

单篇论著学术迹与影响矩比较研究^{*}

唐继瑞 叶 鹰

摘 要 本文对新近提出的学术迹和影响矩指标应用于单篇论著的评价效果进行比较研究,以 JASIST 2005—2014 年间发表的 25 篇高被引论文和 *ASTROPHYSICAL JOURNAL* 2011—2014 年间发表的 24 篇高被引论文为两个样本集,研究学术迹和影响矩相对于总被引数和 h 指数等学术评价指标的异同,发现各指标排序既具有一定相关性,也呈现出一定独立性,说明这些指标均有独立存在价值,而学术迹和影响矩能提供更全面的测度信息。图 3。表 6。参考文献 25。附录 2。

关键词 学术迹 影响矩 学术指标 学术评价 h 指数

分类号 G353.1

A Comparative Study of Academic Trace and Impact Torque on Single Publication

Jimmy R. TANG & Fred Y. YE

ABSTRACT

There have been many academic indicators, but few indicators can be applied to appraise single publication, in which counting total citations becomes the simplest approach. Recently, academic trace and impact torque have been proposed and applied, where academic subjects such as authors and institutes are considered, and academic objects particularly single publications are also concerned. In the article, authors compare and analyze the effects of academic trace and impact torque in the applications for single publications, on which further related studies are expected.

The authors focus on comparing academic trace and impact torque, with referring to total citation and citing h-index. The academic trace originates from mathematical trace of academic matrix and is determined by 5 parameters: total citing papers P , citing h-index h , total citing citations C , uncited citing papers P_z and citations in h-core C_h . The impact torque refers to physical torque, which is defined as the multiply of impact author-force A and impact time-span D . Two datasets 25 highly cited papers in *JASIST* 2005–2014 and 24 highly cited ones in *ASTROPHYSICAL JOURNAL* 2011–2014 are used to compare the effects of academic trace and impact torque and to reveal their correlations and differences.

It is concluded that academic trace and impact torque possess independences and correlations in comparing with pure citations and citing h-index. When the total citations and citing h-index cannot differentiate the impacts of single publications, academic trace and impact torque provide better differentiation. On overall effects, academic trace and impact torque work better than pure citations and citing h-index, and academic trace works better than impact torque.

Comparing with pure citations and citing h-index, the new indicators, academic trace and impact torque, display both in-

^{*} 本文系国家自然科学基金项目“专利 h 指数与专利信息网络测度研究”(编号:71173187)和江苏省数据工程与知识服务重点实验室建设资金资助的研究成果之一。(This article is an outcome of the project “Studies on the patent h-index and measures of patent information networks”(No. 71173187) supported by National Nature Science Foundation of China and the fund for Jiangsu Key Laboratory of Data Engineering and Knowledge Service.)

通信作者:叶鹰,Email:yye@nju.edu.cn (Correspondence should be addressed to Fred Y. Ye, Email: yye@nju.edu.cn)

dependences and correlations and thus supply rich measuring information and new appraising view for application. 3 figs. 6 tabs. 25 refs. 2 appendixes.

KEY WORDS

Academic trace. Impact torque. Academic indicator. Academic appraisalment. h-index.

0 引言

学术评价可针对学术主体(作者、机构、国家等)进行,也可针对学术客体(论著、期刊、专利等)展开。学术主体作为创造学术作品的个人或团体,固然是学术评价的对象,学术客体作为学术主体创作出来的作品,同样是学术评价的对象。任何伟大的创造发明,都必然要落实到具体的论著里。因此,针对单篇论著的学术评价具有重要意义,尤其是对那些作品不多但影响深远的作者,如果其单篇论著影响力明显,就不应忽视其对学术的贡献。

学术评价包括定性评价和定量评价,前者主要通过同行评议进行,后者则通过评价指标实现。已有评价指标众多,可概括为以影响因子等为代表的均值测度指标、以 h 指数等为代表的高影响特征测度指标和新近发展的整体综合测度指标^[1]。这些指标多用来衡量学术主体,或以期刊为主的学术客体,较少关注单篇论著的学术影响力,导致定量衡量单篇论著学术影响力基本采用最简单的方法即统计其被引次数。虽然近年来 Mendeley 学术社交网络平台的单篇论文阅读统计(reader counts)和英国创新型生物医学文献学术评价系统 F1000 因子^[2]开始进入国内学者研究视野,但原创性明显不足。

在国内外不断改进学术评价指标的背景下,学术迹^[3]和影响矩^[4]的概念先后被提出,这不仅对评价学术主体有意义,而且为单篇论著的评价增添了新的视角。本文对单篇论著的学术迹和影响矩进行比较,参照的基本指标是应用较广的被引数和被引 h 指数,期望能进一步推进相关研究。

1 发展脉络

评价学术客体最著名的指标莫过于 Garfield 提出的期刊影响因子(Impact Factor,简称 IF),其原始定义是^[5]:某刊 n 年度的影响因子等于该刊第 n 年度获得的 $n-1$ 和 $n-2$ 两年的总被引数除以该刊 $n-1$ 和 $n-2$ 两年的总发文数,用 C 代表被引数, P 代表发文数,写成公式是:

$$IF_n = \frac{C_{n-1} + C_{n-2}}{P_{n-1} + P_{n-2}} \quad (1)$$

此外也有五年影响因子(IF_5)或更一般的多年影响因子,当年影响因子则称当年指标或即年指数(Immediacy Index)。本质上,影响因子是篇均被引数 $C/P = CPP$,即一种期刊所发表论文的平均被引率。后来,荷兰莱顿大学科技研究中心(简称 CWTS)以 CPP 为核心构建了著名的特色指标“皇冠指数” CI (Crown Indicator)^[6-7],定义如下:

$$CI = \frac{CPP}{FCSm} = \frac{\sum_{i=1}^n (c_i/n)}{\sum_{i=1}^n (e_i/n)} = \frac{\sum_{i=1}^n c_i}{\sum_{i=1}^n e_i} \quad (2)$$

式(2)中 $FCSm$ 为所测领域的平均被引数, c_i 是论文 i 获得的引文数, e_i 是论文 i 所在领域的平均被引数, n 为论文总数。此外还有 $CPP/JCSm$ 等指标($JCSm$ 是所测期刊集发表论文的平均被引数)。

然而,2010 年以来,以 CPP 为核心的评价指标引发学术争议^[8-10],争议的焦点是数学平均在科学计量中是个逻辑不自恰的指标。为此,Lundberg 等提出了改良指标 $MNCS$ (mean normalized citation score)^[11],现正被瑞典著名科学机构 Karolinska Institute 使用:

$$MNCS = \frac{1}{n} \sum_{i=1}^n \frac{c_i}{e_i} \quad (3)$$

进一步的研究表明:指标 CI 和 $MNCS$ 的数学区别本质上只是平均率 (RoA) 和率平均 (AoR) 的差别, CI 是 RoA 而 $MNCS$ 是 AoR , 研究表明 AoR 仅略优于 RoA [12]。

另一指标是 h 指数 [13]; Hirsch 于 2005 年提出 h 指数, 将其作为针对学术主体的评价指标, 旋即产生巨大影响 [14]。 h 指数提出后不久, Braun 等著名信息计量学家即把原来针对学者的 h 指数概念推广应用于期刊 [15], 提出期刊的 h 指数指该期刊所发表的全部论文中至多有 h 篇论文被引用了 h 次。这就把 h 指数从评价学术主体拓展到了评价学术客体。

但期刊是论文集合, 一种期刊的 h 指数高并不是发表在其中的所有论文 h 指数都高, 就像一种期刊的影响因子高并不代表其中的所有论文影响力都高一样。为此, Schubert 等通过考察单篇论著的被引情况, 把被引文献中至多有 h 篇论文被引用了 h 次作为该单篇论著的 h 指数, 这样就能把 h 指数应用于评价单篇论著 [16-17]。被引 h 指数越高的论著, 影响越大。

具体操作是: 在特定数据库里, 将引用某篇论著的文献按其被引次数降序排列, 序次设为 r , TC_r 是文献 r 在该数据库里的被引总数, 则有以下序列:

$$r = (1, 2, \dots, r, \dots) \quad (4)$$

$$TC = (TC_1, TC_2, \dots, TC_r, \dots, TC_z);$$

$$TC_1 \geq TC_2 \geq \dots \geq TC_r \geq \dots \quad (5)$$

被引 h 指数即为:

$$h = \max \{ r; r \leq TC_r \} \quad (6)$$

可是, h 指数也被认为存在逻辑不自恰问题 [18], 尽管在动态情形下该问题并不存在 [19], 还是需要考虑改善。于是, 测度整体综合分布的评价指标 $I3$ 出现 [20], 引导出方法创新。

将 h 指数与 $I3$ 相结合, 可构成学术迹并启发产生影响矩。

2 方法与数据

2.1 方法介绍

本文主要涉及学术迹和影响矩两个评价指标, 由于是对单篇论著进行评价, 故其定义及公式变通如下。

2.1.1 学术迹

学术迹是学术矩阵 (Academic Matrices) 的迹, 而学术矩阵是引文曲线按发文的被引量排布构成的三阶矩阵。引文曲线以 h 指数与零被引为分界点, 即产生 h 核- h 尾分布 [21-22] 的概念。用 $P = P_e + P_t + P_z$ 代表发文总量, $C = C_e + C_t + C_z$ 代表引文总量, 其中 P_e 是 h 核中的论文量, P_t 是 h 尾中的论文量, P_z 是零引 (未被引) 论文量, $C_h = C_e + C_e = h^2 + e^2$ 是 h 核中的引文量, $C_t = t^2$ 是 h 尾中的引文量, $C_e = e^2$ 是超引区引文量, e 为 e 指数 [23], 其构成如图 1 所示。

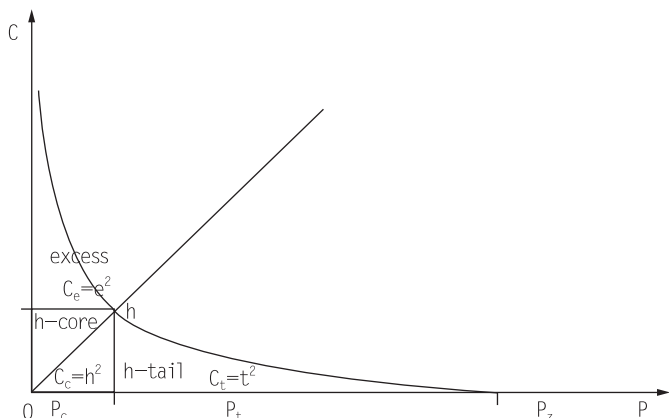


图 1 按发文被引量排布的引文曲线

结合 13 思想, 引进

$$I3X = x_c P_c + x_t P_t + x_z P_z = X_1 + X_2 + X_3; \\ x_c + x_t + x_z = 1 \quad (7)$$

$$I3Y = y_c C_c + y_t C_t + y_e C_e = Y_1 + Y_2 + Y_3; \\ y_c + y_t + y_e = 1 \quad (8)$$

于是可用两个向量 $X = (X_1, X_2, X_3)$ 、 $Y = (Y_1, Y_2, Y_3)$ 来标记以上所有发文量和引文量分布, 即

$$X = (X_1, X_2, X_3) = (P_c^2/P, P_t^2/P, P_z^2/P) \\ = \left(\frac{h^2}{P}, \frac{(P-h-P_z)^2}{P}, \frac{P_z^2}{P} \right) \quad (9)$$

$$Y = (Y_1, Y_2, Y_3) = (C_c^2/C, C_t^2/C, C_e^2/C) \\ = \left(\frac{h^4}{C}, \frac{(C-C_h)^2}{C}, \frac{(C_h-h^2)^2}{C} \right) \quad (10)$$

考虑到 $C-P$ 是一个自恰量^[24], 并考虑三阶矩阵对称性, 引进 Z 向量:

$$Z = (Z_1, Z_2, Z_3) \\ = (Y_1 - X_1, Y_2 - X_2, Y_3 - X_3) \quad (11)$$

构成如下学术矩阵^①:

$$V = \begin{pmatrix} Y_1 & Y_2 & Y_3 \\ X_1 & X_2 & X_3 \\ Z_1 & Z_2 & Z_3 \end{pmatrix} \\ = \begin{pmatrix} Y \\ X \\ Z \end{pmatrix} = (Y \ X \ Z)^T \quad (12)$$

该矩阵的迹就是学术迹(用于测量学术客体时也可称为影响迹):

$$T = tr(V) = Y_1 + X_2 + Z_3 \\ = \frac{C_c^2}{C} + \frac{P_t^2}{P} + \left(\frac{C_e^2}{C} - \frac{P_z^2}{P} \right) \quad (13)$$

由于 $C_c = h^2$, $P_c = h$, $P_t = P - P_c - P_z$, $C_e = C_h - C_c$, 上式可归并并:

$$T = tr(V) \\ = \frac{h^4 + (C_h - h^2)^2}{C} + \frac{(P - h - P_z)^2 - P_z^2}{P} \quad (14)$$

因此, 测算单篇论著学术迹需采集五个基

本参数: 单篇论著评价论文数 P (实际等于该篇论著的被引数), 被引 h 指数 h , 单篇论著评价引文数 C (即被引文献被引数), 评价论文中零被引论文数 P_z 和被引 h 核中评价引文数 C_h ^[25]。

2.1.2 影响矩

影响矩是类比物理概念力矩 (Torque) 建立的测度量, 通过考察单篇论著的影响规模和影响时效两个层面来测评其影响力。前者定义为影响学者力 A : 一篇论著的学术影响力 A 被定义为引用该论著的作者数; 后者定义为影响跨度, 其定义公式为:

$$D = \sum_{j=1}^n t_j \times \frac{P_j}{P} \quad (15)$$

其中, t_j 为论文从发表算起到第 j 年被引时经历的年数, P_j 为第 j 年的评价论文数 (被引数), P 为评价论文总数, n 为该文发表时间至今的总年数。

单篇论著的影响矩 M 定义为影响学者力 A 和影响跨度 D 的乘积:

$$M = A \times D = A \times \sum_{j=1}^n \left(t_j \times \frac{P_j}{P} \right) \\ = \frac{A}{P} \times \sum_{j=1}^n (t_j \times P_j) \quad (16)$$

2.2 数据来源

本文数据来自 ISI Web of Knowledge (WoK) 平台下 Web of Science (WoS) 数据库, 取两个样本。

样本一文献集选自信息科学和图书馆学 (Information Science & Library Science) 重要期刊 JASIST 2005—2014 年期间发表的被引次数最高的 25 篇学术论文 (文献类型为 article)。

样本二文献集选自天文学与天体物理学学科 (ASTRONOMY & ASTROPHYSICS) 中被引频次最高的期刊 ASTROPHYSICAL JOURNAL, 其影响因子与五年影响因子分别为 6.282 和 5.882 (据 Journal Citation Reports 2013 版), 样本文献

① 这里采用改良的 $(Y \ X \ Z)^T$ 排布, 与原始排布 $(X \ Y \ Z)^T$ 相比在数值上更合理。

为该期刊 2011—2014 年被引最高的 24 篇论文。
检索时间均为 2014 年 11 月。数据参见附录。

2.3 测算过程

以样本二中第 1 篇论文为例(记为样例 2-1,下文类同),其篇名为 A 3% solution; Determination of the Hubble constant with the Hubble space telescope and wide field camera 3,发表自

2011 年,截至统计时被引共 428 次(仅 WoK 中有记录的条目)。

关于学术迹,样例 2-1 的 5 个基本参数见表 1,在 WoS 数据库中各参数均可直接或间接查询得到。评价论文数 P 可直接查到;通过对 P 篇被引文献整体进行分析,可得被引文献零引论文数 P_z 、被引 h 指数 h 和评价引文数 C ;对 P 篇论文中 h 核论文分析得 h 核论文评价引文数 C_h 。

表 1 样例 2-1 的 5 个学术迹基本参数

评价论文数 P	零引论文数 P_z	被引 h 指数 h	评价引文数 C	h 核评价引文数 C_h
428	81	35	6 153	3 720

通过公式(9)(10)(11)即可得到该论文 X 、 Y 和 Z 向量,此时可得到其经改良的学术矩阵 V 与学术迹 T :

$$V_{2-1} = \begin{pmatrix} Y \\ X \\ Z \end{pmatrix}$$
$$= \begin{pmatrix} 243.885 & 962.049 & 1011.706 \\ 2.862 & 227.439 & 15.329 \\ 241.023 & 734.610 & 996.376 \end{pmatrix}$$
$$T_{2-1} = Y_1 + X_2 + Z_3 = 1467.700$$

关于影响矩,依据定义需要查询两个基本

参数。

影响学者力 A 在 WoS 数据库中通过分析功能可直接获得。对样例 2-1 的 428 篇被引文献进行作者分析,可得到有 2 563 位学者以样例 2-1 为研究基础,并有学术产出(仅指被收录于 WoS 中),此时样例 2-1 的影响学者力 A 为 2 563。需要指出的是,同一作者对于一篇论文的多次引用在 A 的测算中只记一次,且自引不计在内。影响跨度 D 则需要进行简单计算,样例的 428 篇被引文献随时间分布情况(WoS 数据库检索可直接获得)见表 2。

表 2 样例 2-1 的 428 篇被引文献时间分布

被引时间(年)	2011	2012	2013	2014
被引经历年限 t_j	1	2	3	4
评价论文 P_j	33	125	158	113

以论文发表时为起点,论文发表当年,历史年限为 $t_1=1$,依次类推至今。根据此测算规则,由公式(15)可得: $D_{2-1} \approx 2.818$,故样例 2-1 影响矩为: $M_{2-1} = A \times D = 7\ 223$ 。

3 结果

根据上述理论与测算方法,以总被引数 C 、被引 h 指数为参照,按公式(14)和(16)计算学

术迹 T 和影响矩 M ,样本一和样本二的结果分别如表 3 和表 4 所示。

表 3 和表 4 中 $h-r$ 、 $M-r$ 、 $T-r$ 分别为按照 h 、 T 、 M 值排序后所得排序号,值相同时取对应最高排名序号。

两个样本中各指标排序既显示出一定相关性,也显示出一定独立性,图 2 和图 3 分别是两样本按 M 降序排列的 T 和 M 值。

表 3 JASIST 25 篇高被引论文的学术评价参数

论文编号	P	h	T	M	$h-r$	$T-r$	$M-r$
1	241	26	417.39	2 314.33	2	5	3
2	230	18	488.13	3 619.83	9	3	1
3	204	18	250.27	2 429.30	9	13	2
4	153	22	410.66	934.28	4	6	12
5	142	13	127.04	1 862.64	15	17	4
6	147	26	622.25	1 199.70	2	2	6
7	134	29	867.73	990.34	1	1	7
8	133	22	361.27	971.97	4	7	8
9	89	9	56.13	1 274.47	24	25	5
10	97	16	261.90	951.13	13	12	11
11	88	17	357.22	956.40	12	8	10
12	84	19	217.05	548.00	7	14	18
13	83	21	420.13	449.35	5	4	22
14	75	10	125.42	969.88	20	18	9
15	71	18	293.38	819.69	9	10	13
16	72	10	96.22	516.43	20	21	20
17	67	12	122.78	679.88	17	19	16
18	70	12	174.72	798.78	17	15	14
19	57	10	71.16	517.07	20	24	19
20	69	15	272.43	407.89	14	11	23
21	66	6	80.60	659.04	25	23	17
22	65	10	82.68	183.74	20	22	25
23	66	19	331.54	320.68	7	9	24
24	63	11	108.01	722.52	19	20	15
25	65	13	148.80	491.08	15	16	21

表 4 ASTROPHYSICAL JOURNAL 24 篇高被引论文的学术评价参数

论文编号	P	h	T	M	$h-r$	$T-r$	$M-r$
1	428	35	1 467.7	7 223.00	2	1	2
2	382	36	758.17	4 294.91	1	4	4
3	377	18	316.09	8 634.97	20	19	1

续表

论文编号	<i>P</i>	<i>h</i>	<i>T</i>	<i>M</i>	<i>h-r</i>	<i>T-r</i>	<i>M-r</i>
4	313	17	210.86	2 541.42	22	24	9
5	270	25	531.13	3 049.06	10	10	6
6	231	26	523.78	4 802.13	8	11	3
7	210	31	725.16	2 416.55	4	5	10
8	210	34	929.01	2 239.02	3	2	14
9	179	28	572.24	2 559.93	6	9	8
10	165	15	297.49	4 148.00	24	20	5
11	156	30	632.18	2 931.29	5	7	7
12	147	17	693.77	2 293.72	22	6	11
13	143	28	590.90	1 741.09	6	8	17
14	140	22	352.55	2 115.44	16	16	15
15	135	24	323.02	771.35	12	18	24
16	132	18	220.04	1 349.96	20	22	22
17	130	26	424.96	1 184.99	8	14	23
18	129	25	346.04	2 267.80	10	17	12
19	125	19	215.27	1 365.79	17	23	21
20	122	24	784.53	1 624.27	12	3	20
21	114	23	388.80	1 636.77	15	15	19
22	115	19	274.68	1 673.63	17	21	18
23	113	24	445.04	1 842.26	12	13	16
24	112	19	449.30	2 249.63	17	12	13

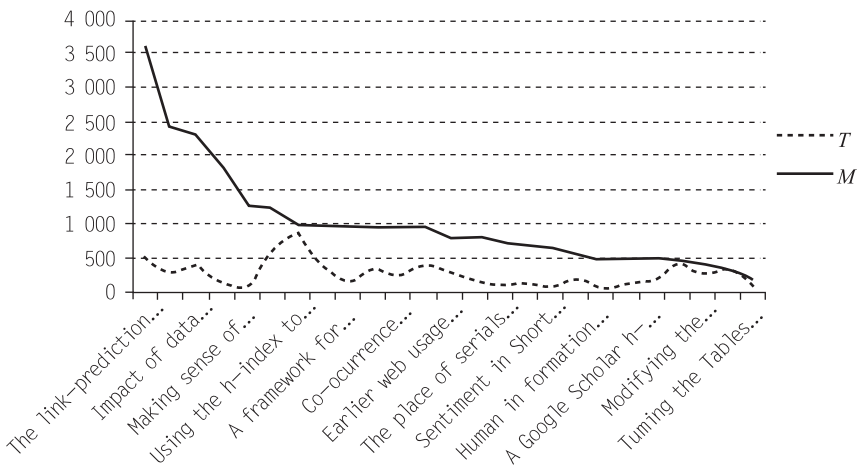


图2 JASIST 25篇高被引论文的*T*和*M*值(按*M*降序排列)

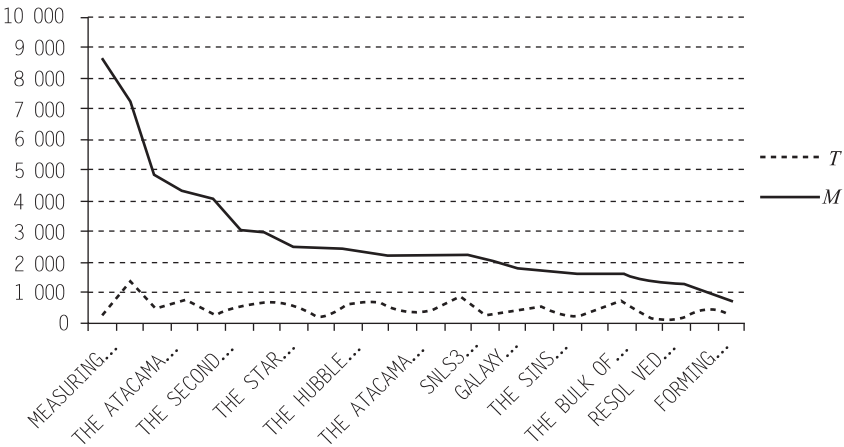


图3 ASTROPHYSICAL JOURNAL 24 篇高被引论文的 T 和 M 值(按 M 降序排列)

4 分析和讨论

由于引文曲线多为幂律分布, P 、 h 和 T 、 M 均不服从正态分布, 用 SPSS 对样本数据进行正态性检验时证实了这一点, 因而选择对样本集采用斯皮尔曼 (Spearman) 检验, 表 5 和表 6 分别为样本一和样本二的 P 、 h 、 T 、 M 的相关性(** 表示在 0.01 置信水平上相关性显著)。

4.1 学术迹和影响矩的独立性

理论和实证层次, 均易发现学术迹和影响矩的独立性。就学术迹和影响矩的构造而言, 相对于评价论文数 P (即被引数) 和 h 指数, T 和

M 均综合考虑了更多影响因素, 学术迹测算时考虑了引文曲线的整体分布, 用五个参数 P 、 h 、 C 、 C_h 和 P_z 覆盖了发文特征和引文特征; 影响矩中则用影响学者力 A 表征一篇论著所影响的并有学术产出的著者总数, 用影响跨度 D 表征被引文献随时间分布的一种加权。由于构造参数不同, 学术迹和影响矩是独立的。相关性检验提供了佐证。

被引 h 指数是单篇论著学术迹的重要构成参数, 故两样本中, 学术迹和被引 h 指数的相关系数为 0.929 和 0.741 的中高等水平(见表 5 和表 6), 而与其他指标间相关性很弱。与此同时, 影响矩的一个重要参数为影响学者力, 显然引用文献的著者总数受文献评价论文数影响, 因此

表 5 JASIST 样本学术评价指标相关性 (N=25)

Spearman		P	h	T	M
P	相关系数	1.000	0.625 **	0.633 **	0.818 **
	Sig. (双尾)	.	0.001	0.001	0.000
h	相关系数	0.625 **	1.000	0.929 **	0.308
	Sig. (双尾)	0.001	.	0.000	0.135
T	相关系数	0.633 **	0.929 **	1.000	0.338
	Sig. (双尾)	0.001	0.000	.	0.099
M	相关系数	0.818 **	0.308	0.338	1.000
	Sig. (双尾)	0.000	0.135	0.099	.

表 6 ASTROPHYSICAL JOURNAL 样本学术评价指标相关性 (N=24)

Spearman		P	h	T	M
P	相关系数	1.000	0.345	0.320	0.784 **
	Sig. (双尾)	.	0.099	0.127	0.000
h	相关系数	0.345	1.000	0.741 **	0.161
	Sig. (双尾)	0.099	.	0.000	0.453
T	相关系数	0.320	0.741 **	1.000	0.263
	Sig. (双尾)	0.127	0.000	.	0.214
M	相关系数	0.784 **	0.161	0.263	1.000
	Sig. (双尾)	0.000	0.453	0.214	.

两样本中影响矩和评价论文数相关性较强,相关性系数分别为 0.818 和 0.784。

学术迹和影响矩相比,由于构造思路完全不同,二者之间相关性很弱。由此可见,学术迹和影响矩在测度单篇论著影响力时是独立的,具有独立存在的价值和意义。

4.2 学术迹与影响矩的相似性

学术迹和影响矩虽然构造迥异,但在测评单篇论著的过程中,二者的测量对象均为单篇论著的被引文献(评价论文数 P 和被引 h 指数也是如此),由于测量角度不同,因而反映了不同信息。

但应该注意到,学术迹和影响矩比评价论文数等具有更高的区分度。以具体样本为例,样例 1-15 和 1-16 评价论文数分别为 71 和 72,评价论文数 P 几乎不具有区分度,而学术迹和影响矩由于考虑了其他因素(如引文质量、影响学者人数、影响时效等),明显地区分了两篇论文的学术影响力。二者在评价论文数、零引论文数均非常接近的前提下,由于样例 1-15 具有更高的被引 h 指数,因此学术迹 $T_{1-15} > T_{1-16}$;与此同时,由于样例 1-15 具有更久的影响跨度,因此影响矩 $M_{1-15} > M_{1-16}$ 。

论文 2-4 和 2-12 则是论文 h 指数不具有可区分度的情形。样例 2-4 的被引文献中,零引论文数明显大于样例 2-12(零引的比例亦

高),故 $T_{2-4} < T_{2-12}$ 。而 2-5 与 2-15 则是 h 无法区分、 M 可以区分的典型。诸如此类的例子不胜枚举。

因此,从相似性的角度来看,在测量单篇论著的学术影响力时,学术迹和影响矩具有相同的测量对象,且比已有指标具有更高的区分度。

4.3 学术迹和影响矩的适用性

从实用角度而言,学术迹和影响矩的测算较评价论文数和被引 h 指数复杂。学术迹的五个基本参数,在带有引文分析功能的数据库中均可直接或间接查到;依学术迹的算法公式编程,可较为便捷地获得一篇论著的学术迹。影响矩的基本参数获取要相对复杂,影响学者力需要通过分析获得,而影响跨度则需要首先采集论著被引文献每年的分布情况,再按公式予以测算。总的来说, WoK 平台下收录的所有西文数据库以及中国科学引文索引数据库(CSCD)收录的中文科学论著,均可通过本文所述方法测算其学术迹和影响矩;而在其他中文引文索引数据库或全文库中,由于引文分析功能不如 WoK 平台全面,相关基本参数的获取需要手工计算或借助其他工具进行分析才可获得(如影响学者力的采集需要对所有被引文献进行作者分析)。

指标的复杂程度与其提供测度信息的多寡是密不可分的。简单指标(如被引数 P 等)测算

虽然简单,但提供的测度信息很少;复杂指标(如本文研究的学术迹和影响矩)虽然测算复杂,但提供的测度信息却很丰富。可以设想随着各种指标的推广,必然会提升各类数据库对于数据的揭示层次和分析深度;而各种指标在此基础上也可进一步优化完善,二者形成良性互动。

5 小结

综上所述,在测量单篇论著的学术影响力时,学术迹综合考察单篇论著的被引文献数量与质量;影响矩则考察单篇论著影响的作者规模与影响时效。作为新近发展的学术评价指标,与纯粹被引数和h指数等相比,学术迹和影

响矩体现出作为独立指标存在的必要性,也表现了与已有指标的相关性,二者均提供了更加全面和丰富的测度信息,具有更加综合的测度意义。此外,学术迹和影响矩均在很大程度上提高了影响力的可区分度,可辅助对单篇论著的影响力进行综合判断。虽然目前学术迹和影响矩的测算相对较为复杂,但随着文献引文信息分析和揭示的进一步深化及程序化,二者的测算会相应简化。

总之,学术迹和影响矩为单篇论著的影响力测评提供了新颖独特的研究视角,可望为后续研究提供参考。建议在实用过程中同时采用多种学术评价指标,当所有指标指向一致时,获得的评判应当具有决定意义。

参考文献

- [1] 叶鹰. 国际学术评价指标研究现状及发展综述[J]. 情报学报, 2014, 33(2): 215-224. (Ye F Y. The research progress and developing perspective of assessment indicators[J]. Journal of the China Society for Scientific and Technical Information, 2014, 33(2): 215-224.)
- [2] 宋丽萍, 王建芳, 王树义. 科学评价视角下 F1000、Mendeley 与传统文献计量指标的比较[J]. 中国图书馆学报, 2014, 40(4): 48-54. (Song Liping, Wang Jianfang, Wang Shuyi. Observation of F1000, Mendeley and traditional bibliometric indicators from the perspective of scientific assessment[J]. Journal of Library Science in China, 2014, 40(4): 48-54.)
- [3] Ye F Y, Leydesdorff L. The "academic trace" of the performance matrix: a mathematical synthesis of the h-index and the Integrated Impact Indicator (I3)[J]. Journal of the Association for Information Science and Technology, 2014, 65(4): 742-750.
- [4] 王术, 叶鹰. 影响矩作为测度单篇论著影响力的评价指标探讨[J]. 大学图书馆学报, 2014, 32(5): 12-18. (Wang Shu, Ye F Y. Impact Torque as an assessment indicator for measuring single publications[J]. Journal of Academic Libraries, 2014, 32(5): 12-18.)
- [5] Garfield E. Citation indexing: its theory and applications in science, technology, and humanities [M]. New York: Wiley, 1979.
- [6] Moed H F, de Bruin R E, van Leeuwen T N. New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications [J]. Scientometrics, 1995, 33(3): 381-422.
- [7] Van Raan A F J. Statistical properties of bibliometric indicators: research group indicator distributions and correlations [J]. Journal of the American Society for Information Science and Technology, 2005, 57(3): 408-430.
- [8] Opthof T, Leydesdorff L. Caveats for the journal and field normalizations in the CWTS ("Leiden") evaluations of research performance [J]. Journal of Informetrics, 2010, 4(3): 423-430.
- [9] Bornmann L. Towards an ideal method of measuring research performance: some comments to the Opthof and Leydesdorff (2010) paper [J]. Journal of Informetrics, 2010, 4(3): 441-443.

- [10] Van Raan A F J, van Leeuwen T N, Visser M S, et al. Rivals for the crown: reply to Ophhof and Leydesdorff [J]. Journal of Informetrics, 2010, 4(3): 431-435.
- [11] Lundberg J. Lifting the crown—citation z-score [J]. Journal of Informetrics, 2007, 1(2): 145-154.
- [12] Egghe L. Averages of ratios compared to ratios of averages; mathematical results [J]. Journal of Informetrics, 2012, 6: 307-317.
- [13] Hirsch J. An index to quantify an individual's scientific research output [J]. Proceedings of the National Academy of Sciences of the USA, 2005, 102(46): 6569-16572.
- [14] Ball P. Index to aims for fair ranking of scientists[J]. Nature, 2005, 436(7053): 900.
- [15] Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals [J]. Scientometrics, 2006, 69(1): 169-173.
- [16] Schubert A. Using the h-index for assessing single publications [J]. Scientometrics, 2009, 78(3): 559-565.
- [17] 杨建林, 严明. 单篇期刊论文 h 指数的实证研究——以图书情报学为例[J]. 图书情报工作, 2010, 54(12): 145-148. (Yang Jianlin, Yan Ming. Empirical research on Schubert's h-index of a single paper—taking the Library and Information Science as examples[J]. Library and Information Service, 2010, 54(12): 145-148.)
- [18] Waltman L, van Eck N J. The inconsistency of the h-index [J]. Journal of the American Society for Information Science and Technology, 2012, 63(2): 406-415.
- [19] Ye F Y. H-inconsistency is not an issue in dynamical systems[J]. ISSI Newsletter, 2012, 8(2): 22-24.
- [20] Leydesdorff L, Bornmann L. Integrated impact indicators compared with impact factors: an alternative research design with policy implications [J]. Journal of the American Society for Information Science and Technology, 2011, 62(11): 2133-2146.
- [21] Ye F Y, Rousseau R. Probing the h-core: an investigation of the tail-core ratio for rank distributions[J]. Scientometrics, 2010, 84(2): 431-439.
- [22] Chen D Z, Huang M H, Ye F Y. A probe into dynamic measures for h-core and h-tail[J]. Journal of Informetrics, 2013, 7(1): 129-137.
- [23] Zhang C T. The e-index, complementing the h-index for excess citations [J]. PLoS ONE, 2009, 4(5): e5429.
- [24] Rousseau R, Ye F Y. A simple impact measure and its evolution over time [J]. Journal of Library and Information Studies, 2011, 9(2): 1-10.
- [25] 薛霏, 叶鹰. 大学的学术矩阵和学术迹探讨[J]. 大学图书馆学报, 2014, 32(1): 25-29. (Xue Fei, Ye F Y. The academic matrices and traces of universities[J]. Journal of Academic Libraries, 2014, 32(1): 25-29.)

附录 1 JASIST 25 篇高被引论文数据

论文编号	P	$P_c = h$	P_z	C	C_h	A	D
1	241	26	42	2 509	1 221	461	5.02
2	230	18	107	1 550	1 133	605	5.98
3	204	18	85	1 252	760	420	5.78
4	153	22	38	1 688	1 100	253	3.69
5	142	13	75	613	416	386	4.83
6	147	26	22	2 453	1 632	267	4.49
7	134	29	14	3 057	2 168	214	4.63
8	133	22	18	1 618	983	236	4.12
9	89	9	42	343	199	207	6.16

续表

论文编号	P	$P_c = h$	P_z	C	C_h	A	D
10	97	16	25	959	657	178	5.34
11	88	17	20	1 013	792	190	5.03
12	84	19	20	893	576	116	4.72
13	83	21	17	1 371	1 035	112	4.01
14	75	10	32	431	309	160	6.06
15	71	18	20	898	711	144	5.69
16	72	10	19	403	243	136	3.80
17	67	12	25	460	328	147	4.63
18	70	12	21	590	417	158	5.06
19	57	10	21	281	194	104	4.97
20	69	15	15	905	648	86	4.74
21	66	6	37	233	179	166	3.97
22	65	10	13	358	204	67	2.74
23	66	19	12	1 039	805	85	3.77
24	63	11	21	397	279	113	6.39
25	65	13	18	519	374	140	3.51

附录 2 *ASTROPHYSICAL JOURNAL* 24 篇高被引论文数据

论文编号	P	$P_c = h$	P_z	C	C_h	A	D
1	428	35	81	6 153	3 720	2 563	2.82
2	382	36	51	5 576	2 443	1 681	2.55
3	377	18	128	1 859	872	2 508	3.44
4	313	17	86	1 730	569	1 188	2.14
5	270	25	44	3 066	1 520	1 114	2.74
6	231	26	21	3 120	1 528	1 906	2.52
7	210	31	29	3 430	2 061	866	2.79
8	210	34	28	4 165	2 610	913	2.45
9	179	28	31	2 693	1 635	847	3.02
10	165	15	47	1 187	717	1 411	2.94
11	156	30	21	2 659	1 731	1 023	2.87
12	147	17	26	1 973	1 361	830	2.76
13	143	28	17	2 612	1 655	624	2.79
14	140	22	23	1 776	1 017	746	2.84

续表

论文编号	P	$P_e = h$	P_z	C	C_h	A	D
15	135	24	30	1 670	947	282	2. 74
16	132	18	24	1 207	627	511	2. 64
17	130	26	23	1 857	1 172	413	2. 87
18	129	25	19	1 804	996	754	3. 01
19	125	19	25	1 306	659	492	2. 78
20	122	24	18	2 484	1 798	701	2. 32
21	114	23	16	1 641	1 059	635	2. 58
22	115	19	29	1 259	780	516	3. 24
23	113	24	18	1 617	1 142	642	2. 87
24	112	19	19	1 769	1 125	857	2. 63

唐继瑞

南京大学信息管理学院硕士研究生。通信地址:南京市仙林大道 163 号。邮编:210023。

叶 鹰

南京大学信息管理学院教授,博士生导师。通信地址同上。

(收稿日期:2014-12-21;修回日期:2015-01-06)