

• 研究方法(Research Method) •

密集追踪数据的中介效应分析*

方 杰¹ 温忠麟² 董育铭³ 王晓洁³

(¹广东财经大学新发展研究院, 广州 510320) (²华南师范大学心理学院/心理应用研究中心, 广州 510631)
(³广东财经大学经济学院, 广州 510320)

摘 要 随着密集追踪数据在社科领域的广泛运用, 如何对密集追踪数据进行中介效应分析吸引了诸多研究者的注意。如果还是按通常追踪数据一样对待, 采用多水平模型和多水平结构方程模型进行中介效应分析, 则既忽略了变量之间的先后顺序, 也无法探究变量之间动态变化的关联。本文以 1-1-1 密集追踪中介模型为例, 详述了基于多水平自回归模型(MAM)及其变式(残差 MAM)、动态结构方程模型(DSEM)及其变式(残差 DSEM、交叉分类的 DSEM)的密集追踪中介效应分析方法, 并总结出一个分析流程。用示例演示如何进行密集追踪数据的中介效应分析, 并给出了相应的 Mplus 和 R 程序。最后讨论了密集追踪数据的中介效应分析的拓展方向。
关键词 密集追踪数据, 中介效应, 多水平自回归模型, 动态结构方程模型, 去趋势
分类号 B841

中介(mediation)是社会科学研究中重要的方法学概念。如果自变量 X 通过某一变量 M 对因变量 Y 产生一定影响, 则称 M 在 X 和 Y 之间起中介作用, 此时称 M 为中介变量。近 30 年来, 中介分析方法和实际应用都得到长足发展。虽然截面数据(cross-sectional data)和追踪数据(也称为纵向数据, longitudinal data)都可以进行中介效应分析, 但使用追踪数据做中介效应分析比使用截面数据更受欢迎, 因为追踪数据更易区分变量的先后顺序(即历时性影响, 温忠麟, 2017)。目前所用的追踪数据主要针对纵向面板数据(longitudinal panel data), 测量时间间隔较长的纵向面板数据允许研究者观察长时间间隔的中介效应, 但难以描述变量在短期内动态变化的关联, 这一不足被测量间隔短(如每天测一次)、频次高(至少 20 次)的密集追踪数据(intensive longitudinal data)所弥补(见表 1)。

表 1 密集追踪数据的示例($t = 20$)				
被试(j)	时间(t)	X	M	Y
1	1	0.33	0.43	-1.72
1	2	-0.48	-0.68	-1.61
⋮	⋮	⋮	⋮	⋮
1	20	0.14	0.26	1.05
2	1	0.31	0.68	0.15
2	2	0.46	0.03	-0.94
⋮	⋮	⋮	⋮	⋮
2	20	-0.86	0.51	-0.56

注: 数据源自 McNeish & Hamaker, 2020

密集追踪数据对统计分析模型提出了很大的挑战。密集追踪数据(如 Y_{ij})可看成时间点(t)嵌套于个体(j)的两水平结构(表 1), 因此研究者常用多水平模型(multilevel model, MLM)和多水平结构方程模型(multilevel structural equation model, MSEM)进行中介建模。但多水平模型和多水平结构方程模型都没有考虑变量之间影响的先后顺序, 也就无法探究变量之间动态变化的关联(这正是密集追踪研究关注的核心问题), 并不适合密集追踪数据的中介(以下简称密集追踪中介)效应建模(Zhang et al., 2018; 方杰 等, 2021)。近年来, 针对

收稿日期: 2024-07-02
* 国家自然科学基金项目(32171091)、教育部人文社会科学规划基金项目(24YJA190003)和青年基金项目(23YJC190026)资助。
通信作者: 温忠麟, E-mail: wenzl@scnu.edu.cn

密集追踪数据中介效应建模,研究者将密集追踪数据的多水平结构和变量之间影响的先后顺序相结合,用两类新模型进行中介效应建模。一类是多水平自回归模型(multilevel autoregressive model, MAM; Zhang et al., 2018)及其变式——残差多水平自回归模型(residual multilevel autoregressive model, RMAM)。另一类是动态结构方程模型(dynamic structural equation model, DSEM)及其变式——残差动态结构方程模型(residual dynamic structural equation model, RDSEM)和交叉分类(cross-classified)的动态结构方程模型(Asparouhov et al., 2018)。

目前,密集追踪数据中介研究较多的是自变量、中介变量和因变量都是密集追踪数据(X_{ij} 、 M_{ij} 和 Y_{ij} ,数据示例见表1)的情况(Zhang et al., 2018; McNeish & MacKinnon, 2022; Fang et al., 2024),记为1-1-1密集追踪中介效应。1-1-1这三个数字依次代表自变量、中介变量和因变量所在的层次(数字1表示变量在层1,数字2表示变量在层2)。1-1-1密集追踪中介模型的优势在于,中介模型的每条路径都存在变量间的时间先后关系,能有效地探究变量之间动态变化的关联。例如,Lischetzke等(2021)收集了313名被试,每名被试重复测量平均18.89次的密集追踪数据,使用多水平结构方程模型考察了每晚的平静(calmness in the evening)在每天的压力(daily stress)和每晚的睡眠质量(nightly sleep quality)之间的1-1-1密集追踪中介效应。本文以1-1-1密集追踪中介效应模型为例,阐述两类模型的密集追踪中介效应的分析方法,并总结出一个分析流程。接着,用示例演示如何进行两类模型的密集追踪中介效应分析,并给出了相应的Mplus程序。最后讨论了密集追踪中介效应分析的拓展方向。

1 多水平自回归模型的密集追踪中介效应分析

1.1 多水平自回归模型

多水平自回归模型是在多水平模型的基础上,加入变量之间影响的先后顺序(例如,图1中的 $X_{(t-2)j} \rightarrow M_{(t-1)j}$),弥补了多水平模型的缺陷。基于多水平自回归模型的1-1-1密集追踪中介效应分析(图1)可表示为(Zhang et al., 2018):

$$\text{层1: } X_{ij} = l_{Xj} + \beta_{Xj}X_{(t-1)j} + \varepsilon_{Xij} \quad (1)$$

$$M_{ij} = l_{Mj} + \beta_{Mj}M_{(t-1)j} + a_jX_{(t-1)j} + \varepsilon_{Mij} \quad (2)$$

$$Y_{ij} = l_{Yj} + \beta_{Yj}Y_{(t-1)j} + b_jM_{(t-1)j} + c'_jX_{(t-1)j} + \varepsilon_{Yij} \quad (3)$$

$$\text{层2: } l_{Xj} = l_X + \mu_{1j} \quad l_{Mj} = l_M + \mu_{2j} \quad l_{Yj} = l_Y + \mu_{3j} \quad (4)$$

$$\beta_{Xj} = \beta_X + \mu_{4j} \quad \beta_{Mj} = \beta_M + \mu_{5j} \quad \beta_{Yj} = \beta_Y + \mu_{6j} \quad (5)$$

$$a_j = a + \mu_{7j} \quad b_j = b + \mu_{8j} \quad c'_j = c' + \mu_{9j} \quad (6)$$

其中, l_{Xj} 、 l_{Mj} 和 l_{Yj} 表示截距项; β_{Xj} 、 β_{Mj} 和 β_{Yj} 表示自回归效应(autoregressive effect),即同一变量前后时间点的效应(图1)。 a_j 、 b_j 和 c'_j 表示滞后效应(lagged effect),即控制了结果变量的自回归效应后,预测变量对结果变量的历时性影响(图1)。中介效应为 a_jb_j ,可以随个体 j 的变化而变化,即每个个体的中介效应都可能不同。所有个体的中介效应的均值为 $E(a_jb_j)$ 。由于 a_j 和 b_j 可能随个体 j 的变化而发生共变,所以随个体变化的中介效应的均值 $E(a_jb_j)$ 为 $ab + \sigma_{a,b_j}$,其中 σ_{a,b_j} 是 σ_{a_j} 和 σ_{b_j} 的协方差(在数值上等于 a_j 和 b_j 的协方差),随个体变化的直接效应 c'_j 的均值为 c' (Kenny et al., 2003)。 ε_{Xij} 、 ε_{Mij} 和 ε_{Yij} 表示层1残差,多水平自回归模型假设层1残差服从正态分布[即 $\varepsilon_{Xij} \sim N(0, \sigma_X^2)$, $\varepsilon_{Mij} \sim N(0, \sigma_M^2)$, $\varepsilon_{Yij} \sim N(0, \sigma_Y^2)$],且残差方差(σ_X^2 、 σ_M^2 和 σ_Y^2)不随个体变化(没有下标 j)。

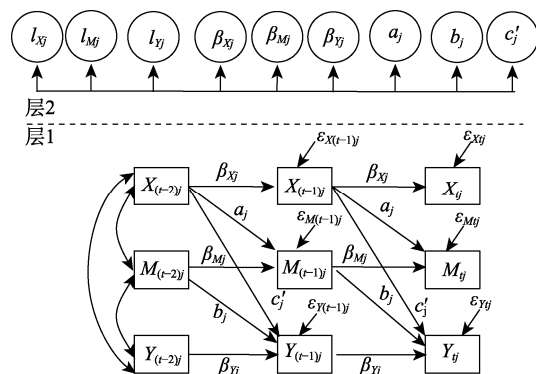


图1 基于多水平自回归模型的1-1-1密集追踪中介模型(改编自Zhang et al., 2018)

另外,在层2,所有层1的参数(l_{Xj} 、 l_{Mj} 、 l_{Yj} 、 β_{Xj} 、 β_{Mj} 、 β_{Yj} 、 a_j 、 b_j 和 c'_j)都可以设为随机效应。也就是说,层1的参数都可以在层2作为因变量,写成均值加残差的形式[即方程(4)~(6)]。其中, l_X 、 l_M 、 l_Y 、 β_X 、 β_M 、 β_Y 、 a 、 b 和 c' 表示所有个体的均

值, $\mu_{1j} \sim \mu_{9j}$ 表示残差, 代表每个个体值偏离均值的程度。设置随机效应就是为了考察参数是否会随个体而变化[见方程(4)~(6)]。当然, 如果研究者假设某些层 1 参数不随个体变化, 也可以将相应的层 1 参数设置为固定效应, 以简化模型。例如, 要将自变量的自回归系数设置为固定效应, 可将方程(5)中的 $\beta_{Xj} = \beta_X + \mu_{4j}$ 删除, 同时将方程(1)的参数 β_{Xj} 改为 β_X (即去掉下标 j) 即可。值得注意的是, 多水平自回归模型虽然允许层 1 参数随个体变化, 但却不允许这些参数随时间变化。多水平自回归模型设定自回归效应(β_{Xj} 、 β_{Mj} 、 β_{Yj})和滞后效应(a_j 、 b_j 、 c'_j)不随时间变化(没有下标 t , 见图 1)的前提条件是重复测量的时间间隔相同。

1.2 残差多水平自回归模型

多水平自回归模型还假设时间序列是平稳(stationarity)的, 即变量的均值、方差以及自相关都不会随着时间发生系统性变化(McNeish & Hamaker, 2020)。但密集追踪数据往往因为存在时间趋势(trend)而违背了平稳性假设, 此时需要对原始数据进行去趋势(detrending)处理。

常用的一种去趋势方法是从变量中剥离随时间的发展趋势(即做每个变量对时间的回归)后, 再利用残差建构密集追踪中介模型(Asparouhov et al., 2018; Fang et al., 2024; McNeish & Hamaker, 2020; 郑舒方等, 2021; 肖悦等, 2024)。本文将这种去趋势的多水平自回归模型称为残差多水平自回归模型。基于残差多水平自回归模型的 1-1-1 密集追踪中介效应分析分两步进行。第一步, 对层 1 密集追踪变量去趋势, 即做每个变量对时间 $time_{ij}$ 的回归[方程(7)~(9), Curran & Bauer, 2011; Wang & Maxwell, 2015],

$$\text{层 1: } X_{ij} = f_{Xj} + D_{Xj}time_{ij} + e_{Xij} \quad (7)$$

$$M_{ij} = f_{Mj} + D_{Mj}time_{ij} + e_{Mij} \quad (8)$$

$$Y_{ij} = f_{Yj} + D_{Yj}time_{ij} + e_{Yij} \quad (9)$$

其中 f_{Xj} 、 f_{Mj} 和 f_{Yj} 表示截距项, D_{Xj} 、 D_{Mj} 和 D_{Yj} 表示变量(X_{ij} 、 M_{ij} 和 Y_{ij})对时间的回归系数, 残差 e_{Xij} 、 e_{Mij} 和 e_{Yij} 表示 X_{ij} 、 M_{ij} 和 Y_{ij} 的方差中不能被时间变量 $time_{ij}$ 解释的部分, 因此 e_{Xij} 、 e_{Mij} 和 e_{Yij} 是去除了趋势的。

第二步, 用残差 e_{Xij} 、 e_{Mij} 和 e_{Yij} 建构基于多水平自回归模型的 1-1-1 密集追踪中介效应模型(Zhang et al., 2018)。具体地, 用残差 e_{Xij} 、 e_{Mij} 和 e_{Yij} 替换方程(1)~(3)的 X_{ij} 、 M_{ij} 和 Y_{ij} , 则方程(1)~(3)

变为,

$$\text{层 1: } e_{Xij} = l_{Xj} + \beta_{Xj}e_{X(i-1)j} + \delta_{Xij} \quad (10)$$

$$e_{Mij} = l_{Mj} + \beta_{Mj}e_{M(i-1)j} + a_j e_{X(i-1)j} + \delta_{Mij} \quad (11)$$

$$e_{Yij} = l_{Yj} + \beta_{Yj}e_{Y(i-1)j} + b_j e_{M(i-1)j} + c'_j e_{X(i-1)j} + \delta_{Yij} \quad (12)$$

其中, l_{Xj} 、 l_{Mj} 和 l_{Yj} 表示截距项。方程(10)~(12)的自回归系数(β_{Xj} 、 β_{Mj} 和 β_{Yj})和滞后系数(a_j 、 b_j 和 c'_j)是不受趋势影响。层 2 方程包括方程(4)~(6)和方程(13)。

$$\text{层 2: } D_{Xj} = D_X + \mu_{43j} \quad D_{Mj} = D_M + \mu_{44j} \quad (13)$$

$$D_{Yj} = D_Y + \mu_{45j}$$

随个体变化的中介效应的均值 $E(a_j b_j)$ 为 $ab + \sigma_{a_j b_j}$, 随个体变化的直接效应 c'_j 的均值为 c' 。残差多水平自回归模型也假设自回归系数和滞后系数不随时间变化, 同时还假设 δ 服从正态分布[即 $\delta_{Xij} \sim N(0, \sigma_X^2)$ 、 $\delta_{Mij} \sim N(0, \sigma_M^2)$ 和 $\delta_{Yij} \sim N(0, \sigma_Y^2)$], 且残差方差(σ_X^2 、 σ_M^2 和 σ_Y^2)不随个体变化。

2 动态结构方程模型的密集追踪中介效应分析

2.1 动态结构方程模型

动态结构方程模型是在多水平结构方程模型的基础上, 加入变量之间影响的先后顺序(见图 2), 弥补了多水平结构方程模型的不足(McNeish & Hamaker, 2020), 基于动态结构方程模型的 1-1-1 密集追踪中介效应分析(图 2)可表示为(Fang et al., 2024):

$$\text{层 1: } X_{ij}^c = \beta_{Xj} X_{(i-1)j}^c + \varepsilon_{Xij} \quad (14)$$

$$M_{ij}^c = \beta_{Mj} M_{(i-1)j}^c + a_j X_{(i-1)j}^c + \varepsilon_{Mij} \quad (15)$$

$$Y_{ij}^c = \beta_{Yj} Y_{(i-1)j}^c + b_j M_{(i-1)j}^c + c'_j X_{(i-1)j}^c + \varepsilon_{Yij} \quad (16)$$

$$\text{层 2: } l_{Xj} = l_X + \mu_{4j} \quad l_{Mj} = l_M + \mu_{5j} \quad l_{Yj} = l_Y + \mu_{6j} \quad (4)$$

$$\beta_{Xj} = \beta_X + \mu_{4j} \quad \beta_{Mj} = \beta_M + \mu_{5j} \quad \beta_{Yj} = \beta_Y + \mu_{6j} \quad (5)$$

$$a_j = a + \mu_{7j} \quad b_j = b + \mu_{8j} \quad c'_j = c' + \mu_{9j} \quad (6)$$

$$\log(\sigma_{Xj}^2) = \omega_X + \mu_{10j} \quad \log(\sigma_{Mj}^2) = \omega_M + \mu_{11j}$$

$$\log(\sigma_{Yj}^2) = \omega_Y + \mu_{12j} \quad (17)$$

与基于多水平自回归模型的中介模型[方程(1)~(6)]相比, 基于动态结构方程模型的中介模型主要有两个变化。第一, 动态结构方程模型利用潜变量进行中介建模[见方程(14)~(16)]。具体地, l_{Xj} 、 l_{Mj} 和 l_{Yj} 表示观测变量 X_{ij} 、 M_{ij} 和 Y_{ij} 的潜个体均值(见图 2)。上标 c (center)表示按潜均值 l_{Xj} 、 l_{Mj}

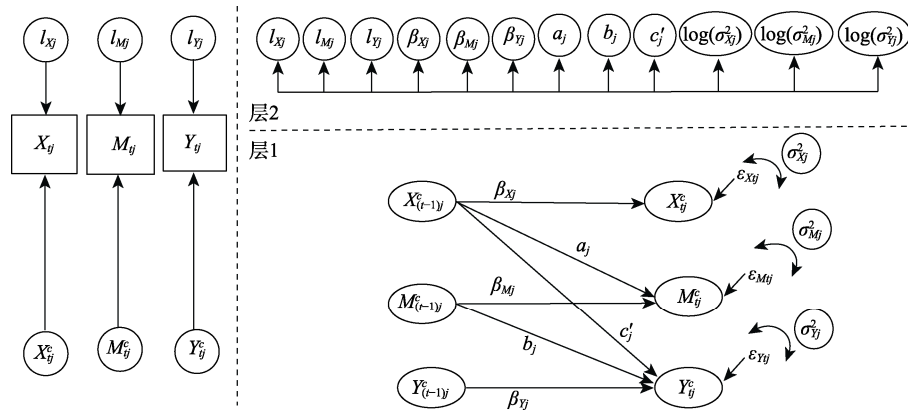


图2 动态结构方程模型的1-1-1密集追踪中介模型图(源自 Fang et al., 2024)

和 l_{Yj} 进行中心化, 称为潜均值中心化, 即

$$X_{ij}^c = X_{ij} - l_{Xj}, \quad M_{ij}^c = M_{ij} - l_{Mj}, \quad Y_{ij}^c = Y_{ij} - l_{Yj} \quad (18)$$

因此, 层1 观测变量 X_{ij} 、 M_{ij} 和 Y_{ij} 被分解为个体间潜变量(即 l_{Xj} 、 l_{Mj} 和 l_{Yj})和个体内潜变量(即 X_{ij}^c 、 M_{ij}^c 和 Y_{ij}^c) (见图2左侧)。由于 X_{ij} 、 $X_{(t-1)j}$ 的潜个体均值都是 l_{Xj} , 因此 $X_{(t-1)j}^c = X_{(t-1)j} - l_{Xj}$ 。同理, $M_{(t-1)j}^c = M_{(t-1)j} - l_{Mj}$, $Y_{(t-1)j}^c = Y_{(t-1)j} - l_{Yj}$ 。随后, 利用个体内潜变量构建密集追踪中介模型(见图2右下侧), 得到随个体变化的中介效应的均值 $E(a_j b_j)$ 为 $ab + \sigma_{a_j b_j}$, 随个体变化的直接效应 c'_j 的均值为 c' 。潜均值中心化的优点在于, 当重复测量10次及以上时, 潜均值中心化能有效克服 Nickell 偏差, 即显均值中心化使自回归效应的估计产生负偏差的现象(Asparouhov et al., 2018; Nickell, 1981; Gistelinck et al., 2021)。

第二, 动态结构方程模型允许层1 残差的方差在个体间变化[方程(17)]。动态结构方程模型假设层1 残差服从正态分布 $\varepsilon_{Xij} \sim N(0, \sigma_{Xj}^2)$ 、 $\varepsilon_{Mij} \sim N(0, \sigma_{Mj}^2)$ 和 $\varepsilon_{Yij} \sim N(0, \sigma_{Yj}^2)$ 。层1 残差的方差 σ_{Xj}^2 、 σ_{Mj}^2 和 σ_{Yj}^2 可在层2 设为随机效应[见方程(17)], ω_X 、 ω_M 和 ω_Y 表示残差方差的对数的个体均值, $\mu_{10j} \sim \mu_{12j}$ 代表每个个体值偏离均值的程度。由于残差方差 σ_{Xj}^2 、 σ_{Mj}^2 和 σ_{Yj}^2 都是非负数, 因此方程(17)将它们对数作为因变量。当然, 如果研究者假设层1 残差的方差不随个体变化, 研究者也可将残差方差设为固定效应[即去掉方程(17), 将 σ_{Xj}^2 、 σ_{Mj}^2 和 σ_{Yj}^2 设为 σ_X^2 、 σ_M^2 和 σ_Y^2], 可简化模型。

此外, 与多水平自回归模型一样, 动态结构方程模型也设定自回归效应和滞后效应不随时间

变化, 也要求重复测量的时间间隔相等。但多水平自回归模型无法处理时间间隔不等的问题, 而动态结构方程模型用卡尔曼滤波(Kalman filter)方法能处理时间间隔不等的问题(McNeish & Hamaker, 2020)。例如, 实际数据收集在上午8点、11点和12点, 这是一个时间间隔不等的现象。如果我们假设每隔1小时测量一次数据(Mplus 软件在 VARIABLE 命令下, 用 Tinterval 语句指定时间间隔为1小时), 此时只需要解决9点和10点没有观测值的问题(卡尔曼滤波可以对9点和10点的观测值进行预测), 就可以将时间间隔不等的难题转化为时间间隔相等了。具体地, 卡尔曼滤波会根据先前的值对后续每个时间点的值进行预测。如果当前时间点有观测值, 卡尔曼滤波会利用当前观测值继续对下一时间点的观测值进行预测。如果当前时间点没有观测值, 卡尔曼滤波就会根据当前时间点的预测值继续对下一时间点的观测值进行预测。继续用上述例子进行说明, 卡尔曼滤波会用8点的观测值预测9点的观测值; 由于没有9点的观测值, 卡尔曼滤波用9点的预测值预测10点的观测值; 由于没有10点的观测值, 卡尔曼滤波用10点的预测值预测11点的观测值; 由于11点有观测值, 卡尔曼滤波用11点的观测值替代11点的预测值, 用11点的观测值预测12点的观测值; 由于12点有观测值, 卡尔曼滤波用12点的观测值替代12点的预测值。至此, 我们得到了一个时间间隔相等的数据(8点的观测值、9点和10点的预测值、11点和12点的观测值)。卡尔曼滤波的优势在于, 无论当前时间点是否有观测值, 卡尔曼滤波都能移动到下一个时间点, 利

用前一个时间点的值来预测下一时间点的观测值, 不会删除任何数据, 也不需要原始数据中添加缺失值。

总之, 相比多水平自回归中介模型, 动态结构方程模型不仅能有效克服 Nickell 偏差(用潜均值中心化), 而且适用范围更广。但是, 动态结构方程模型比多水平自回归模型更为复杂, 模型不收敛的风险也在增大。因此我们建议优先使用能克服 Nickell 偏差的动态结构方程模型进行中介效应分析, 收敛不了, 再考虑在假设层 1 残差不随个体变化且重复测量的时间间隔相等时, 换用多水平自回归模型进行中介效应分析。

2.2 残差动态结构方程模型

与多水平自回归模型一样, 动态结构方程模型也假设时间序列是平稳(stationarity)的。当密集追踪数据存在时间趋势(trend)时, 推荐使用残差动态结构方程模型进行去趋势(detrending)处理。残差动态结构方程模型首先对层 1 密集追踪变量去趋势(即层 1 的去趋势部分), 然后将残差用于中介建模(即层 1 的残差部分和层 2 部分, Fang et al., 2024)。具体地, 去趋势部分用 $time_{ij}$ 为自变量, 实现层 1 变量去趋势:

$$\text{层 1: } X_{ij}^c = D_{Xj}time_{ij} + e_{Xij} \quad (19)$$

$$M_{ij}^c = D_{Mj}time_{ij} + e_{Mij} \quad (20)$$

$$Y_{ij}^c = D_{Yj}time_{ij} + e_{Yij} \quad (21)$$

相比残差多水平自回归模型[方程(7)~(9)], 残差动态结构方程模型[方程(19)~(21)]的变化在于使用潜变量进行建模。然后, 残差部分用前面得到的残差 e 来建构中介模型, 其中层 1 的残差部分

是方程(22)~(24),

$$\text{层 1: } e_{Xij} = \beta_{Xj}e_{X(t-1)j} + \delta_{Xij} \quad (22)$$

$$e_{Mij} = \beta_{Mj}e_{M(t-1)j} + a_j e_{X(t-1)j} + \delta_{Mij} \quad (23)$$

$$e_{Yij} = \beta_{Yj}e_{Y(t-1)j} + b_j e_{M(t-1)j} + c'_j e_{X(t-1)j} + \delta_{Yij} \quad (24)$$

相比残差多水平自回归模型[方程(10)~(12)], 残差动态结构方程模型[方程(22)~(24)]没有截距项, 因为所有变量都被中心化了(Asparouhov & Muthén, 2020)。另外, 残差动态结构方程模型假设残差 δ 服从正态分布 $\delta_{Xij} \sim N(0, \sigma_{Xj}^2)$, $\delta_{Mij} \sim N(0, \sigma_{Mj}^2)$ 和 $\delta_{Yij} \sim N(0, \sigma_{Yj}^2)$, 允许残差方差 σ_{Xj}^2 、 σ_{Mj}^2 和 σ_{Yj}^2 在个体间变化。层 2 方程包括方程(4)~(6)、方程(13)和(17)。随个体变化的中介效应的均值 $E(a_j b_j)$ 为 $ab + \sigma_{a_j b_j}$, 随个体变化的直接效应 c'_j 的均值为 c' 。

2.3 交叉分类的动态结构方程模型

动态结构方程模型和残差动态结构方程模型都只能将密集追踪数据看成时间点嵌套于个体的两水平结构, 所以这两种模型的层 2 只有个体方程。如果将密集追踪数据既看成时间点嵌套于个体, 又看成是个体嵌套于时间点的交叉分类结构(Luo, 2017; Kim et al., 2022), 则层 2 既有个体方程, 也有时间方程(见图 3 右上), 因此需要换用交叉分类的动态结构方程模型进行密集追踪中介效应分析(见图 3)。

基于交叉分类的动态结构方程模型的密集追踪中介效应分析(图 3)只需要对动态结构方程的中介模型[方程(14)~(17)和(4)~(6)]进行如下调整即可(McNeish & MacKinnon, 2022)。具体地, 将方程(15)、(16)变为方程(25)和(26)

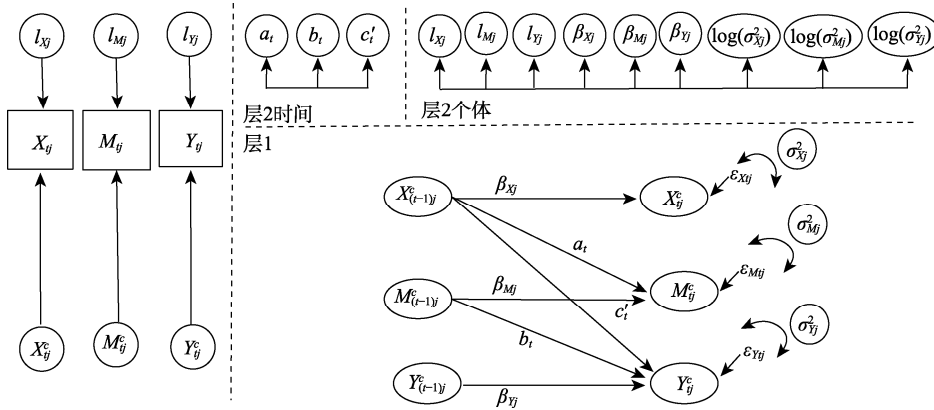


图 3 交叉分类的动态结构方程模型的 1-1-1 密集追踪中介模型图
(整合自 Hamaker et al., 2018 和 McNeish & MacKinnon, 2022)

$$\text{层 1: } M_{ij}^c = \beta_{Mj} M_{(t-1)j}^c + a_t X_{(t-1)j}^c + \varepsilon_{Mij} \quad (25)$$

$$Y_{ij}^c = \beta_{Yj} Y_{(t-1)j}^c + b_t M_{(t-1)j}^c + c_t' X_{(t-1)j}^c + \varepsilon_{Yij} \quad (26)$$

即将方程(15)和(16)中的滞后效应由随个体 j 变化 (a_j 、 b_j 和 c_j') 变为随时间 t 变化 (a_t 、 b_t 和 c_t')。另外, 增加层 2 时间方程, 即将方程(6)变为方程(27)

$$\begin{aligned} \text{层 2 时间: } a_t &= a + \mu_{7t} & b_t &= b + \mu_{8t} \\ c_t' &= c' + \mu_{9t} \end{aligned} \quad (27)$$

其中, a 、 b 和 c' 表示所有时间点的均值, $\mu_{7t} \sim \mu_{9t}$ 代表每个时间点的具体值偏离均值的程度。

相比动态结构方程模型和残差动态结构方程模型, 交叉分类的动态结构方程模型有两个优势。第一, 交叉分类的动态结构方程模型在层 2 既有个体方程[见方程(4)、(5)和(17)], 也有时间方程[见方程(27)], 而动态结构方程模型和残差动态结构方程模型在层 2 只有个体方程。第二, 交叉分类的动态结构方程模型得到的中介效应 ($a_t b_t$) 可以随时间变化而变化, 即每个时间点的中介效应都可能不同。所有时间点的中介效应的均值为 $E(a_t b_t)$ 。由于 a_t 和 b_t 可能随时间 t 的变化而发生共变, 所以随时间变化的中介效应的均值 $E(a_t b_t)$ 为 $ab + \sigma_{a,b}$, 随时间变化的直接效应 c_t' 的均值为 c' 。而动态结构方程模型和残差动态结构方程模型得到的中介效应 ($a_j b_j$) 却只能随个体变化, 不能随时间变化。Fang 等(2024)建议, 如果研究者想研究随时间变化的中介效应, 就将密集追踪数据看成交叉分类结构, 用交叉分类的结构方程模型进行中介效应分析; 如果研究者想研究随个体变化的中介效应, 就将密集追踪数据看成时间点嵌套于个体的两水平结构, 用动态结构方程模型或残差动态结构方程模型进行中介效应分析。

需要说明的是, 第一, 目前仅有 *Mplus* 软件可进行基于交叉分类的动态结构方程模型的中介效应分析, 但 *Mplus* 软件无法将所有层 1 参数设置为随时间 t 变化, 且用时远超动态结构方程模型的中介效应分析。为了简化模型和缩短用时, 因此模型中只允许 a 、 b 和 c' 随时间 t 变化。第二, McNeish 和 MacKinnon(2022)将方程(25)和(26)中的 a_t 、 b_t 和 c_t' 变为 a_{ij} 、 b_{ij} 和 c_{ij}' , 同时将方程(27)变为

$$\begin{aligned} \text{层 2 时间: } a_{ij} &= a + \mu_{7t} + \mu_{7j} & b_{ij} &= b + \mu_{8t} + \mu_{8j} \\ c_{ij}' &= c' + \mu_{9t} + \mu_{9j} \end{aligned} \quad (28)$$

即在方程(27)中增加随个体 j 变化的残差 (μ_{7j} 、 μ_{8j}

和 μ_{9j}), 这种改进后的模型被称为完全交叉分类 (fully cross-classified) 的中介模型。与交叉分类的中介模型[用方程(27)]相比, 完全交叉分类的中介模型[用方程(28)]仅多得到 3 个随个体变化的残差, 却得不到随个体变化的中介效应, 改进的实际意义并不大。第三, 图 3 的左边将观测变量 (X_{ij} 、 M_{ij} 和 Y_{ij}) 分解为潜个体均值 (I_{Xj} 、 I_{Mj} 和 I_{Yj}) 和个体内潜变量 (X_{ij}^c 、 M_{ij}^c 和 Y_{ij}^c), 而没有分解为潜时间点的均值 (I_{Xt} 、 I_{Mt} 和 I_{Yt}) 和时间点内的潜变量 (例如, 此时的 $X_{ij}^c = X_{ij} - I_{Xt}$), 这是因为 *Mplus* 软件默认进行潜个体均值中心化 (考虑潜个体均值)。第四, 纵向面板数据无法进行随时间变化的中介效应 $a_t b_t$ 分析。因为纵向面板数据一般只重复测量数次 (只有几个时间点), 无法确保层 2 时间点达到多水平数据分析的基本要求。

3 密集追踪数据的中介效应分析的流程

3.1 密集追踪中介效应的分析模型的发展

密集追踪中介效应的分析模型呈现出两条平行的发展路径 (图 4)。第一, 基于显变量的密集追踪中介效应分析。多水平模型 (MLM) 和变量的时间先后顺序相结合, 发展出了多水平自回归模型 (MAM); 多水平自回归模型结合去趋势的需要, 发展出了残差多水平自回归模型 (RMAM)。第二, 基于潜变量的密集追踪中介效应分析。多水平结构方程模型 (MSEM) 和变量的时间先后顺序相结合, 发展出了动态结构方程模型 (DSEM); 动态结构方程模型结合去趋势的需要, 发展出了残差动态结构方程模型 (RDSEM)。

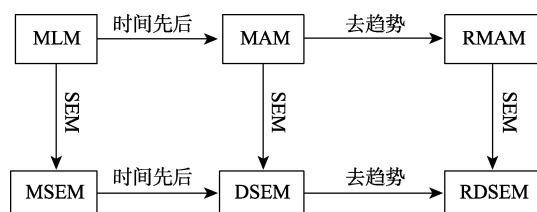


图 4 随个体变化的密集追踪中介效应的分析方法发展

此外, 两条平行的发展路径之间存在联系, 即多水平模型 (MLM)、多水平自回归模型 (MAM)、残差多水平自回归模型 (RMAM) 分别和结构方程模型 (SEM) 相结合, 发展出多水平结构方程模型 (MSEM)、动态结构方程模型 (DSEM) 和残差动态

结构方程模型(RDSEM)。和结构方程模型(SEM)相结合,是指在模型中增加潜变量,由显变量建模改为潜变量建模。例如,对于层1的自变量 X_{ij} ,多水平模型默认 t 次重复测量的 X_{ij} 的均值 $\bar{X}_{\cdot j}$ 就是组均值或个体均值(将 X_{ij} 的组均值看成是直接观测的显变量),从而将层1变量 X_{ij} 被分解为个体内显变量($X_{ij} - \bar{X}_{\cdot j}$)和个体间显变量 $\bar{X}_{\cdot j}$,利用显变量($X_{ij} - \bar{X}_{\cdot j}$)和 $\bar{X}_{\cdot j}$ 进行后续建模。这种建模方法的不足在于,从 N 次重复测量中抽取部分重复测量结果($t \leq N$)来求组均值,会产生抽样误差(方杰等, 2014)。而多层结构方程模型则是将组均值看成是无法直接观测的潜变量(用 I_{Xj} 表示),则 X_{ij} 的个体内部分($X_{ij}^c = X_{ij} - I_{Xj}$)也是一个潜变量,这样层1变量 X_{ij} 被分解为个体内潜变量 X_{ij}^c 和个体间潜变量 I_{Xj} ,利用潜变量 X_{ij}^c 和 I_{Xj} 进行后续建模。同理,多水平自回归模型可以对变量 X_{ij} 、 M_{ij} 和 Y_{ij} 进行显均值中心化,即用 $(X_{ij} - \bar{X}_{\cdot j})$ 、 $(M_{ij} - \bar{M}_{\cdot j})$ 和 $(Y_{ij} - \bar{Y}_{\cdot j})$ 替换方程(1)-(3)中的 X_{ij} 、 M_{ij} 和 Y_{ij} ,这同样会造成抽样误差。而动态结构方程模型则将组均值看成潜变量(I_{Xj} 、 I_{Mj} 和 I_{Yj}),用个体内潜变量 X_{ij}^c 、 M_{ij}^c 和 Y_{ij}^c [见方程(18)]替代 $(X_{ij} - \bar{X}_{\cdot j})$ 、 $(M_{ij} - \bar{M}_{\cdot j})$ 和 $(Y_{ij} - \bar{Y}_{\cdot j})$ 进行建模。

3.2 密集追踪中介效应分析的流程图

面对一个密集追踪数据的中介效应分析任务,研究者应当如何进行呢?根据前面的讨论,我们总结出一个密集追踪数据的中介效应分析的流程图(图5)如下:

(1)判断是否要分析随时间变化的中介效应。如果是,则使用交叉分类的动态结构方程模型进行中介效应分析。如果不是,则进入步骤2。

(2)判断是否要去趋势。如果是,则利用残差动态结构方程模型(RDSEM)进行随个体变化的中介效应分析。如果不是,则利用动态结构方程模型(DSEM)进行随个体变化的中介效应分析。目前暂时只能依据理论或者经验推断密集追踪数据是否存在时间趋势,如果不确定数据是否存在时间趋势,可参考示例的做法。

(3)判断模型是否收敛。如果是,则报告残差动态结构方程模型或动态结构方程模型的中介效应分析结果。如果不是,则在假设层1残差不随个体变化且重复测量的时间间隔相等时,换用多水平自回归模型(MAM)或残差多水平自回归模型(RMAM)进行中介效应分析。

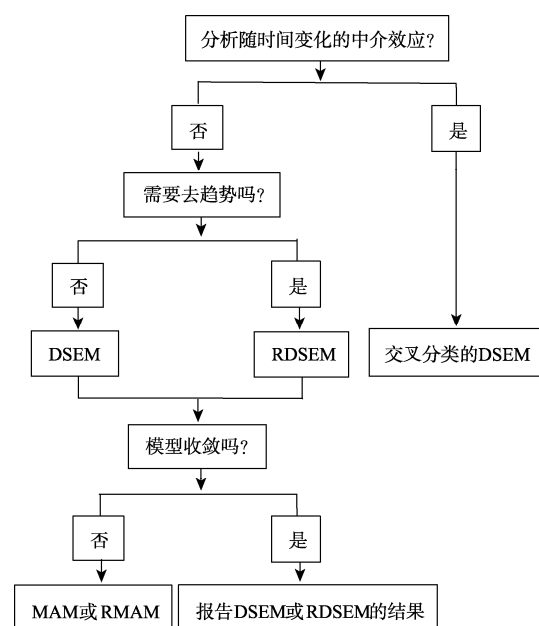


图5 密集追踪数据的中介效应分析的流程图

4 密集追踪数据的中介效应分析示例

示例数据(<https://osf.io/c4dqr/>下载)包括100名被试($j = 100$),每名被试都重复测量50次($t = 50$)。示例包含三个变量,自变量 X_{ij} 为积极取向(positive orientation)、中介变量 M_{ij} 为积极情绪(positive affect)、因变量 Y_{ij} 为亲社会行为(prosocial behavior)。本例假设积极情绪在积极取向和亲社会行为之间起中介作用,即每天的积极取向越多,每天的积极情绪就越多,从而每天的亲社会行为就越多。本例的变量和中介关系都源自Laguna等(2022)的研究。

由于有多个随机效应需要估计,本例采用贝叶斯法进行参数估计(更多贝叶斯法的内容详见:方杰,温忠麟,2023)。参数的先验分布使用无信息先验分布。用PSR(Proportional Scale Reduction)值来检查Markov链的收敛性,当PSR值小于1.1时,表示Markov链收敛。采用2条Markov链,每条Markov链各迭代2万次,每条Markov链都抛弃前一半的迭代作为预热期或燃烧期(burn in period),用后一半的迭代作为待估参数 ab 的后验分布,如果中介效应 ab 的可靠区间(credible interval)不包含0,则表示中介效应显著。偏差信息准则(deviance information criterion, DIC)被用来

进行模型比较, DIC 越小, 模型和数据拟合越好。

4.1 随个体变化的中介效应分析

4.1.1 动态结构方程模型

由于不确定数据是否存在时间趋势, 用残差动态结构方程模型[方程(4)~(6)、(13)、(17)和(19)~(24)]和动态结构方程模型[方程(14)~(17)和(4)~(6)]进行潜变量的中介效应分析(<https://osf.io/c4dqr/>下载 Mplus 程序和运行结果)。结果发现, 去

趋势部分[方程(19)~(21)]的 $time_{ij}$ 的回归系数的均值 D_X 、 D_M 和 D_Y 都不显著, 残差动态结构方程模型和动态结构方程模型的中介效应分析结果也相近, 表明去趋势对中介效应的检验结果没有产生显著影响。此外, 动态结构方程模型的 DIC 值(3663.73)比残差动态结构方程模型的 DIC 值(4247.62)小得多, 说明动态结构方程模型和数据拟合的更好, 因此以下仅报告动态结构方程模型的中介效应分

表2 密集追踪数据的中介效应分析结果

参数估计	DSEM		MAM		交叉分类 DSEM	
	估计值	95% CI	估计值	95% CI	估计值	95% CI
a	0.618	[0.525, 0.713]	0.621	[0.540, 0.703]	0.576	[0.551, 0.598]
b	0.725	[0.594, 0.851]	0.762	[0.638, 0.879]	0.734	[0.684, 0.780]
c'	0.044	[0.005, 0.085]	0.037	[-0.017, 0.092]	0.010	[-0.019, 0.039]
$\sigma_{a,b}$	0.158	[0.093, 0.253]	0.110	[0.061, 0.171]		
$\sigma_{a,h}$					0.001	[-0.002, 0.004]
β_X	0.301	[0.265, 0.338]	0.277	[0.227, 0.336]	0.302	[0.269, 0.335]
β_M	0.293	[0.263, 0.323]	0.285	[0.245, 0.324]	0.335	[0.305, 0.363]
β_Y	0.280	[0.247, 0.314]	0.256	[0.214, 0.298]	0.324	[0.292, 0.355]
l_X	-0.074	[-0.220, 0.071]	0.000	[-0.027, 0.027]	-0.076	[-0.206, 0.050]
l_M	0.075	[-0.061, 0.210]	0.000	[-0.021, 0.022]	0.073	[-0.046, 0.195]
l_Y	1.926	[1.776, 2.075]	0.000	[-0.028, 0.027]	1.925	[1.794, 2.055]
ω_X	-1.922	[-2.167, -1.680]			-1.922	[-2.136, -1.712]
ω_M	-4.172	[-4.483, -3.861]			-3.259	[-3.531, -2.994]
ω_Y	-1.875	[-2.104, -1.650]			-1.667	[-1.853, -1.481]
$\text{var}(\mu_{1j})$	0.519	[0.381, 0.732]	0.014	[0.010, 0.019]	0.402	[0.304, 0.544]
$\text{var}(\mu_{2j})$	0.458	[0.334, 0.649]	0.011	[0.008, 0.015]	0.354	[0.268, 0.479]
$\text{var}(\mu_{3j})$	0.526	[0.378, 0.756]	0.014	[0.011, 0.019]	0.409	[0.307, 0.561]
$\text{var}(\mu_{4j})$	0.015	[0.007, 0.029]	0.032	[0.022, 0.046]	0.007	[0.001, 0.016]
$\text{var}(\mu_{5j})$	0.016	[0.011, 0.025]	0.027	[0.020, 0.038]	0.012	[0.008, 0.020]
$\text{var}(\mu_{6j})$	0.016	[0.009, 0.027]	0.028	[0.020, 0.040]	0.012	[0.007, 0.019]
$\text{var}(\mu_{7j})$	0.213	[0.155, 0.300]	0.167	[0.124, 0.222]		
$\text{var}(\mu_{8j})$	0.331	[0.227, 0.501]	0.286	[0.197, 0.407]		
$\text{var}(\mu_{9j})$	0.009	[0.003, 0.019]	0.038	[0.025, 0.056]		
$\text{var}(\mu_{7t})$					0.002	[0.001, 0.006]
$\text{var}(\mu_{8t})$					0.008	[0.003, 0.018]
$\text{var}(\mu_{9t})$					0.003	[0.001, 0.008]
$\text{var}(\mu_{10j})$	1.455	[1.058, 2.060]			1.098	[0.825, 1.500]
$\text{var}(\mu_{11j})$	2.377	[1.740, 3.354]			1.811	[1.367, 2.457]
$\text{var}(\mu_{12j})$	1.233	[0.897, 1.743]			0.838	[0.628, 1.142]

注: DSEM 表示动态结构方程模型, MAM 表示多水平自回归模型, CI 表示可靠区间(credible interval)。

析结果。首先, Markov 链迭代 700 次之后, PSR 值始终小于 1.1, 表明 Markov 链稳定收敛。其次, 中介效应分析结果表明(见表 2), 除了 l_X 和 l_M 之外, 其余参数的可靠区间都不包含 0, 表明这些参数都显著不等于 0。随个体变化的中介效应的均值 $E(a_j b_j) = 0.606$, $E(a_j b_j)$ 的 95% 可靠区间[0.478, 0.772]不包含 0, 因此中介效应显著不等于 0。由于直接效应($c' = 0.044$)仍然显著, 因此积极情绪 $M_{(t-1)j}$ 在积极取向 $X_{(t-2)j}$ 和亲社会行为 Y_{tj} 之间起部分中介作用。再次, 层 2 的残差方差 [$\text{var}(\mu_{1j}) \sim \text{var}(\mu_{12j})$] 的分析结果见表 2, 由于 p 值都小于 0.001, 表明相应的参数都存在显著的个体差异。需要说明的是, 残差方差是非负数, 因此残差方差的可靠区间往往都不包含 0, 用可靠区间不包含 0 来判断残差方差显著不等于 0 是不合适的(McNeish & Hamaker, 2020)。此时可用 p 值小于设定的显著性概率(如 $\alpha = 0.05$)来判断残差方差显著不等于 0。

4.1.2 多水平自回归模型

利用多水平自回归模型(用方程(1)~(6)进行显变量的密集追踪中介效应分析(<https://osf.io/c4dqr/>下载 R 程序和运行结果)。中介效应分析结果表明(见表 2), 除了 l_X 、 l_M 、 l_Y 和 c' 之外, 其余参数的可靠区间都不包含 0, 表明这些参数都显著不等于 0。随个体变化的中介效应的均值 $E(a_j b_j) = 0.584$, $E(a_j b_j)$ 的 95% 可靠区间[0.468, 0.716]不包含 0, 因此中介效应显著不等于 0。相比动态结构方程模型的中介效应分析结果(见表 2), 仅发现 l_M 和 c' 的显著性检验结果从显著变为不显著了。接着用残差多水平自回归模型[方程(4)~(13)]进行中介效应分析(<https://osf.io/c4dqr/>下载 R 程序和运行结果)。结果发现, 残差多水平自回归模型和多水平自回归模型的中介效应分析结果相近, 表明去趋势对中介效应的检验结果没有产生显著影响, 在此不再赘述。

4.2 随时间变化的中介效应分析

利用交叉分类的动态结构方程模型[方程(4)、(5)、(14)、(17)、(25)~(27)]进行中介效应分析(<https://osf.io/c4dqr/>下载 Mplus 程序和运行结果)。第一, Markov 链迭代 200 次后, PSR 值始终小于 1.1, 表明 Markov 链稳定收敛。第二, 中介效应分析结果表明(见表 2), 除了 c' 、 l_X 和 l_M 之外, 其余参数的可靠区间都不包含 0, 表明这些参数都显著不等于 0。随时间变化的中介效应的均值 $E(a_j b_j) =$

0.423, $E(a_j b_j)$ 的 95% 可靠区间[0.388, 0.457]不包含 0, 因此中介效应显著不等于 0。第三, a_i 、 b_i 和 c'_i 的残差方差 $\text{var}(\mu_{7i}) \sim \text{var}(\mu_{9i})$ 的分析结果见表 2, 由于 p 值都小于 0.001, 表明相应的参数都存在显著的时间差异。其余层 2 残差方差的分析结果见表 2, 由于 p 值都小于 0.001, 表明相应的参数都存在显著的个体差异。

5 局限和展望

本文以 1-1-1 密集追踪中介模型为例, 详述了基于多水平自回归模型、残差多水平自回归模型、动态结构方程模型、残差动态结构方程模型、交叉分类的动态结构方程模型的密集追踪中介效应分析方法, 并总结出一个密集追踪中介分析流程。用示例演示如何进行密集追踪数据的中介效应分析, 并给出了相应的 Mplus 和 R 程序。但本文仍存在不足, 尚需进一步拓展。

第一, 本文只涉及跨一个时间点的一阶滞后中介效应模型。虽然, 目前研究者更多使用一阶滞后的密集追踪模型, 但本文的方法和步骤可拓展到跨两个时间点的($X_{(t-2)} \rightarrow M_{(t-1)} \rightarrow Y_t$)的二阶滞后中介效应模型, 仅需将方程(3)、(16)和(26)的 $X_{(t-1)j}^c$ 改为 $X_{(t-2)j}^c$ 即可, 将方程(12)和(24)的 $e_{X(t-1)j}$ 改为 $e_{X(t-2)j}$ 即可。

第二, 根据变量所在的水平(1 或 2), 可能有 $2 \times 2 \times 2 = 8$ 种密集追踪中介模型, 本文重点考察了 1-1-1 密集追踪中介模型, 而 2-2-2 中介模型不属于密集追踪中介模型(所有变量都不是密集追踪数据), 因此还有 2-1-1、2-2-1、2-1-2、1-2-2、1-2-1 和 1-1-2 共 6 种密集追踪中介模型。如何分析这 6 种密集追踪中介效应呢? 目前, 仅有 Hamaker 等(2018)介绍了动态结构方程模型的 2-1-2 密集追踪中介效应分析, Fang 等(2024)介绍了动态结构方程模型和残差动态结构方程模型的 2-1-1 和 2-2-1 密集追踪中介效应分析。

基于动态结构方程模型的 2-1-2、2-2-1、2-1-1 密集追踪中介效应分析的共同点在于, 这三种中介效应都发生在层 2。以 2-1-1 密集追踪中介效应分析为例, 基于动态结构方程模型的 2-1-1 密集追踪中介效应分析(见图 6)可表示为:

$$\text{层 1: } M_{ij}^c = \beta_{Mj} M_{(t-1)j}^c + \varepsilon_{Mij} \quad (29)$$

$$Y_{ij}^c = \beta_{Yj} Y_{(t-1)j}^c + b_{Yj} M_{(t-1)j}^c + \varepsilon_{Yij} \quad (30)$$

$$\text{层 2: } l_{Mj} = l_M + aX_j + \mu_{1j} \quad (31)$$

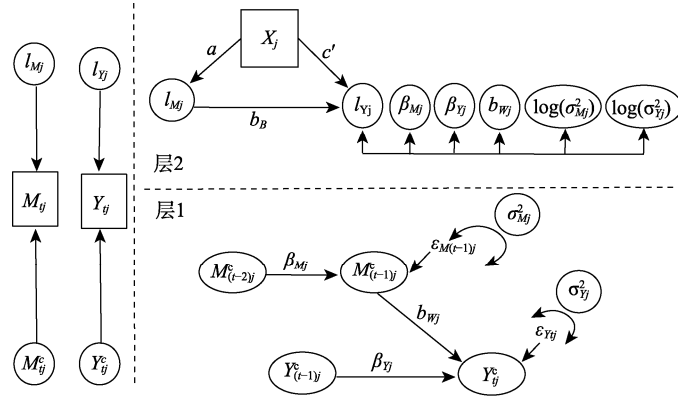


图6 动态结构方程模型的2-1-1密集追踪中介模型图
(改编自 Fang et al., 2024, X_j 也可以是潜变量)

$$l_{Yj} = l_Y + c'X_j + b_B l_{Mj} + \mu_{2j} \quad (32)$$

$$\beta_{Mj} = \beta_M + \mu_{3j} \quad \beta_{Yj} = \beta_Y + \mu_{4j} \quad b_{Wj} = b_W + \mu_{5j} \quad (33)$$

$$\log(\sigma_{Mj}^2) = \omega_M + \mu_{1j} \quad \log(\sigma_{Yj}^2) = \omega_Y + \mu_{2j} \quad (34)$$

此时, b_{Wj} 表示 $M_{(t-1)j}^c$ 对因变量 Y_{ij}^c 的个体内效应。2-1-1 密集追踪中介效应本质上是发生在水平 2 (即 $X_j \rightarrow l_{Mj} \rightarrow l_{Yj}$), 中介效应为 ab_B , 直接效应为 c' 。中介效应和直接效应既不随个体变化, 也不随时间变化。

第三, 本文使用的去趋势方法是用每个变量对时间的回归的残差建构密集追踪中介模型(本文称为残差法)。但是, 在基于交叉分类的动态结构方程模型的中介效应分析中, 暂时无法使用残差法。实际上, 还有一种去趋势方法是在描述变量自身(即自回归效应)和变量间关系(即中介效应)的同时, 将时间 $time_{ij}$ 作为协变量加入模型以控制时间效应(本文称为协变量法, 肖悦等, 2024; Fang et al., 2024)。例如, 基于动态结构方程模型的以协变量法去趋势的 1-1-1 密集追踪中介效应分析仅需将方程(14)-(16)变化为方程(35)-(37),

$$\text{层 1: } X_{ij}^c = \beta_{Xj} X_{(t-1)j}^c + D_{Xj} time_{ij} + \varepsilon_{Xij} \quad (35)$$

$$M_{ij}^c = \beta_{Mj} M_{(t-1)j}^c + a_j X_{(t-1)j}^c + D_{Mj} time_{ij} + \varepsilon_{Mij} \quad (36)$$

$$Y_{ij}^c = \beta_{Yj} Y_{(t-1)j}^c + b_j M_{(t-1)j}^c + c_j' X_{(t-1)j}^c + D_{Yj} time_{ij} + \varepsilon_{Yij} \quad (37)$$

层 2 方程仍为方程(4)-(6)、(13)和(17) (Fang et al., 2024)。更进一步, 我们提出, 当密集追踪数据存在时间趋势时, 就需要在交叉分类的动态结构方程模型的 1-1-1 密集追踪中介效应分析中, 使用协变量法去趋势, 以满足动态结构方程模型的平

稳性假设。具体做法是, 将方程(36)和(37)变化为方程(38)和(39)即可。

$$\text{层 1: } M_{ij}^c = \beta_{Mj} M_{(t-1)j}^c + a_i X_{(t-1)j}^c + D_{Mj} time_{ij} + \varepsilon_{Mij} \quad (38)$$

$$Y_{ij}^c = \beta_{Yj} Y_{(t-1)j}^c + b_i M_{(t-1)j}^c + c_i' X_{(t-1)j}^c + D_{Yj} time_{ij} + \varepsilon_{Yij} \quad (39)$$

我们用示例数据进行了交叉分类的动态结构方程模型的去趋势的 1-1-1 密集追踪中介效应分析[层 1 方程是(35)、(38)和(39), 层 2 方程是(4)、(5)、(13)、(17)和(27)]。结果(<https://osf.io/c4dqr/>下载 Mplus 程序和运行结果)表明, 去趋势后, 系数 a 和 b 都显著($a = 0.569$, 95%可靠区间[0.540, 0.594]; $b = 0.724$, 95%可靠区间[0.672, 0.772]), a_i 和 b_i 的协方差 $\sigma_{a,b_i} = 0.001$ 。中介效应的均值 $E(a_i b_i) = 0.412$, $E(a_i b_i)$ 的 95%可靠区间[0.374, 0.450]不包含 0, 因此中介效应显著。

在心理学和其他社科领域研究中越来越多研究者收集密集追踪数据, 希望进行密集追踪数据的中介效应分析, 本文为应用研究者提供了方法指引。实际上, 密集追踪数据的中介效应分析方法研究仍处于起步阶段(Fang et al., 2024; Hamaker et al., 2018; McNeish & MacKinnon, 2022; Zhang et al., 2018), 还有诸多问题函待解决。例如, 本文仅涉及连续变量, 如何进行类别变量的密集追踪中介效应分析呢? 再如, 在本文中, 残差法和协变量法去除的都是线性趋势, 如何去除非线性的趋势呢? 又如, KPSS (Kwiatkowski-Phillips-Schmidt-Shin, Kwiatkowski et al., 1992)检验可用于判断单个体的时间序列数据是否存在时间趋势, 但还缺

乏统计方法判断多个体的密集追踪数据是否存在时间趋势。方法的进步给研究者提供了一个深入理解和应用两类模型(即多水平自回归模型和动态结构方程模型)进行密集追踪中介效应分析的机会,相信随着密集追踪中介效应分析方法研究的深入,会不断增加我们对两类模型和密集追踪中介效应问题的理解。

参考文献

- 方杰, 温忠麟. (2023). 中介效应和调节效应模型进阶. 北京: 教育科学出版社.
- 方杰, 温忠麟, 邱皓政. (2021). 纵向数据的中介效应分析. *心理科学*, 44(4), 989-996.
- 方杰, 温忠麟, 张敏强, 任皓. (2014). 基于结构方程模型的多层中介效应分析. *心理科学进展*, 22(3), 530-539.
- 温忠麟. (2017). 实证研究中的因果推理与分析. *心理科学*, 40, 200-208.
- 肖悦, 刘红云, 徐永泽. (2024). 密集追踪成对数据分析的模型建构探索. *心理科学进展*, 32(9), 1450-1462.
- 郑舒方, 张沥今, 乔欣宇, 潘俊豪. (2021). 密集追踪数据分析: 模型及其应用. *心理科学进展*, 29(11), 1948-1969.
- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 359-388.
- Asparouhov, T., & Muthén, B. (2020). Comparison of models for the analysis of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(2), 275-297.
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 583-619.
- Fang, J., Wen, Z., Hau, K-T. (2024). Mediation analyses of intensive longitudinal data with dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 31(4), 728-741.
- Gistelinck, F., Loeys, T., & Flamant, N. (2021). Multilevel autoregressive models when the number of time points is small. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(1), 15-27.
- Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, 53(6), 820-841.
- Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, 8(2), 115-128.
- Kim, M., Winkler, C., Uanboro, J., Peri, J., & Lochman, J. (2022). Using a cross-classified multilevel mediation model (CC-M3) with longitudinal data having changes in cluster membership. *Structural Equation Modeling: A Multidisciplinary Journal*, 29(2), 218-232.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1), 159-178.
- Laguna, M., De Longis, E., Mazur-Socha, Z., & Alessandri, G. (2022). Explaining prosocial behavior from the inter- and within-individual perspectives: A role of positive orientation and positive affect. *Journal of Happiness Studies*, 23, 1599-1615.
- Lischetzke, T., Schemer, L., Glombiewski, J. A., In-Albon, T., Karbach, J., & Könen, T. (2021). Negative emotion differentiation attenuates the within-person indirect effect of daily stress on nightly sleep quality through calmness. *Frontiers in Psychology*, 12, 684117.
- Luo, W. (2017). Testing mediation effects in cross-classified multilevel data. *Behavior Research Methods*, 49(2), 674-684.
- McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, 25(5), 610-625.
- McNeish, D., & MacKinnon, D. P. (2022). Intensive longitudinal mediation in Mplus. *Psychological Methods*. <https://dx.doi.org/10.1037/met0000536>
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417-1426.
- Wang, L., & Maxwell, S. E. (2015). A note on disaggregating betweenperson and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, 20(1), 63-83.
- Zhang, Q., Wang, L. J., & Bergeman, C. S. (2018). Multilevel autoregressive mediation models: Specification, estimation, and applications. *Psychological Methods*, 23(2), 278-297.

Mediation analysis of intensive longitudinal data

FANG Jie¹, WEN Zhonglin², DONG Yuming³, WANG Xiaojie³

(¹ Institute of New Development, Guangdong University of Finance & Economics, Guangzhou 510320, China)

(² Center for Studies of Psychological Application & School of Psychology, South China Normal University, Guangzhou 510631, China)

(³ The School of Economics, Guangdong University of Finance & Economics, Guangzhou 510320, China)

Abstract: With the widespread use of intensive longitudinal data in the social sciences, how to analysis intensive longitudinal mediation (ILM) effect has attracted the attention of many researchers. A conventional approach is using multilevel models or multilevel structural equation models. In that case, the temporal sequence of variables is ignored, with the dynamic relationship between variables remaining unexplored. In this paper, we summarize five types of ILM analysis approaches: 1) multilevel autoregressive model (MAM); 2) residual multilevel autoregressive model (RMAM); 3) dynamic structural equation model (DSEM); 4) residual dynamic structural equation model (RDSEM); 5) cross-classified dynamic structural equation model (cross-classified DSEM). ILM effects with time changes can only be analyzed by cross-classified DSEM; other models can only obtain ILM effects with individual changes. After introducing the new method, we proposed a procedure for analyzing ILM effects. Then, this paper exemplifies how to conduct the proposed procedure and provides corresponding Mplus and R codes. Directions for future study on mediation analysis of intensive longitudinal data were discussed at the end of the paper.

Keywords: intensive longitudinal data, mediation, multilevel autoregressive model, dynamic structural equation model, detrending