

刘瑞平,李建新.个体导向方法的理论缘起、实例应用及评述[M]//赵联飞,赵锋.社会研究方法评论:第3卷.重庆:重庆大学出版社.

个体导向方法的理论缘起、实例应用及 评述

刘瑞平 李建新^①

摘要: 定量社会研究长期被变量导向方法所主导,这种看似如此自然的研究范式从假设、推断逻辑到研究结果可能存在潜在风险。近些年来,以整体互动论和总体异质性为理论基础逐渐兴起的个体导向方法,通过个体间关系识别同质性的子群体,对于研究个体发展轨迹以及解决总体异质性问题发挥着重要作用。本文系统地介绍了个体导向方法的理论缘起以及具体统计分析技术,包括聚类分析、潜类别分析、潜剖面分析、组基轨迹模型和增长混合模型等。同时还通过对个体导向方法与变量导向方法在截面数据和纵向数据中的实例应用进行比较分析,展示二者在实际研究中的思路、过程和结果方面的差异。最后,本文就个体导向方法使用的条件以及局限性进行了讨论。在实证研究中,个体导向和变量导向两种方法同等重要,将二者相结合的研究思路能够更好地发挥它们的优势。

关键词: 变量导向方法;个体导向方法;总体异质性

^①作者简介:刘瑞平,中国人民公安大学治安学院讲师,联系方式:liuruiping6@163.com。李建新,北京大学社会学系教授,联系方式:ljx@pku.edu.cn。

Abstract: Quantitative social research has long been dominated by variable-oriented approaches. This seemingly natural research paradigm may have potential risks from hypothesis, inference logic to research results. In recent years, person-oriented approaches that have gradually emerged are based on holistic-interactionistic theory and population heterogeneity, which identifies the homogeneous subgroups through relationships between individuals. This plays an important role in analyzing individual growth trajectory and solving the problem of population heterogeneity. This paper systematically introduces the theoretical basis of the person-oriented approaches, as well as specific statistical analysis techniques, including cluster analysis, latent class analysis, latent profile analysis, group-based trajectory modeling and growth mixture modeling. This paper compares and analyzes the application of person-oriented approaches and traditional variable-oriented approaches in cross-sectional data and longitudinal data, and shows the differences in ways of thinking, processes and results of two approaches in practical research. Finally, this paper discusses the conditions and limitations of the use of person-oriented approaches. Person-oriented approaches and variable-oriented approaches are equally important, and the research ideas that combine the two can better exert their own advantages in empirical research.

Key words: Variable-Oriented Approaches; Person-Oriented Approaches; Population Heterogeneity

一、引言

在定量社会研究中，变量导向/中心方法（Variable-Oriented Approaches 或 Variable-Centered Approaches）的使用最为普遍，它奠基基于个体同质性的假设，即总体中的个体是同质的，或者至少是足够同质的（von Eye & Bogat, 2006; Laursen & Hoff, 2006; Morin et al., 2018）。变量导向方法的定义包含理论和方法论两个层面：在理论层面，其主要概念和分析单位是变量，关注变量间的（因果）关系，除了以变量为主要单位进行统计分析外，还以变量和假设检验来构建或发展理论，根据观察到的变量关系来解释来自理论假设的变量关系；在方法论层面，关注相关变量的测量并研究这些变量在时间上的关系，通常使用线性统计模型来完成（Bergman & Magnusson, 1997; Bergman & Trost, 2006）。在统计分析方法中经常应用的平均值、相关分析、方差分析、因子分析、多元回归、增长曲线模型等，都属于变量导向方法。该方法能够客观、精确地测量研究对象信息，利用统计模型控制混淆变量从而达到更准确的因果推论，以及在探讨关系机制、检验和构建理论等方面发挥着独特优势（Bergman & Andersson, 2010）。然而，以广义线性模型为代表的变量导向方法的基本假设与社会现实解释之间存在巨大鸿沟（Abbott, 1988）。近些年来，不断发展的个体导向/中心方法（Person-Oriented Approaches 或 Person-Centered Approaches）形成了从理论到统计应用的比较完整的研究体系，通过识别总体中的不同子群体来减少异质性，其在社会科学中的应用也越来越广泛，与变量导向的统计方法优势互补，二者的结合为多角度认识社会现实、

改善因果推断提供了新的思路。本文基于以往研究对变量导向方法的反思，系统地对个体导向方法的理论基础、统计方法、实际应用以及存在的问题等进行梳理和讨论，这有助于我们在社会科学研究中对该方法的使用及其所存在的问题有清晰的认识。

二、个体导向方法的缘起和理论基础

（一）变量导向方法的局限性

阿伯特很早就发现了变量导向方法的弊端，他认为广义线性模型以广义线性实在（General Linear Reality）假设来解释社会世界，这不仅不符合社会现实，还与社会理论相矛盾，他以研究实例对社会世界由可变属性的固定实体组成、单调因果流、单一的意义、缺乏的序列效应、个案独立性与相关性假设、情境独立性等广义线性模型的假设进行一一驳斥（Abbott, 1988）。虽然目前变量导向方法不断改进，并开发出更复杂的统计模型放松了以上假设，但始终未能从根本上改变认识现实世界的方式，其在理论和方法方面存在一定的局限性。通过对以往相关文献的梳理，其被归纳为以下三个主要方面：

第一，研究中的推断逻辑存在风险。变量导向方法的研究基于个体同质性假设，把个体组成的总体看作单一类型或模式，重点在于总体的平均值和变化发展的平均轨迹，以一组“平均”参数来描述总体，探讨所感兴趣变量的平均因果效应，关注个体的量的变化而非质的变化。然而一旦存在异质性的子群体，如果仍以描述和解释个体所集合的总体层面的关系、特征或模式等来推断个体，则会

出现逻辑错误，造成生态谬误（von Eye & Bogat, 2006; Bergman & Trost, 2006; Sterba & Bauer, 2010）。因为以集合数据结果推断个体存在较大风险，施密兹（Schmitz, 2000）将其总结为四个定理：（1）关于时间序列参数估计，即个体轨迹不能由集合轨迹来确定；（2）相关性分析，即截面的、个体的相关关系不能描述个体内部相关性和过程；（3）非共时性关系，即在追踪研究中，交叉滞后的面板分析可能无法捕捉任何给定个体的变量间关系；（4）操作顺序，即集合和预测的顺序应从个人而不是群体开始。变量导向方法是从集合到个体的逻辑分析，除非假设数据是完美的（例如，集合层级和个体层级数据具有相同的参数估计），否则都会产生这些风险，然而，我们并非生活在一个完美的世界里（Bogat et al., 2016）。以往学者对变量导向方法的假设可总结为“组成总体的个体是同质性的”（von Eye & Bogat, 2006; Morin et al., 2018），也就是假设个体或组内同质性。也有学者认为需要区分对组内异质性的忽视和组内同质性的假设，经济学、社会学、人口学等社会科学学科的定量研究并不再假设组内同质性，然而在因果推断中，个体层面变异的普遍存在决定了不可能研究个体层面上的因果效应，因此出于实际原因，我们可以选择不分析研究背景中的组内和个体层面的异质性，因果推断只能在总体层面进行统计分析，这也就忽略了个体层面的变异性（Xie, 2013）。不管是否承认变量导向方法是否建立在个体同质性假设的基础之上，其在总体层面上所得出的结果仍抹去了个体差异和组间差异，生态谬误的风险依然存在。

第二，忽略个体差异、缺乏整体互动视角。变量导向方法的研究不仅只见变量不见人，而且倾向于孤立地描述某一自变量的平均特征或探讨自变量对因变量的平均影响效应，忽视了个体差异。通

过孤立地研究对象的各个组成部分来理解研究对象，这是变量导向方法的关键特征（Kelso，2000）。生物个体是一个有机整体，个体的各个组成部分及其发展的各个阶段都是互相联系的整体，整体都是通过所涉及的各因素之间的相互作用来获得其特征的，而不是通过分析每个孤立因素对整体的影响来获得的（Magnusson，2003）。因此，影响个体或总体的因素是互相关联的整体，有时候并非单一因素发挥作用，而是多个因素互相关联，甚至同时发生作用才能够对结果产生真正的影响，这种例子在生活中并不少见。例如，冗余系统装置的机器，只有当两个组件同时故障时才会导致整个机器发生故障，然而每个组件故障率的异质性不同，机器的故障率则会出现多个状态，包括整个系统比单系统具有更高故障率的状态（Vaupe1，1985）。再如，在医学诊断的相关治疗中，观察高烧是疾病的一个重要指标，然而，只有它与患者其他症状同时发生并一起评估，才能在疾病诊断中做出最后的选择（Magnusson，2003）。变量导向方法的统计模型缺乏这种整体互动视角，主要表现在以下几个方面。首先，在进行因果推断时，建立的回归模型解释需要控制其他因素不变，而现实社会中不存在改变关键自变量而不改变控制变量的结果变量。这种单独来看某一自变量，只见树木不见森林，就如同盲人摸象，某一特征下的因变量的个体是比较笼统的，甚至是有偏的。其次，在截面数据分析中，变量导向方法总是孤立地分析每个变量对因变量的影响，很少纳入变量之间的互动作用或交互作用，其统计模型无法纳入多个因素同时发生的情况，而个体却是由多个因素互动作用影响的功能整体。最后，在追踪数据分析中，分析发展趋势，变量导向方法是以平均趋势来拟合总体，忽视了不同子群体和不同个体的发展轨迹的差异性。

第三，存在理论和方法的匹配问题，尤其是关注于个体发展的纵向研究，变量导向方法的统计模型和指导其研究的理论之间往往不匹配。理论是对现实世界或现象的认识，并以抽象概念进行知识性的概括，尤其是在社会科学中，构建的理论更为抽象、宏观、全面，其包含了现实非线性关系、复杂的因果解释（如多因多果、互为因果等）、各个因素的动态交互作用等特征。但在实证研究中，理论被分解为变量来进行检验，盲目地应用变量导向方法可能会在研究过程中违背理论的关键特性，得出与现实世界不符的结论（Bergman et al., 2009）。在变量导向的统计模型中，虽然现在已经发展出其他模型来处理理论中的非线性问题，但是线性回归模型是应用中最常见的模型，也是很多统计方法应用的起点。并且，我们所建立的因果模型很多都是建立在线性假设基础上的，所得到的也是线性因果关系，除了影响不显著，变量的系数或为正或为负，这也是一种线性假设。因为大多数人类行为并非线性，例如，两个变量只有在某些条件下才可能存在关系，或者干预因素对结果变量并不一定一直产生正向或负向作用，有的是早期对结果变量产生负向影响，后期变成了正向影响，或者反之（Bogat et al., 2016）。

社会科学研究中的非线性和高阶交互作用较为常见，也并不能用简单的线性关系来表示，即使在基本线性回归模型加入二次项或交互作用，也无法实现模型和理论现实间的拟合，反而会使其分析过程变得更为复杂，结果难以解释（Bergman, 2001）。另外，在对同一类型的多个自变量对因变量影响的研究中，侧重于单一自变量对因变量的独特作用或这些自变量的累积作用，每个变量对所有个体具有相同的权重，反映了一般人的特征（Bergman et al., 2003）。例如，过去探讨童年逆境与老年健康关系研究的经常应用的两种变

量导向方法，第一种是关注于单一童年逆境或同一性质的童年逆境，如童年饥饿（刘亚飞，2018）等。但经历一种逆境的人通常会经历更多的逆境，因此仅关注单一类型的童年逆境可能会高估单一童年逆境对健康的影响，低估其他童年逆境对健康的影响（Anda et al., 2010）。第二种是对多种童年逆境的数量进行累加，以最终童年逆境得分作为自变量考察对健康的累积作用（石智雷、吴志明，2018），这种方法不仅假设不同种类/类型的童年逆境对健康的影响具有同等作用，忽视了童年逆境对健康影响的异质性信息，并且它还假设童年逆境数量与健康是一种正相关关系，这种典型的线性假设思维在某些情况下是站不住脚的（Barboza, 2018）。

（二）个体导向方法的理论基础

一直以来变量导向方法在定量社会研究中占据主导地位，越来越复杂的变量导向的统计分析方法被引入社会学、教育学、人口学等社会科学领域，这些方法背后的假设却忽视了异质性的社会世界，研究中表现出过分追求解决内生性问题的高级统计方法的使用倾向，甚至不去思考该方法的假设条件、适用性和局限性，导致结果仅关注特定变量对因变量的平均因果效应，忽略了个体层面的因果效应和社会事物的动态发展过程。近几十年里，学者们开始不断反思变量导向方法，尤其是关注个体的发展心理学家，他们以整体互动主义为理论基础，强调个体应被视为相互作用的有机整体，提倡将个体导向方法作为一种新的概念、思维方式和分析方法，并推动该方法逐渐兴起和发展完善（Bergman, 2001）。布洛克（Block, 1971）在心理学领域最早明确使用现代意义的个体方法并进行开创性研究，之后学者对其进一步拓展（Bergman & Magnusson, 1997；

Magnusson, 2003; von Eye & Bogat, 2006; Sterba & Bauer, 2010; O'Donnell et al., 2017)。由于心理学研究主体是个体,学者不仅关注个体层面的因果关系,还关注个体的动态发展过程(Bergman et al., 2003)。又因为每个个体的功能和发展模式可能是不同的,而变量导向方法忽略了个体差异性,因此心理学领域开始深入反思变量导向方法,倡导要以个体导向方法作为重要补充。与变量导向方法的假设不同,个体导向方法是基于总体异质性假设,即一个总体存在多个不同的子群体,强调个体无论从横向还是纵向发展都是以复杂的动态系统为基础的,它们包括多种不同层次的因素。在实证研究中把具有这些特征的个体作为分析单位,重点关注个人特征模式、个体之间的关系等,分析过程中看似仍然以变量为载体,但实际上这时的变量值本身不再重要。总体而言,个体导向方法是以“模式”和“过程”为分析单位的,因此有时也被称为“模式或过程导向方法”(Bergman & Magnusson, 1997; Bergman, 2001)。

心理学领域把整体互动论作为个体导向方法的基本理论框架,其把个体概念化为一个有机整体而不是变量的总和,强调个体生命发展的各个阶段都是相互联系的,个体功能的心理、行为和生物学因素以及社会环境、文化等都应该被纳入理论框架中,个体发展是由各相关要素之间的相互作用而非各个孤立因素所决定的(Magnusson, 2003; Bergman & Magnusson, 1997)。个体导向方法的基本原则是需要以整体和互动的观点来全面了解个体的功能和发展,在此基础上,以往学者提出多个个体导向方法的理论原则(Bergman & Magnusson, 1997; Bergman, 2001; von Eye & Bergman, 2003; von Eye & Bogat, 2006; Bergman & Andersson, 2010),有学者将它们汇总为六条:“(1)个体独特性,在某种程度上,个体的

功能、过程和行为发展是独特的；(2) 复杂互动性，这个过程是复杂的，它涉及许多不同层次上相互关联的因素；(3) 个体内变化的个体间差异性，个体内和个体间的恒常性和变化具有规律性和结构性；(4) 模式概括性，这个过程以规律的方式发展，可以被描述为相关因素的模式；(5) 整体论，所涉及因素的意义是由这些因素的相互作用决定的；(6) 模式有限性，虽然从理论上讲，在详细层次上，所观察的个体的过程和状态特征具有无穷无尽的差异，但在整体层次上，能够被概括为更为频繁的少数模式。”（详见 Sterba & Bauer, 2010）个体导向方法的理论框架、概念发展和理论原则成为实证研究的基础，尤其是进一步细化的六条原则使这一方法能够具体化到操作技术层面，并形成两个互相关联的命题：第一，在某一现象的多维空间中，由独特个体系统的所有运行元素来定义，但某一特定总体中的个体属于有限个数量的不同子群体；第二，属于不同子群体的个体之间的主要区别在于子群体的因素运行模式，可以根据个体的相似性特征归类为同质性子群体（Magnusson, 2003）。这两个假设为个体导向方法研究范式运用到统计模型提供了可能性，利用这一方法进行实证研究的目的一般包括关注个体、典型模式或非典型模式，认识非线性和交互作用、动态发展过程等（Bergman, 2001；Bergman & Andersson, 2010）。

人口学中的总体异质性（Population Heterogeneity）理论为个体导向方法假设的合理性以及解决总体异质性问题的必要性奠定了基础。异质性是社会科学的本质，异质性总体由多种同质性的子群体组成，其模式（由于个体随时间不断地死亡、退出，或者事件发生的风险率不同）可能与组成该总体的子群体模式存在质的不同。而一旦由几个同质性子群体组成复合型总体，这一异质性总体的所有

特征几乎都会显现出来，无论在这一总体上考察多少个不同属性，个体或子群体在各种被忽略的维度上将可能会与总体所表现出的属性有所不同，因为一旦存在异质性，选择性将会发生，那些被选择机制筛选之后幸存的总体将不同于原来最初的总体（Vaupel et al., 1979; Vaupel & Yashin, 1985）。综合来看，总体异质性体现在静态和动态两个方面，静态总体异质性表现在不同类别的同质性子群体组成的总体模式不同于子群体的单独模式，可能会导致辛普森悖论（Simpson, 1951）。动态总体异质性强调个体发展的时间效应，一般包括两种情况，第一种是总体的发展轨迹由多个同质性子群体发展轨迹组成，而子群体的发展轨迹并不符合总体平均发展轨迹。第二种是随着时间的发展，多个异质性子群体会产生选择效应，而幸存的总体将不同于最初的总体。在统计分析中，不管是运用截面数据还是纵向数据对某一问题进行研究，这两个方面的总体异质性将会误导我们对事实的判断，例如，美国黑人和白人在老年时期出现死亡率交叉现象（Manton & Stallard, 1981; Johnson, 2000）。因此假设总体是异质的个体导向方法为破解总体异质性难题提供了一种思路。

三、个体导向的常用统计方法

（一）截面数据中个体导向的常用统计方法

个体导向方法的思想有着古老的根源和基本形式，通常被称为“类型学方法”（Bergman et al., 2009），也就是要对总体进行分类或分组。谢宇（2012）把变异性、社会分组和社会情境作为社会科学研究的三个基本定理，其中变异性是社会科学研究的真正本质，

这是研究的理论指导，而具体的操作通过社会分组减少组内差异，增加组间差异，这样能够减少总体异质性和内生性问题所带来的偏差。在定量社会研究中，目前已经发展出多个将异质性总体分类成不同类型的子群体的个体导向的统计方法，截面数据中常用的方法包括传统的聚类分析和基于模型的聚类分析，后者包括潜类别分析（Latent Class Analysis, LCA）和潜剖面分析（Latent Profile Analysis, LPA）。

传统的聚类分析通过识别特定变量上的相似值来对个体进行分组，主要方法包括分层聚类分析和非分层聚类分析。群组数量通过最小化组内差异和最大化组间差异来确定（Mun et al., 2008）。但是聚类分析具有一定的局限性：首先，最终模型的选择没有可供比较的统计标准指标，因此分类结果往往比较主观（Meyer et al., 2013）；其次，模型也没有考虑对分类结果比较敏感的测量误差（Bergman & Andersson, 2010）；再次，聚类分析有很强的假设，即条件独立性和类不变方差（Morin et al., 2011）；最后，因为传统的聚类分析没有基于模型，无论异质性子群体是否真的存在，都可以从数据中识别出子群体，但基于模型的聚类分析（LCA 和 LPA）能够解决这种缺陷（Jobe-Shields et al., 2015）。

LCA 主要应用于二分类变量指标，描述一组观测到的类别变量或指标的概率在未观察到组成员身份的子群体之间如何变化，将未观察到的子群体作为潜在类别。LCA 的目标是寻找能描述一组观察到的分类变量之间关系的最小数量的潜在类别，模型参数是存在于每个类中的概率和满足给定类成员身份的条件概率，潜类别模型还估计了每个个体的潜在类别概率，即后验概率（Muthén & Muthén, 2000）。另外一个与 LCA 方法建模过程比较相似的是 LPA，其应用

于连续性的观测变量，有学者认为没必要对LPA和LCA进行严格的区分，因为它们都是基于一个潜在变量混合模型，只是针对两种特征不同的变量使用（Pastor et al., 2007）。相比于传统的分类方法，LCA和LPA的优势在于使用最大似然估计、迭代运行和模型比较来确定观测变量的不同子群体和由观测变量组成的潜在变量，最优数量的潜在类别的确定是基于比较客观的模型与数据之间的统计拟合指数，包括Akaike信息标准（Akaike Information Criterion, AIC）、贝叶斯信息标准（Bayesian Information Criterion, BIC）、调整样本规模的贝叶斯信息标准（Sample-Size Adjusted Bayesian Information Criterion, SABIC）等（O'Donnell et al., 2017），这也将有助于研究者评估给定模型的统计价值以及在给定样本中使用个体导向方法的优点（Muthén, 2002）。

（二）追踪数据中个体导向的常用统计方法

利用追踪数据绘制和解释发展轨迹是社会科学领域中基础和重要的研究主题，而大多数传统的发展轨迹的统计方法主要以总体平均趋势来解释个体差异，最常用的是增长曲线模型（Growth Curve Modeling, GCM），这也是常用的变量导向方法。GCM假设所有个体的增长轨迹能够通过单一的增长参数来表示，即假定所有个体只遵循一种平均水平轨迹，子群体之间是同质的。传统的GCM并不能够识别和预测总体中未观察到的子群体，而子群体的不同潜在轨迹可能是检验理论或解释与先前研究结论不一致的焦点，因此研究总体内部的子群体轨迹是必要的，个体同质性假设的应用显然是不现实的（Muthén, 1989；Wang & Bodner, 2007）。鉴于此，以个体导向方法识别异质性总体中同质性子群体的发展轨迹显得尤为重

要，目前组基轨迹模型（Group-Based Trajectory Modeling, GBTM）与增长混合模型（Growth Mixture Modeling, GMM）成为最常用也是影响最大的处理总体异质性增长的模型（王孟成等，2014）。

组基轨迹模型又称“潜类别增长分析”（Latent Class Growth Analysis, LCGA）（Muthén & Muthén, 2000），该模型能够较好地拟合异质性的总体，它的主要目的是探索识别总体中不同发展趋势的子群体，并确定各个子群体的平均发展轨迹，该模型详细的基本原理请参见纳金的专著（Nagin, 2005）。组基轨迹模型的操作步骤主要包括确定结果变量的函数分布的特定形式（分为删截分布、泊松分布和逻辑斯谛分布）、估计并确定组数和轨迹形状、计算组成员身份概率、计算组成员身份后验概率、诊断模型充分性；另外，还可以通过连接组成员身份与协变量、对估计本身添加协变量、双轨迹分析和多轨迹分析等来对组基轨迹模型进行扩展应用（Nagin, 2005；Jones & Nagin, 2007）。增长混合模型又被称为“潜在增长混合模型”（Latent Growth Mixture Model, LGMM），该方法放松了增长曲线模型中的同质性假设，允许在未观测到的子群体之间存在参数差异，通过使用潜在轨迹类来实现（Muthén, 2004），它不考虑单个类别的平均增长曲线周围的个体变化，而是允许不同类别的个体围绕不同的平均增长曲线变化。

GBTM是GMM的一种特殊类型，但二者的统计差异在于个体变异性，即组间和组内变异性。GBTM拟合每一类的平均轨迹水平，但忽略组内变异性，而GMM结合了传统的GCM和GBTM的特点，既可以估计每类的平均增长曲线，也可捕捉这些增长曲线周围的个体差异（Muthén & Muthén, 2000；Frankfurt et al., 2016）。并且相比于GBTM，GMM的主要优势在于能够以较少的类来建模，但

是其也存在一定的劣势，如GMM模型的复杂性很难将其扩展到其他类型数据的模型（如计数型数据）。另外，虽然一个组的参数定义能够表示组的平均轨迹，也能够表示组成员关于这个平均轨迹的可变性，但是这些参数过于复杂（Nagin, 2005），最终模型的收敛结果在统计上可能是不合理的，甚至根本不收敛（Meyer & Morin, 2016）。相比于GMM，GBTM分析使用的参数更少，计算速度更快，分类准确性更高，结果也更容易解释，因此在研究中，GBTM是更为实际的选择（Frankfurt et al., 2016）。

四、变量导向和个体导向方法统计应用的实例比较

（一）截面数据中的实例分析比较

1. 数据来源

这一部分将利用具体的实例比较个体导向方法和变量导向方法在研究问题以及分析过程中的区别。分析所采用的数据来自中国健康与养老追踪调查（China Health and Retirement Longitudinal Study, CHARLS），基线调查于2011年开展，采用分层（依据区县的人均GDP）多阶段（县/区—村/社区—户）抽样方法，以县区（包括城镇和农村）作为初级抽样单位（PSU），对45岁及以上中老年住户人群进行调查，样本覆盖了不包括西藏在内的中国内地28个省150个县区的450个村、居，并于2013、2015、2018年对基线样本进行第二至四期的追踪调查（赵耀辉等，2019）。此外，2014年CHARLS对中老年人进行了生命历程专项调查，用回溯方法记录受

访者自从出生以来的生活经历。本部分分别采用变量导向和个体导向方法分析中老年人所经历的童年逆境状况，选取父母缺失、虐待等10个童年逆境指标，每个变量指标都被处理为0/1的虚拟变量。使用数据的样本来自2013年和2014年的45岁及以上的中老年人群，剔除了缺失或其他无效样本，最终得到有效样本量13233个。

2. 变量导向方法实例：因子分析

因子分析关注多个变量之间的关系，把这些变量简化为几个因子，然后以这几个因子代替原始变量。通过采用最优斜交法，对10个童年逆境变量进行因子分析，发现有四个因子的特征值大于1，累积方差贡献率约为52.5%，说明提取的这四个因子对原始变量信息解释力并不高（本文只展示方法，并不过分强调理论和统计意义）。分析结果的因子载荷如表1所示，因子1的主要载荷变量有身体虐待、父母肢体冲突和父母经常吵架；因子2的主要载荷变量有丧失父母、相对贫困和挨饿；因子3的主要载荷变量有自评健康较差和出现严重的健康问题；因子4的主要载荷变量有父母长期卧病在床、父母身体残疾/精神失常。根据以上主要载荷的变量含义，可以把这四个因子概括为中老年人群所经历的童年逆境类别的四个维度，分别是身体/情感虐待（因子1）、家庭功能失调（因子2）、个体健康较差（因子3）和父母身心障碍（因子4）。然后，可以计算出个体在每个维度上的因子得分，根据这些因子得分，可以计算出不同特征群体的因子得分平均值。

表1 变量导向的探索性因子分析结果（旋转以后的因子载荷矩阵）

变量	因子1	因子2	因子3	因子4
丧失父母	-0.150	0.644	0.032	-0.133
身体虐待	0.531	0.130	0.039	-0.045
父母肢体冲突	0.774	-0.066	-0.013	0.009
父母经常吵架	0.797	-0.024	-0.002	-0.019
父母长期卧病在床	-0.017	0.069	0.108	0.693
父母身体残疾/精神失常	-0.029	-0.024	-0.074	0.812
相对贫困	0.104	0.680	0.049	0.076
挨饿	0.071	0.585	-0.116	0.106
自评健康较差	0.033	0.073	0.767	-0.092
出现严重的健康问题	-0.015	-0.103	0.775	0.115

注：加粗体表示主要变量的载荷。

3. 个体导向方法实例：LCA

采用与因子分析一样的数据，使用 Mplus8.3（Muthén & Muthén, 1998—2017），通过潜类别分析方法识别出经历相同组合风险的同质性子群体。为了选择最佳拟合模型和分类数量，分别估计了1类至7类的模型结果，最优数量的潜类别的确定基于比较客观的统计拟合指数，主要包括 AIC、BIC、ABIC（O'Donnell et al., 2017）、调整后的罗-门德尔-鲁宾似然比检验（Lo-Mendell-Rubin Adjusted Likelihood Ratio Test, LMR）（Lo et al., 2001）、自举似然比检验（Bootstrap Likelihood Ratio Test, BLRT）（Nylund et al.,

2007) 以及分类准确性指数 Entropy (Ramaswamy et al., 1993), 以上每类模型输出的所有指标如表 2 所示。虽然第 6 类模型的 AIC、ABIC 最小, 并且该分类相对于第 5 类模型仍然有显著的意义 (BLRT 和 LMR 的 p 值小于 0.01)。但是第 5 类的 BIC 最小, 且 Entropy 最大, 分类结果中不存在过小比例, 而第 6 类中有两类小于 4%。因此, 综合而言, 第 5 类的分类准确性更高, 分类结果更有意义。

表 2 中老年人的童年逆境潜类别分析的模型拟合参数

类别 数量	Log(L)	AIC	BIC	ABIC	BLRT(p)	LMR(p)	Entropy
1	-49662.7	99345.3	99420.2	99388.5	—	—	—
2	-48431.8	96905.6	97062.9	96996.1	<0.001	<0.001	0.474
3	-47890.1	95844.3	96084.0	95982.3	<0.001	<0.001	0.575
4	-47760.5	95607.0	95929.1	95792.4	<0.05	<0.05	0.599
5	-47647.6	95403.3	95807.7	95636.1	<0.001	<0.001	0.656
6	-47608.2	95346.5	95833.4	95626.8	<0.01	<0.01	0.637
7	-47588.413	95328.826	95898.101	95656.581	0.212	0.215	0.579

表 3 为中老年人的童年逆境项目响应概率和潜类别比例。结果显示, 大多数人都属于类型 1, 其比例为 60.6%, 经历各个童年逆境的可能性都较低, 该类型被标记为常规型。类型 2 为父母身心障碍型, 其比例为 12.0%, 相较于其他类型, 该类型中童年时期经历父母长期卧病在床或身体残疾/精神失常的可能性较大。类型 3 为家庭功能失调型, 比例约为 16.8%, 其特征为童年时期经历父母死

亡、家庭相对贫困和挨饿的可能性最大。类型4为个体健康差型，其比例为4.1%，其特征为自评童年时期自身的健康状况较差，且曾出现过严重的健康问题的可能性更大。类型5为身体/情感虐待型，比例约为6.5%，相较于其他类型，该类型中童年时期遭受身体虐待和情感虐待（父母肢体冲突和父母经常吵架）的可能性较大。

表3 中老年人的童年逆境项目响应概率和潜类别比例

变量	类型1： 常规型 (60.6%)	类型2： 父母身心 障碍型 (12.0%)	类型3： 家庭贫困 型(16..8%)	类型4： 个体健康 差型 (4.1%)	类型5： 身体/情感 虐待型 (6.5%)
丧失父母	0.14	0.31	0.37	0.31	0.10
身体虐待	0.03	0.08	0.06	0.09	0.32
父母肢体冲突	0.03	0.04	0.02	0.70	0.65
父母经常吵架	0.01	0.00	0.01	0.03	0.58
父母长期卧病在床	0.08	0.69	0.08	0.31	0.29
父母身体残疾/精神失常	0.03	0.30	0.05	0.11	0.14
相对贫困	0.04	0.42	0.70	0.43	0.41
挨饿	0.62	0.87	0.91	0.80	0.80
自评健康较差	0.06	0.12	0.13	1.00	0.23
出现严重的健康问题	0.04	0.18	0.05	0.54	0.16

注：加粗体表示 $p<0.05$ 。

比较变量导向方法和个体导向方法对中老年人群童年逆境的分析过程及其结果,可以发现,二者分析的思路不同,变量导向方法是先对多个变量进行降维,而个体导向方法根据个体在所有变量方面的差异划分同质性子群体。从结果来看,因子分析中的四个因子和潜类别分析中的四个类别(除常规型外)的概念含义基本是相似的。变量导向方法提供的四个因子展现了中老年人群童年逆境的四个维度,可以看作生成新的连续性变量,由于因子分析把多重童年逆境降维成四个因子,忽视了中老年人群经历不同童年逆境的状况以及经历多重组合童年逆境的状况,也不能进一步分析中老年人在这四个维度上的比例和特征。而个体导向方法既提供了不同童年逆境的组合变量,又区分了不同童年逆境的子群体。个体导向方法识别了中老年人群经历5类童年逆境组合的子群体,但并非所有的中老年人都经历了童年逆境,常规型中老年人群经历每种童年逆境的风险都较低。另外,个体导向方法的潜类别分析计算了子群体在每个童年逆境的项目响应概率,这样能够了解每个子群体经历童年逆境种类的细微差异,所以个体导向方法比变量导向方法能提供更加丰富的信息。从这两种方法的实例比较能够清晰地看出,虽然分析都是以变量为基础的,但是变量导向方法的分析目标和结果仍然是变量,而个体导向方法的分析目标和结果是个体或子群体,通过对子群体的识别减少了总体异质性。将总体分类之后,根据研究问题的需要,个体导向方法所识别的子群体也可以赋予变量意义,进而比较各个特征/变量上的子群体状况,还可以继续采用变量导向方法进行机制分析。个体导向方法通过变量组合这种整体的方式来关

注个体，可以检验变量之间复杂的交互作用等（Meyer & Morin, 2016）。

（二）纵向数据中的实例分析比较

1. 数据来源

这一部分实例分析的数据采用的是中国老年健康影响因素跟踪调查（Chinese Longitudinal Healthy Longevity Survey, CLHLS），该项目于1998年进行基线调查，分别于2000年、2002年、2005年、2008年、2011年、2014年和2018年进行了追踪调查，样本涵盖中国22个省/市/自治区，大约覆盖了全国85%的人口。CLHLS数据的年龄自报比较准确可靠，主要健康指标的信度和效度较高，被证明具有较高的质量（Gu, 2008；Zeng et al., 2017）。本部分分析的问题是老年人日常生活自理能力（Activities of Daily Living, ADL）的发展轨迹，选择2002至2014年五期追踪数据，删除5次调查中死亡、失访和缺失样本后，最终选取年龄为65岁及以上的老年人口的有效样本量为1659。问卷中对ADL的测量包括洗澡、穿衣、上厕所、室内活动、控制大小便、吃饭六个指标，如果老年人每项活动需要别人帮助则说明该项ADL受损，六项指标受损的数量取值范围为0~6，数值越大说明老年人日常生活自理能力受损的严重程度越高。以下统计分析模型均采用Mplus8.3（Muthén & Muthén, 1998—2017）统计软件来实现。

2. 变量导向方法实例：GCM分析

根据ADL受损项目这一变量特征，观测样本中出现大量的0值，故采用零膨胀泊松模型的线性增长模型。根据65岁及以上老年人的ADL受损数量，对其发展轨迹进行五次追踪分析，结果如图

1所示,老年人的ADL受损程度越来越严重,尤其是进入70岁以上的群体,其ADL恶化速度加快。并且结果显示,计数部分的斜率增长因子和膨胀部分的斜率因子和截距增长因子均在0.01水平上显著,说明老年人个体在ADL受损数量的初始水平及增长速度上存在显著的个体间差异;计数部分的斜率增长因子和截距增长因子的相关系数为 -0.589 ($p<0.001$),膨胀部分的斜率增长因子和截距增长因子的相关系数为 -0.911 ($p<0.001$),说明老年人个体ADL受损数量初始状态越低,其恶化的速度则越快。根据研究需要,增长曲线模型可以加入协变量,进一步分析其对初始ADL受损数量以及ADL受损数量发展轨迹随时间变化的影响,既能够描述出单一特征或不同特征的老年人的ADL受损数量的发展轨迹,也可以展示出老年人ADL受损数量与其他相关结果构成的双轨迹增长模型,这里仅作方法比较的简单实例介绍,更加复杂的应用分析详见其他文献(如Liu, 2007)。

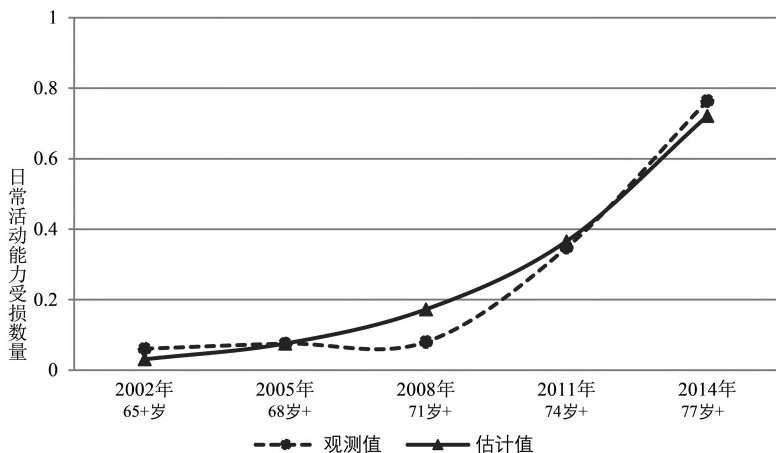


图1 老年人ADL受损数量平均发展轨迹

3. 个体导向方法实例：GBTM 分析

根据 ADL 变量的特征，采用零膨胀泊松分布函数形式的 GBTM 分析。GBTM 实质应用于纵向数据，进行不同群体增长趋势的潜类别分析，在选择最佳拟合模型和分类数量方面，同以上截面数据的潜类别分析方法的过程相似。本文分别估计了 1 类至 6 类的模型结果，其拟合统计指标如表 4 所示，虽然分类越多，AIC、BIC 和 ABIC 值的值越小，但第 4 类之后的分类的模型的 BLRT 和 LMR 的 p 值已经大于 0.05，说明分类结果不再有显著的意义，相较于其他类，第 4 类的分类准确性指数 Entropy 较低，并且存在类型所占比例低于 3% 的情况。因此，综合而言，第 3 类是最佳选择，其在具有统计显著意义中的分类结果中的 AIC、BIC 和 ABIC 值最小，分类准确性又相对较高。

表 4 不同分类的 GBTM 拟合参数

类别 数量	Log(L)	AIC	BIC	ABIC	BLRT (p)	LMR (p)	Entropy
1	-4297.8	8601.6	8617.8	8608.3	—	—	—
2	-3921.7	7855.3	7887.8	7868.7	<0.000	<0.000	0.707
3	-3816.6	7651.3	7700.0	7671.4	<0.000	<0.000	0.674
4	-3794.3	7612.6	7677.5	7639.4	0.009	0.010	0.489
5	-3779.6	7589.1	7670.3	7622.7	0.500	0.500	0.679
6	-3766.5	7569.1	7666.5	7609.4	0.707	0.710	0.487

图 2 展示了样本观测值中不同群体的老年人 ADL 受损数量的平均发展轨迹和模型的发展轨迹，这三类群体的发展轨迹特征呈现出

明显的差异。在样本观测值中,类型1的老年人群体ADL受损程度的起点最高,并且在波动增加,到2008年(即71岁以上的群体)以后,其ADL恶化程度急速加重;到2014年,即进入77岁以上的老年人群体,他们的ADL受损的平均数量已经高于2.5,这一群体的成员占总体的8.5%。类型2的老年人群体的ADL起点基本是完好的,随着时间的推移,ADL受损程度在缓慢上升,直到2008年以后,即进入71岁以上群体,其ADL受损水平加速上升,到2014年达到接近类型1的老年人ADL受损水平,这一类型的老年人比例约为8.7%。类型3的老年人ADL受损程度一直维持着较低水平,这类群体老年人口数量最多,占比约为82.3%。根据样本观测值,我国65岁及以上老年人群体中ADL恶化速度较快、受损程度较高的人口占17.2%(即类型1和类型2比例之和),但根据模型估计值,这一比例高达26.3%。

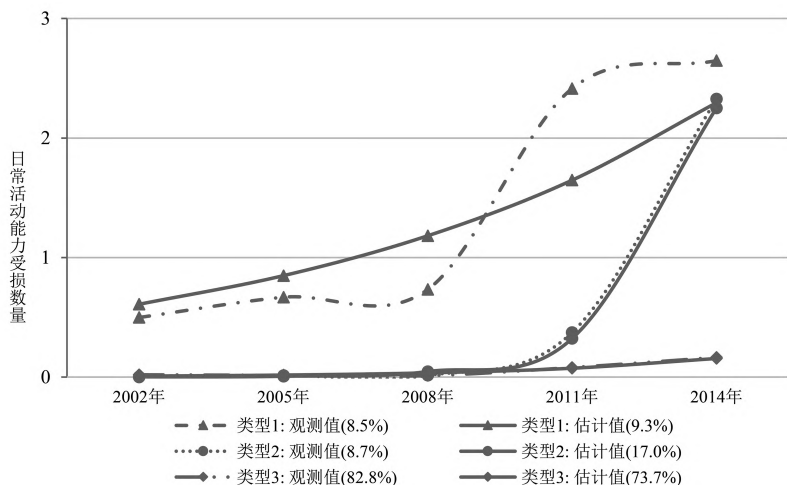


图2 三个潜类别群体的老年人ADL受损数量平均发展轨迹

以上通过两种方法对同一研究问题进行分析比较,可以看出,以变量为导向的增长曲线模型描述了老年人ADL受损数量随时间变化的平均发展轨迹,虽然能够检验总体中老年人ADL受损数量的发展轨迹是否存在个体差异,分析个体起始状态和增长速度差异及二者的关系,但却忽视了总体中的个体或同质性子群体的发展轨迹差异,并不能展示个体或子群体发展轨迹差异的具体形态。采用个体导向的统计方法GBTM,识别出老年人ADL受损的发展轨迹存在显著差异的3个子群体,并且估计出子群体的平均发展轨迹,该方法能够直观地展示出老年人总体中不同的子群体ADL随时间变化的受损程度和发展轨迹的特征差异。比较这两种方法的结果,变量导向方法的平均发展轨迹掩盖了总体中的群体异质性,个体导向方法识别的三个子群体的发展轨迹与总平均轨迹差异较大,类型1和类型2的老年人在77岁以上群体中的ADL受损数量平均超过了2,远高于总平均数量,并且类型3的老年人ADL受损情况在五次追踪期间一直维持着较低水平。探究这三个老年人子群体的ADL受损数量的发展轨迹的特征差异,有利于实施有针对性的健康干预政策。

五、个体导向的统计分析方法的挑战

(一) 解决总体异质性问题的有限性

总体异质性被邓肯认为是最难克服的问题,如何处理总体异质性成为量化社会科学和统计学研究所面临的最严峻挑战(Xie, 2007)。虽然个体导向方法以总体异质性为假设,对于解决社会科学中的总体异质性这一挑战以及减少由总体异质性造成的因果推断

的内生性问题具有重要作用，但该方法并不能彻底解决总体异质性问题。具体而言，个体导向分析是利用模型把总体分类成子群体，这有利于解决静态总体异质性和动态异质性中的不同发展轨迹，但对于随时间变化的样本选择性的总体异质性是无法处理的。个体在经历某一事件上的风险不同，人口学称其为“虚弱度”（Frailty），不同虚弱度的个体从总体中逐渐退出，留下更为强壮的个体，最后的异质性总体的构成可能与原来最初的总体有质的不同（Vaupel et al., 1979）。因此，即使使用个体导向方法对总体发展轨迹进行分类，也不能够解决动态中的这种异质性，因为所分析的样本已经是经过异质性总体选择之后的样本。例如，对老年健康的横截面研究中，研究对象仅限于已经进入老年的人群，他们已经通过了生命不同阶段的不同虚弱度的个体选择，那些比较虚弱的个体可能在进入老年之前便已经死亡，而最终进入观察期的老年人反而更健康，基于该群体的个体导向方法所分析的老年健康便是这种已经经过健康或死亡选择过的样本。另外，基于追踪数据，利用个体导向方法能够识别出老年人的健康发展轨迹类型同样会面临总体异质性的问题，因为追踪时期内不同虚弱度的老年人从样本中退出，如果不考虑退出老年人的健康状况，最终对幸存老年人的健康发展轨迹的识别将是有偏的。鉴于此，未来在追踪数据中使用个体导向方法对不同群体的发展轨迹进行识别，理想的情况是把该方法应用于那些死亡率较低的年轻人群（Zimmer et al., 2012），以及对某一人群的某一特征的分析应尽可能不受或较少受“虚弱度”的影响。

（二）发展轨迹分析中时间维度的混淆性

个体导向方法被越来越多地运用到社会科学研究中，尤其是分

析动态发展状况。但利用追踪数据对个体发展轨迹进行分析存在重大的操作困境，即时间间隔的意义和有效性。时间对追踪数据中发展轨迹的拟合具有较大影响，当个体在第一次调查时的年龄存在较大差异或个体间测量时间间隔不固定时，估计结果是有偏的，不考虑年龄异质性，会把年龄效应、时期效应和队列效应完全混淆（Meyer & Morin, 2016），尤其是时间变动对研究对象的结果变量产生较大影响时更需要注意这一问题。例如，对于老年健康的研究，国内外一些学者采用GBTM方法对我国老年健康的发展轨迹进行分析，对于时间的处理主要有两种方式。第一种是以调查波次作为发展轨迹建模的时间指标，关注一定时间内老年健康的发展，如本文纵向数据分析实例与其他文献（巫锡炜，2009）。第二种是以年龄为时间指标，关注不同年龄的老年人的健康状况（胡晓茜等，2019；Zimmer et al., 2012）。第一种模型虽然能够反映出该时期的研究对象的发展轨迹，但其时间指标混杂了年龄、时期和队列效应。第二种情况的时间指标并不是真正意义上的时间，分析中的年龄是所有波次追踪样本的年龄整合，并不是按照第一次调查的特定年龄老年人群体的追踪下的年龄，这种时间指标的处理会产生比第一种情况更大的风险。首先，各年龄（尤其是研究对象有更宽的年龄区间）的总样本量以及进行分类之后的子群体样本量比波次的样本量少很多，会影响分类结果的稳定性。其次，其仍然混淆了年龄、时期和队列效应，并且不同时间间隔上的老年健康状况是整个追踪时期的平均值，只见横向的时间（不同年龄）上的健康状况，不见纵向的（不同时期）健康状况，并没有发挥出追踪数据的个体发展轨迹的动态优势，也就失去了发展轨迹的真正意义。另外，追踪数据的波次越多，时间越长，时期和队列效应越强，不同出生队

列的老年人在相同年龄的时代背景具有较大的差异性，调查间隔时间越长的老年人的健康状况的变化越大，因此不同调查时点所汇集样本中的相同年龄的老年人健康状况并没有可比性，估计结果对正确认识老年人的健康发展轨迹的意义不大。因此，未来在这一方面的研究，应尽可能分离出不同队列的样本，并对其发展轨迹进行估计，或采用多波次追踪的年龄间隔比较窄或同一年龄的人群，这样才能使研究中的时间指标更有意义。

（三）其他方面的问题

以往有学者总结个体导向分析方法在具体应用中面临的三个主要的问题，即确定类别或亚组的效度、个体导向的理论原则在个体导向数据分析方法中的检验、个体导向方法与变量导向方法的选择与整合（纪林芹、张文新，2011）。前文笔者提出个体导向方法在解决总体异质性以及处理时间维度方面存在一定的挑战。除此以外，个体导向方法还面临一些其他的问题。在理论方面，虽然个体导向方法的假设更加符合社会现实，如异质性、个体独特性、整体性、相互作用的动态发展等，也期望能够关注到异质性总体中的个体状况，但将其应用于统计分析中，仅是将总体分成有限的子总体，最终分析结果只不过是模式或类型层面上的“平均值”，仍然不能关注到个体。另外，这些现实性的假设更为复杂，意味着用于拟合观测数据的数学模型需要估计更多的参数，那么模型拟合将变得更加困难，这也限制了个体导向方法在社会研究中的统计应用。个体导向的具体统计应用也并非完美。首先，个体导向的一些统计分析方法仍存在一定主观性，如传统的聚类分析没有基于统计标准指标，分类结果主观，GBTM的分析过程有一些指导模型选择

的经验规则和衡量模型充分性的客观标准 (Nagin, 2005),但是在确定轨迹类型和形状时仍然涉及较多的主观判断,这也导致了所选择的模型可能无法很好地符合现实。例如,在本文GBTM实例分析中,根据模型拟合统计指标将总体分成3个类别看似是最优选择,但图2显示出类型1的估计值曲线与观察值曲线拟合得并不是很好。如果要继续调试出更好的结果,研究者可能需要考虑测量指标(本文是童年逆境)数量或组合、减少类型数量等方面,这种调试仅是追求高度拟合指标,却失去了理论和现实意义。这正反映了个体导向方法的主观性和类别确定的困难。其次,个体导向的统计分析方法仍然需要假定条件,如GBTM需要假定条件独立性、组间同方差等,违反这些假设可能会导致模型效果不佳 (Davies et al., 2017)。再次,个体导向方法需要足够多的样本量,较小的样本量可能会影响到多个变量的分类结果和发展轨迹类型划分的稳定性。最后,仍需要强调的是,个体导向方法所分析的变量间关系并非因果关系,更多的是一种对总体内部的特征描述,并不具有对事物发展变化的预测作用。

六、结语

定量社会研究中的变量导向方法已经取得巨大的进步,但由于社会科学比自然科学的运行机制更复杂,使用这种研究范式所造成的问题-方法不匹配的问题越来越明显 (Bergman et al., 2009)。在研究中所采用的个体集合数据可能会造成更多的风险,如高阶交互作用的存在可能要对个体导向方法中的低阶交互作用有所限定,总体概率可能与特定群体的概率相矛盾,以及出现生态谬误等,这将

使个体导向分析成为必然 (von Eye & Bogat, 2006)。尤其是在追踪数据分析中, 变量导向方法对个体发展轨迹的描述显得苍白无力。异质性总体由同质性的子群体组成, 所观察到的总体人口水平的动态发展将可能偏离子群体中个体层面的潜在模式, 掩盖了子群体的差异性, 甚至可能误导我们对某一结果的事实判断 (Vaupel et al., 1979), 几个子群体的线性趋势可能会形成整个群体的曲线趋势。因此, 揭示子群体发展轨迹模式有助于了解现象背后的本质, 有利于构建或检验理论, 并且对正确预测政策干预的影响有重要作用 (Vaupel & Yashin, 1985)。而异质性总体中各个子群体遵循不同的发展轨迹, 运用变量导向方法 (某些个体特征, 如性别、社会经济地位等变量) 很难识别不同发展轨迹, 也不容易估计某些事件的干预是否改变个体的发展轨迹类型, 因此以个体导向方法来识别不同发展轨迹类型对于减少总体异质性发挥着重要作用。

虽然近些年来兴起的个体导向方法出于自身独特的优势不断发展, 但由于只能用于描述性分析, 进一步深入探讨影响机制受到限制。因此, 个体导向方法与变量导向方法同等重要, 二者并不是竞争、对立、二分的关系, 而是互补关系, 应该根据研究问题, 充分发挥二者优势。在实践研究中, 也可以将这两种方法进行融合, 通常有三种研究思路。第一, 利用个体导向方法为定量导向方法提供新的潜变量, 比如首先通过个体导向方法把异质性总体分类成有意义的子群体, 然后采用变量导向方法分析子群体之间的变量关系、子群体特征以及其他因素对某一结果变量的影响等。第二, 利用个体导向方法可以通过分析个体间关系来对多个变量进行组合, 然后利用变量导向方法探讨高阶交互作用 (Bergman & Wangby, 2014)。第三, 个体导向的统计分析方法为变量导向统计分析方法的发展完

善提供了新思路,对因果推断的改进起到了重要作用。例如,将GBTM和倾向值匹配方法相结合,通过对轨迹群的分层分析,使观测数据的关键协变量更加平衡(Haviland & Nagin, 2007),将GBTM引入中断时间序列分析,通过比较各个轨迹组的结果,识别可以比较的未干预单位,从而能够在中断时间序列分析结果中进行控制,帮助改善因果推断(Linden, 2018)。同时应用这两种方法能够有效突破数据和统计模型的障碍,成为未来研究中越来越重要的趋势。

参考文献

- 胡晓茜、高奇隆、赵灿等, 2019,《中国高龄老人失能发展轨迹及死亡轨迹》,《人口研究》第5期。
- 纪林芹、张文新, 2010,《发展心理学研究中个体定向的理论与方法》,《心理科学进展》第11期。
- 刘亚飞, 2018,《童年饥饿经历会影响老年健康吗?》,《经济评论》第6期。
- 石智雷、吴志明, 2018,《早年不幸对健康不平等的长远影响: 生命历程与双重累积劣势》,《社会学研究》第3期。
- 王孟成、毕向阳、叶浩生, 2014,《增长混合模型: 分析不同类别个体发展趋势》,《社会学研究》第4期。
- 巫锡炜, 2009,《中国高龄老人残障发展轨迹的类型: 组基发展建模的一个应用》,《人口研究》第4期。
- 谢宇, 2012,《社会学方法与定量研究》,北京: 社会科学文献出版社。
- 赵耀辉、王亚峰、陈欣欣等, 2019,《中国健康与养老报告》,北京: 北京大学。
- Abbott, Andrew 1988, "Transcending General Linear Reality." *Sociological Theory* 6(2).
- Anda, Robert F. Bendahan, Alexander Butchart, Vincent J. Felitti, David W. Brown 2010, "Building a Framework for Global Surveillance of the Public Health Implica-

- tions of Adverse Childhood Experiences.” *American Journal of Preventive Medicine* 39(1).
- Barboza, Gia Elise 2018, “Latent Classes and Cumulative Impacts of Adverse Childhood Experiences.” *Child Maltreatment* 23 (2).
- Bergman, Lars R. 2001, “A Person Approach in Research on Adolescence: Some Methodological Challenges.” *Journal of Adolescent Research* 16 (1).
- Bergman, Lars R. & David Magnusson 1997, “A Person-Oriented Approach in Research on Developmental Psychopathology.” *Development and Psychopathology* 9 (2).
- Bergman, Lars R. & Kari Trost 2006. “The Person-Oriented Versus the Variable-Oriented Approach: Are they Complementary, Opposites, or Exploring Different Worlds?” *Merrill-Palmer Quarterly* 52 (3).
- Bergman, Lars R. & Margit Wångby 2014, “The Person-Oriented Approach: A Short Theoretical and Practical Guide.” *Eesti Haridusteaduste Ajakiri. Estonian Journal of Education* 2 (1).
- Bergman, Lars R., David Magnusson & Bassam M. El-Khoury 2003, *Studying Individual Development in an Interindividual Context: A Person-Oriented Approach (Vol. 4)*. Mahwah: Psychology Press.
- Bergman, Lars R. & Håkan Andersson 2010, “The Person and the Variable in Developmental Psychology.” *Journal of Psychology* 218 (3).
- Bergman, Lars R., Henrik Andershed & Anna-Karin Andershed 2009, “Types and Continua in Developmental Psychopathology: Problem Behaviors in School and their Relationship to Later Antisocial Behavior.” *Development and Psychopathology* 21 (3).
- Block Jack 1971, *Lives through time*. Berkeley. CA: Bancroft Books.
- Bogat, G. A., A. von Eye & L. R. Bergman 2016, “Person-Oriented Approaches.” In D. Cicchetti (ed.), *Developmental Psychopathology (vol. 1)*. Hoboken: John Wiley & Sons, Inc.
- Davies, Christopher E., Gary F. V. Glonek & Lynne C. Giles 2017, “The Impact of Covariance Misspecification in Group-Based Trajectory Models for Longitudinal Data with Non-Stationary Covariance Structure.” *Statistical Methods in Medical Research* 26 (4).
- Frankfurt, Sheila, Patricia Frazier, Moin Syed & Kyoung Rae Jung 2016, “Using

- Group-Based Trajectory and Growth Mixture Modeling to Identify Classes of Change Trajectories.” *The Counseling Psychologist* 44 (5).
- Gu, Danan 2008, “General Data Quality Assessment of the CLHLS.” In Z. Yi, D. L. Poston, D. A. Vlosky & D. Gu (ed.) *Healthy Longevity in China* (Vol. 20). Dordrecht: Springer Netherlands.
- Haviland, Amelia M. & Daniel S. Nagin 2007, “Using Group-Based Trajectory Modeling in Conjunction with Propensity Scores to Improve Balance.” *Journal of Experimental Criminology* 3 (1).
- Jobe-Shields, Lisa, Arthur R. Andrews, Gilbert R. Parra & Natalie A. Williams 2015, “Person-Centered Approaches to Understanding Early Family Risk.” *Journal of Family Theory & Review* 7 (4).
- Johnson, Nan E. 2000, “The Racial Crossover in Comorbidity, Disability, and Mortality.” *Demography* 37(3).
- Jones, Bobby L. & Daniel S. Nagin 2007, “Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating them.” *Sociological Methods & Research* 35 (4).
- Kelso, J. A. Scott 2000, “Principles of Dynamic Pattern Formation and Change for a Science of Human Behaviour.” In L. R. Bergman, R. B. Cairns, L.-G. Nilsson, & L. Nystedt (ed.), *Developmental Science and the Holistic Approach*. Mahwah: Erlbaum.
- Laursen, Brett & Erika Hoff 2006, “Person-Centered and Variable-Centered Approaches to Longitudinal Data.” *Merrill-Palmer Quarterly* 52 (3).
- Linden, Ariel 2018, “Using group-based Trajectory Modelling to Enhance Causal Inference in Interrupted Time Series Analysis.” *Journal of Evaluation in Clinical Practice* 24 (3).
- Liu, Hui 2007, “Growth Curve Models for Zero-Inflated Count Data: An Application to Smoking Behavior.” *Structural Equation Modeling: A Multidisciplinary Journal* 14 (2).
- Lo, Yungtai, Nancy R. Mendell & Donald B. Rubin 2001, “Testing the Number of Components in a Normal Mixture.” *Biometrika* 88(3).
- Magnusson, David 2003, “The Person Approach: Concepts, Measurement Models, and Research Strategy.” *New Directions for Child and Adolescent Development* 2003 (101).

- Manton, Kenneth G. & Eric Stallard 1981, "Methods for Evaluating the Heterogeneity of Aging Processes in Human Populations using Vital Statistics Data: Explaining the Black/White Mortality Crossover by a Model of Mortality Selection." *Human Biology* 53 (1).
- Meyer, John P. & Alexandre J. S. Morin 2016, "A person-centered Approach to Commitment Research: Theory, Research, and Methodology." *Journal of Organizational Behavior* 37 (4).
- Meyer, John P., Laura J. Stanley & Robert J. Vandenberg 2013, "A Person-Centered Approach to the Study of Commitment." *Human Resource Management Review* 23 (2).
- Morin, Alexandre J. S., Aleksandra Bujacz & Marylène Gagné 2018, "Person-Centered Methodologies in the Organizational Sciences: Introduction to the Feature Topic." *Organizational Research Methods* 21 (4).
- Morin, Alexandre J. S., Julien Morizot, Jean-Sébastien Boudrias & Isabelle Madore 2011, "A Multifoci Person-Centered Perspective on Workplace Affective Commitment: A Latent Profile/Factor Mixture Analysis." *Organizational Research Methods* 14 (1).
- Mun, Eun Young, Michael Windle & Lisa M. Schainker 2008, "A Model-Based Cluster Analysis Approach to Adolescent Problem Behaviors and Young Adult Outcomes." *Development and Psychopathology* 20 (1).
- Muthén, Bengt & Linda K. Muthén 2000, "Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling with Latent Trajectory Classes." *Alcoholism, Clinical and Experimental Research* 24 (6).
- Muthén, Bengt O. 1989, "Latent Variable Modeling in Heterogeneous Populations." *Psychometrika* 54 (4).
- 2002, "Beyond Sem: General Latent Variable Modeling." *Behaviormetrika* 29 (1).
- 2004, "Latent Variable Analysis: Growth Mixture Modeling and Related Techniques for Longitudinal Data." In D. Kaplan (ed.), *Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks: Sage.
- Muthén, Linda K. & Bengt O. Muthén 1998–2017, *Mplus User's Guide*. Los Angeles: Muthén & Muthén.
- Nagin, Daniel 2005, *Group-Based Modeling of Development*. Cambridge: Harvard University Press.

- Nylund, Karen L., Tihomir Asparouhov & Bengt O. Muthén 2007, “Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study.” *Structural Equation Modeling* 14 (4).
- O'Donnell, Meaghan L., Ingo Schaefer, Tracey Varker, Dzenana Kartal, David Forbes, Richard A. A. Bryant & Derrick Silove, et al. 2017, “A Systematic Review of Person-Centered Approaches to Investigating Patterns of Trauma Exposure.” *Clinical Psychology Review* 57.
- Pastor, Dena A., Kenneth E. Barron, B. J. Miller & Susan L. Davis. 2007. “A Latent Profile Analysis of College Students' Achievement Goal Orientation.” *Contemporary Educational Psychology* 32 (1).
- Ramaswamy, Venkatram, Wayne S. Desarbo, David J. Reibstein & William T. Robinson. 1993. “An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data.” *Marketing Science (Providence, R.I.)* 12 (1).
- Schmitz, Bernhard 2000, “Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen [Searching for the Lost Individual: Four Theorems on the Aggregation of Processes].” *Psychologische Rundschau* 51(2).
- Simpson, E. H. 1951, “The Interpretation of Interaction in Contingency Tables.” *Journal of the Royal Statistical Society (Series B, Methodological)* 13 (2).
- Sterba, Sonya K. & Daniel J. Bauer 2010, “Matching Method with Theory in Person-Oriented Developmental Psychopathology Research.” *Development and Psychopathology* 22 (2).
- Vaupel, James W. & Anatoli I. Yashin 1985, “Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics.” *American Statistician* 39 (3).
- Vaupel, James W., Kenneth G. Manton & Eric Stallard 1979, “The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality.” *Demography* 16 (3).
- von Eye, Alexander & G. Anne Bogat 2006, “Person-Oriented and Variable-Oriented Research: Concepts, Results, and Development.” *Merrill-Palmer Quarterly* 52 (3).
- von Eye, Alexander & Lars R. Bergman 2003, “Research Strategies in Developmental Psychopathology: Dimensional Identity and the Person-Oriented Approach.” *Development and Psychopathology* 15 (3).
- Wang, Mo & Todd E. Bodner 2007, “Growth Mixture Modeling: Identifying and Predicting Unobserved Subpopulations with Longitudinal Data.” *Organizational Re-*

search Methods 10 (4).

Xie, Yu 2007, "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." *Research in Social Stratification and Mobility* 25 (2).

—2013, "Population Heterogeneity and Causal Inference." *Proceedings of the National Academy of Sciences of the United States of America* 110 (16).

Zeng, Yi, Qiushi Feng, Therese Hesketh, Kaare Christensen & James W. Vaupel 2017, "Survival, Disabilities in Activities of Daily Living, and Physical and Cognitive Functioning among the Oldest-Old in China: A Cohort Study." *The Lancet (British Edition)* 389 (10079).

Zimmer, Zachary, Linda G. Martin, Daniel S. Nagin & Bobby L. Jones 2012, "Modeling Disability Trajectories and Mortality of the Oldest-Old in China." *Demography* 49 (1).