

学术期刊评价中主成分分析法应用悖论研究¹

俞立平 潘云涛 武夷山

摘要：本文依托中国科学技术信息研究所 CSTPC 数据库，采用主成分分析对医学期刊进行评价，然后用评价值作为被解释变量，评价指标作为解释变量进行回归分析，发现部分指标的回归系数为负，删除这几个指标后继续进行主成分分析及回归分析，发现可以删除绝大多数评价指标，得出违背常理的结论。本文认为，由于学术期刊的特点，不能采取主成分分析进行学术期刊的评价；主成分分析法自身提供的方法适用性检验是不完备的，必须再增加主成分分析回归系数为正这一条件；在评价条件完全具备的情况下，利用本文提出的主成分结果回归法（PCARR）可以进行评价指标的筛选；因子分析法由于和主成分分析原理类似，也存在同样的问题；本文提供的方法可广泛应用于某些系统评价方法的检验。

关键词：学术期刊 评价 主成分分析 回归

1 引言

期刊评价是文献计量学研究的重要组成部分，它通过对学术期刊的发展规律和增长趋势进行量化分析，揭示学科文献数量在期刊中的分布规律，为优化学术期刊的使用提供重要参考，同时可以提高学术期刊的内在质量，促进学术期刊的健康成长和发展。美国著名情报学家加菲尔德博士在 20 世纪 60 年代对期刊文献的引文进行了大规模统计分析，得到了大量被引用文献集中在少数期刊上，而少量被引用文献散布在大量期刊中的结论，这可以被认为是国外期刊评价理论的起源。

学术期刊评价主要有两大类方法，一种是直接根据期刊评价指标进行对比，二是采用多指标综合评价方法，该方法克服了单指标评价容易带来的片面性，比较适合期刊的综合评价，因此得到了较广泛的应用。学术期刊多指标综合评价是一项复杂的系统工程，牵涉到评价原则、指标选取、数据归一化、评价方法选择等诸多方面，国内外学者在该领域进行了大量广泛的研究。从评价方法的角度，Weiping Yue、Concepcion S. Wilson (2004)^[1]利用结构方程的原理建立了一个期刊影响力的分析框架。苏新宁 (2008)^[2]采用指标体系赋权进行中国人文社会科学期刊的评价。邱均平、张荣等 (2004)^[3]提出了期刊评价指标体系的三维层次结构图，并利用灰色关联法进行评价。庞景安、张玉华等 (2000)^[4]及李凯扬、贾玉萍 (2005)^[5]利用层次分析法对期刊进行评价。王小唯、杨波等 (2003)^[6]将期刊以往状态的评价结果作为各期刊基础条件的一种度量，再运用数据包络分析方法 (DEA) 测算出它们的二次相对评价值。李修杰、陈景武 (2006)^[7]运用判别分析法建立的期刊评估指标体系。王玖、徐天和 (2003)^[8]运用秩和比法进行医学学术期刊学术质量综合评价。陈汉忠 (2004)^[9]等数位学者^[10-15]应用主成分分析对学术期刊进行评价。凌春艳、莫琳 (2004)^[16]提出自然科学学术期刊质量指标体系的属性数学综合评价模型并进行了评价。李继晓、蔡成瑛 (2006)^[17]对各种核心期刊的评价方法进行了介绍和分析，认为选择核心期刊的评价方法并不是一件一劳永逸的事，只能通过不断实践、比较、分析，才能使核心期刊的评价日臻完善。此外，还有一些学者将两种或两种以上的评价方法进行有机融合，采用复合方法进行评价，如根据层次分析法确定权重，再采取加权 TOPSIS 法进行评价，由于复合评价方法只有一个评价结果，因此本质上仍然可以认为是一种多指标综合评价方法。

¹国家十一五支撑计划项目 (2006BAH03B05)；国家自然科学基金资助 (70673019)

作者简介：俞立平 (1967-) 男，江苏泰州人，博士，教授，主要从事信息经济、科学计量领域的教学科研工作。Email: chinayangzhou@yahoo.com.cn

潘云涛 (1967-) 女，中国科学技术信息研究所研究员，主要从事期刊评价领域的研究。

武夷山 (1958-) 男，中国科学技术信息研究所研究员，主要从事科学计量学与科技管理领域的研究。

主成分分析是一种多元统计方法,可以将众多指标简化浓缩为少数几个甚至一个综合评价指标,使简化的指标既能基本包括全部指标具有的信息,又使指标之间相互无关。它是一种客观评价方法,在经济、管理、环境、教育等评价领域得到了较广泛的应用,也是学术期刊评价中应用较多的一种评价方法,此外专家评分法也是用得较多的一种评价方法。本文依托中国科学技术信息研究所医学期刊文献计量学指标数据,采用主成分分析进行评价,然后将评价结果作为被解释变量,评价指标作为解释变量进行多元回归分析,提出了一种新的主成分结果回归法(PCARR, Principle Components Analysis Result Regression),对学术期刊评价应用主成分分析的适用条件和注意事项做深入探讨。

2 研究方法

2.1 主成分分析

主成分分析(Principle Components Analysis)的概念首先由 Karl Parson 在 1901 年提出,当时只对非随机变量进行了讨论,1933 年 Hotelling 将这个概念推广到随机变量。在多数实际问题中,由于指标较多及指标间有一定的相关性,势必增加分析问题的复杂性。主成分分析就是设法将原来指标重新组合成一组新的互相无关的几个综合指标来代替原来指标,同时根据实际需要从中选取几个较少的综合指标尽可能多地反映原来的指标的信息。

2.2 多元回归分析

为了分析主成分分析的评价结果与各指标间的关系,采用常见的多元线性回归模型进行分析,其基本形式如下:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (2)$$

Y 为评价值,用其作为被解释变量, $X_1 \dots X_n$ 为期刊评价指标即被解释变量, $a_1 \dots a_n$ 为系数,实际上该系数反映了各指标的贡献大小,某种程度上相当于权重。

3 数据

本文数据来自于中国科学技术信息研究所 CSTPC 数据库,以医学类期刊为例进行分析。中国科学技术信息研究所从 1987 年开始对中国科技人员在国内外发表论文数量和被引情况进行统计分析,并利用统计数据建立了中国科技论文与引文数据库,同时出版《中国学术期刊引证报告》。本文数据是 2006 年的医学学术期刊数据,共 518 种医学期刊,表 1 为数据的描述统计量。

表 1 变量描述统计量

变量名	内容	均值	最大值	最小值	标准差
X1	总被引频次	794.02	5805.00	17.00	811.65
X2	他引率	0.81	1.00	0.11	0.14
X3	扩散因子	31.95	78.43	3.67	15.74
X4	学科影响指标	0.54	1.00	0.02	0.26
X5	学科扩散因子	5.14	31.20	0.17	3.81
X6	被引半衰期	4.27	10.00	1.75	1.13
X7	影响因子	0.44	1.86	0.06	0.28
X8	即年指标	0.05	0.33	0.00	0.04
X9	基金论文比	0.26	0.97	0.01	0.19
X10	平均作者数	4.04	8.16	2.04	0.87
X11	平均引文数	9.25	34.72	3.12	3.88
X12	引用半衰期	5.63	8.52	2.78	0.86

X13	地区分布数	22.61	31.00	2.00	5.81
X14	海外论文比	0.02	0.78	0.00	0.05
n	期刊数量	518			

数据来源：中国科学技术信息研究所 CSTPC
数据库 2006 年数据

由于要对期刊进行评价，所以必需对数据进行标准化处理，每项指标最大值设为 100，然后按比例进行调整。此外，被引半衰期和引用半衰期是两个反向指标，必须进行适当处理，方法是用 100 减去其标准化后的结果后再做标准化，这种处理方式是线性处理方式，比反向指标取倒数的非线性处理方式要科学一些。

4 实证结果

在进行主成分分析时，必须首先进行 KMO 检验与 Bartlett 检验，这是主成分分析法的适用条件。KMO 是对样本充分度进行检验的指标，一般要大于 0.5。本文采用 SPSS 进行数据处理，KMO 值为 0.680>0.5，通过了统计检验；Bartlett 值为 3319.828， $P<0.000$ ，也通过了统计检验。换句话说，期刊评价采用主成分分析（因子分析）的条件全部具备。前 4 个主成分的累计贡献率为 67.49%，因此采用前 4 个主成分进行评价。这也从另外一个侧面说明评价期刊质量的视角是多方面的，至少有 4 个互相独立的因素。表 2 是主成分结果回归结果。

从主成分分析回归结果可以看出，除了扩散因子指标在 10%的水平上没有通过统计检验外，其它所有指标在 1%的水平上都通过了统计检验， R^2 为 0.9496，具有较高的拟合优度。但是他引率、扩散因子、学科影响指标、被引半衰期、地区分布数的系数为负数，也就是说，这 5 个指标越低，总体评分越高，这与人们的常识是完全相反的。

假设主成分分析结果代表了医学期刊的总体水平，下面对回归系数为负数的指标进行进一步分析：

他引率指论文被其他期刊引用次数占总引用次数的百分比，该指标的设计是为了防止部分期刊为了提高自己的影响因子而人为增加自引。通常质量高的期刊一般不会这么做，质量低的期刊才会采取如此措施，因此如果期刊人为增加自引情况较普遍，那么他引率与总评分值必然呈负相关，当然还有学科特点、数据等问题。

许多医学期刊的专业化程度较高，其在学科内以及总体统计源期刊内被引水平有一定的限度。扩散因子是期刊每被引 100 次所涉及的期刊数，学科影响指标是学科内引用该刊的期刊占全部期刊的比例，这两个指标的含义是类似的。一般来说，其他学科引用医学期刊较少，而医学期刊由于分工很细，除了少部分综合医学期刊外，总是学科内引用较多。比如眼科期刊少，儿科期刊多，那么眼科期刊与儿科期刊的被引情况必然大不相同。因此扩散因子、学科影响指标与总评分值负相关。

被引半衰期通常不是针对个别文献或某一组文献，而是对某些学科或专业领域的文献总和而言的，在医学学科内，各子学科的发展速度并不一样，那些发展较快的子学科，被引半衰期可能较短，发展较慢的子学科，被引半衰期可能较长。但是如果将各子学科放在一起进行回归分析，该指标有可能为负。

如果全国各地区所有医学子学科均衡发展，那么质量较好的期刊必然地区分布数较高，而现在地区分布数系数为负，可能与国内不同医学子学科发展不均衡有关。

基于以上分析，删除他引率、扩散因子、学科影响指标、被引半衰期、地区分布数 5 个指标，重新进行主成分分析，当然要继续进行 KMO 和 Bartlett 检验，将分析结果和剩下的 9 个指标继续进行回归，如仍然见表 2。

从 9 个指标主成分分析的回归结果看， R^2 为 0.7874，降低很多，属于中等水平的相关。总被引频次没有通过统计检验，且回归系数为负，其他所有指标在 1%的水平上都通过了统

计检验。学科扩散因子、平均作者数、引用半衰期的系数为负数，也就是说，这4个指标越低，总体评分越高，问题依旧。

总被引频次的系数为负，这是最出人意料之外的，可能原因是不同医学期刊办刊历史不同，有的期刊办刊较长，有的较短，当然办刊较长的期刊总被引频次较高。学科扩散因子的系数为负，原因和上述扩散因子、学科影响指标相同。平均作者数的系数为负，原因可能是多方面的，比如一些作者在论文写作时署名比较随意。引用半衰期的系数为负，原因和被引半衰期相同，与不同医学子学科的发展速度有关。

基于以上分析，继续删除总被引频次、学科扩散因子、平均作者数、引用半衰期4个指标，用剩下的5个指标做主成分分析，然后在此基础上进行回归，见表2。

所有5个指标在1%的水平上都通过了统计检验， R^2 为0.6560，继续降低，但拟合优度可以接受。影响因子、平均引文数的回归系数为负，尤其是期刊评价中用得最多的指标影响因子为负，这是难以解释的，仅剩下3个指标已经没有必要采取主成分分析，当然继续采取主成分结果回归已经没有必要。

即使是剩下的3个指标（即年指标、基金论文比、海外论文比）也并非毫无问题。某些学科由于发展较慢，也许即年指标不能说明问题；基金论文比存在的问题是部分作者可能会在论文写作时人为增加基金支持；海外论文比显然也比较片面。

那么问题究竟出在什么地方？撇开主成分分析与回归分析，如果站在期刊的角度，提高这14个指标中的任何一个（只要不是人为的），都是值得鼓励的，并没有任何问题。主要原因在于，在期刊评价中，有一个大前提，就是同类可比，但是与其他评价方法相比，主成分分析发似乎对同类可比这个大前提比较敏感，即使在同一医学期刊评价中也是如此。或者说，主成分分析结果回归系数为正是采用主成分分析法进行期刊评价的前提条件。

那么是否是由于多重共线性问题导致了回归系数本应为正的指标出现出负数呢？本文重点并不在于通过回归分析解决多重共线性问题，而是要找到主成分分析中那些指标值增加，评价结果反而下降的指标。为了进一步进行验证，本文选了两种期刊，将他引率和扩散因子值（这两个指标回归系数为负）都适当加大了，其他所有值都不变，然后再用主成分分析进行评价，发现这两种期刊的排名都后退了，这是不可以解释的。

表2 各种主成分结果回归结果

变量	含义	14个指标主成分 分析后回归	9个指标主成分 分析后回归	5个指标主成 分分析后回归
X1	总被引频次	0.0030** (6.3026)	-0.000952 (-0.9178)	--
X2	他引率	-0.0034*** (-10.2138)	--	--
X3	扩散因子	-0.0006 (-1.6329)	--	--
X4	学科影响指 标	-0.0012*** (-5.3378)	--	--
X5	学科扩散因 子	0.0039*** (7.6585)	-0.0039*** (-4.1248)	--
X6	被引半衰期	-0.0060*** (-18.6412)	--	--
X7	影响因子	0.0051*** (10.6334)	0.0081*** (8.0907)	-0.0081*** (-7.0813)
X8	即年指标	0.0069***	0.0090***	0.0041***

		(17.0723)	(10.1618)	(3.0245)
X9	基金论文比	0.0034***	0.0064***	0.0072***
		(13.9702)	(12.2397)	(9.6860)
X10	平均作者数	0.0054***	-0.0079***	--
		(12.6233)	(-11.5879)	--
X11	平均引文数	0.0064***	0.0073***	-0.0042
		(13.8040)	(7.3317)	(-3.8117)
X12	引用半衰期	0.0045***	-0.0051***	--
		(15.0577)	(-9.6263)	--
X13	地区分布数	-0.0044***	--	--
		(-16.7668)	--	--
X14	海外论文比	0.0097***	0.0268***	0.0596***
		(12.7174)	(15.7801)	(25.0172)
KMO/Bartlett	主成分检验	0.685/3319.828	0.677/1593.265	0.627/646.158
R ²	拟合优度	0.9496	0.7874	0.6560
n	期刊数	518	518	518

注：***表示在 1%的水平上统计检验显著。

5 结论与讨论

5.1 学术期刊不能用主成分分析法进行评价

虽然主成分分析法提供了一定的统计检验方法,用来检验某种评价是否可以采用主成分分析进行,但这种检验是不可靠的。在学术期刊评价中,不同期刊的办刊历史不同,不同学科的地区发展不同,不同学科的发展速度不同,所有这些导致不同学科期刊文献计量学指标表现不同,因此学术期刊不适合采用主成分分析法进行评价。

5.2 主成分结果回归系数为正是其方法适用的前提条件

采用主成分分析,除了同类可比及主成分分析自己提供的统计检验外,必须再加一个前提条件,那就是主成分分析结果与所有评价指标回归后的指标系数为正,不然很容易出现某个评价对象指标值增加,但是总排名反而会下降的异常结果。当然,在评价对象较少的情况下,考虑到自由度不够,是无法运用这种方法进行检验,但并不说明就可以采用主成分分析进行评价。

5.3 主成分分析结果回归可以用来进行评价指标的选择

如果评价对象被很好地分类,也通过了主成分分析适用条件的 KMO 检验和 Bartlett 检验,仍然出现某个指标的主成分结果回归系数为负数的情况,那么要对该指标的内涵进行深入分析,慎重选取该指标进行评价。主成分结果回归可以作为评价指标筛选的一种方法。

5.4 基于影响因素分析的主成分回归与主成分分析结果回归不一样

在一些研究中,如果因变量之间相关比较严重,有时为了消除多重共线性,先将自变量做主成分或因子分析,提取若干主成分(因子),然后用这些主成分(因子)与因变量进行回归,可以有效地消除多重共线性,并且克服数据较少的影响,显然,这与本文中的主成分分析结果与指标间的回归是不一样的。

5.5 因子分析法与主成分分析类似

由于因子分析法的原理与主成分分析类似,因此本文的结论同样适用于因子分析。

5.6 评价结果回归可广泛用于系统评价方法的检验

本文提供的方法可广泛应用于某些系统评价方法的检验,如 TOPSIS、突变理论、证据理论等。

参考文献

- [1] Weiping Yue, Concepcion S. Wilson. Measuring the citation impact of research journals in clinical neurology: a structural equation modeling analysis[J]. *Scientometrics*, 2004 (3): 317-334
- [2] 苏新宁. 构建人文社会科学学术期刊评价体系[J]. *东岳论丛*, 2008 (1): 35-42
- [3] 邱均平、张荣等. 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
- [4] 庞景安、张玉华等. 中国学术期刊综合评价指标体系的研究[J]. *中国学术期刊研究*, 2000(11): 217-219
- [5] 李凯扬、贾玉萍. 基于 AHP 的期刊全文数据库的模糊综合评价[J]. *情报科学*, 2005 (11): 1688-1703
- [6] 王小唯、杨波等. 学术期刊质量评估的二次相对评价方法[J]. *编辑学报*, 2003 (6): 231-232
- [7] 李修杰、陈景武. 运用判别分析法建立的期刊评估指标体系[J]. *江西图书馆学刊*, 2006 (3): 48-50
- [8] 王玖、徐天和. 秩和比法在医学学术期刊学术质量综合评价中的应用[J]. *数理医药学杂志* 2003 (3): 266-267
- [9] 陈汉忠. 主成分分析在学术期刊评价中的应用[J]. *中国学术期刊研究*, 2004 (6): 658-660
- [10] 王引斌. 测定核心期刊的新方法——主成分分析法[J]. *情报学报*, 1998, 17(5): 13.
- [11] 贺颖. 2001-2004 年中国管理类期刊学术影响力综合评价[J]. *中国软科学*, 2007 (1): 107-112
- [12] 管进, 陈文凯等. 外文核心期刊的综合评价——主成分分析法的应用[J]. *图书情报工作*, 2004(1): 13-16
- [13] 张弘、赵惠祥等. 基于主成分分析法的学术期刊评价方法[J]. *编辑学刊*, 2008 (2): 87-90
- [14] 杨文燕, 刘亚民等. 利用主成分分析法对中国肿瘤类期刊学术影响力的综合评价[J]. *中国肿瘤*, 2008 (1): 79-81
- [15] 周玲、张玲玲. 利用因子分析法对国内主要数学期刊进行评价[J]. *淮北煤炭师范学院学报* 2006 (6): 67-70
- [16] 凌春艳、莫琳. 自然科学学术期刊质量指标体系的属性数学综合评价模型[J]. *数学的实践与认识*, 2004 (5) :1-7
- [17] 李继晓、蔡成瑛. 对各种核心期刊评价方法的分析[J]. *中国学术期刊研究*, 2006 (2): 253-256