

现代情报
Journal of Modern Information
ISSN 1008-0821, CN 22-1182/G3

《现代情报》网络首发论文

题目： 基于轨迹聚类分析的中国人文领域核心学者成长模式研究
作者： 彭贤哲，张瑜，孟德杰，石进
网络首发日期： 2024-09-19
引用格式： 彭贤哲，张瑜，孟德杰，石进. 基于轨迹聚类分析的中国人文领域核心学者成长模式研究[J/OL]. 现代情报.
<https://link.cnki.net/urlid/22.1182.G3.20240919.0837.002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于轨迹聚类分析的中国人文领域核心学者成长模式研究

A Study on the Growth Pattern of Core Scholars in the Chinese Humanities Field Based on Trajectory Cluster Analysis

彭贤哲 张瑜 孟德杰 石进*
Peng Xianzhe Zhang Yu Meng Dejie Shi Jin*

(南京大学信息管理学院, 江苏 南京, 210023)

(School of Information Management, Nanjing University, Nanjing 210023, China)

摘要：[目的/意义] 明晰中国人文领域核心学者的成长模式，有利于为该领域人才培养奖励政策的制定或调整提供切实的参考依据，从人才发展角度出发，促进中国新文科建设的落地与推进。[方法/过程] 为挖掘中国人文领域核心学者潜在共同的成长模式，本文以回顾的视角构建他们的科研生产力、学术影响力和综合能力的时序变化记录，通过动态时间规整和轨迹聚类算法，展现了他们在3个方面普遍的学术成长模式。[结果/结论] 结果表明，中国人文领域核心学者的科研生产力的成长模式具有多样性，哲学和文学学科的核心学者的学术影响力、综合能力的成长发生在后期，但历史学学科却集中在早期阶段，3个方面学术能力的成长通常表现出同步效应。由此得出的学者成长模式的阶段性特征、学科性差异以及组合规律，有助于拟定中国人文领域人才的评价方式、评价周期、评估机制。

关键词：轨迹聚类分析；中国人文领域；核心学者；成长模式；CSSCI

分类号：G353

Abstract: [Purpose/Significance] Clarifying the growth pattern of core scholars in the Chinese humanities is conducive to providing practical references for the formulation or adjustment of Chinese policies on cultivating and rewarding talents in this field, and promoting the construction of Chinese new liberal arts from the perspective of talent development. [Method/Process] In order to excavate the potentially common growth patterns of core scholars in the Chinese humanities, this study constructed a record of the temporal changes in their research productivity, academic influence, and comprehensive competence from a retrospective perspective, and demonstrated their generalized academic growth patterns in three aspects through Dynamic Time Warping and trajectory clustering algorithms. [Results/Conclusion] The results indicate that the growth patterns of research productivity among core scholars in Chinese humanities field are diverse. The academic influence and comprehensive capabilities of core scholars in philosophy and literature tend to develop in the later stages, but the discipline of history tends to concentrate in the early stages. The growth of academic capabilities in these three aspects typically shows a synchronicity effect. The phased characteristics, disciplinary differences, and combination rules of scholars' growth patterns provide guidance for the formulation of evaluation methods, evaluation cycles, and evaluation mechanisms for talents in this field.

Key words: trajectory clustering algorithms; Chinese humanities field; core scholars; growth pattern; CSSCI

2021年，《关于学习贯彻习近平总书记给〈文史哲〉编辑部全体编辑人员重要回信精神

基金项目：国家社科基金一般项目“面向国家安全的科技情报态势感知研究”，项目编号 21BTQ012。

作者简介：彭贤哲（1995-），男，博士研究生，研究方向：科学计量、大数据分析技术。张瑜（1999-），女，硕士研究生，研究方向：科技情报、大数据分析与技术。孟德杰（1999-），男，硕士，研究方向：文献计量、科学评价。

通信作者：石进（1976-），男，博士，教授，博士生导师，研究方向：大数据分析与技术、情报学。

的通知》明确指出，当前需要大力推进新文科建设，全面提升高校文科人才培养质量与核心竞争力^[1]，由此表明了人才培养在新文科建设过程中的核心地位。然而，长久以来，关于人文领域人才培养的研究多停留在理论层面或个别案例研究，或缺乏实证数据支撑而不具有可操作性，或数据范围有限而无法推广。已有的从人才培养角度展开的实证分析研究也多关注于自然科学、工程技术领域的学者，即使有涉及人文领域的学者，也并非主体研究对象。为此，本研究将研究对象锁定为中国人文领域的核心学者，通过回顾他们在不同时间阶段的学术能力变化情况，分析他们是否具有普遍的特定成长模式，具体在不同学科又有何差别性，以及不同方面的成长模式如何组合出现，在细节微观层面展现学者在各个方面的详细成长过程，进而为该领域人才的培育奖励政策提供参考。

具体而言，本研究将依托中国社会科学引文索引库（Chinese Social Sciences Citation Index，简称 CSSCI），以哲学、历史学、文学 3 个人文领域的学科的核心学者为例，构建他们在科研生产力、学术影响力和综合能力的时序变化记录，即学术成长轨迹。通过动态时间规整算法计算不同学者之间成长轨迹的相似性和最佳匹配路径，之后采用轨迹聚类分析算法识别轨迹的形态变化类别，观察他们的学术能力成长过程是否具有共性特征，不同学科之间的成长模式是否具有差异性，以及各个方面的成长模式的共现特征。至此，可从不同类别的学术成长轨迹的形态变化、学科分布、组合情况，深入了解中国人文领域的核心学者的成长模式，可从人才成长发展的视角为该领域学者的培养奖励提供参考依据。

1 相关研究

关于学者成长过程的研究涉及多个方面的内容，学者们基于不同的角度对此展开了探讨，并总结得出了有关该过程的多种表现模式。

1.1 学者成长过程的研究内容

学者的顺利成长虽然直观表现为工作职位的高升、荣誉奖项的获得等，但最终均脱离不开科研生产力或影响力的提升^[2]。为此，现已有不少研究从文献计量学的角度出发，衡量他们的生产力、影响力等计量指标的变化，展现了他们的学术能力的成长过程。

由于学者的科研生产力可由发文量直接表示，从文献计量学角度切入科技人才成长规律的早期研究多聚焦于分析他们的生产力的变化特征，如 Horner K 等认为科学家生产力的变化有一定的规律可循，他们的创造力在职业生涯中存在巅峰时期^[3]。具体来看，大部分研究均表明科研人员的生产力的变化模式一般都是迅速上升到峰值，然后缓慢下降^[3-7]，但 Way

S F 等以计算机科学领域美国和加拿大的教师信息为研究对象, 结果发现仅五分之一的教师的生产力变化表现出“快速崛起, 逐渐衰落”的一般模式, 其余五分之四表现出丰富多样的生产力变化模式^[8]。此外, 不同科研人员生产力变化的具体细节特征仍然存在一定的区别, 如 Györfy B 等通过分析不同学科科研人员的学术表现与年龄的相关性, 结果表明大多数学者需要几十年才能达到生产力的顶峰, 而生产力最佳表现期的时间长度因学科而异^[9]; Yair G 等发现大多数科学家的发文量峰值仅有一个, 少数科学家有几个高峰阶段^[10]。

另有学者主要应用科研人员的引文量、h 指数等计量评价指标的变化, 用以反映科研人员的影响力变化。其中, 科研人员的 h 指数变化的时间序列, 作为量化科学家职业生涯发展的一种方式而受到广泛关注^[11-13]。Liang L M 首先提出 h 指数时间变化序列的概念, 但其认为 h 指数的时间变化序列应按从晚到早的顺序排列^[14], 这与惯常更合乎逻辑的由早到晚的排列方式相反, 这一点在之后的研究中也得到了体现^[15, 16]。基于此, 另有学者展开了对科研人员的 h 指数时间序列的形态学的研究^[17, 18], 展现了他们的学术影响力在不同阶段的状态变化。虽然另有一些诸如 p 指数^[19]、z 指数^[20]的计量指标可用于科研人员学术能力的评估, 并综合利用各类指标用于审视某一领域学者的成长过程^[21], 但关于这些指标变化的时间序列的研究相对较少。

学者的成长作为一个动态变化的过程, 与其相关的评价衡量指标在不同时间节点也通常是存在变化的。为此, 在有关于此的多项研究中往往可以借由他们的计量评价指标的时间变化序列表示^[22], 进而通过时间序列分析方法剖析成长过程的细节特征。如 Symonds M R E 等构建生命科学领域 168 名科研人员的发文量时间序列, 发现性别的影响在他们的学术生涯的早期就已形成^[23]; 郭新艳依据某个高校的科研人员的发文量时间序列的变化形态, 将高校科研人员的成长过程分为 4 个阶段^[24]; Sinatra R 等通过分析 2 887 位物理学家发表的出版物的时间序列数据, 结果表明他们发表最高影响力的论文的时间在他们的职业生涯中是随机分布的^[25]; Way S F 等综合分段线性回归模型和时间序列分析方法, 用以分析获得终身教职前后科研人员产出时间序列的变化情况^[8]。

1.2 学者成长过程的表现模式

不同学者的成长过程虽然难免存在个体差异性, 但通过分析不同个体在不同大环境下的成长特点后发现, 他们的成长过程具备群体共性特征^[26], 这说明学者的成长遵循一系列共同的规律, 表现为师承效应规律、扬长避短规律、最佳年龄规律、马太效应规律、期望效应规

律、共生效应规律、积累效应规律、综合效应规律、时势造就规律、优势积累规律、竞争选择规律、实践成才规律、过程转换规律、协同合作规律、环境磨练规律、范式效应规律、自我修炼规律等^[27-31]。这些规律性现象在现有研究中均有体现，如李祖超等^[32]、宋晓欣等^[33]、蔡翔等^[34]分别对国家科学技术奖得主、长江学者和院士进行分析，发现他们多师从名师，这极大促进了他们的成长速度，由此论证了师承效应规律的存在；科技人才成长过程中的最佳年龄规律，体现在诺贝尔自然科学奖获得者取得获奖成果的年龄和中国科学院新增院士与诺贝尔奖得主平均年龄的集中分布现象中^[35,36]；在积累效应规律（或称马太效应）方面，中国科学精英成长存在显著的“光环叠加”效应，就读名校、师承名师、初始任职机构、职业流动、多重身份等因素均对科学精英成长有重要影响^[37-39]。

学者的成长过程并非一蹴而就，在不同阶段存在独立的特征，为此，现有研究根据人才成长过程的阶段性差异，拟定了关于科技人才成长路径的分段分期方案。如中国科学院人力资源管理研究组从科技创新人才年龄和能力出发，提出了科技人才“三波段理论”，将科技人才的成长过程划分为35岁以下、36~55岁、55岁以上3个阶段^[40]；刘少雪从科技领军人才能力形成过程出发，将科技领军人才成长分为基本素质养成、专业能力形成、创新能力激发和领军人才完型4个阶段^[41]；林崇德等通过分析拔尖创新人才的深度访谈资料，将创造性人才的成长过程分为自我探索期、集中训练期、才华展露与领域定向期、创造期、创造后期5个阶段^[42]；曹晓丽等综合考虑科技人才成长关键事件、成长周期、成长推动因素，将创新型科技人才成长过程划分为成长基础期、成长上升期和成长成熟期3个阶段^[43]；瞿群臻等依据人才能力成长的速度和表现不同，将科技创新人才的成长过程分为4个阶段，包括累积阶段、快速成长阶段、成熟阶段和衰退阶段^[44]。

总结现有关于学者成长过程的研究，一者在于多关注自然科学和工程技术领域的学者，且多是其领域的领军人物，而忽视了更为广泛的学者群体的成长模式，针对人文领域的学者的研究更是匮乏；二者在于仅聚焦于某一指标的变化，对学者的成长过程缺乏一定的系统性把控，使用的时间序列分析方法仍有提升优化的空间。为此，本项研究将视角移向中国人文领域的核心学者，引入更适于研究他们成长模式的动态时间规整和轨迹聚类分析方法，综合分析多个指标的时间变化序列，用以挖掘他们成长过程的各项特征以及这些特征的分布情况，为该领域人才培育激励政策的优化完善提供实证参考。

2 研究设计

2.1 数据来源

CSSCI 作为由南京大学投资建设、南京大学中国社会科学研究评价中心开发研制的人文社会科学引文数据库,过滤收录了人文社会科学领域的代表性中文期刊论文的题录和引文数据,相比国内其他数据库更具针对性、权威性和完整性^[45],得到了学界的广泛认可和应用。根据教育部 2022 年发布的《研究生教育学科专业目录》,以学科大类为单位,可将人文领域锁定为哲学、历史学、文学 3 个学科。对于这 3 个学科而言,其大多集中于在中文期刊发表论文,CSSCI 数据库基本完整收录有上述学科的学者的代表性文献成果记录。为此,本研究将研究对象锁定在人文领域的学者群体,该数据库收录的学者成果信息适用于分析人文领域学者在学术上的成长过程。

本研究在不考虑同一学者使用多个姓名的情况下,初步假定以“机构+姓名”的方式确定唯一学者,从 CSSCI 数据库统计同名不同机构的学者的合作者列表,以两个合作者列表中的重合学者数量至少超过 3 个为合并条件,可将不同机构但相同姓名的学者视为同一个人,这一步的合并条件主要是通过人工随机抽样实验的方式确定的,即从数据库中随机抽取 100 个相同姓名的样本进行人工审查,以他们至少具有 3 个同机构同姓名的合作者为合并条件,可保证在不同机构工作的同一学者可被正确地视为同一人。基于上述步骤,通过人工随机抽检进一步校核数据,基本确保了同名但不同机构的同一学者的准确融合,至此,本研究从 CSSCI 数据库共计获取人文领域 168 116 位学者的 547 367 条文献成果信息。其中,哲学学科包括有 32 295 位学者发表的 102 702 条文献,历史学学科包括有 34 725 位学者发表的 100 084 条文献,文学学科包括有 101 096 位学者发表的 344 581 条文献。由于并非每一位发表文献的学者均会在该领域进行持续长足的学术研究,特别是大量的学生群体,为此,本研究应用洛特卡定律,将研究对象进一步聚焦到人文领域 3 个学科的核心学者,再者缘于关注的是学者较为完整的成长过程,以他们具备至少 20 年的职业生涯跨度为限制条件^[46],获得具有较为稳定且完整的学术成长路径的核心学者群体。

经统计,哲学、历史学、文学 3 个学科学者的最大发文量依次为 289、176 和 229,如图 1 所示,核心作者发表论文数量的阈值分别为 12.733、9.937 和 11.334,由此筛选获得的核心学者数量分别为 1 399、1 896、5 375,进一步限制核心学者需至少具备 20 年的职业生涯,最终在 3 个学科确定的核心学者数量分别为 702、801 和 2 188,共计筛选出 3 691 名学

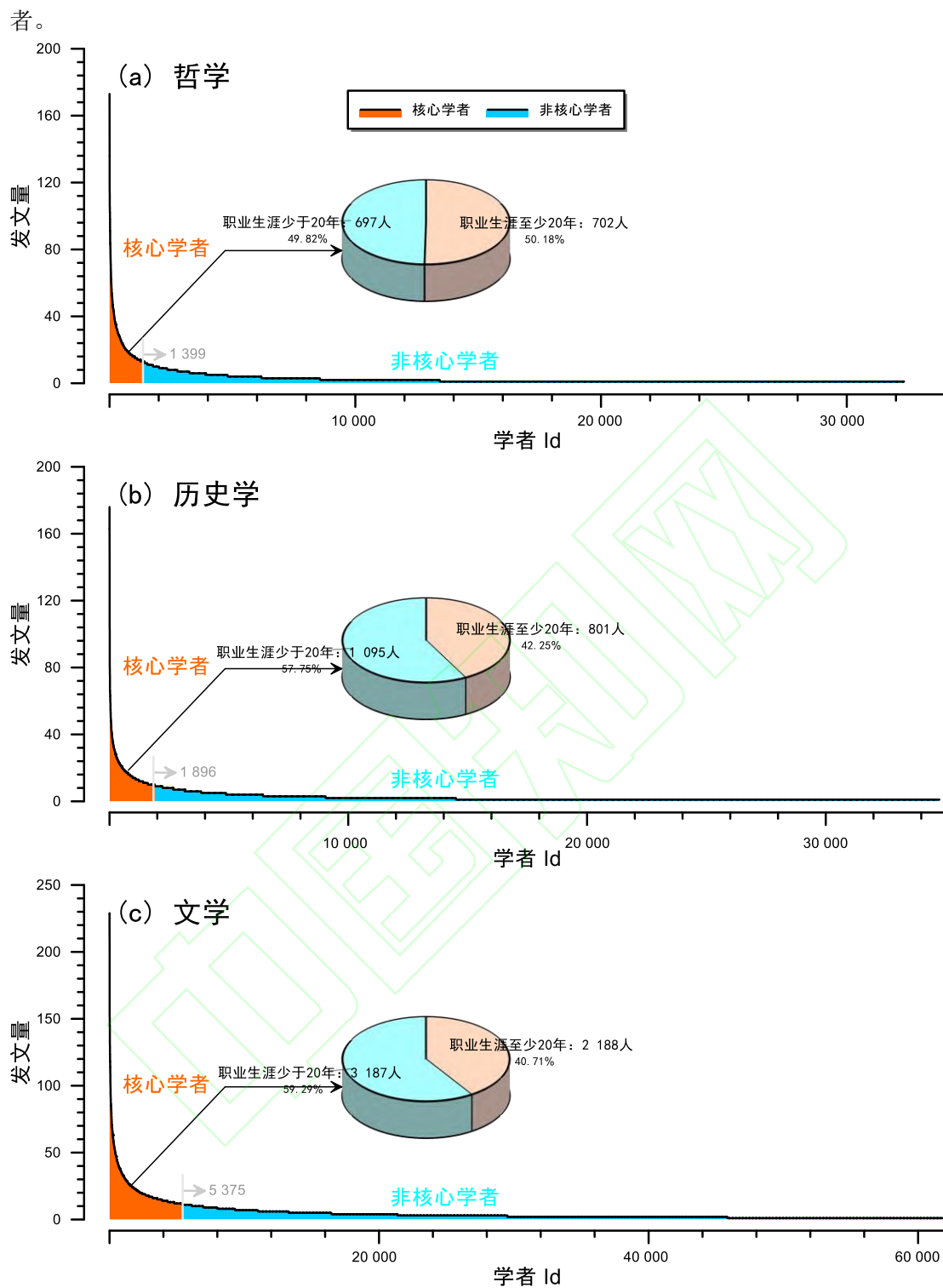


图 1 人文领域不同学科的核心学者的筛选过程

Fig.1 Selection of Core Scholars from Different Disciplines in the Field of Humanities

2.2 研究方法

参照 Peng X Z 等的研究^[21]，在学者成长过程的表示模型中，可通过计算学者在每个时刻近 5 年内计算的发文量、h 指数和 p 指数，并结合调谐算法评估学者在每篇文献中的贡献^[47]，进而对计量指标进行调整，由此构成各项指标的时序变化记录，分别表示他们在科研究生

产力、学术影响力和综合能力 3 个方面的学术能力的成长轨迹。

本研究关注的是成长轨迹的变化形态,其重点不在于分析比较不同学者的学术能力大小,而在于探析不同学者的成长过程是否具有类似的变化模式。为此,在计算成长轨迹之间的相似性之前,需要将不同学者的计量指标进行均一化处理,避免因他们学术能力的上限不同而影响轨迹之间相似性的计算,均一化处理可通过公式(1)实现,具体如下:

$$Index_{norm} = \frac{Index - Index_{min}}{Index_{max} - Index_{min}} \quad (1)$$

其中, $Index$ 指原始的计量指标, $Index_{norm}$ 指均一化之后的计量指标, $Index_{min}$ 为在不同时间窗口下计算的最小计量指标, $Index_{max}$ 为在不同时间窗口下计算的最大计量指标。

在将各类指标时间序列归一化至同一变化范围后,需要通过动态时间规整方法计算他们的学术能力衡量指标的时序曲线的距离和匹配路径,既可以构建他们的学术成长轨迹之间的距离矩阵,又可对齐不同时间跨度的时间序列。之后,可采用 K-means 聚类分析和轨迹聚类分析方法,确定时间序列的最优聚类个数,由此呈现学者在 3 个方面的学术能力的潜在成长模式。

2.2.1 动态时间规整

动态时间规整(Dynamic Time Wrap, 简称 DTW)方法可计算轨迹之间的相似性。DTW 方法作为用来测量两个点之间距离的一种办法,对于 n 维空间中的两点 $x = (x_1, x_2, \dots, x_n)$ 和 $y = (y_1, y_2, \dots, y_n)$, 可以通过欧氏距离计算它们之间的距离:

$$dist(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \quad (2)$$

但是,如果 x 的长度与 y 不等,即无法使用上述公式测量其距离。DTW 方法的目标则是利用动态规划(Dynamic Programming, DP)来找到点与点之间最好的匹配方式,继而获得点与点之间的距离。这个方法之所以被称为时间规整,主要因为在一般情况下 x 和 y 是时间序列向量,需要将其压缩或者伸展用以获得最好的匹配方法^[48]。

具体而言,给定两个时间序列 t 和 r , 长度分别为 m 和 n , 即代表了时间序列的跨度, DTW 方法的目标是获得一个最佳匹配路径 $\{(p_1, q_1), (p_2, q_2), \dots, (p_k, q_k)\}$ 使得匹配路径上的距离和 $\sum_{i=1}^k |t(p_i) - r(q_i)|$ 最小。约束条件如下:

(1) 端点条件: $(p_1, q_1) = (t(1), r(1))$, $(p_k, q_k) = (t(m), r(n))$, 确保时间序列之间是“头对头、尾对尾”的匹配方式;

(2) 局部条件: 对路径中的任意点 $(t(i), r(j))$, 其下一步只有 3 种可能: $(t(i+1), r(j))$,

$(t(i), r(j+1))$ 和 $(t(i+1), r(j+1))$ 。该条件保证了时间序列向量的匹配点 $t(i)$ 和 $r(j)$ 是单调递增的。

在给出约束条件之后，以下是找到最佳路径的算法：

(1) 定义目标函数：定义 $D(t(1:i), r(1:j))$ 是 $t(1:i), r(1:j)$ 之间的 DTW 距离，对应的最佳路径是由 $(t(1), r(1))$ 走到 $(t(m), r(n))$ ；

(2) 定义目标函数的巡回关系：

$$D(t(1:i), r(1:j)) = |t(i) - r(j)| + \min\{D(t(1:i-1), r(1:j)), D(t(1:i-1), r(1:j-1)), D(t(1:i), r(1:j-1))\} \quad (3)$$

端点条件为 $D(t(1), r(1)) = |t(1) - r(1)|$ ；

(3) 得到最优匹配距离 $D(t(1:m), r(1:n))$ 。

在实际计算过程中，首先建立一个矩阵 $D(m \times n)$ ，先根据端点条件填入 $D(t(1), r(1))$ ，然后根据巡回关系，逐行逐列算出 $D(t(1:i), r(1:j))$ ，最后得到 $D(t(1:m), r(1:n))$ 。

2.2.2 轨迹聚类分析

在使用 DTW 方法计算不同学者的指标变化序列之间的距离和最佳匹配路径之后，可分别构建不同学者的 3 个指标变化序列的距离矩阵，并以时间序列长度居中的某位学者的成长轨迹为标准，根据最佳匹配路径将其他学者与这个学者的时间序列对齐，便于后续的轨迹聚类分析工作的展开。此后，在一定范围内遍历设置聚类个数，通过主成分分析和 K-Means 聚类获得不同学者的成长轨迹的类别，并根据聚类结果计算簇内平方和（Cluster Sum of Square，简称 CSS）及轮廓系数，兼顾 CSS 尽量小和轮廓系数尽量接近 1 的原则，判断最优聚类个数。

在通过最佳匹配路径将表示不同轨迹的时间序列对齐后，本项研究借鉴生物信息学中适用于挖掘时间序列变化特征的轨迹聚类分析方法^[49, 50]，其通过引入模糊 c 均值聚类算法降低噪音数据的干扰，可确定各个学者的成长轨迹的类别。模糊 c 均值聚类作为最广泛使用的模糊聚类算法之一，与 K-Means 聚类方法较为相似，通过不断迭代使簇内平方和最小化。二者最主要的区别在于，K-Means 聚类方法中每个对象只由一个聚类中心约束，模糊 c 均值聚类中对象与所有聚类中心都有关。在模糊 c 均值聚类过程中，簇内平方和计算如下：

$$CSS = \sum_{j=1}^k \sum_{x_i \in C_j} u_{ij}^m (x_i - \mu_j)^2 \quad (4)$$

其中， u_{ij}^m 通过下式计算：

$$u_{ij}^m = \frac{1}{\sum_{l=1}^k \left(\frac{|x_i - c_j|}{|x_i - c_l|} \right)^{\frac{2}{m-1}}} \quad (5)$$

u_{ij} 是对象 x_i 属于簇 C_j 的程度, 与对象 x_i 离 C_j 质心 (聚类中心) 的距离成反比; c_j 是簇 C_j 的质心。在模糊 c 均值聚类过程中, 一个对象理论上可以被分配为所有簇, 并通过给定对象在各簇中的成员值描述对象属于某一类的程度或概率。成员值介于 0 和 1 之间, 接近 0 代表对象远离该聚类中心, 接近 1 代表对象靠近该聚类中心, 每个对象在各类中成员值的总和为 1。聚类中心计算为组内所有对象的质心, 并根据对象的成员度进行加权, 具体计算如下:

$$c_j = \frac{\sum_{x \in C_j} u_{ij}^m x}{\sum_{x \in C_j} u_{ij}^m} \quad (6)$$

m 是模糊参数, 理论上可以取大于或等于 1 的任何值, 当接近于 1 时聚类结果将与 K-Means 聚类结果相似, 值接近无限将导致完全模糊, 通常使用 $m=2$ 。总结而言, 模糊 c 均值聚类过程分为五步: ①指定聚类数 k ; ②为每个对象任意分配一组初始成员值; ③使用上述公式计算每个聚类簇的质心; ④使用上述公式重新计算每个对象在各聚类簇中的成员值; ⑤重复③、④过程, 直到成员值收敛或达到最大迭代次数为止。通常来看, 最小质心距离随聚类数的增加而减少, 当聚类数较大时, 最小质心距离的减小速度放缓, 需兼顾聚类数和最小质心距离尽量小的原则确定最优聚类个数。结合利用 K-Means 聚类算法计算的聚类个数与簇内平方和、轮廓系数的变化对应关系, 以及使用模糊 c 均值聚类算法获得的最小质心距离与聚类个数的关系, 可精准确定最优聚类个数, 得到正确的成长轨迹聚类结果。

3 结果分析

在获取表示学者在学术上的成长过程的指标一时间变化序列后, 利用序列之间的距离矩阵和最佳匹配路径, 判别序列的最优聚类个数, 并通过轨迹聚类分析方法呈现成长轨迹普遍或特定的形态特征, 进而捕捉学者潜在的学术能力成长模式以及模式的分布组合特征。

3.1 聚类结果分析

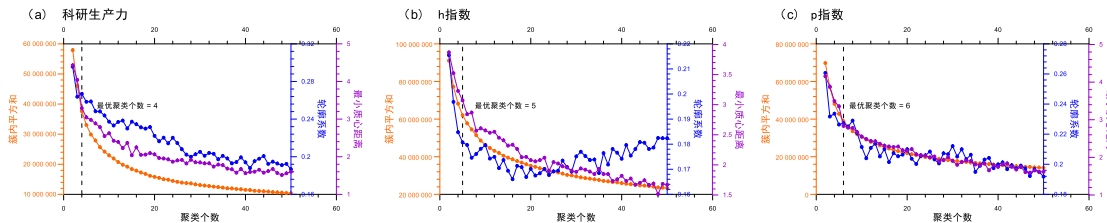


图 2 不同指标变化序列的最优聚类结果判别依据

Fig.2 The Basis for Determining the Optimal Clustering Results for Different Indicator Time Series

通过设定聚类个数在 2~50 这个范围内变化, 基于预先设定不同聚类个数下的聚类结果计算出簇内平方和、轮廓系数、最小质心距离, 如图 2 所示。秉持“尽量小的聚类个数、簇

内平方和、质心距离，尽量大的轮廓系数”的原则，本研究分别设定 4、5、6 分别为科研生产力、h 指数、p 指数的成长轨迹的最优聚类个数，保证聚类结果中同类时间序列的相似性足够大，不同类别序列的相似性足够小。

在确定不同时间序列的最优聚类个数之后，鉴于核心学者的职业生涯跨度的中位数为 21 年，可将不同学者的指标变化序列对齐至这个时间长度，进而构成序列面板数据，应用轨迹聚类分析方法分析成长轨迹的形态特征，得到聚类结果实现学者在学术上成长模式的划分。

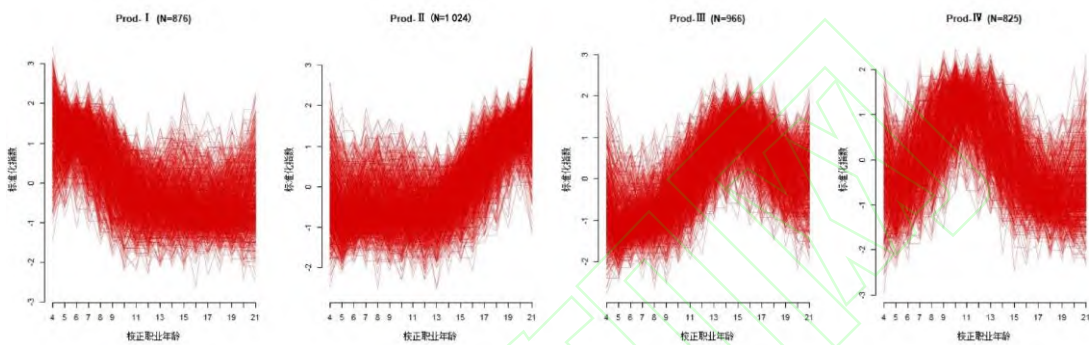


图 3 科研生产力成长轨迹的聚类分析结果

Fig.3 Results of Cluster Analysis of Growth Trajectories of Scientific Productivity

首先，观察核心学者的科研生产力成长轨迹的聚类分析结果，如图 3 所示，可将他们科研生产力的成长过程分为 4 种模式，根据人数由多至少的顺序分别为 Prod- II、Prod-III、Prod- I 和 Prod-IV。其中，Prod- II 和 Prod-III 的科研生产力在早中期表现为总体增长的趋势，二者的差别在于后期的科研生产力的变化形式略有不同，Prod- II 表现为进一步的增长，Prod- III 表现为开始急速下降。在人数相对较少的 Prod- I 和 Prod-IV 模式中，他们的科研生产力的高峰期均集中于前中期，后期的发文量通常处于一个较低的水平，二者的不同在于前期的科研生产力的变化情况，Prod- I 表现为从极高值持续下降，Prod-IV 表现为先上升至极高值后迅速下降的对称形态。

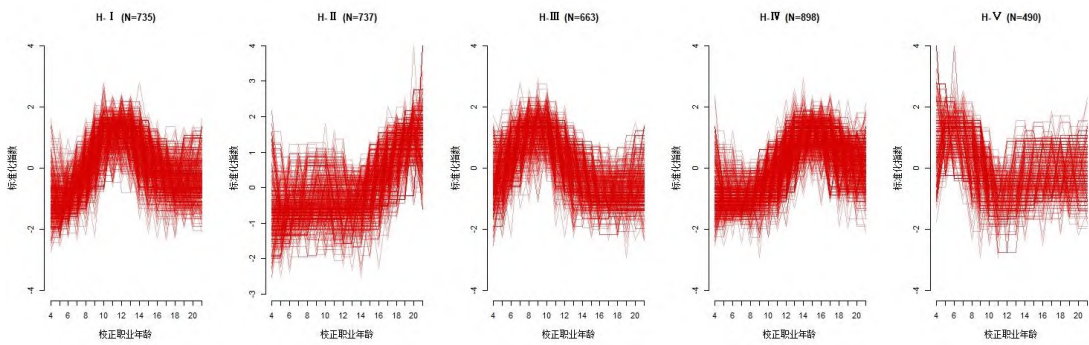


图 4 h 指数成长轨迹的聚类分析结果

Fig.4 Results of Cluster Analysis of H-index Growth Trajectories

其次，分析核心学者的 h 指数成长轨迹的聚类分析结果，如图 4 所示，根据轨迹形态特征的不同可将其分为 5 种模式，根据人数由多至少排序为 H-IV、H-II、H-I、H-III、H-V。详细而论，人数最多的 H-IV 和 H-II 的共同点在于 h 指数的职业生涯在前期波动较少，基本保持在一个低值状态，但在中期开始上升至最大值，高值状态集中在后期阶段；二者的区别在于 H-IV 在末期出现小幅下降，而 H-II 表现为持续性的上升。人数相对次之的 H-I 的 h 指数变化表现为早期上升、后期下降的近似对称形态，最高值出现在中期阶段；而人数最少的 H-III 和 H-V 的最高值出现在前期，尤其以 H-V 类为最，二者均是 h 指数前期较大而后后期较小，且波动幅度也表现出此特点，但 H-V 的 h 指数总体上表现为持续性的下降，而 H-III 在 h 指数持续下降开始前的早期阶段还存在一段快速上升状态。

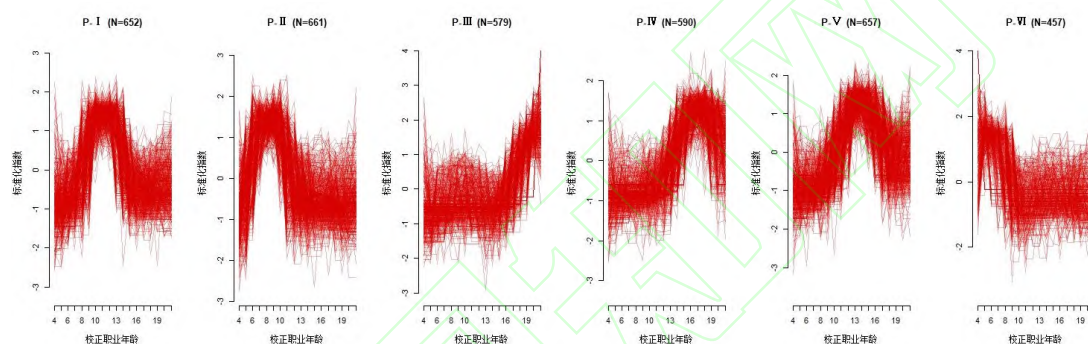


图 5 p 指数成长轨迹的聚类分析结果

Fig.5 Results of Cluster Analysis of P-index Growth Trajectories

最后，通过轨迹聚类分析方法，可获得综合考量发文量和引文量的 p 指数的 6 类变化形态，如图 5 所示，即 p 指数的 6 种成长模式，根据人数由多至少降序排列，依次为 P-II、P-V、P-I、P-IV、P-III、P-VI，除却 P-VI 外，其他模式的人数相差并不大。进一步观察 p 指数各类成长模式的形态特征，P-III、P-IV、P-V 在前中期的 p 指数均表现为总体性增长的趋势，最高值均出现在后期，三者的区别在于末期 p 指数的变化情况不同，P-III 表现为 p 指数持续性增长，P-IV 则在末期表现为末期短暂小幅下降，P-V 为 p 指数后期稳步下降的形态，这三类模式合计包括 1 826 人。另外，P-I、P-II、P-VI 在后期 p 指数的变化形态均为波动性不大的低值状态，其中 P-I 的 p 指数在前期表现为稳步上升至最高值，该值出现在中期阶段，而 P-II 和 P-VI 的 p 指数的最高值出现在早期阶段，在最高值出现之后迅速下降，但 P-II 的 p 指数在最高值出现之前尚有一段快速上升的阶段，以上三类模式共计包括 1 770 人，P-I 包括 652 人，P-II 和 P-V 共包括 1 118 人。

3.2 类别学科分布

为了解各个人文学科学者在学术上的成长过程的共性和独特性，可从学科的角度出发，

分析根据轨迹聚类分析方法得到的不同类别的学术成长模式在哲学、历史学、文学 3 个学科的学者人数分布情况。

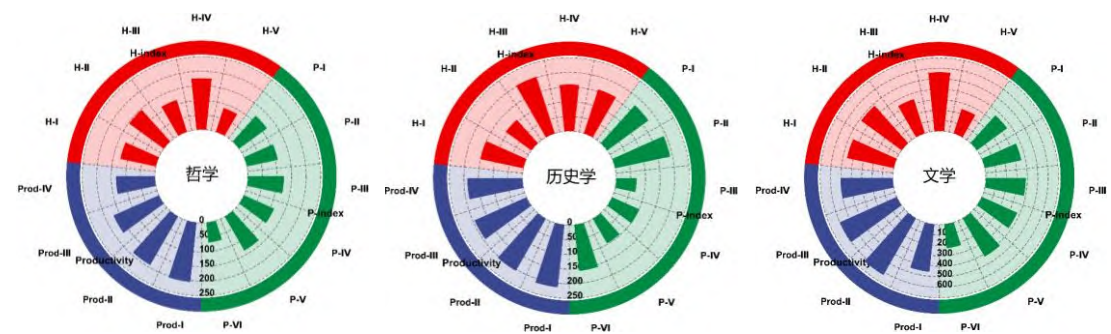


图 6 不同学科学者成长模式的人数分布情况

Fig.6 Distribution of the Number of Scholars in Different Disciplines in Terms of Their Growth Patterns

分析哲学学科的核心学者在科研生产力、h 指数、p 指数 3 个方面的成长模式的分布情况，如图 6 所示，该学科核心学者的科研生产力的成长模式根据人数降序排列为 Prod- I 、Prod- II 、Prod-III、Prod-IV，人数相差不大，基本在 100~200 这个范围变化。科研生产力最高值出现在早期的 Prod- I 、Prod-IV，和科研生产力最高值出现在早期的 Prod- II 、Prod-III 的人数相当。再者，该学科核心学者的 h 指数的成长模式多属于峰值出现在后期的 H- II 和 H-IV，峰值出现在早期的 H-III和 H- V 包括的核心学者人数相当有限。类似地，该学科核心学者的 p 指数的成长模式也多属于峰值出现在后期的 P-III、P-IV和 P- V，峰值出现在早期的 P- II 和 P-VI的人数同样较少。

以同样的方式审视历史学学科 3 个方面学术能力的成长模式的人数分布情况，如图 6 所示，该学科核心学者的科研生产力成长模式极为分散，并不集中表现为某一类成长模式，4 个类别的人数分布极为均衡。与哲学学科核心学者的 h 指数成长模式的人数分布情况相反，历史学学科 h 指数的成长模式多属于峰值出现在早期的 H-III和 H- V 包括的核心学者，峰值出现在后期的 H- II 和 H-IV的人数并不多。同样地，历史学学科核心学者的 p 指数的成长模式的人数分布情况与哲学学科相反，该学科核心学者的 p 指数成长模式大多属于峰值出现在早期的 P- II 和 P-VI，仅有相当少的学者的 p 指数成长模式属于峰值出现在后期的 P-III、P-IV和 P- V。

观察文学学科核心学者在 3 个方面的学术成长模式的人数分布情况，如图 6 所示，该学科科研生产力成长模式的人数分布情况与哲学、历史学不同，表现出显著的集中现象，即它们大多属于峰值出现在后期的 Prod- II 、Prod-III，仅有较少的核心学者的科研生产力成长模式属于峰值出现在早期的 Prod- I 、Prod-IV。与哲学学科核心学者的 h 指数成长模式的人数

分布情况相同，文学学科 h 指数的成长模式多属于峰值出现在后期的 H- II 和 H-IV，峰值出现在早期的 H-III 和 H- V 的人数并不多。文学学科核心学者的 p 指数的成长模式也多属于峰值出现在后期的 P-III、P-IV 和 P- V，峰值出现在早期的 P- II 和 P-VI 的人数相对较少，但也占据了一定数量。

3.3 类别组合特征

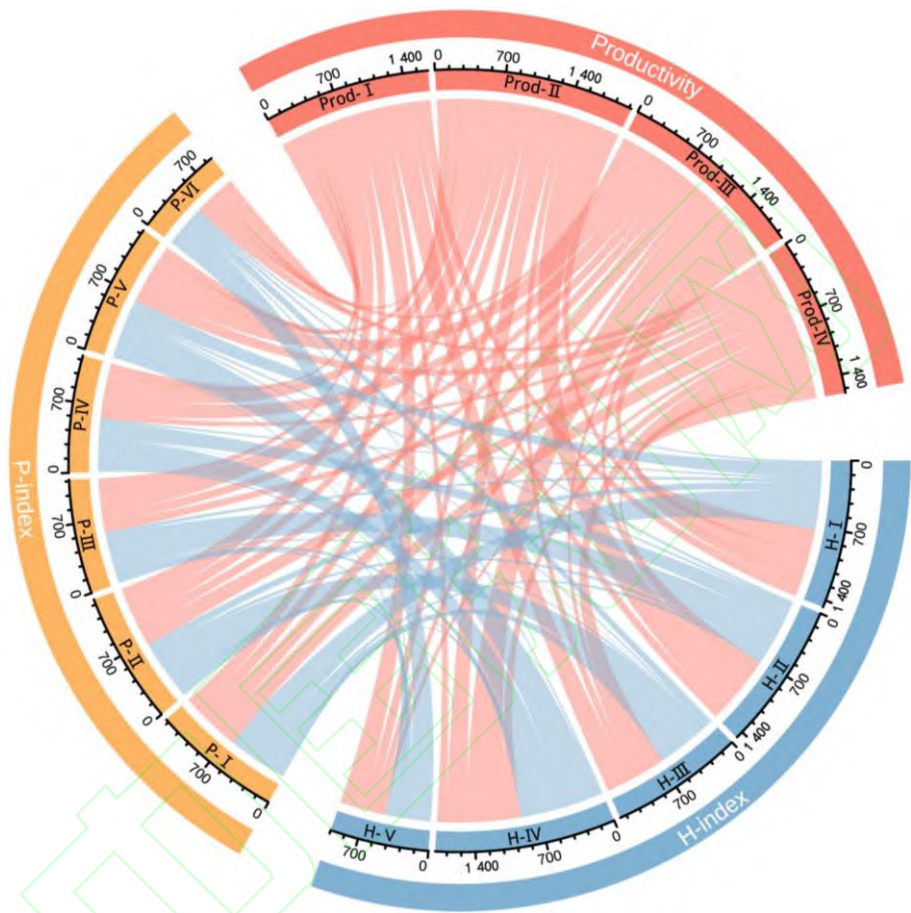


图 7 核心学者在 3 个方面学术能力的成长模式之间的共现频次统计

Fig.7 Frequency Statistics of Co-occurrence Between the Growth Patterns of Three Kinds of Academic Ability of Core Scholars

核心学者在 3 个方面学术能力的成长模式之间是否具有相关性，可通过统计它们之间的共现频次，分析科研生产力、h 指数、p 指数的成长轨迹类别的搭配组合规律，如图 7 所示。

分析科研生产力与 h 指数、p 指数的成长模式之间的共现频次，Prod- I 、Prod-IV 多与 H-III、H- V 组合出现，Prod- II 、Prod-III 多与 H- II 、H-IV 共现，这说明科研生产力最大值出现在早期的核心学者，他们的 h 指数的最大值同样也出现在早期，反之，当科研生产力最大值出现在后期时，他们的 h 指数的最大值通常也在后期。此外，Prod- I 、Prod-IV 多与 P- I 、

P-II、P-VI组合出现, Prod-II、Prod-III多与 P-III、P-IV、P-V 共现, 科研生产力和 p 指数的最大值出现的阶段也表现出同步效应。再者, 观察 h 指数和 p 指数的成长模式之间的共现情况, H-II、H-IV多与 P-III、P-IV、P-V 组合出现, H-III、H-V 的常见组合类别为 P-II、P-VI, H-I 多与 P-I 组合出现, h 指数和 p 指数的最大值和升降形态的出现阶段也较为一致。

根据人文领域核心学者的科研生产力、h 指数、p 指数的成长模式之间的常见共现组合, 他们在这 3 个方面的成长轨迹形态通常表现出同步效应, 具体包括同步增长或下降、峰值出现阶段的一致性。

4 学者成长模式的探讨与应用

中国人文领域核心学者的科研生产力、h 指数、p 指数成长轨迹的聚类分析结果, 反映了他们在学术上的成长模式。为此, 可结合指标所代表的学术能力及相关研究, 对该领域的学者成长模式展开分析探讨, 并进一步将分析结果拓展应用至中国人文领域人才相关制度的优化与完善。

4.1 学者成长模式的分析探讨

通过轨迹聚类分析方法, 对核心学者的科研生产力、h 指数和 p 指数的成长轨迹进行分类, 可揭示出中国人文领域核心学者在学术能力上的成长模式。结果表明, 科研生产力、h 指数、p 指数的成长模式个数分别为 4、5、6。其中, 中国人文领域核心学者的科研生产力的成长模式大多为最高值在后期阶段的 Prod-II 和 Prod-III, 表现为早中期稳步性增长, 说明他们的高产期多发生在后期, 该领域可能更注重前期的学术积累, 因而能陆续产出丰富的学术成果, 这与高产期多在职业生涯早期的自然科学和工程技术领域的学者并不相同^[21, 51, 52]。再者, 该领域核心学者的 h 指数的成长模式多属于早期维持低值状态、后期稳步增长的 H-IV 和 H-II, 说明该领域学者的学术影响力通常需要积累至后期方变得愈发显著。此外, 存在大量的核心学者的 p 指数的成长模式属于前期波动不大、后期快速增长的 P-III、P-IV、P-V, 但仍有相当部分的学者的 p 指数的成长模式属于最大值出现在早中期的其他类别, 综合发文量和引文量的 p 指数的成长模式表明, 他们的综合能力和科研生产力、学术影响力相似, 均是多在职业生涯的后期阶段达到最佳状态, 3 个方面的学术能力的成长都需要前期的积累。

由于人文领域可进一步细分为哲学、历史学和文学 3 个学科, 可比较这 3 个学科核心学者的学术能力成长模式的普遍性和差异性。其中, 哲学和文学学科的核心学者的学术成长模

式与总体样本的人数分布特征大致相同,即科研生产力、学术影响力、综合能力的峰值和增长过程多发生在他们职业生涯的后期,但哲学学科的学者的科研生产力成长模式并没有固定集中于某一类。与之形成对比的是,历史学学科的核心学者的科研生产力的成长模式虽不集中于某一类,但他们的学术影响力、综合能力的峰值和增长过程却多发生在他们职业生涯的早期。综合来看,说明文学学科核心学者的高产期稳定分布在他们职业生涯的后期,科研创作更需要前期的学术积累,而哲学和历史学学科的核心学者的科研生产力变化表现出多样化的特点。从学术影响力和综合能力的成长模式来看,哲学和文学学科的核心学者的学术能力衡量指标的上升多发生在后期阶段,历史学学科却集中于早期阶段,优势积累效应并不显著。

根据科研生产力、学术影响力、综合能力的成长模式的共现组合关系来看,三者之间的峰值出现位置 and 变化过程表现出一定的同步性。但是,细究哲学、历史学、文学 3 个学科核心学者成长模式的人数分布情况,文学学科的核心学者在 3 个方面的学术能力的成长过程的同步性十分显著,即大多表现学术能力衡量指标峰值和上升阶段同步出现在后期,说明人文领域核心学者在 3 个方面的学术能力的成长可能具有一定的相关性。然而,哲学和历史学核心学者的科研生产力并没有如他们的学术影响力、综合能力的成长过程表现出同步性,虽然他们的科研生产力的成长模式多样不一,但他们的学术影响力、综合能力的成长模式却表现出集中分布于某一类的现象,说明这两个学科核心学者仅在学术影响力和综合能力两个方面表现出一定的相关性,但这两个方面的学术能力的变化与科研生产力变化基本不相关。

4.2 学者成长模式的应用拓展

人文领域的知识具有弥散性、积累叠加性,学者必须付出时间和精力来回溯经典、领悟精髓,无捷径可走,决定了新文科的人才评价必须将过程评价和结果评价结合起来,以诊断性和动态性评价为主^[53]。这一点在哲学和文学领域的核心学者的成长模式中表现得十分显著,即他们的学术影响力和综合能力的快速成长通常发生在后期。为此,《关于分类推进人才评价机制改革的指导意见》明确指出,现有的人才评价体系有必要遵循人才成长规律,科学设置评价周期,注重过程与结果、短期与长期评价的结合,避免急功近利和功利化倾向。具体可结合本项研究中的人文领域核心学者的主流成长模式,拟定不同学科学者的评价周期,如文学学科的核心学者的高产期通常发生在职业生涯的后期阶段,可适当放长该学科人才的评价周期;而如历史学学科的核心学者,他们的学术影响力和综合能力的高峰期集中在

职业生涯早期阶段，因此，可适当缩短评价周期，避免“等年龄、熬资历”等负面影响，激励青年人才尽早脱颖而出。

此外，人文领域应建立符合人才个性化和多样化特点的评价体系，应考虑学者成长模式的差异性，避免“一刀切”“一把尺子量到底”，促进人才培养的持续健康发展。本项研究表明，不同于高产期集中在早期的自然科学和工程技术领域的学者，除却文学学科学者的高产期集中于后期阶段外，其他人文领域核心学者的科研生产力成长模式多样且发散。为此，需要了解学者成长的不同模式，确保“用不同的尺子量不同的人才”，针对各类人才特征建立成长成效全过程跟踪监测和动态评估机制，改变以静态评价结果给人才贴上“永久牌”标签的做法^[54]。详细而言，可通过分析青年学者在3个方面学术能力的成长轨迹形态特征，判别其成长模式的类别属性，根据所在类别的轨迹变化特点，进而评估其未来可能的发展趋势，如人文领域某一青年学者的科研生产力的成长轨迹形态在早期更接近 Prod-I 这类模式，其更有可能在中后期的发文量呈现出衰减的趋势，具体可根据其与 Prod-I 前期成长轨迹的 DTW 距离及其匹配长度衡量可能性大小。

综合来看，人文领域的人才评价应结合过程评价与结果评价，重视学者成长的动态性和积累性。评价体系需科学设置周期，避免急功近利。不同学科应根据其核心学者的成长模式设定合适的评价周期，以促进人才的持续健康发展。此外，应建立个性化和多样化的评价体系，避免“一刀切”，根据学者成长模式的差异性进行动态评估，以更好地激励和培养人才。

5 结语

新文科建设背景下，分析中国人文领域核心学者在学术上的成长模式，有助于提高中国新文科人才培养过程的科学性，推动人才强国战略布局的展开。本项研究以 CSSCI 为研究基础，通过 DTW 和轨迹聚类分析方法，分析了他们在科研生产力、学术影响力、综合能力3个方面成长轨迹的形态特征。结果表明，不同于自然科学和工程技术领域，人文领域中文学核心学者的高产期大多稳定在他们职业生涯的后期阶段，而哲学和历史学学者的高产期分布不具有显著的规律性。再者，哲学和文学的核心学者的学术影响力和综合能力的快速成长通常发生在职业生涯的后期，这个成长更依赖于他们早期的学术积累，而历史学核心学者在这两方面学术能力的成长多发生在早期，说明该学科核心学者的成长往往更需要他们早期充沛的精力投入。此外，中国人文领域核心学者的学术影响力和综合能力的成长过程具有显著的同步性，而这个同步性表现在文学学者3个方面学术能力的成长过程上。

中国人文领域核心学者的成长模式研究对人才培养评价相关政策的导向提供了参考依据,如过程与结果相结合的评价方式、动态可调整的考核评价周期、全过程跟踪式监测的评估机制等,根据本项研究得出的该领域学者成长模式的阶段性特征、学科性差异以及组合规律,拟定人文领域详实的人才培育评价相关政策的优化完善方案。然而,学术论文虽然一定程度上可表示学者的学术能力,但如需更为细致、具体、完整地反映他们在学术上的成长轨迹,有必要将专著、报纸等其他形式的文献成果或履历信息纳入评价参考体系;再者,学者的成长模式不仅包括科研生产力、学术影响力、综合能力的变化,同时也涉及其他方面的学术能力以及科研行为的改变,如创新力、合作模式、研究兴趣等,未来可拓展研究范围,更加全面地挖掘分析该领域学者成长模式。

参 考 文 献

- [1] 致公党中央教育委员会,钟周,浦天龙.关于“加强数字人文人才培养大力推进新文科建设”的调研报告[J].中国发展,2024,24(1): 71-75.
- [2] Nielsen, M. W., Andersen, J. P. Global citation inequality is on the rise[J]. Proceedings of the national academy of sciences, 2021,118(7): e2012208118.
- [3] Horner, K. L., Rushton, J. P., Vernon, P. A. Relation between aging and research productivity of academic psychologists[J]. Psychology and aging, 1986,1(4): 319 - 324.
- [4] Lehman, H. C. Men's Creative Production Rate at Different Ages and in Different Countries[J]. Scientific monthly, 1954,78(5): 321-326.
- [5] Dennis, W. Age and Productivity among Scientists[J]. Science, 1956,123(3200): 724-725.
- [6] Bayer, A. E., Dutton, J. E. Career age and research-professional activities of academic scientists: Tests of alternative nonlinear models and some implications for higher education faculty policies[J]. The Journal of higher education, 1977,48(3): 259-282.
- [7] Diamond, A. M. The Life-cycle Research Productivity of Mathematicians and Scientists1[J]. Journal of Gerontology, 1986,41(4): 520-525.
- [8] Way, S. F., Morgan, A. C., Clauset, A. et al. The misleading narrative of the canonical faculty productivity trajectory[J]. Proceedings of the national academy of sciences, 2017,114(44): E9216-E9223.
- [9] Györfy, B., Csuka, G., Herman, P. et al. Is there a golden age in publication activity?—an analysis of age-related scholarly performance across all scientific disciplines[J]. Scientometrics, 2020,124(2): 1081-1097.
- [10] Yair, G., Goldstein, K. The Annus Mirabilis paper: years of peak productivity in scientific careers[J]. Scientometrics, 2020,124(2): 887-902.
- [11] Egghe, L. The Hirsch index and related impact measures[J]. Annual Review of Information Science and Technology, 2010,44(1): 65-114.
- [12] Hirsch, J. E. An index to quantify an individual's scientific research output[J]. Proceedings of the national academy of sciences, 2005,102(46): 16569-16572.
- [13] Liu, Y., Rousseau, R. Definitions of time series in citation analysis with special attention to the h-index[J].

Journal of Informetrics, 2008,2(3): 202-210.

[14] Liang, L. h-index sequence and h-index matrix: Constructions and applications[J]. Scientometrics, 2006,69(1): 153-159.

[15] Egghe, L. Mathematical study of h-index sequences[J]. Information Processing & Management, 2009,45(2): 288-297.

[16] Egghe, L. Comparative study of h-index sequences[J]. Scientometrics, 2009,81(2): 311-320.

[17] Guns, R., Rousseau, R. Simulating growth of the h-index[J]. Journal of the American Society for Information Science and Technology, 2009,60(2): 410-417.

[18] Wu, J., Lozano, S., Helbing, D. Empirical study of the growth dynamics in real career h-index sequences[J]. Journal of Informetrics, 2011,5(4): 489-497.

[19] Prathap, G. The 100 most prolific economists using the p-index[J]. Scientometrics, 2010,84(1): 167-172.

[20] Prathap, G. A three-class, three-dimensional bibliometric performance indicator[J]. Journal of the Association for Information Science and Technology, 2014,65(7): 1506-1508.

[21] Peng, X., Xu, H., Shi, J. Are the bibliometric growth patterns of excellent scholars similar? From the analysis of ACM Fellows[J]. Journal of Informetrics, 2024,18(3): 101543.

[22] 高志, 张志强. 个人学术影响力与年龄的变化规律研究进展——基于自然科学领域[J]. 情报杂志, 2016,35(2): 105-109+156.

[23] Symonds, M. R. E., Gemmell, N. J., Braisher, T. L. et al. Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance[J]. PloS one, 2006,1(1): e127.

[24] 郭新艳. 科技人才成长规律研究[J]. 科技管理研究, 2007 (9): 223-225.

[25] Sinatra, R., Wang, D., Deville, P. et al. Quantifying the evolution of individual scientific impact[J]. Science, 2016,354(6312): aaf5239.

[26] 陈姗姗, 林春培, 张向前. 面向 2035 年我国青年科技人才发展机制研究[J]. 科学管理研究, 2023,41(1): 129-138.

[27] 李维平. 论人才成长[C]. 中国人力资源开发研究会第九次会员代表大会暨学术研讨会, 2006, 中国北京.

[28] 王通讯. 人才成长的八大规律[J]. 决策与信息, 2006 (5): 53-54.

[29] 叶忠海. 人才成长规律和科学用人方略[J]. 中国人才, 2007 (5): 31-32.

[30] 杨淞月. 高校拔尖创新人才成长规律及培养策略研究[D]. 中国地质大学, 2012.

[31] 刘琳琳. 创新型人才成长的规律与路径研究[J]. 科学管理研究, 2014,32(1): 82-85.

[32] 李祖超, 李蔚然, 王天娥. 24 位国家最高科学技术奖获得者成才因素分析[J]. 教育研究, 2014,35(12): 61-71.

[33] 宋晓欣, 马陆亭, 赵世奎. 教育学科高层次人才成长规律探究——以 22 位长江学者为例[J]. 中国高教研究, 2018 (3): 51-55, 87.

[34] 蔡翔, 谢梅英, 程发新等. 中国院士群体师承效应的量化研究[J]. 科技管理研究, 2021,41(23): 133-140.

[35] 刘俊婉. 从诺贝尔奖现象看科学创造的特征[J]. 科学学研究, 2009,27(9): 1289-1297.

[36] 徐飞, 陈仕伟. 中国杰出科学家年龄管理策略的新思考——从近十年(2001—2010)中国科学院新增院士与诺贝尔奖获得者年龄比较的反差谈起[J]. 科学学研究, 2012,30(7): 976-982.

[37] Merton, R. K. The Matthew Effect in Science[J]. Science, 1968,159(3810): 56-63.

[38] R.K.默顿.科学社会学[M]. 北京: 商务印书馆, 2003: 9.

[39] 周光礼, 罗睿. 光环叠加: 中国科学精英的生成机制[J]. 教育研究, 2021,42(10): 138-153.

[40] 中国科学院人力资源管理研究组. 关于我院创新三期人力资源管理的若干思考[J]. 中国科学院院刊,

2007 (5): 355-373.

[41] 刘少雪.面向创新型国家建设的科技领军人才成长研究[M]. 北京: 中国人民大学出版社, 2009: 77-178.

[42] 林崇德, 胡卫平. 创造性人才的成长规律和培养模式[J]. 北京师范大学学报(社会科学版), 2012 (1): 36-42.

[43] 曹晓丽, 王森芑, 姚妞妞. 创新型科技人才成长路径分析[J]. 信息系统工程, 2020 (11): 4-9.

[44] 瞿群臻, 王嘉吉, 唐梦雪等. 基于逻辑增长模型的科技人才成长规律及影响因素研究——以海洋领域科技人才为例[J]. 科技管理研究, 2021,41(12): 157-164.

[45] 马费成, 陈柏彤. 我国人文社会科学学科多样性研究[J]. 情报科学, 2015,33(1): 3-8.

[46] Fernandes, J. M., Costa, A., Cortez, P. Author placement in Computer Science: a study based on the careers of ACM Fellows[J]. Scientometrics, 2022,127(1): 351-368.

[47] Hagen, N. T. Harmonic Allocation of Authorship Credit: Source-Level Correction of Bibliometric Bias Assures Accurate Publication and Citation Analysis[J]. PloS one, 2008,3(12): e4021.

[48] 王星, 郑滨彬, 朱枫怡等. 基于非参数回归和卷积神经网络的在线手写签名身份认证模型研究[J]. 数理统计与管理, 2018,37(4): 610-623.

[49] Futschik, M. E., Carlisle, B. Noise-robust Soft Clustering of Gene Expression Time-course Data[J]. Journal of Bioinformatics and Computational Biology, 2005,03(4): 965-988.

[50] Kumar, L., M, E. F. Mfuzz: a software package for soft clustering of microarray data[J]. Bioinformation, 2007,2(1): 5-7.

[51] Barabási, A.-L., Wang, D. (2021). The Science of Science. In A.-L. Barabási & D. Wang (Eds.), *The Science of Science* (pp. i-ii). Cambridge University Press.

[52] Schilling, M. A., Green, E. Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences[J]. Research policy, 2011,40(10): 1321-1331.

[53] 公钦正, 张海生. 新文科人才培养何以实现——基于中英两所大学 PPE 专业人才培养方案的案例研究[J]. 重庆高教研究, 2022,10(2): 87-98.

[54] 习近平. 努力成为世界主要科学中心和创新高地[J]. 共产党员, 2021 (8): 4-7.