

背景と目的

課題

識別可能条件の
概要

数学的基盤

定理説明

定理説明

数値実験

数値実験

未観測交絡因子の存在下での総合効果の識別可能条件

佐藤 力

富山県立大学
u220029@st.pu-toyama.ac.jp

November 26, 2024

1. 線形構造方程式モデルを用いた因果分析が主題。
2. 処理変数と反応変数の間に未観測交絡因子がある場合、総合効果（処理変数を操作した場合の反応変数の変化量）を推定する識別可能条件を探る.

背景と目的

課題

識別可能条件の
概要

数学的基盤

定理説明

定理説明

数値実験

数値実験

- ・ 未観測の交絡因子は因果推論における大きな障壁
- ・ これらの因子が変数間の真の関係を歪める可能性がある
- ・ 既存の識別可能条件である「バックドア基準」「フロントドア基準」や「条件付き操作変数法」が適用できない状況に対応するため、新たな条件を提案。

背景と目的

課題

識別可能条件の
概要

数学的基盤

定理説明

定理説明

数値実験

数値実験

バックドア基準

処理変数から反応変数への直接的な因果経路を遮断する変数集合を利用。

フロントドア基準

処理変数と反応変数の間の中間変数を用いて識別

条件付き操作変数法

条件付き独立性を利用した識別方法。

提案

因子モデルに基づき、未観測交絡因子が複数存在する場合にも対応可能な新たな識別可能条件を提示。

背景と目的

課題

識別可能条件の
概要

数学的基盤

定理説明

定理説明

数値実験

数値実験

提案作

1. パスダイアグラム（有向非巡回グラフ）と線形構造方程式モデルの導入。
2. 未観測交絡因子が存在する場合の共分散行列の表現。
3. 総合効果の識別可能条件を定理として形式化。

定理 1: 単一未観測因子を仮定。

定理 2-3: 複数の未観測因子に対応。

定理 1: 単一未観測因子を仮定

背景

未観測交絡因子 U が 1 つだけ存在すると仮定します。

処理変数 X 、反応変数 Y 、および他の観測可能な変数 W 、 Z を含むパスダイアグラムで、未観測因子の影響を考慮。

ポイント

未観測因子が直接影響を及ぼしている場合でも、観測変数間の条件付き独立性を利用して識別が可能。

観測される変数 W, Z を適切に選択すれば、未観測因子の情報を間接的に取り込むことで、処理変数 X から反応変数 Y への総合効果 T_{yx} を計算できます。

定理 2-3: 複数の未観測因子に対応

背景

実際の問題では、複数の未観測因子が存在することが多く、単一因子モデルでは不十分

線形構造方程式モデルを、複数の未観測因子を含む形に拡張。

ポイント

各未観測因子について、条件付き独立性を利用することで、交絡の影響を統計的に取り除けます。

モデルの概要

線形構造方程式モデルを採用。

処理変数 X 、反応変数 Y 、未観測交絡因子 U 、および観測可能な補助変数 Z が存在。

総合効果 T_{xy} を推定することが目標

パラメータ設定

未観測交絡因子 U の影響を反映するため、2つのパラメータを変
 $a_{x,y} \ a_{z,x} \in (0.1, 0.1), (0.1, 0.5), (0.5, 0.1), (0.5, 0.5)$ と設定

既存の識別可能条件と比較し、推定精度を検証。
提案条件が一定の条件下で他の手法より優れることを確認。

表 3. 数値実験結果

	$(\alpha_{xu}, \alpha_{xz})$	$n = 500$	$n = 1000$	$n = 1500$	$n = 2000$
バックドア基準	(0.1,0.1)	0.0018	0.0009	0.0006	0.0005
	(0.1,0.5)	0.0019	0.0010	0.0006	0.0005
	(0.5,0.1)	0.0022	0.0011	0.0007	0.0005
	(0.5,0.5)	0.0025	0.0012	0.0008	0.0006
条件付き操作変数	(0.1,0.1)	3.6228	0.2800	0.0761	0.0519
	(0.1,0.5)	0.0104	0.0051	0.0035	0.0025
	(0.5,0.1)	3.1562	0.1154	0.0641	0.0477
	(0.5,0.5)	0.0056	0.0025	0.0018	0.0013
定理 1	(0.1,0.1)	0.0119	0.0014	0.0009	0.0006
	(0.1,0.5)	0.1919	0.0155	0.0027	0.0016
	(0.5,0.1)	4.5311	0.0993	0.0118	0.0075
	(0.5,0.5)	6.3753	1.1822	0.0347	0.0125

バックドア基準

未観測交絡因子が存在しない状況では適切に動作することを示している。

条件付き操作変数法

未観測交絡因子の影響が強い場合、推定誤差が非常に大きくなる
サンプルサイズが増加すると誤差は減少するが、依然として提案条件より
精度が劣る。

定理 1

サンプルサイズが増加するにつれ、推定精度がさらに向上