	7	6
综合值排序	8	10	4	2	3	7	1	9	5	6

由表5可以看出,两种方法的计算结果有着一定的相似性,即两种方法在体现某一时期的对比城市宏观发展状态的优劣程度上具有一致性。但是,本文方法与动态综合评价法计算所得排名结果上仍然存在一定差异。具体来说:成都与杭州排列次序上在两种评价方法中前后有变化;青岛在本文方法中位列中灰类且仅次于杭州,而在动态综合评价法结果中排名第7略高于天津、苏州和重庆;无锡在本文方法中排名属于低灰类,但在动态综合评价法结果中位列第5。对于不同方法在评价结果上的一些出入主要是由评价方法的构造机理不同所导致的,这也反映了不同方法的建模思想的侧重不同。而且灰色面板可能度函数聚类法相较于动态综合评价法而言,虽然两者都考虑了指标发展趋势,但是灰色可能度聚类法同时考虑了评测对象间的综合比较,在排名上产生了灰类划分,使得评价结果在聚类效果上更加丰富。

综上可得,对比于其他综合评价模型,本文提出的面板灰色可能度聚类模型可以有效评价面板数据,且效果良好,充分、有效地扩展了灰色可能度聚类模型针对面板数据的适用范围。

### 3 结论

针对原有静态灰色可能度聚类方法没有考虑指标的发展趋势以及不能综合评价一个时间段内评价对象发展情况的问题,本文构建了基于发展趋势的灰色面板可能度聚类模型,解决了传统灰色可能度聚类模型只适用截面数据的问题,实现了灰色可能度聚类模型对于面板数据的研究,进一步完善了灰色可能度聚类方法的适用范围。最后,运用该方法对2011~2015年中国10个城市的宏观经济社会发展情况进行分析,通过实例体现了该模型的应用价值。但是,该模型仍然存在一定的局限性:首先,时间权重的时间度选取具有人为性,这对于各年度的评价内容的量化会产生影响;其次,该方法是基于发展趋势实现面板数据灰色可能度聚类。然而,对于面板数据的研究不仅仅局限于发展趋势,这也是未来值得进一步研究的内容。

### 参考文献(References)

[1] 刘思峰, 党耀国, 方志耕, 等. 灰色系统理论及其应用[M]. 第7版. 北京: 科学出版社, 2014: 113-140.

(Liu S F, Dang Y G, Fang Z G, et al. Grey system theory and its applications[M]. 7th ed. Beijing: Science Press, 2014: 113-140.)

[2] 刘思峰, 杨英杰. 灰色系统研究进展(2004~2014)[J]. 南京航空航天大学学报, 2015, 47(1): 1-18.

(Liu S F, Yang Y J. Advanced in grey system research(2004~2014)[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2015, 47(1): 1-18.)

[3] 党耀国, 刘思峰, 刘斌, 等. 聚类系数无显著性差异下的灰色综合聚类方法研究[J]. 中国管理科学, 2005, 13(4): 69-73.

(Dang Y G, Liu S F, Liu B, et al. Study on the integrated grey clustering method under the clustering coefficient with non-distinguished difference[J]. Chinese Journal of Management Science, 2005, 13(4): 69-73.)

[4] 董一哲, 党耀国. 基于离差最大化的灰色聚类方法[J]. 系统工程理论与实践, 2009, 29(9): 141-146.

(Dong Y Z, Dang Y G. Grey clustering method based on maximizing deviation[J]. Systems Engineering—Theory & Practice, 2009, 29(9): 141-146.)

[5] 钱丽丽, 刘思峰, 谢乃明. 基于熵权和区间灰数信息的灰色聚类模型[J]. 系统工程与电子技术, 2016, 38(2): 352-356.

(Qian L L, Liu S F, Xie N M. Grey clustering method based on entropy-weight and grey numbers[J]. Systems Engineering and Electronics, 2016, 38(2): 352-356.)

[6] 刘勇, 王冬冬, 周婷. 基于决策粗糙集的多属性灰色关联聚类方法[J]. 控制与决策, 2017, 32(11): 2034-2038.

(Liu Y, Wang D D, Zhou T. Decision-theoretic rough set based multi-attribute grey incidence clustering method[J]. Control and Decision, 2017, 32(11): 2034-2038.)

[7] 韩敏, 张瑞全, 许美玲. 一种基于改进灰色关联分析的变量选择算法[J]. 控制与决策, 2017, 32(9): 1647-1652.

(Han M, Zhang R Q, Xu M L. A variable selection algorithm based on improved grey relational analysis[J]. Control and Decision, 2017, 32(9): 1647-1652.)

[8] 宋捷, 党耀国, 花增木. 基于灰色聚类的群决策方法研究[J]. 控制与决策, 2010, 25(10): 1593-1597.

(Song J, Dang Y G, Hua Z M. Study on group decision-making method based on grey clustering model[J]. Control and Decision, 2010, 25(10):

- 1593-1597.)
- [9] 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.)
- [10] 罗党, 叶莉莉, 韦保磊, 等. 面板数据的灰色矩阵关联模型及在旱灾脆弱性风险中的应用[J]. 控制与决策, 2018, 33(11): 2051-2056.  
(Luo D, Ye L L, Wei B L, et al. Grey matrix incidence model for panel data and its application in drought vulnerability risk[J]. Control and Decision, 2018, 33(11): 2051-2056.)
- [11] 李雪梅, 党耀国, 王俊杰. 面板数据下的灰色指标关联聚类模型与应用[J]. 控制与决策, 2015, 30(8): 1447-1452.  
(Li X M, Dang Y G, Wang J J. Grey relational clustering model for panel data clustering on indicators and its application[J]. Control and Decision, 2015, 30(8): 1447-1452.)
- [12] 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.)
- [13] 党耀国, 朱晓月, 丁松, 等. 基于灰关联度的面板数据聚类方法及在空气污染分析中的应用[J]. 控制与决策, 2017, 32(12): 2227-2232.  
(Dang Y G, Zhu X Y, Ding S, et al. Panel data clustering method based on grey correlation degree and its application in air pollution analysis[J]. Control and Decision, 2017, 32(12): 2227-2232.)
- [14] Wang J J, Keith W H, Dang Y G. An improved grey dynamic trend incidence model with application to factors causing smog weather[J]. Expert Systems with Applications, 2017, 87: 240-251.
- [15] 刘微微, 石春生, 赵圣斌. 具有速度特征的动态综合评价模型[J]. 系统工程理论与实践, 2013, 33(3): 705-710.  
(Liu W W, Shi C S, Zhao S B. Dynamic comprehensive evaluation model with the feature of speed[J]. Systems Engineering — Theory & Practice, 2013, 33(3): 705-710.)
- [16] 郭亚军, 姚远, 易平涛. 一种动态综合评价方法及应用[J]. 系统工程理论与实践, 2007, 27(10): 154-158.  
(Guo Y J, Yao Y, Yi P T. A method and application of dynamic comprehensive evaluation[J]. Systems Engineering — Theory & Practice, 2007, 27(10): 154-158.)
- [17] 易平涛, 周义, 郭亚军, 等. 一种体现发展趋势的动态综合评价方法[J]. 运筹与管理, 2016, 25(6): 175-180.  
(Yi P T, Zhou Y, Guo Y J, et al. A dynamic comprehensive evaluation method embodying development tendency[J]. Operations Research and Management Science, 2016, 25(6): 175-180.)

#### 作者简介

耿率帅(1987—), 男, 博士生, 从事灰色系统理论及应用的研究, E-mail: GENGshuaishuai1987@163.com;

党耀国(1964—), 男, 教授, 博士生导师, 从事灰色系统理论、数量经济等研究, E-mail: iamdangyg@163.com;

丁松(1992—), 男, 讲师, 博士, 从事灰色系统理论、数量经济的研究, E-mail: dingsong1129@163.com;

尚中举(1993—), 男, 硕士生, 从事灰色系统理论的研究, E-mail: shangzhongju@163.com.

(责任编辑: 李君玲)