



开放科学
(资源服务)
标识码
(OSID)

图情学多属性评价中无量纲化方法综述

俞立平 沈洁

浙江工商大学统计与数学学院 杭州 310018

摘要: 多属性评价是图情学学术评价中重要的评价方法,而指标无量纲化方法是多属性评价过程中重要的环节,不同的指标无量纲化方法会对指标权重及多属性评价结果产生重大影响,因此,对多属性评价中无量纲化方法的系统梳理具有重要意义。本文从多属性评价中无量纲化方法的分类出发,对常见的线性无量纲化方法与非线性无量纲化方法的特点进行系统阐述,并对反向指标、适度指标的无量纲化处理进行特别说明,厘清了一些误用的方法。在此基础上,本文详细介绍了对多属性评价中无量纲化选择的标准,并提出研究的展望。通过本文的梳理以期对图情学学术评价中无量纲化方法的选择提供指导。

关键词: 多属性评价; 指标标准化; 无量纲化; 综述

中图分类号: C812; G35

Summary of Dimensionless Methods in Multi-attribute Evaluation

YU Liping SHEN Jie

Zhejiang Gongshang University, School of Statistics and Mathematics, Hangzhou 310018, China

Abstract: Multi-attribute evaluation is an important evaluation method in the academic evaluation of library and information science, and the non-dimensional index method is an important link in the multi-attribute evaluation process. Different non-dimensional index methods will have a significant impact on the index weight and the results of the multi-attribute evaluation. Therefore, it is of great significance to systematically comb the dimensionless methods in multi-attribute evaluation. Starting from the classification of dimensionless methods in multi-attribute evaluation, this paper systematically describes the characteristics of common linear dimensionless methods and nonlinear dimensionless methods, specializes in the dimensionless processing of reverse indicators and moderate indicators, clarified some misused methods. On this basis, this paper introduces in detail the criteria for non-dimensional method selection in multi-attribute evaluation, and puts forward a research outlook. This article hopes to provide guidance for the choice of dimensionless methods in the academic evaluation of library and information science.

Keywords: Multi-attribute comprehensive evaluation; index normalization; dimensionless; summary

基金项目: 国家社科基金项目“学术评价与创新绩效评价问题研究”(19FTQB011)。

作者简介: 俞立平(1967-), 博士, 教授, 博导, 研究方向: 技术经济、科技评价领域的研究, E-mail: yvliping@126.com; 沈洁(1997-), 硕士生, 研究方向: 科技评价领域的研究。

引言

多属性评价是指在多属性体系结构的对象系统中,采用一定的方法给每个评价对象赋予一个评价值,再据此进行择优或排序^[1]。它从不同的侧面选取刻画系统某种特征的评价指标,建立指标体系,并通过一定的数学模型(或算法)将多个评价指标值合成一个整体性的评价值^[2]。利用多属性评价方法可以综合各指标所蕴含的信息,一个科学合理的多属性评价指标体系可以有效提高评价和决策水平。因此这一方法在经济、管理等领域得到了广泛应用,同时也成为了社会科学与自然科学中的重要研究内容。

情报学及人文社科研究中广泛应用了多属性评价方法,其主要应用方向包括期刊评价、科技评价、竞争情报评价、舆情信息评价等。期刊评价中对多属性评价的主要研究方向包括对不同评价方法的对比研究^[3,4],期刊评价指标的改进^[5],对指标权重的分析^[6]等。科技评价中则侧重于对科技评价指标体系的构建^[7-9]、多属性评价中权重的确定^[10,11]、中国科技评价制度的研究^[12]等方面的研究。竞争情报评价中邱均平^[13]对面向绩效评价的竞争情报评价体系构建展开研究,夏立新^[14]分析了网络情况下竞争情报综合评价体系的构建,糜万俊^[15]则对不完全信息下竞争情报综合评价方法进行了论述。

多属性评价涉及多个影响因素或指标,这些指标之间往往量纲差距较大,存在着不可公度性, Ma Fengmei 等^[16]为了减少或消除由于各指标的单位及数量级间的差别所带来的影响,需要在多属性评价前先进行指标无量纲化,也称为指标标准化、规范化。研究中常将指标标

准化与指标无量纲化等同起来,但指标无量纲化的范围显然更大。量纲是对物理量量度的概念或范畴,无量纲化就是去除量纲对数据的影响,经过无量纲化处理的数据具有公度性,在关联分析、聚类分析、多属性评价等方面都有应用。而指标标准化则是指通过一定的数学变换把性质、量纲各异的指标转化为可以综合的一个相对数即标准化值,利用标准化值进行多属性评价。此外值得注意的是,这里的指标标准化不同于作为无量纲化中的标准化处理,后者只是一种具体的处理方法。因此,在本文中提及的无量纲化方法特指在多属性评价中所使用的无量纲化方法。

指标无量纲化并非多属性评价中独立的一个环节,其起因于数据本身量纲不同而产生的需求,又作用于最后评价的结果,不同的无量纲化方法对指标权重及多属性评价结果均产生重大影响^[17,18],而且这种影响及其作用机制并不体现在结论中,往往会被忽视。长期以来,对于多属性评价体系的构建、指标选择和权重设置的研究较为广泛、深入,而对于数据无量纲化方面的研究相对薄弱。因此,对现有的无量纲化方法的研究成果与进展进行总结,既有利于丰富多属性评价的理论基础,也可以为在评价过程中选择合适的指标无量纲化方法提供思路。

1 无量纲化方法分类

应用于图情学多属性评价中的指标无量纲化方法不仅要考虑将指标转变为固定区间内的值,更需要考虑获取可以进行定量的优劣比较

的值,要实现这两个目标,指标的无量纲化要做到下列三个步骤:同度量、同方向、同基准。同度量是值将同一指标的数据统一单位,这一步骤通常比较简单;同方向是指将反向指标、适度指标进行正向化处理;同基准是指将不同指标的正向化数据转变为标准化值,以实现不同指标的综合与比较^[19]。

无量纲化处理后的指标值通常被称为标准化值,按照指标值与标准化值之间的关系,可以将指标无量纲方法分为线性无量纲方法与非线性无量纲方法。线性无量纲方法假设指标值与标准化值之间呈线性关系。非线性无量纲方法是假设二者关系为非线性关系,指标值同等变动时,对应的标准化值变动不等^[20]。本文所提及的线性无量纲方法主要包括极值处理法、线性比例法、功效系数法及其改进方法,而非线性无量纲方法包括经典的对数型、幂函数型及改进的折线型、曲线型、复合函数型无量纲化方法。其中极值处理法、线性比例法是图情学中最常用的两种指标线性无量纲化方法^[21],而对于非线性无量纲化方法来说,复合函数型的非线性无量纲化处理法有较好的性质^[10]。

对指标进行无量纲化操作,还应考虑指标的类型及性质,按照指标的属性可以将指标分为三类,即正指标、反向指标和适度指标^[22]。正指标是指越大越有利的指标,无量纲化方法应使指标值越大,标准化值越大;反向指标是指越小越好的指标,无量纲化方法应使小的指标值获得较大的标准化值;适度指标是指不宜过大,也不宜过小的,以适度值最好的指标,无量纲化方法应做到将适度值两侧的值分别转

化为正指标和反向指标。对反向指标以及适度指标的正向化处理可以起到同方向的作用,这一步可以与同基准分开处理,也可以同时实现。

实际上线性与非线性无量纲化方法的分类方式已经囊括所有的无量纲化方法,但由于反向指标与适度指标的处理较为特殊且易于被忽视,特将其处理方法单独列举进行更详细的描述。因此本文提及的线性无量纲化方法与非线性无量纲化方法针对的是正指标和进行正向化后的指标,并对反向指标与适度指标的特殊处理方法单列进行详细说明。此外,无量纲化方法不仅在多属性评价中是必要的处理步骤,在其他领域也多有应用。因此,本文也对其在其他领域的应用进行介绍以充实研究内容并为多属性评价提供参考。

2 各种无量纲方法特点

2.1 线性无量纲化方法

线性无量纲方法是应用最为广泛、最为简便的无量纲化方法。这里先讨论正向指标的无量纲方法。

设多属性评价中共有 n 个单位, m 个指标, x_{ij} 表示第 i 个单位第 j 个指标值, y_{ij} 表示第 i 个单位第 j 个标准化值。

1) 极值处理法

$$y_{ij} = [x_{ij} - \min_i(x_{ij})] / [\max_i(x_{ij}) - \min_i(x_{ij})] \quad (1)$$

2) 线性比例法

$$y_{ij} = x_{ij} / x_j^* \quad (2)$$

其中, x_j^* 为特殊点,通常可以取 $\max_i(x_{ij})$ 、 $\min_i(x_{ij})$ 或 $\bar{x}_j$ 。

3) 功效系数法

$$y_{ij} = a + \frac{x_{ij} - m_j}{M_j - m_j} b \quad (3)$$

其中, M_j 、 m_j 分别为指标的满意值和不容许值, a 、 b 为已知正常数。

极值处理法被用于学术期刊评价^[23,24]、科技评价^[7]等方面, 使用这一无量纲化方法进行数据预处理的多属性评价方法包括了 TOPSIS、因子分析、熵值法、主成分分析等方法。而线性比例法则更常见于期刊评价^[25-27]中, 其应用范围也包括了 TOPSIS、熵值法、线性加权汇总等多属性评价方法。可以看到, 无量纲化方法在图情学的多属性评价中是必不可少的一部分, 研究中较为常用的仍然是极值处理法和线性比例法。

然而, 一些研究发现现有的几种线性去量纲化方法应用于图情学的多属性评价时, 在性质上存在一定的缺陷。俞立平、潘云涛等^[28]认为在期刊评价中, 指标的无量纲化应满足: 同一指标内部相对差距不变、不同指标间相对差距不变、标准化后极大值相等条件, 而上述方法中, 只有线性比例法满足这三个条件。韩明彩^[21]指出现有的应用于期刊评价的无量纲化方法存在不同样本数据不可比、结果不能反映价值认知标准等问题。郭亚军、易平涛^[29]发现经典的线性无量纲化方法均不能完全满足其提出的六大性质。

由于经典线性无量纲化方法存在的不足, 在此基础上提出了一些改进的方法, 这些方法大多基于经典方法, 但仍保留了标准化值与原数据之间线性的关系。韩明彩^[21]分析了在期刊综合评价中存在的问题, 提出一种新的指标标准化方法: 价值评估法。而针对其提出的在期

刊评价中无量纲化方法存在的问题其实也普遍存在于任何需要进行无量纲化处理的综合评价场景中, 因此, 在这些领域中无量纲化方法的改进可以对图情学提供借鉴的意义。例如, 对于个别数据增减的敏感性问题, 郭亚军、易平涛^[29]构造了一种复合无量纲化方法“极标复合法”。对于处理动态评价中的增量信息, 易平涛、李伟伟等^[30]提出了标准序列法、全序列法和增量权法。类似地, Gao L^[31]认为全序列法可以很好地处理动态综合评价。徐林明、李美娟^[32]对于同一问题的研究提出了全局归一化方法。

2.2 非线性无量纲化方法

线性无量纲化方法相较于非线性无量纲化方法较为简单, 虽然专门针对图情学领域对其改进研究不多, 但线性无量纲化方法应用于综合评价中的改进研究却较为全面。即便如此, 线性无量纲化方法仍有其适用范围: 一方面, 在使用线性无量纲化方法时, 要确保标准化值与原指标值之间是呈线性关系; 另一方面, 线性无量纲化方法易受个别极端数据的影响, 转换结果具有较大的偶然性, 且当各评价对象的属性值分布不均匀时, 大部分数据的转换结果均集中在区间的低端或高端, 会降低分辨率。

针对线性无量纲化方法的使用条件及存在的缺陷, 非线性无量纲化方法应运而生, 经典的非线性无量纲化方法^[33]包括指数型、对数型、幂函数型, 其用于综合评价中的主要是对数型和幂函数型, 其具体计算公式如下:

1) 对数型

$$y = \frac{\log_a x - \log_a x_0}{\log_a x_1 - \log_a x_0} \quad (4)$$

$$y = \frac{\log(x)}{\max[\log(x)]} \quad (5)$$

其中, y 为标准化值, x 为原指标值, x_0 为不容许值, x_1 为满意值。

2) 幂函数型

$$y = \frac{x^p - [\min(x)]^p}{[\max(x)]^p - [\min(x)]^p} \quad (6)$$

其中, y 为标准化值, x 为原指标值, p 为参数, 且 $p > 1$ 。

在传统的非线性无量纲化方法基础上, 也涌现了创新的研究, 其主要形式包括折线型无量纲化方法、曲线型无量纲化方法^[34]及复合函数型无量纲化方法等。

3) 折线型

折线型无量纲化方法通过寻找使标准化值发生变动的转折点, 对转折点前后的指标值采取线性无量纲化处理, 本质上就是进行分段的线性无量纲化处理。一般情况下, 非线性关系有两种情况: 一是指标值的变动对应于逐渐增大的标准化值, 此类指标值后期的变动对评价对象更重要; 二是指标值的变动对应于逐渐减小的标准化值, 此类指标值前期变动对评价对象更重要。蔡辉、丁昌慧^[35]针对第一种情况, 提出其折线无量纲化的公式为:

$$y = y_0 + \frac{y_m - y_0}{x_m - x_0} \times (x - x_0), x_0 \leq x < x_m \quad (7)$$

$$y = y_m + \frac{y_m - y_{\max}}{x_m - x_{\max}} \times (x - x_m), x_m \leq x < x_{\max} \quad (8)$$

其中, 将指标值 x , 标准化值 y 分别作为横纵坐标轴, 则 (x_m, y_m) 为转折点的坐标, (x_0, y_0) 为基础点的坐标, $(x_{\max}, y_{\max})$ 为最大指标值及其对应的标准化值。

针对第二种情况则只需在上述式子上进行些许改动。蒋维杨、赵嵩^[36]正结合大样本综合

评价的特点, 提出了阶梯型归一化处理方法, 将指标值划分为多个独立的子区间, 并分别确定其归一化的公式, 是对折线型无量纲化方法的一种改进方法。

4) 曲线型

曲线型无量纲化方法通过假设指标值与标准化值之间的曲线关系求解曲线方程。任传鹏、刘建国^[37]提出了一种基于曲线拟合的无量纲化方法, 该方法具体实施步骤为: (1) 确定统计数据的分布曲线 $f(x)$, 将指标值映射到所需的统一的值域。这一步包括绘制指标散点图, 初步确定分布曲线, 利用待定系数法、假设检验、回归分析等方法确定拟合曲线。(2) 求取标准化值, 即将指标值代入所得的拟合曲线得到标准化值。

5) 复合函数型

复合函数型无量纲化方法通过效用函数或强“奖惩”算子将指标值转化到一个合理的区间内。Ting Chen 等^[38]使用梯形隶属度函数确定每个测试值相对于不同安全级别的隶属度进行无量纲化处理。戴文战、邹立华、汪建章等^[39]提出的效用函数为:

$$y = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (9)$$

该效用函数以指标的平均值为参考点, 突出了“奖优罚劣”原则, 提高了分辨精度。同样, 基于“奖优罚劣”的原则, 刘学之、杨泽宇等^[40]根据 Sigmoid 函数构造了 S 型曲线:

$$y_{ij} = \frac{1}{1 + e^{-\alpha(x_{ij} - \beta)}} \quad (10)$$

其中 $\beta = \text{medi}x_{ij}$, $\alpha = \frac{\ln 9}{t}$, t 为 90% 分位点到中点的距离和 10% 分位点到中点的距离中的较小值。俞立平^[10]指出, 在科技评价中, 对

于存在自然权重问题的指标体系, 基于 Sigmoid 函数的非线性无量纲化方法能改善存在自然权重的问题。

以上研究均利用实例证明了利用具有“奖惩”功能的效用函数进行无量纲化处理, 可以有效地提高了评价对象的区分度和分辨率, 并且保持了评价对象的排列序位基本不发生变化, 更为符合图情学中对于多属性评价结果的要求。

在无量纲化处理中也涉及一类特别的研究, 即无量纲化中的异常点处理方法。李伟伟等^[41]对异常值进行判断与识别, 并提出了一种基于极值处理法的分段无量纲化处理方法。郭亚军、易平涛^[29]认为异常点是导致无量纲化结果稳定性降低的主要原因, 并分析了如何确定异常点数值, 将异常点处理成除异常点外的最大值或最小值再进行去量纲化。何乃强、惠晓滨等^[42]基于正态区间将 98% 置信区间外的数值确认为异常值, 使其等于正常值区间的上、下限, 在此基础上进行无量纲化。

2.3 反向指标与适度指标无量纲

反向指标与适度指标对评价结果所起的作用不同于正指标, 不加以处理所得结果会与实际情况产生较大差距。刘新华^[43]利用实际数据进行不同指标正向化方法处理, 并进行处理前后的对比, 发现利用原始数据所做评价与实际情况严重不符, 从而说明了正向化处理的必要性。

(1) 处理方法一

一种处理反向指标及适度指标的方法是先对原指标进行正向化再进行无量纲化的处理, 具体正向化方法如下:

设综合评价中共有 n 个单位, m 个指标, x_{ij} 表示第 i 个单位第 j 个指标值, x'_{ij} 表示第 i 个单位第 j 个正向化的指标值。

对反向指标正向化处理的计算公式如下:

$$x'_{ij} = \frac{1}{x_{ij}} \quad (11)$$

$$x'_{ij} = a - bx_{ij}, \quad b > 0, \quad a \text{ 为任意常数} \quad (12)$$

$$x'_{ij} = \frac{1}{\max_i |x_{ij}| + x_{ij} + a}, \quad a \geq 0 \quad (13)$$

设 a 为适度指标的适度值, 则对适度指标正向化处理的计算公式如下:

$$x'_{ij} = \frac{1}{K + |a - x_{ij}|}, \quad K \geq 0 \quad (14)$$

$$x'_{ij} = K - |a - x_{ij}|, \quad K > 0 \quad (15)$$

反向指标正向化的三种方法中, 式(11)^[44,45]和式(12)^[46]使用较为广泛式(13)可以看成是对式(13)的改进, 在文献[47]中有所应用。对适度指标的正向化处理, 文献[48,49]采用了式(14), 文献[50,51]采用了式(15)。

选择反向指标正向化方法的标准在理论上并未得到统一。一些研究认为式(12)稳定性更高^[52], 不改变指标的分布规律^[53]及相对位置^[54]。因此更适合用于反向指标的正向化处理, 但文献[55]认为式(12)进行处理前后的综合评价结果不变, 体现不出正向化处理的效果, 因此是无效的, 式(11)与式(13)更合理。此外, 当数据中出现 0 时, 可以使用式(13)进行正向化。

选择适度指标正向化方法的标准也并不统一。当考虑处理后的数据与原始数据分布的一致性时, 选择式(14)较为合理^[52-54]。但实际应用中, 选择式(15)的并不在少数, 这一现

象可能由指标正向化处理不受重视,或对指标正向化处理的理论研究较少所导致,学者研究时直接采取了较为广泛应用的方法。

(2) 处理方法二

另一种处理方法是,进行指标无量纲化处理时同时考虑反向指标及适度指标正向化的问题。常用的反向指标无量纲化方法有以下几种:

$$y = \frac{\max(x) - x}{\max(x) - \min(x)} \quad (16)$$

$$y = 1 - \frac{x}{\max(x)} \quad (17)$$

$$y = Ae^{\frac{\max(x) - x}{\max(x) - \min(x)}} \quad (18)$$

其中式(16)(17)为反向指标的线性无量纲化方法,俞立平、武夷山^[56]研究发现反向指标经典线性无量纲化方法会影响评价结果的排序,而以上两种方法均不满足其提出的三大原则。因此在式(17)基础上提出了一种新的线性反向指标标准化方法,其改进后的公式为:

$$y = 1 - \frac{x}{\max(x)} + \left\{ 1 - \max \left[1 - \frac{x}{\max(x)} \right] \right\} \quad (19)$$

从对指标正向化方法的研究中可以看到,根据不同的选择标准,合理的指标正向化方法是多样性的,但不改变指标分布特征是更为公认且符合统计规范的标准。在实际应用中,不应追求便利,理所当然选择使用较多的方法,而应根据数据特征及数据处理标准选择不同的正向化方法,使综合评价的结果更为合理。

2.4 其他无量纲化方法

无量纲化方法作为数据处理的一种方法,有众多形式,但并非所有的无量纲化方法都适用于综合评价中。本部分主要介绍一些方法不

适用于多属性评价的原因以及其他领域应用的无量纲化方法,一方面可以减少无量纲化方法的误用,另一方面也为多属性评价中无量纲化方法的改进提供思路。

首先,一些较为常用的无量纲化方法并不适用于多属性评价中,比如下面两种方法:

1) 标准化处理法

$$y_{ij} = (x_{ij} - \bar{x}_j) / s_j \quad (20)$$

其中, $\bar{x}_j$ 、 s_j 分别为第 j 项指标值的平均值和标准差。

2) 归一化处理法

$$y_{ij} = x_{ij} / \sum_{i=1}^n x_{ij} \quad (21)$$

3) 向量规范法

$$y_{ij} = x'_{ij} / \sqrt{\sum_{i=1}^n x'^2_{ij}} \quad (22)$$

根据俞立平、潘云涛等^[28],根据标准化后极大值相等原则,在多属性评价中的无量纲化必须保证标准化后的极大值全部相等,否则就失去了标准化的意义。因此归一化处理法和向量规范法并不适用于多属性评价,需要进行相应的改进。

其次,在其他领域对无量纲化方法的研究对于多属性评价中的无量纲化应用会带来新的思考及改进方法。比如灰色关联度是用以描述因素间关联度强弱的方法,张月、肖峰^[57]比较了极值处理法、标准化处理法、归一化处理法三种无量纲化方法对灰色关联分析的影响,结果表明标准化方法的结果更符合实际情况。张戈^[58]利用高校选课数据构建高校学生选课预测模型时,探讨了数据分别利用归一化和标准化的数据预处理对预测结果准确性的影响,结果显示归一化可解决数据收敛程度的问题,标准

化避免了模型拟合时某些特征向量系数归于零的问题。

3 无量纲化方法的选择

(1) 线性无量纲化方法的选择

用于线性无量纲化的方法众多,有部分研究针对性地探讨了选择线性无量纲化方法的标准。这些标准可应用于任何使用多属性评价的领域。He Y 等^[59]、糜万俊^[60]从方差角度考虑无量纲化方法的选择原则,他们的研究考虑了无量纲化方法对属性方差的影响以及这种影响对属性权重的传导机制,其研究说明了可通过方差最小优化方法进行无量纲化方法的选择。Chakraborty S 等^[61]从权重敏感性和从序一致性角度评估评价无量纲化方法的优劣。郭亚军、易平涛^[29]从线性无量纲化方法的本身性质出发,提出理想无量纲化方法应满足的性质,包括单调性、差异比不变性、平移无关性、缩放无关性、区间稳定性、总量恒定性。李玲玉、郭亚军等^[62]探讨了无量纲化方法的三个选取原则,即变异性、差异性、稳定性。其中,变异性要保证原有数据的离散程度,这一原则是前提,可利用变异系数衡量;差异性则要突出对象间的差异,寻找离差平方和最大的方法;稳

定性则衡量极端对象对评价的影响,通过相关系数衡量。樊红艳,刘学录^[63]认为保留原始数据整体的一致性和关联系数一致性,并且不以个别极值来进行数据的标准化方法是合理的标准化方法。

除了从方法的本身性质出发外,有相关文献是从不同的多属性评价方法的无量纲化方法适用性角度出发进行研究的。朱喜安、魏国栋^[64]依据郭亚军、易平涛^[29]提出的理想无量纲化方法的六种性质,发现利用熵权法进行多属性评价时,利用极值处理法所得的评价结果最优。李玲玉、郭亚军等^[62]结合提出的三大原则,认为对于拉开档次法的评价来说,使用线性比例法中选取最小值的无量纲化方法最优。

(2) 非线性无量纲化方法的选择

一方面对于经典的非线性无量纲化方法,苏为华^[33]指出,公式的繁多并不利于无量纲化方法实现其效果,在进行选择时,可从无量纲化方法的方向上进行判断,包括递增型、递减型、常数型、混合型。结合不同类型的图像特征,可以选择出最适合用于某具体数据的方法。

另一方面在非线性无量纲化三种改进的处理方法的选择中,可结合下表各自的优缺点进行选择:

表 1 三种改进非线性无量纲化方法优缺点对比

非线性无量纲化方法	优点	缺点
折线型	直线函数的确定比较容易;直线形式较为直观,计算方便,便于理解和应用	难点在于如何确定转折点的个数及位置;不适用于缓慢变化的非线性过程
曲线型	反映的信息较为客观、全面;具有“奖优罚劣”的功能	一般需要已知较详细的坐标点,而在无量纲化中,标准化值往往是未知的,所以曲线方程很难确定;曲线拟合一般仅适用于本身所处理的指标,适用范围非常有限。
复合函数型	分辨精度高,具有“奖优罚劣”的功能,适用范围较广	操作较为繁琐,可选择的效用函数较少

4 研究结论与展望

4.1 研究结论

第一, 指标无量纲化方法对多属性评价的权重及评价结果都有重大影响, 应科学选择合理的无量纲化方法。对于标准化值与原指标值呈线性关系的情形, 选择线性无量纲化方法, 当标准化值与指标值呈现非线性关系或具有异常点时, 需选择非线性无量纲化方法。

第二, 常用于图情学多属性评价的线性无量纲化方法包括极值处理法、线性比例法、功效系数法等。对于不同的评价方法, 每种无量纲化方法处理的性能表现有所差异, 并没有万能的处理公式。此外, 对于经典线性无量纲化的一些缺陷, 如不同样本数据结果不可比、标准化极大值不一致等问题, 研究中也给出了针对性的改进方法。这些方法具有不同的数学性质, 且会产生不同的评价结果, 可以根据多属性评价的具体方法以及对于评价结果的具体标准进行选择。

第三, 非线性无量纲化的方法在原有的指数型、对数型及幂函数型上有新的改进, 主要形式包括折线型无量纲化方法、曲线型无量纲化方法及复合函数型无量纲化方法等。非线性无量纲化方法虽然在操作上更为复杂, 但其具有奖优罚劣、消除自然权重的功能, 能显著提高多属性评价结果的区分度、分辨精度, 能满足图情学多属性评价结果的要求。因此, 针对性的改进及应用也在逐渐增多。

第四, 对指标进行无量纲化的过程通常包括三个方面的内容: 同度量、同方向、同基准, 其中需要引起重视的是同方向与同基准两步,

同方向即针对反向指标、适度指标的正向化处理。通常的处理方法是先进行反向指标与适度指标的正向化处理, 再进行同基准的操作, 也有一些无量纲化方法可以将这两步在一次处理中完成。

第五, 图情学多属性评价中的无量纲化方法选择有多种标准, 对于线性无量纲化方法可从方差最小化、无量纲化方法的稳定性、标准值与原始数据的一致性角度进行选择, 而且一些常见的无量纲化方法是不适用于图情学评价中的。非线性无量纲化方法常需结合数据的具体特征加以选择, 可以从图像、处理效率等角度进行考虑。

4.2 研究展望

针对现有研究的整理分析, 可以发现进一步研究的方向主要包括:

首先, 虽然在多属性评价中对于无量纲化方法的性质研究及改进方法研究都较为全面, 但专门针对图情学的研究还是相对较少, 在这一方面可以借鉴其他领域的改进方法做相应处理。

其次, 对指标无量纲化方法在不同评价方法中的适用性研究。这一问题可以从两个角度考虑, 一是根据一定的选择标准, 对常用的图情学多属性评价方法选择其最适合的无量纲方法, 而且实际上针对这一问题的研究并不多; 二是通过对无量纲化方法的改进提高其鲁棒性, 使其可以运用于多种评价方法中。

最后, 在图情学多属性评价的实际应用中, 指标无量纲化方法的选择仍没有受到应有的重视。一方面部分研究会采用较为常用的方法而

忽视应用前提, 另一方面则是对使用的无量纲化方法及选择的原因不加以阐述。因此, 在运用多属性评价方法时, 应重视对这一方面的说明以便于后续研究。

参考文献

- [1] 王宗军. 综合评价的方法、问题及其研究趋势[J]. 管理科学学报, 1998(1):75-81.
- [2] 李远远, 云俊. 多属性综合评价指标体系理论综述[J]. 武汉理工大学学报(信息与管理工程版), 2009, 31(2):305-309.
- [3] 俞立平. 期刊影响力指标的时间异质性及其重构研究——基于多属性评价的视角[J]. 图书情报工作, 2016, 60(12):109-114.
- [4] 邱均平, 张荣, 赵蓉英. 期刊评价指标体系及定量方法研究[J]. 现代图书情报技术, 2004(7):23-26.
- [5] 俞立平, 潘云涛, 武夷山. 学术期刊多属性评价方法的可比性研究[J]. 编辑学报, 2010, 22(5):381-384.
- [6] 熊国经, 熊玲玲, 陈小山. 基于因子分析与TOPSIS法在学术期刊评价中的改进研究[J]. 情报杂志, 2016, 35(7):196-200.
- [7] 顾雪松, 迟国泰, 程鹤. 基于聚类-因子分析的科技评价指标体系构建[J]. 科学学研究, 2010, 28(4):508-514.
- [8] 谈毅, 仝允桓. 中国科技评价体系的特点、模式及发展[J]. 科学学与科学技术管理, 2004(5):15-18.
- [9] 章穗, 张梅, 迟国泰. 基于熵权法的科学技术评价模型及其实证研究[J]. 管理学报, 2010, 7(1):34-42.
- [10] 俞立平, 宋夏云. 评价型无量纲法、自然权重与线性科技评价[J]. 现代情报, 2020, 40(4):120-127.
- [11] 张爱琴, 俞立平, 赵公民. 科技评价中加权TOPSIS的权重可靠吗?——基于分子加权TOPSIS法的改进[J]. 现代情报, 2018, 38(11):48-54.
- [12] 程津培, 李晓轩, 徐芳. 行进中的中国科技评价制度改革[J]. 中国科学基金, 2019, 33(2):105.
- [13] 邱均平, 谢辉, 李进华. 竞争情报系统绩效综合评价研究[J]. 情报科学, 2011, 29(4):481-485+498.
- [14] 夏立新, 王俊. 网络环境下竞争情报多主体模糊综合评价方法研究[J]. 情报科学, 2008(6):860-865.
- [15] 糜万俊, 江文奇. 不完全信息下面向过程的竞争情报评价研究[J]. 情报杂志, 2017, 36(1):132-135+186.
- [16] Ma Feng-Mei, Guo Ya-Jun. Density-induced ordered Weighted averaging operators [J]. International Journal of Intelligent Systems, 2011, 26(9):866-886.
- [17] 江文奇. 无量纲化方法对属性权重影响的敏感性和方案保序性[J]. 系统工程与电子技术, 2012, 34(12):2520-2523.
- [18] 张卫华, 赵铭军. 指标无量纲化方法对综合评价结果可靠性的影响及其实证分析[J]. 统计与信息论坛, 2005(3):33-36.
- [19] 詹敏, 廖志高, 徐玖平. 线性无量纲化方法比较研究[J]. 统计与信息论坛, 2016, 31(12):17-22.
- [20] 周德群. 也谈无量纲方法——兼与蒋萍、刘秀梅等同志商榷[J]. 中国统计, 1995(10):36-37.
- [21] 韩明彩. 期刊综合评价指标标准化方法研究: 价值评估法——以图书情报学期刊为例[J]. 情报理论与实践, 2012, 35(10):30-32.
- [22] 王晓军. 多指标综合评价中指标无量纲化方法的探讨[J]. 人口研究, 1993(4):47-51.
- [23] 熊国经, 熊玲玲, 陈小山. 基于因子分析与TOPSIS法在学术期刊评价中的改进研究[J]. 情报杂志, 2016, 35(7):196-200.
- [24] 曾伟, 田时中, 田家华. 科技期刊学术影响力综合评价模型与实证[J]. 中国科技期刊研究, 2016, 27(3):316-323.
- [25] 徐新华, 俞立平, 王作功. 学术期刊评价中不同指标之间互补研究[J]. 图书情报工作, 2018, 62(14):98-105.
- [26] 俞立平, 张全, 刘爱军. 不同学科多属性评价横向比较研究——以数学、物理学、生物学期刊为例[J]. 图书情报工作, 2014, 58(20):100-105.
- [27] 俞立平. 基于学术期刊视角的学科发展水平评价研究[J]. 图书情报工作, 2016, 60(4):5-11+28.
- [28] 俞立平, 潘云涛, 武夷山. 学术期刊综合评价数据标准化方法研究[J]. 图书情报工作, 2009, 53(12):136-139.
- [29] 郭亚军, 易平涛. 线性无量纲化方法的性质分析[J]. 统计研究, 2008(2):93-100.
- [30] 易平涛, 李伟伟, 郭亚军. 线性无量纲化方法的结构稳定性分析[J]. 系统管理学报, 2014, 23(1):104-110.
- [31] Gao Liqun. Study on Dimensionless Methods in Dynamic Comprehensive Evaluation[J]. Journal of Northeastern University, 2009, 30(6):889-892.

- [32] 徐林明, 李美娟. 动态综合评价中的数据预处理方法研究[J]. 中国管理科学, 2020, 28(1):162-169.
- [33] 苏为华. 多指标综合评价理论与方法问题研究[D]. 厦门: 厦门大学, 2000.
- [34] Hwang C L, Yoon K P. Multiple attribute decision making: methods and applications[M]. Berlin: Springer - Verlag, 1981:1-50
- [35] 蔡辉, 丁昌慧. 综合效益评价中数据的非直线化无量纲化方法[J]. 中国医院统计, 2003(1):46-48.
- [36] 蒋维杨, 赵嵩正, 刘丹, 王莉芳. 大样本评价的定量指标无量纲化方法[J]. 统计与决策, 2012(17):4-9.
- [37] 任传鹏, 刘建国. 一种基于曲线拟和的无量纲化方法[J]. 山东煤炭科技, 2005(6):33-34.
- [38] Ting Chen, Yiyang Jin, Xiaopeng Qiu, Xin Chen. A hybrid fuzzy evaluation method for safety assessment of food-waste feed based on entropy and the analytic hierarchy process methods[J]. Expert Systems with Applications, 2014;41(16):7328-7337.
- [39] 戴文战, 邹立华, 汪建章, 朱金刚, 周融. 一种基于奖优罚劣原则的多阶段多目标决策模型[J]. 系统工程理论与实践, 2000(6):32-36.
- [40] 刘学之, 杨泽宇, 沈凤武, 尚玥佟, 刘嘉. 基于S型曲线的指标非线性标准化研究[J]. 统计与信息论坛, 2018, 33(2):17-21.
- [41] 李伟伟, 易平涛, 李玲玉. 综合评价中异常值的识别及无量纲化处理[J]. 运筹与管理, 2018, 27(4):173-178.
- [42] 何乃强, 惠晓滨, 周璇. 基于正态区间估计的改进型无量纲化方法[J]. 计算机工程与应用, 2012, 48(5):236-238.
- [43] 刘新华. 因子分析中数据正向化处理的必要性及其软件实现[J]. 重庆工学院学报(自然科学版), 2009, 23(9):152-155.
- [44] 杜心灵. 河南省城市经济发展水平的实证分析[J]. 经济师, 2003(12):225-226.
- [45] 郑成德, 尹德松. 主成分分析法在城市大气环境质量评价中的应用[J]. 大连铁道学院学报, 2002(3):39-41.
- [46] 陆琦, 郭宗楼, 姚杰. 灌区灌溉管理质量的综合评价指标研究[J]. 农业工程学报, 2005(S1):15-19.
- [47] 陈迪红, 李华中, 杨湘豫. 行业景气指数建立的方法选择及实证分析[J]. 系统工程, 2003(4):72-76.
- [48] 李燕. 基于因子分析法的城市商业银行财务风险预警体系研究[J]. 金融理论与实践, 2017(4):31-34.
- [49] 徐国祥, 檀向球, 胡穗华. 上市公司经营业绩综合评价及其实证研究[J]. 统计研究, 2000(9):44-51.
- [50] 王军. 商业银行核心竞争力评价的指标体系构建与运用[J]. 统计与决策, 2011(5):137-140.
- [51] 谢赤, 钟赞. 熵权法在银行经营绩效综合评价中的应用[J]. 中国软科学, 2002(9):109-111+108.
- [52] 张立军, 袁能文. 线性综合评价模型中指标标准化方法的比较与选择[J]. 统计与信息论坛, 2010, 25(8):10-15.
- [53] 叶宗裕. 关于多指标综合评价中指标正向化和无量纲化方法的选择[J]. 浙江统计, 2003(4):25-26.
- [54] 范坤, 冯长焕. 因子分析中指标数据如何正确预处理[J]. 财会月刊, 2013(6):85-88.
- [55] 陈军才. 主成分与因子分析中指标同趋势化方法探讨[J]. 统计与信息论坛, 2005(2):19-23.
- [56] 俞立平, 武夷山. 科技评价中标准化方法对评价结果的影响研究[J]. 现代图书情报技术, 2011(9):66-71.
- [57] 张月, 肖峰. GRA方法中的无量纲化比较[J]. 统计与决策, 2005(4):120-121.
- [58] 张戈. 课程推荐预测模型中的数据预处理方法研究[J]. 中国新通信, 2019, 21(19):185.
- [59] He Y, Zhang Y, Pang Y. Comprehensive Evaluation of Global Clean Energy Development Level Based on Compatibility and Difference Degree[J]. Mathematical Problems in Engineering, 2019:1-16.
- [60] 糜万俊. 无量纲化对属性权重影响的传导机制及调权研究[J]. 统计与决策, 2013(4):11-14.
- [61] Chakraborty S, Yeh C. A simulation comparison of normalization procedures for TOPSIS[C]. International Conference on Computers and Industrial Engineering, 2009: 1815-1820.
- [62] 李玲玉, 郭亚军, 易平涛. 无量纲化方法的选取原则[J]. 系统管理学报, 2016, 25(6):1040-1045.
- [63] 樊红艳, 刘学录. 基于综合评价法的各种无量纲化方法的比较和优选——以兰州市永登县的土地开发为例[J]. 湖南农业科学, 2010(17):163-166.
- [64] 朱喜安, 魏国栋. 熵值法中无量纲化方法优良标准的探讨[J]. 统计与决策, 2015(2):12-15.