

個人内変化に基づく因果推論のための 方法論的展開

宇佐美 慧

(東京大学)

Email: usami_s@p.u-tokyo.ac.jp

HP: <http://www.satoshiusami.com/>

量的研究（データサイエンス）における研究目的の分類

Hamaker, Mulder, & van Ijzendoorn (2020)

1. 記述的研究（descriptive research）

測定値に関する集団や個人の実態を把握することが主目的。

→ 幼児の語彙力は時間と共にどう推移する？ 個人差・集団差は？

2. 予測的研究（predictive research）

（選抜・選択・管理等を意図した）将来値の予測が主目的。

→ 言語発達遅滞がある幼児をどのようにスクリーニングするか？

3. 説明的研究（explanatory research）

因果のメカニズムを理解することが主目的。

→ 保育園の通園は幼児の語彙力を高めるか？

- ・ 発達研究においては記述的・説明的研究の割合が高い。
- ・ 説明的研究では交差遅延パネルモデル（後述）が良く利用される。

説明的研究：変数間の関係を調べる

横断データの場合 (xからyへの因果に関心)

$$y_i = \alpha + \beta x_i + e_i$$

(iは個人を表す)

y_i : 従属変数 x_i : 独立変数 e_i : 残差

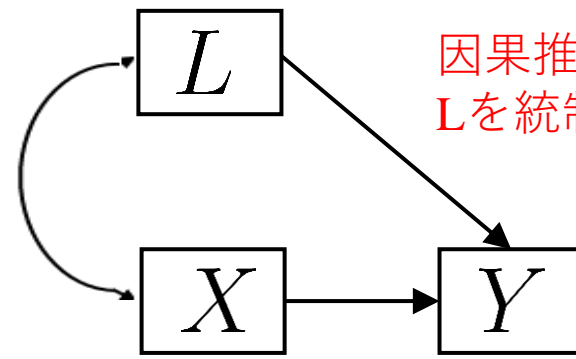
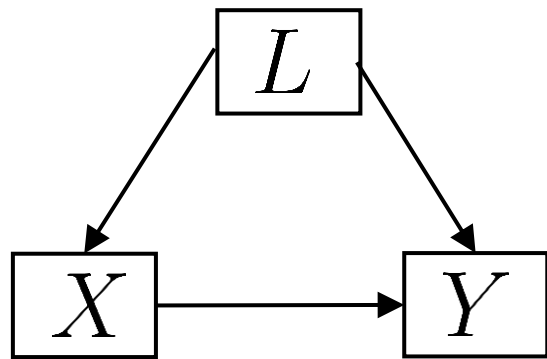
- 交絡変数 l があるときの対処 (回帰モデル・共分散分析モデル)

$$y_i = \alpha + \beta x_i + \gamma l_i + e_i$$

交絡変数 l の影響を除いても、 x 独自の効果があるか？

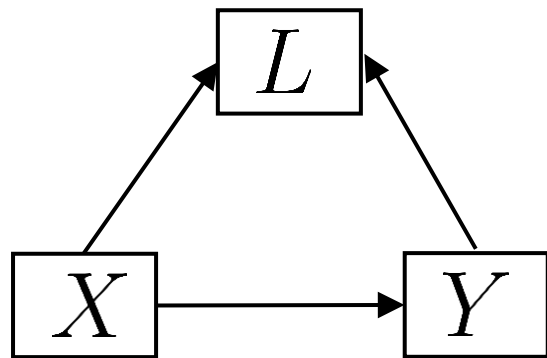
交絡変数(confounder, confounding variable)とは？

- 従属変数 y と独立変数 x の双方に影響を与える（または、 x については相関する）変数。

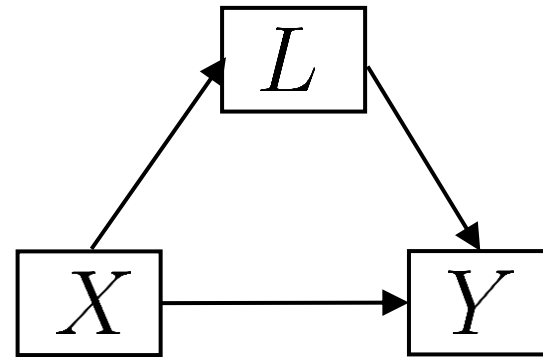


因果推論の為に
 L を統制する必要。

有向非巡回グラフ (directed acyclical graphs: DAGs)に基づく交絡変数の例



collider



媒介変数 (mediator)

X から Y への効果を検証する上では、 L (collider, mediator)は統制する必要はない。

縦断デザイン（結果変数が一つ）の場合

縦断データの場合

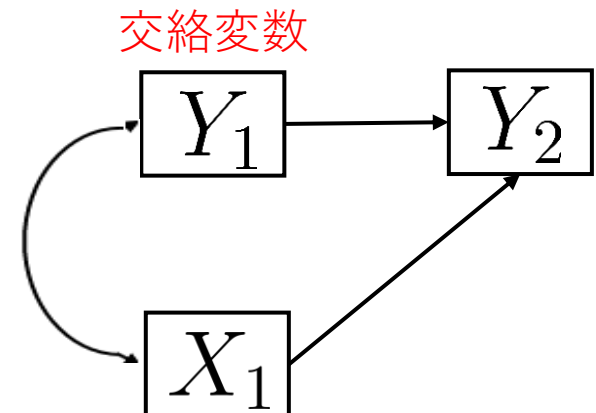
$$y_{i2} = \alpha + \beta x_{i1} + \gamma y_{i1} + e_{it}$$

一時点前の自分自身の変数 y_1 （自己回帰項）が主要な交絡変数。

- （ y 以外の）交絡変数 l がある場合の回帰（共分散分析）モデル

$$y_{i2} = \alpha + \beta x_{i1} + \gamma y_{i1} + \eta l_{i1} + e_{it}$$

単に x_1 から y_2 を説明する回帰分析を行うだけでは因果関係や個人内関係について推測する上では不十分であり、 y_1 （自己回帰項）を投入することが有効。→縦断デザインのメリット



様々な「関係」

- ・ 集団間の相関関係 **group-level relation**

x が大きい人ほど、 y も大きい→交絡が起こりうる。

- ・ 個人内の（共変）関係 **within-person relation**

（ある人について） x が大きくなると y も大きくなる→交絡が起こりうる。

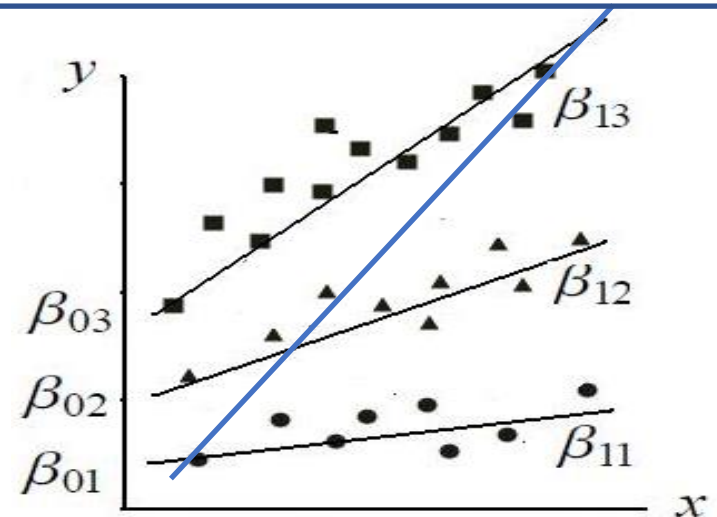
- ・ 因果関係 **causal relation**

x が大きくなるから y も大きくなる。

南風原(2002)

- ・ 集団の相関関係と個人（または下位集団）内の関係は同一でない（e.g., シンプソンパラドックス）。

- ・ 個人内関係も因果関係と同一ではない。



黒線が各個人の関係、青線が集団の関係

個人内の関係と因果推論

・ x を体重、 y を身長とする。通常、体重 x_2 と身長 y_2 の間の（集団間の）相関関係は高いので、 x_1 と y_2 や x_2 と y_1 の間の（集団間の）相関関係が高いのは不思議でない。

→では、例えば体重が重くなれば身長が高くなる（ $x_1 \rightarrow y_2$ ）？

・ y_1 （自己回帰項）を投入することで、**過去（一時点前）の身長が同じ人たちのなかで**、体重が重いほど身長が高くなるか検討できる。

→（集団ではなく）個人内の共変関係への着目の必要性。

→自己回帰項の投入で個人内の共変関係に「接近できる」。

→個人内の共変関係に着目することは因果推論に不可欠。

→存在しうる他の交絡変数 l を適切な統計モデルの下で統制することで、因果推論に更に迫った検討が可能になる（後述）。

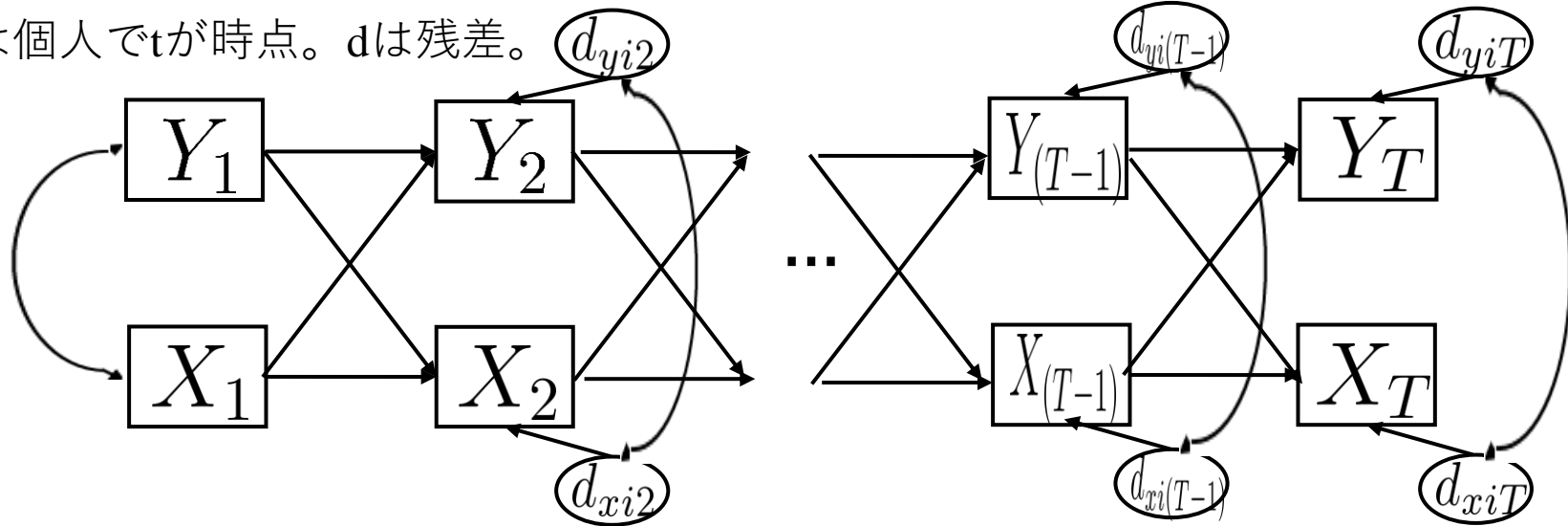
参考：宇佐美慧 (2016). 臨床疫学研究のための統計学—縦断データと個人内相関から探る因果関係— 精神科, 29. 513-518. 個人内関係への着目の重要性についてはHamaker(2012)も参考になる。

縦断デザイン（二つの結果変数）の場合

- 交差遅延パネルモデル(cross-lagged panel model : CLPM)

$$\begin{aligned} y_{it} &= \alpha_{yt} + \beta_{yt}y_{i(t-1)} + \gamma_{yt}x_{i(t-1)} + d_{yit} \\ x_{it} &= \alpha_{xt} + \beta_{xt}x_{i(t-1)} + \gamma_{xt}y_{i(t-1)} + d_{xit} \end{aligned} \quad \begin{pmatrix} d_{yit} \\ d_{xit} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega_{yt}^2 & \\ & \omega_{xt}^2 \end{pmatrix} \right)$$

iは個人でtが時点。dは残差。



α : 切片 β : 自己回帰係数 γ : 交差遅延（クロスラグ）係数

- 2つの γ を通して変数間の相互関係 (reciprocal relation) が推定される。通常SEM(構造方程式モデリング) を用いて推定される。*訳語は定まっていない

CLPM

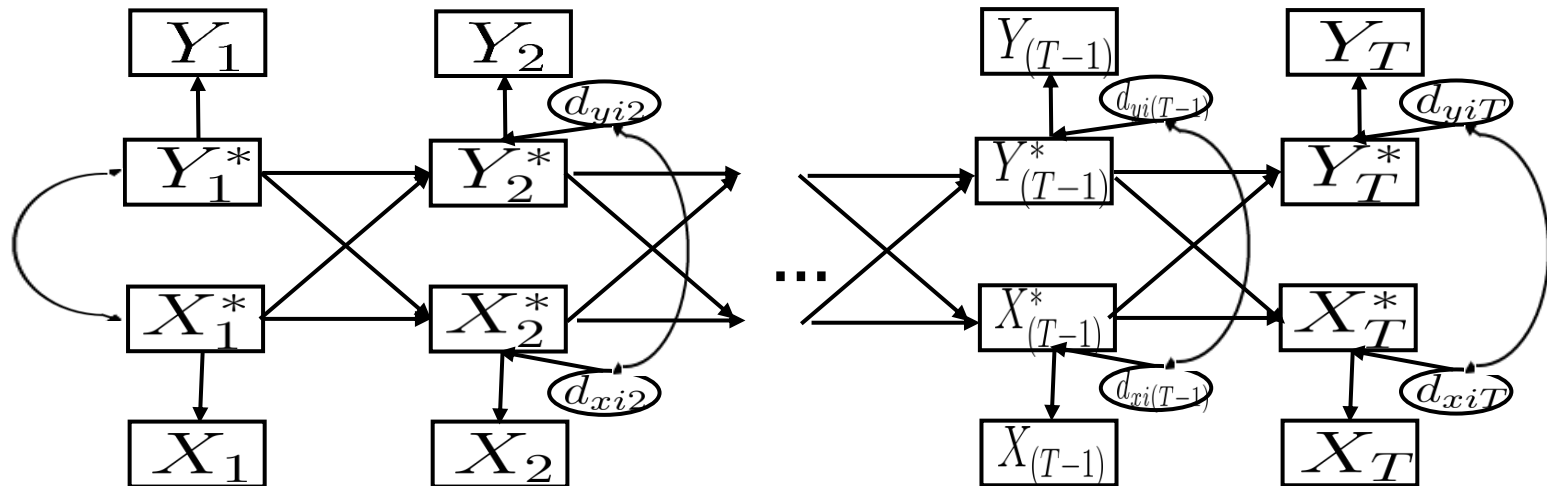
・ CLPMはReciprocal relationを推定する方法として、1980-90年代以降社会科学を中心に急速に普及し、一つのgold standardとなっている(Hamaker et al., 2015; Usami et al., 2019a)。

CLPMの別表現：

$$y_{it} = \mu_{yt} + y_{it}^* \quad y_{it}^* = \beta_{yt}y_{i(t-1)}^* + \gamma_{yt}x_{i(t-1)}^* + d_{yit}$$

$$x_{it} = \mu_{xt} + x_{it}^* \quad x_{it}^* = \beta_{xt}x_{i(t-1)}^* + \gamma_{xt}y_{i(t-1)}^* + d_{xit}$$

μ_{xt}, μ_{yt} : 時点 t の平均 x_{it}^*, y_{it}^* : 個人 i の平均 μ からの偏差



CLPMへの批判 (Hamaker et al., 2015)

Psychological Methods
2015, Vol. 20, No. 1, 102–116

© 2015 American Psychological Association
1082-989X/15/\$12.00 <http://dx.doi.org/10.1037/a0038889>

A Critique of the Cross-Lagged Panel Model

Ellen L. Hamaker and Rebecca M. Kuiper
Utrecht University

Raoul P. P. P. Grasman
University of Amsterdam

The cross-lagged panel model (CLPM) is believed by many to overcome the problems associated with the use of cross-lagged correlations as a way to study causal influences in longitudinal panel data. The current article, however, shows that if stability of constructs is to some extent of a trait-like, time-invariant nature, the autoregressive relationships of the CLPM fail to adequately account for this. As a result, the lagged parameters that are obtained with the CLPM do not represent the actual within-person relationships over time, and this may lead to erroneous conclusions regarding the *presence*, *predominance*, and *sign* of causal influences. In this article we present an alternative model that separates the within-person process from stable between-person differences through the inclusion of random intercepts, and we discuss how this model is related to existing structural equation models that include cross-lagged relationships. We derive the analytical relationship between the cross-lagged parameters from the CLPM and the alternative model, and use simulations to demonstrate the spurious results that may arise when using the CLPM to analyze data that include stable, trait-like individual differences. We also present a modeling strategy to avoid this pitfall and illustrate this using an empirical data set. The implications for both existing and future cross-lagged panel research are discussed.

Keywords: cross-lagged panel, reciprocal effects, longitudinal model, trait–state models, within-person dynamics

- ・ CLPMで推定されているクロスラグ係数は、個人内プロセス (within-personal process) と個人間差 (between-person differences) が混在しており、個人内の関係を正しく評価できないと批判。
- 自己回帰項の統制だけでは個人内関係を捕捉する上で不十分。
- 一つの対処法として random-intercepts CLPM (RI-CLPM) を提案。

RI-CLPM

$$\begin{aligned} y_{it} &= \mu_{yt} + I_{yi} + y_{it}^* & y_{it}^* &= \beta_{yt}y_{i(t-1)}^* + \gamma_{yt}x_{i(t-1)}^* + d_{yit} \\ x_{it} &= \mu_{xt} + I_{xi} + x_{it}^* & x_{it}^* &= \beta_{xt}x_{i(t-1)}^* + \gamma_{xt}y_{i(t-1)}^* + d_{xit} \end{aligned}$$

μ_{xt}, μ_{yt} : 時点tの平均

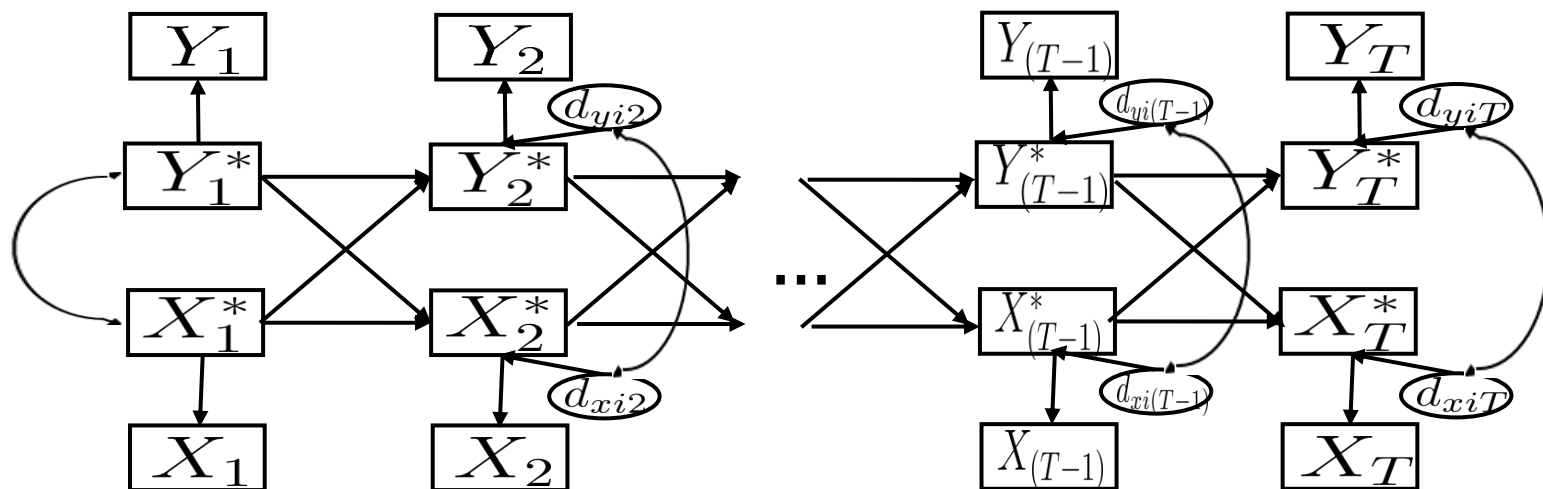
I_{xi}, I_{yi} : 個人iの特性因子(stable trait factor)

x_{it}^*, y_{it}^* : 個人iの期待値 $\mu_{xt} + I_{xi}, \mu_{yt} + I_{yi}$ からの偏差
→ 個人内変化を表す成分！

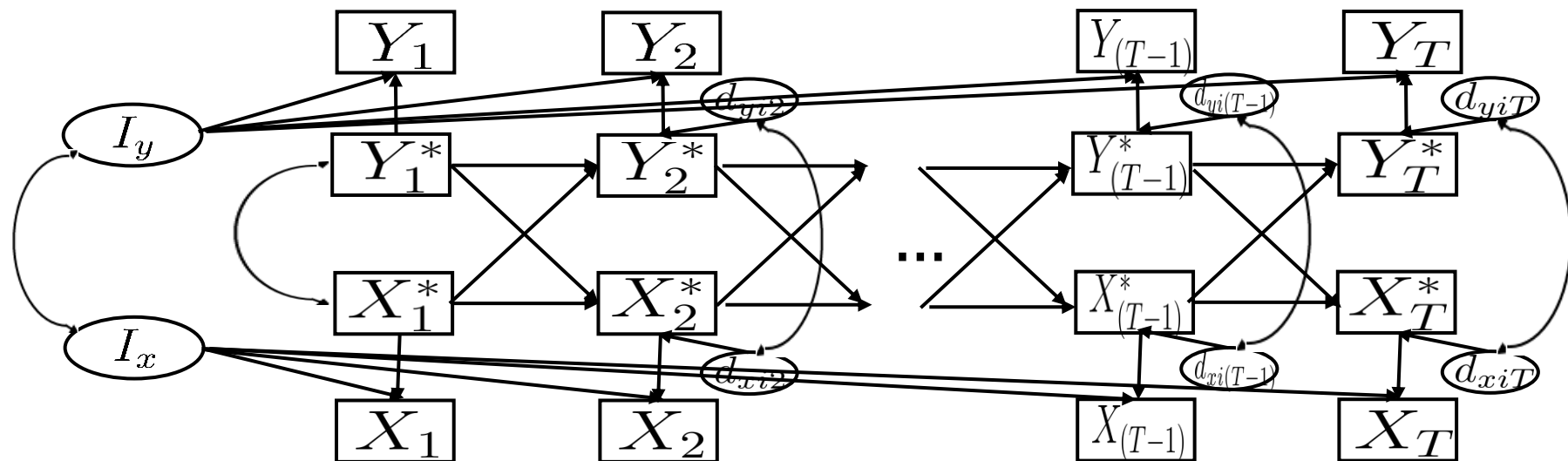
- ・ 測定期間全体を通してみられる個人の測定値の平均的な高低を、その個人の安定した特性と考え、因子得点で表現する。
- ・ 特性因子の平均は0であり、各時点の偏差とは無相関と仮定。

パス図

CLPM



RI-CLPM



RI-CLPM

- ・ 因子負荷を1に固定した2因子モデルと捉えられる。
- ・ 特性因子による統制は、時間不変的(time-invariant)な交絡変数による統制を暗に意味する (Usami et al., 2019a)。したがって、もし時間変動的な交絡変数があればそれも考慮する必要。
- ・ 欧州を中心に急速に利用が拡大（2021年3月時点で引用回数900以上）。
- ・ RI-CLPMの母数の識別には、3時点以上必要。CLPMは2時点で識別可能（しかし、実際の研究の多くが2,3時点; Usami et al., 2019b）。
- ・ RI-CLPMに測定誤差の存在を仮定したモデルはSTARTSと呼ばれる (Usami et al., 2019a)。ただし、これらのモデルはしばしば不適解が生じる→ベイズ推測は一つの対処法 (Ludtke et al., 2018)

データ分析

- ・ 同じ（2変数の）縦断データに対してCLPMとRI-CLPMを当てはめて分析。
- ・ 分析結果

省略

分析結果

- ・最尤法(MLE)ではRI-CLPMで不適解が生じたため、二つのモデルそれぞれでベイズ推定 (blavaanを利用) を行った。
- ・ここでは説明の便宜として、時点間でパス係数・残差分散が同じという (強い) 仮定のもとで分析。
- ・情報量規準(BIC,DIC,WAIC)の観点からはRI-CLPMの方が良い。
- ・CLPMではいずれの交差遅延係数 (γ) も負の推定値で、統計学的にも有意になる。
- ・RI-CLPMでも負の推定値が見られるが、交差遅延係数の確信区間はいずれも0を含む。 このようにしばしば結論がモデル間で変わる！
- ・最初の時点では、各変数の分散のうち特性因子の分散が占める割合は40-50%ほど。 ←個人の特性を統制することの重要性。

分析の実行 (lavaan & blavaan)

- SEMによる分析は、Mplus, OpenMx, R(lavaan, semパッケージ), Amos, Onyxなど様々なソフトウェアから実行可能。

- lavaan (Rosseel, 2012)のHP (インストールやコード例)

<https://lavaan.ugent.be/>

- マニュアル・コード例

<https://lavaan.ugent.be/tutorial/tutorial.pdf> (英語)

<http://www.ec.kansai-u.ac.jp/user/arakit/documents/lavaanTutorial20170124.pdf> (日本語)

豊田秀樹編 (2014). 共分散構造分析R編 東京図書

- blavaan (Merkle & Rosseel, 2018)によるSEMのベイズ推測も実行可。

lavaan

latent variable analysis

About lavaan Tutorial Resources Version History

Tutorial

Overview

Before you start

Installation

Model syntax 1

A CFA example

A SEM example

Model syntax 2

Meanstructures

Multiple groups

Growth curves

Categorical data

Covariance matrix input

Another important type of latent variable models are latent growth curve models. Growth modeling is often used to analyze longitudinal or developmental data. In this type of data, an outcome measure is measured on several occasions, and we want to study the change over time. In many cases, the trajectory over time can be modeled as a simple linear or quadratic curve. Random effects are used to capture individual differences. The random effects are conveniently represented by (continuous) latent variables, often called *growth factors*. In the example below, we use an artificial dataset called *Demo.growth* where a score (say, a standardized score on a reading ability scale) is measured on 4 time points. To fit a linear growth model for these four time points, we need to specify a model with two latent variables: a random intercept, and a random slope:

```
# linear growth model with 4 timepoints
# intercept and slope with fixed coefficients
i =~ 1*t1 + 1*t2 + 1*t3 + 1*t4
s =~ 0*t1 + 1*t2 + 2*t3 + 3*t4
```

In this model, we have fixed all the coefficients of the growth functions. To fit this model, the lavaan package provides a special `growth()` function:

しかし、似たようなモデルは沢山ある

・他にも、交差遅延係数を含む統計モデルとして、例えば

- (1) latent change score (**LCS**) model (Hamagami & McArdle, 2001; McArdle & Hamagami, 2001),
- (2) autoregressive latent trajectory (**ALT**) model (Bollen & Curran, 2004; Curran & Bollen, 2001),
- (3) (a bivariate version of) stable trait autoregressive trait and state (**STARTS**) model (Kenny & Zautra, 1995, 2001).
- (4) Latent curve model with structured residuals (**LCM-SR**) (Curran et al., 2013)

*これら引用文献についてはUsami et al (2019a)を参照。

・これらのモデルは異なる領域から提案され、中には個人内関係の推測を意図して開発されたものもあるが、これらのモデル間の異同は、（心理）統計学者の中でもあまり明確に意識されていない。そもそも、このような他のモデルの存在自体、必ずしも良く知られていない。

・実際の応用研究では、単一のクロスラグモデル（主に**CLPM**）のみが明確な根拠なく利用されることが多く、実際の目的（=個人内変化の推測や因果推論）とは乖離した実践となっている。

・しかも、これら他のモデルの解は、**CLPM**とはしばしば全く異なった値を示すことが経験的に知られている(Usami et al., 2019a; Orth et al., in press)。

A Unified Framework of Longitudinal Models to Examine Reciprocal Relations

Satoshi Usami
University of Tokyo

Kou Murayama
University of Reading and Kochi University of Technology

Ellen L. Hamaker
Utrecht University

Abstract

Inferring reciprocal effects or causality between variables is a central aim of behavioral and psychological research. To address reciprocal effects, a variety of longitudinal models that include cross-lagged relations have been proposed in different contexts and disciplines. However, the relations between these cross-lagged models have not been systematically discussed in the literature. This lack of insight makes it difficult for researchers to select an appropriate model when analyzing longitudinal data, and some researchers do not even think about alternative cross-lagged models. The present research provides a unified framework that clarifies the conceptual and mathematical similarities and differences between these models. The unified framework shows that existing longitudinal models can be effectively classified based on whether the model posits unique factors and/or dynamic residuals and what types of common factors are used to model changes. The latter is essential to understand how cross-lagged parameters are interpreted. We also present an example using empirical data to demonstrate that there is

Unified Framework

$$x_{it} = f_{xit} + \epsilon_{xit} \quad y_{it} = f_{yit} + \epsilon_{yit}$$

$$f_{xit} = [\mu_{xt} + \{I_{xi} + (t-1)S_{xi}\}] + f_{xit}^*$$

$$f_{yit} = [\mu_{yt} + \{I_{yi} + (t-1)S_{yi}\}] + f_{yit}^*$$

$$f_{xit}^* = \{A_{xi} + (t-1)B_{xi}\} + \beta_x f_{xi(t-1)}^* + \gamma_x f_{yi(t-1)}^* + d_{xit}$$

$$f_{yit}^* = \{A_{yi} + (t-1)B_{yi}\} + \beta_y f_{yi(t-1)}^* + \gamma_y f_{xi(t-1)}^* + d_{yit}$$

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, 24, 637-657.

<https://psycnet.apa.org/fulltext/2019-21491-001.pdf>

- ・ 様々なモデル間の数理的・概念的関係を整理。
- ・ 個人内変化を捕捉する上では特性因子**Iのみ**含めることが有効であり、（ALT, LCS, LCM-SR のように）**S・A・B**など他の共通因子を含めることは却って過剰調整(overadjustment)を誘発することを指摘。
- ・ **RI-CLPM**は個人内変化を捉える上で有効な方法ではあるが、絶対的ではない。交絡変数を適切にモデリングできることが重要。

他のモデルの問題点は？（１）

LCM-SR

$$\begin{aligned} y_{it} &= I_{yi} + (t-1)S_{yi} + y_{it}^* & y_{it}^* &= \beta_{yt}y_{i(t-1)}^* + \gamma_{yt}x_{i(t-1)}^* + d_{yit} \\ x_{it} &= I_{xi} + (t-1)S_{xi} + x_{it}^* & x_{it}^* &= \beta_{xt}x_{i(t-1)}^* + \gamma_{xt}y_{i(t-1)}^* + d_{xit} \end{aligned}$$

・ μ の代わりに、LGM（潜在成長モデル）のように成長因子 (growth factor) I, S を入れる。→ x^*, y^* は個人内変動を表し、 γ は個人内変化の関係を表すという主張がされている。

・ しかし $(t-1)S$ の項によって（研究者が関心のある）個人の変化・成長の成分を表現（統制）した後の量である x^*, y^* の関係をモデリングしているため、最早 γ に注目することの意義は乏しいのでは？

→ “throwing the baby out with the bathwater”, resulting from wrongly controlling S (Usami et al., 2019a).

他のモデルの問題点は？（２）

LCS（測定誤差を仮定しない場合; unified frameworkによる表現）

$$y_{it} = y_{it}^* \quad y_{it}^* = A_{yi} + \beta_{yt}y_{i(t-1)}^* + \gamma_{yt}x_{i(t-1)}^* + d_{yit}$$

$$x_{it} = x_{it}^* \quad x_{it}^* = A_{xi} + \beta_{xt}x_{i(t-1)}^* + \gamma_{xt}y_{i(t-1)}^* + d_{xit}$$

・観測値の式（前半）ではなく、ラグ付きの回帰の式（後半）に共通因子（A）が含まれている。

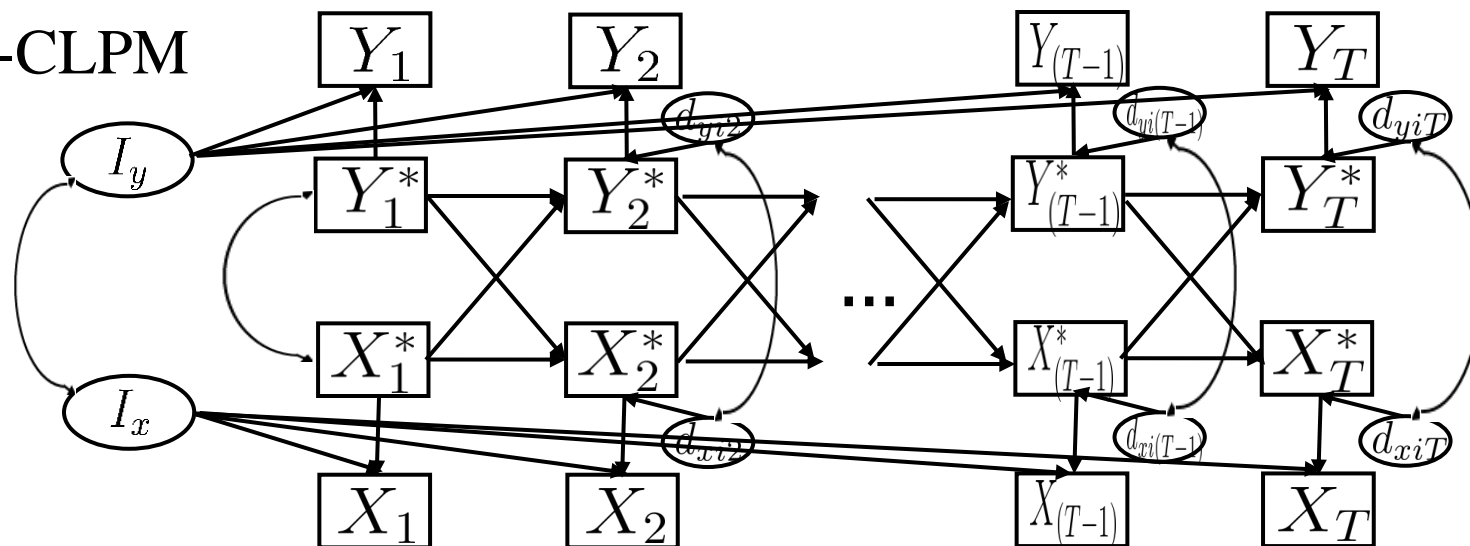
・特性因子（I）と共通因子Aの概念的・数理的役割は異なる。
Aは**累積因子(accumulating factor)**と呼ばれる(Usami et al., 2019a).

・特性因子は各時点の観測値に対して**直接効果**の形でのみ寄与するのに対し、accumulating factorはある時点への寄与が後続の時点まで累積しながら影響する。つまり**間接効果**がある。

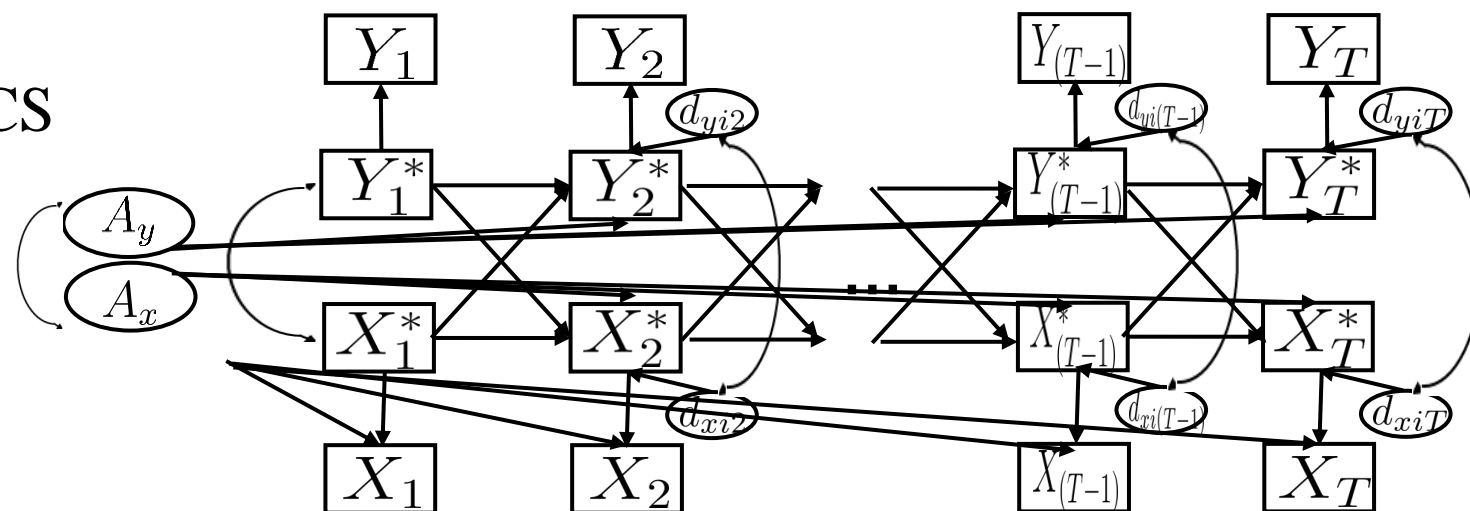
→“throwing the baby out with the bathwater”問題がLCSでも生じる。

パス図

RI-CLPM



LCS



- ALTでもLCSと同様の問題が生じ、個人内変化の関係を推測する目的においては適切ではない。
- 経済学でよく利用される動的モデル(dynamic model)に含まれる個人効果の項もaccumulating factorと解釈できる。
- (ラグ変数を伴う) 階層線形モデル内のランダム切片も (適切なセンタリングをしない限り) accumulating factorを意味し、特性因子とは異なる。

他のモデルの問題点は？（３）

GCLM (Zyphur et al., 2020ab; unified frameworkによる表現)

$$\begin{aligned} y_{it} &= y_{it}^* & y_{it}^* &= (t-1)B_{yi} + \beta_{yt}y_{i(t-1)}^* + \gamma_{yt}x_{i(t-1)}^* + \delta_{yt}d_{yi(t-1)} + \zeta_{yt}d_{xi(t-1)} + d_{yit} \\ x_{it} &= x_{it}^* & x_{it}^* &= (t-1)B_{xi} + \beta_{xt}x_{i(t-1)}^* + \gamma_{xt}y_{i(t-1)}^* + \delta_{xt}d_{xi(t-1)} + \zeta_{xt}d_{yi(t-1)} + d_{xit} \end{aligned}$$

- ・時变的な効果をもつ **accumulating factor (B)** と移動平均項($d_{(t-1)}$)をモデリング。

- ・このモデルは **RI-CLPM** のように特性因子(I)を統制していない。

- ・アウトカムの予測モデルとしては機能しうるが、個人内変化の関心の捕捉や因果推論の目的からは適さない。 (Usami, in press, *Structural Equation Modeling*)

STRUCTURAL EQUATION MODELING: A MULTIDISCIPLINARY JOURNAL
<https://doi.org/10.1080/10705511.2020.1821690>

 **Routledge**
Taylor & Francis Group

 OPEN ACCESS 

On the Differences between General Cross-Lagged Panel Model and Random-Intercept Cross-Lagged Panel Model: Interpretation of Cross-Lagged Parameters and Model Choice

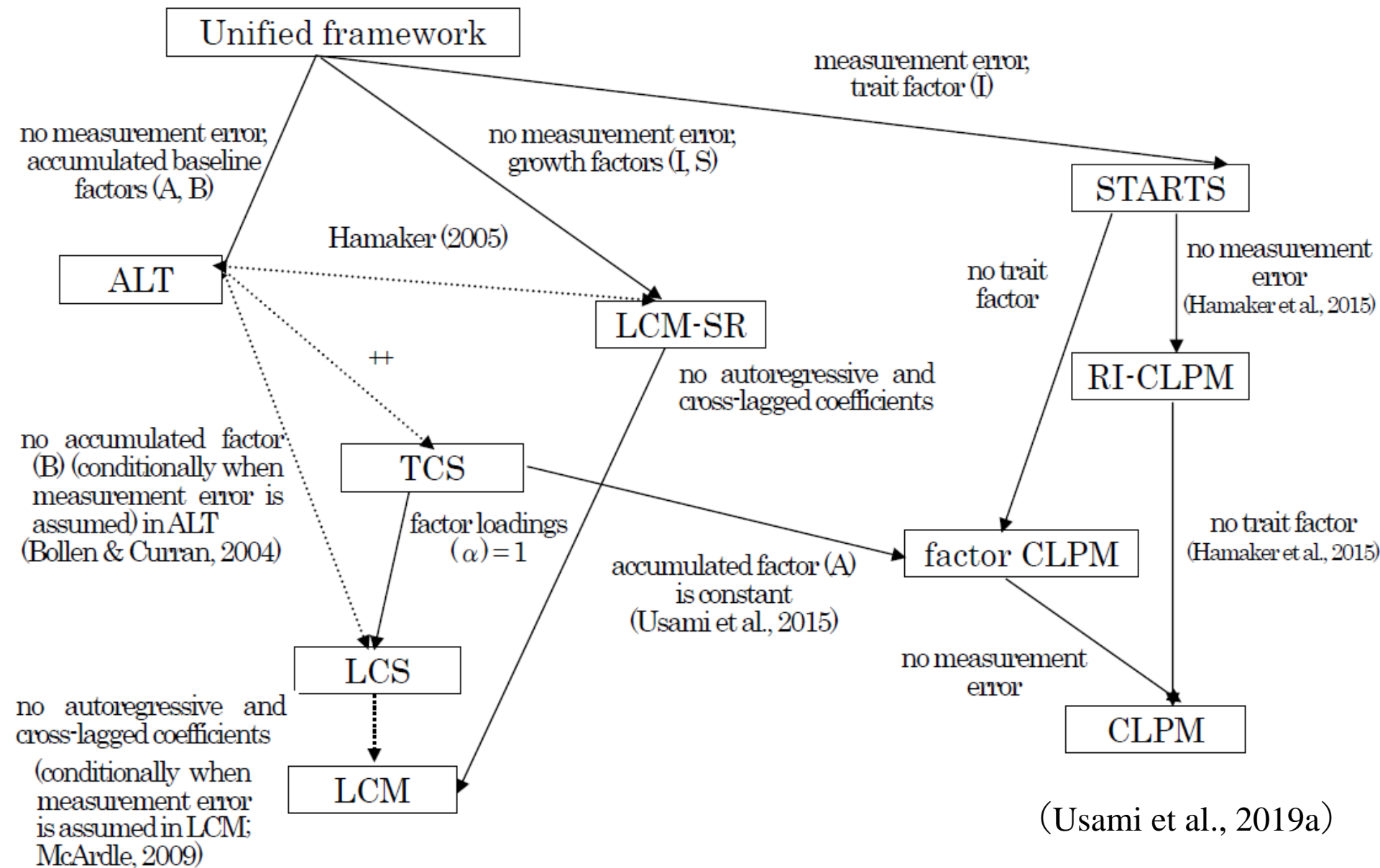
Satoshi Usami 

University of Tokyo

ABSTRACT

Many methods have been developed to infer reciprocal relations between longitudinally observed variables. Among them, the general cross-lagged panel model (GCLM) is the most recent development as a variant of the cross-lagged panel model (CLPM), while the random-intercept CLPM (RI-CLPM) has rapidly become a popular approach. In this article, we describe how common factors and cross-lagged parameters included in these models can be interpreted, using a unified framework that was recently developed. Because common factors are modeled with lagged effects in the GCLM, they have both direct and indirect influences on observed scores, unlike stable trait factors included in the RI-CLPM. This indicates that the GCLM does not control for stable traits as the RI-CLPM does, and that there are interpretative differences in cross-lagged parameters between these models. We also explain that including such common factors as well as moving-average terms in the GCLM makes this interpretation very complicated.

Unified frameworkに基づくモデル間の概念的関係



ここまでのまとめ

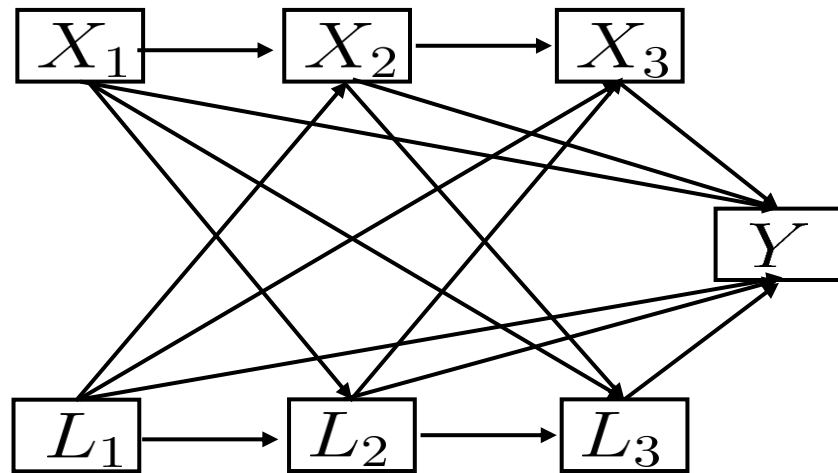
- 因果関係に接近する上で個人内関係に着目することが重要であり、そのためには自己回帰項や特性因子による統制・モデリングが有効。

→縦断デザインのメリット

- 2変数の系列から変化の関係を調べる目的で利用されてきたCLPMは、特性因子を統制しておらず個人内の関係を捕捉する上で不十分。
- 個人内関係の捕捉を意図したモデルや類似のモデルも数多く提案されてきたが、概念的・数理的整理は最近までなされてこなかった。
- 時不変的な交絡変数と考えられる特性因子をモデリングしたRI-CLPMは個人内関係を捕捉する上で一つの有効な方法。近年爆発的に応用が進み、二次分析の試みも活発化している(Orth et al., in press)。
- RI-CLPMの適用には $T \geq 3$ の縦断データが必要。
- RI-CLPMであっても、ありうる交絡変数を適切に特定・測定・収集しモデリングする必要性は変わらない。縦断デザインにおける（特に、個人内変化に基づく）因果推論の方法はまだ発展途上（後述）。

縦断研究における因果推論の難しさ

- 縦断研究では、存在する交絡変数を**時点別**に考え、これらの関係について正しくモデル設定をする必要がある。



* Yが最終時点でのみ観測される場合

- 交絡変数Lは将来の独立変数(X)とアウトカム(Y)の双方に影響。
- 不適切なモデリングの例（このモデルでは各時点のxに影響する交絡変数の種類の違いを考慮できていない）

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 l_{i1} + \gamma_2 l_{i2} + \gamma_3 l_{i3} + e_{it}$$

- ・ さきのDAGに基づくSEMを適用する方法は一案。
- ・ ただし、変数間の関係が基本的に線形関数（線形モデル）で記述できるという強い仮定が通常のSEM（と共分散分析モデル・パスモデル、RI-CLPM含む）にはあり、これはしばしば統計的因果推論の文献では批判的に捉えられる（Hong, 2015）。
- ・ 疫学・医学領域では、**周辺構造モデル**(Marginal Structural Model：MSM Robins, 1999; Robins et al., 2000, Robins & Hernan, 2009)や**構造ネスト平均モデル**(Structural Nested Mean Model: SNMM, Robins, 1994)と呼ばれるSEM以外の方法論が良く利用される。
- ・ 海外の心理学・発達科学研究ではまだこれらの方法の適用例は少ないが、MSMについては手続きが比較的単純であるため、その利用が増えてきている。

交絡変数Lを含むRI-CLPMの一表現

$$Y_{ik} = \mu_k^{(Y)} + I_i^{(Y)} + Y_{ik}^*, \quad X_{ik} = \mu_k^{(X)} + I_i^{(X)} + X_{ik}^*, \quad L_{ik} = \mu_k^{(L)} + I_i^{(L)} + L_{ik}^*$$

$$Y_{ik}^* = \alpha_k^{(Y)} Y_{i(k-1)}^* + \beta_k^{(Y)} X_{i(k-1)}^* + \gamma_k^{(Y)} L_{i(k-1)}^* + d_{ik}^{(Y)}$$

$$X_{ik}^* = \alpha_k^{(X)} Y_{ik}^* + \beta_k^{(X)} X_{i(k-1)}^* + \gamma_k^{(X)} L_{ik}^* + d_{ik}^{(X)}$$

$$L_{ik}^* = \alpha_k^{(L)} Y_{i(k-1)}^* + \beta_k^{(L)} X_{i(k-1)}^* + \gamma_k^{(L)} L_{i(k-1)}^* + d_{ik}^{(X)},$$

- 周辺構造モデル(Robins et al., 2000; VanderWeele et al., 2011)

$$E(Y|X_1, X_2, X_3) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}$$

- Yが最終時点でのみ観測される場合に主に利用される。
- 各時点のXに関わる交絡変数Lのデータを基に各対象の重みを時点別に算出して、それを基に重み付き回帰を行い、
(様々な交絡変数の背景をもつ集団が居る中での) 平均的なXの効果
を推定する。

- ただし、重みの推定結果が（特に時点数が多いときに）不安定になりやすいことが一般に知られている。

- Usami (2020)では、各対象の個人内変動部分を因子分析モデルを用いて予測してから、それを用いてMSMやSNMMを適用する方法を提案。→個人内変化に基づく因果推論。

- 縦断デザインを活用した（特に、個人内変化に基づく）因果推論の方法の開発はまだ発展途上と言える。

引用文献

- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B. (2014). *Bayesian Data Analysis 3rd edition*. Boca Raton, Chapman and Hall–CRC
- 南風原朝和 (2002). 心理統計学の基礎 有斐閣
- Hamaker, E.L., (2012). Why researchers should think "within-person": a paradigmatic rationale. *Handbook of Research Methods for Studying Daily Life*. The Guilford Press, New York, NY, pp. 43–61.
- Hamaker, E.L., Kuiper, R.M., Grasman, R.P.P.P., 2015. A critique of the cross-lagged panel model. *Psychological Methods*, 20, 102–116. <https://doi.org/10.1037/a0038889>.
- Hamaker, E.L., Mulder, J.D., & van IJzendoorn, M.H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development, *Developmental Cognitive Neuroscience*, 46, 100867, ISSN 1878-9293, <https://doi.org/10.1016/j.dcn.2020.100867>.
- Hong, G. (2015). *Causality in a social world: Moderation, mediation and spill-over*. John Wiley & Sons.
- Lüdtke, O., Robitzsch, A., & Wagner, J. (2018). More stable estimation of the STARTS model: A Bayesian approach using Markov Chain Monte Carlo techniques. *Psychological Methods*, 23, 570–593.
- Merkle, E. C. & Rosseel, Y. (2018). blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software*, 85, 1-30.
- Orth, U.D., Clark, A.M., Donnellan, B., & Robins, R. W. (in press). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*. <http://dx.doi.org/10.1037/pspp0000358>
- Robins, J.M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics -Theory and Methods-*, 23, 2379-2412.
- Robins, J.M. (1999). Marginal structural models versus structural nested models as tools for causal inference. *Epidemiology*, 116, 95-134.

引用文献

- Robins, J.M., Hernán, M.A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11, 550-560.
- Robins, J.M., & Hernan, M.A. (2009). Estimation of the causal effects of time-varying exposures. In G. Fitzmaurice et al. (Eds.), *Handbooks of modern statistical methods: Longitudinal data analysis* (pp. 553-599). Boca Raton: CRC Press.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48, 1-36.
- 豊田秀樹編 (2014). 共分散構造分析R編 東京図書
- 宇佐美慧 (2016). 臨床疫学研究のための統計学—縦断データと個人内相関から探る因果関係—精神科, 29. 513-518.
- Usami, S., Murayama, K., & Hamaker, E.L. (2019a). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, 24, 637-657
- Usami, S., Todo, N., & Murayama, K. (2019b). Modeling reciprocal effects in medical research: Critical discussion on the current practices and potential alternative models. *PLOS ONE*. 14(9): e0209133.
- Usami, S. (2020). Within-person variability score-based causal inference: A two-Step semiparametric estimation for joint effects of time-varying treatments. arXiv.
http://satoshiusami.com/Arxiv_20200703_ALL.pdf
- Usami, S. (in press). On the differences between general cross-lagged panel model and random-intercept cross-lagged panel model: Interpretation of cross-lagged parameters and model choice. *Structural Equation Modeling*.
<https://www.tandfonline.com/doi/pdf/10.1080/10705511.2020.1821690?needAccess=true>

引用文献

- VanderWeele, T. J., Hawkley, L. C., Thisted, R. A., & Cacioppo, J. T. (2011). A marginal structural model for loneliness: Implications for intervention trials and clinical practice. *Journal of Consulting and Clinical Psychology, 79*, 225–235.
- Zyphur, M.J., Allison, P.D., Tay, L., Voelkle, M.C., Preacher, K.J., Zhang, Z., Hamaker, E.L., Shamsollahi, A., Pierides, D.C., Koval, P., & Diener, E. (2020a). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods, 23*(4), 651–687.
- Zyphur, M.J., Voelkle, M.C., Tay, L., Allison, P.D., Preacher, K.J., Zhang, Z., Hamaker, E.L., Shamsollahi, A., Pierides, D.C., Koval, P., & Diener, E. (2020b). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods, 23*(4), 688–716.