

操作変数法への応用

機械学習

川田恵介
東京大学
keisukekawata@iss.u-tokyo.ac.jp

2026-01-05

1 操作変数法: 復習

1.1 因果/経済モデル推定への応用

- ・機械学習は、より複雑なモデルにより定義されるパラメタ推定へ応用できる
 - (線型/非線形)方程式体系により定義される因果効果/経済モデル
 - 本スライドでは、操作変数法への応用を紹介

1.2 操作変数法

- ・局所的な実験的状況を活用する因果効果の識別方法
 - ランダム化された変数(操作変数)により、介入のみが”誘発”される方法

1.3 実例 Twin experiment

- ・Griffen, Nakamuro and Inui (2015)
- ・「女性の労働供給」の議論の中で、子供が与える影響は大きな論点
 - 子育て負担が女性に集中しがち → 労働供給減少
 - 生活/教育費の増大 → 労働供給増加
- ・子供数の決定には、無数の背景変数
- ・操作変数: 双子が生まれたかどうか

1.4 推定と識別の区別

- ・識別: 無限大の事例数があるランダムサンプリング・データから、因果効果をどのように計算すれば良いのか?
 - ≈ 母集団のどのような特徴と因果効果が一致するのか
- ・推定: 限られた事例から、母集団のどのような特徴を推定すれば良いのか?

1.5 Twin experiment の識別

- ・「双子が生まれるかどうか」はランダムに決まっているならば
 - ▶ 労働供給への双子の効果 = 母集団における「双子がいる家計といない家計の労働供給の差」
 - ▶ 子供の数への効果 = 「双子がいる家計といない家計の子供の数の差」

1.6 Twin experiment の識別

- ・双子は子供の数のみを通じて、労働供給に影響を与えるのであれば(除外制約)、

▶ 子供の数が労働供給に与える効果

$$= \frac{\text{双子の労働供給への効果}}{\text{双子の子供の数への効果}}$$

2 構造方程式モデル

2.1 構造モデル

- ・前期では、潜在結果モデルに基づく、因果効果の定義と識別を紹介
 - ▶ 他の枠組みとしては、構造モデル (Structural Equation Model; SEM)が有力
 - ▶ 変数の決定過程について、複雑な仮定を導入できる
 - 経済学における伝統的なアプローチであり、DAG を利用した表現も注目されている

2.2 潜在結果モデルとの関係性

- ・潜在結果モデルと補完的な関係にあり、両方理解すると便利
 - ▶ Wang, Richardson and Robins (2025)
 - ▶ Econometrics Journal の特集号

2.3 Y の構造モデル

- ・労働供給 (Y) は、以下のモデルに従って決まるとする

$$Y = \underbrace{\tau_D \times D}_{\text{子供の数}} + \underbrace{f_Y(X)}_{\text{観察可能な決定要因}} + \underbrace{U_Y}_{\text{観察不可能な決定要因}}$$

- ▶ D, X, U_Y が同一の個人内では、 Y は同じ
- ▶ U_Y には、データから観察されない要因も含む

2.4 D の構造モデル

- ・ D は、以下のモデルに従って決まるとする

- $\underbrace{D}_{\text{子供の数}} = f_D(X) + \underbrace{U_D}_{\text{観察不可能な決定要}}$
- U_D には、データから観察されない要因も含む

2.5 因果効果の定義

- U_Y, X を一定にしたままで、 Y を変化させる
- 因果効果 =

$$\begin{aligned} \tau_D \times (D + 1) + U_Y - (\tau_D \times (D + 1) + U_Y) \\ = \tau_D \end{aligned}$$

2.6 因果効果の識別

- 計量経済学の核となるアイディア
- 観察できない要因 (U_Y, U_D)を確率変数として扱う (ノーベル記念経済学賞のプレスリリース)
- もし U_Y, U_D が存在しないのであれば、データが異なっても $E[Y | D, X]$ は一定
 - サンプルング誤差 (データが異なると結果が違う) の原因は、 U_Y, U_D

2.7 因果効果の識別

- Y の条件付き母平均 $E[Y | d, x]$ と構造式の間には、以下の関係がある:

$$\begin{aligned} E[Y | d, x] &= \tau_D \times d + f(x) + E[U_Y | d, x] \\ &E[Y | d + 1, x] - E[Y | d, x] \\ &= \tau_D + E[U_Y | d + 1, x] - E[U_Y | d, x] \end{aligned}$$

2.8 注意

- 統計モデルと構造モデルは厳密に区別する必要がある
- Y を D, X で回帰する際には、以下の統計モデルを推定する

$$Y = \beta_0 + \beta_D \times D + f(X) + u$$

- 基本的には、 $u \neq U_Y$ と $\beta_D \neq \tau_D$

2.9 注意

- $\beta_D = E[Y | d + 1, x] - E[Y | d, x]$
- $= \tau_D$
- $+ E[U_Y | d + 1, x] - E[U_Y | d, x]$

2.10 識別の条件

- もし現実の社会(母集団)において、 $E[U_Y | d+1, x] = E[U_Y | d, x]$ ならば

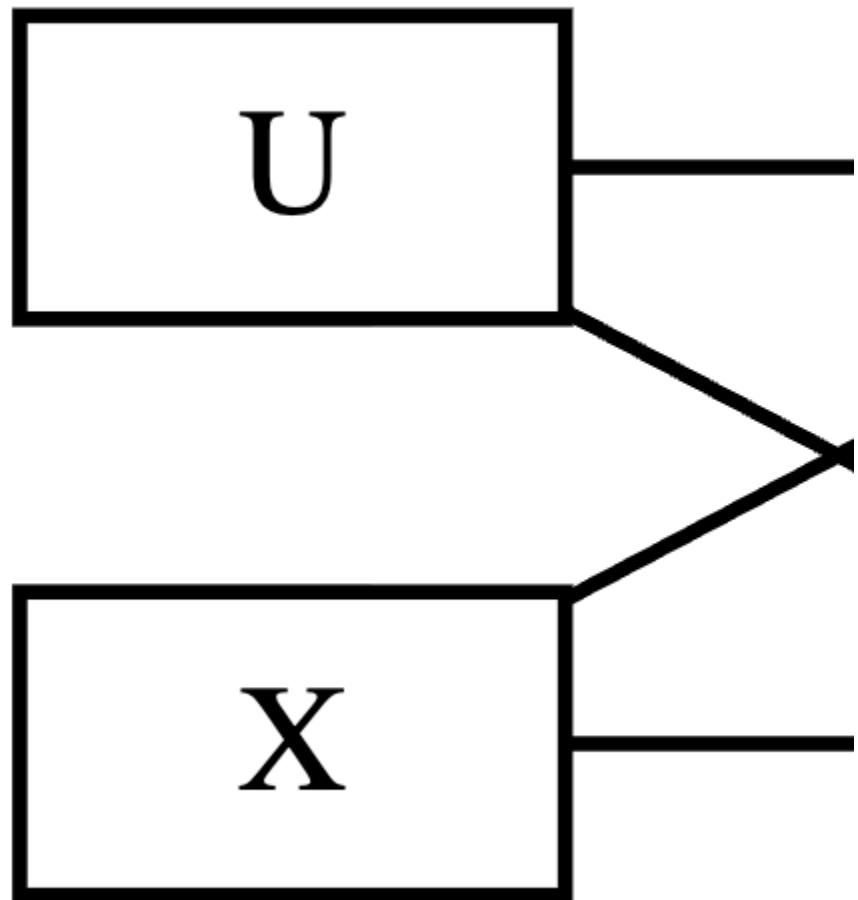
$$\tau_D = \beta_D$$

- U_D と U_Y が独立であれば、保証できる
 - D にも Y にも影響を与える要因 U がない

2.11 ランダム化対象実験による識別

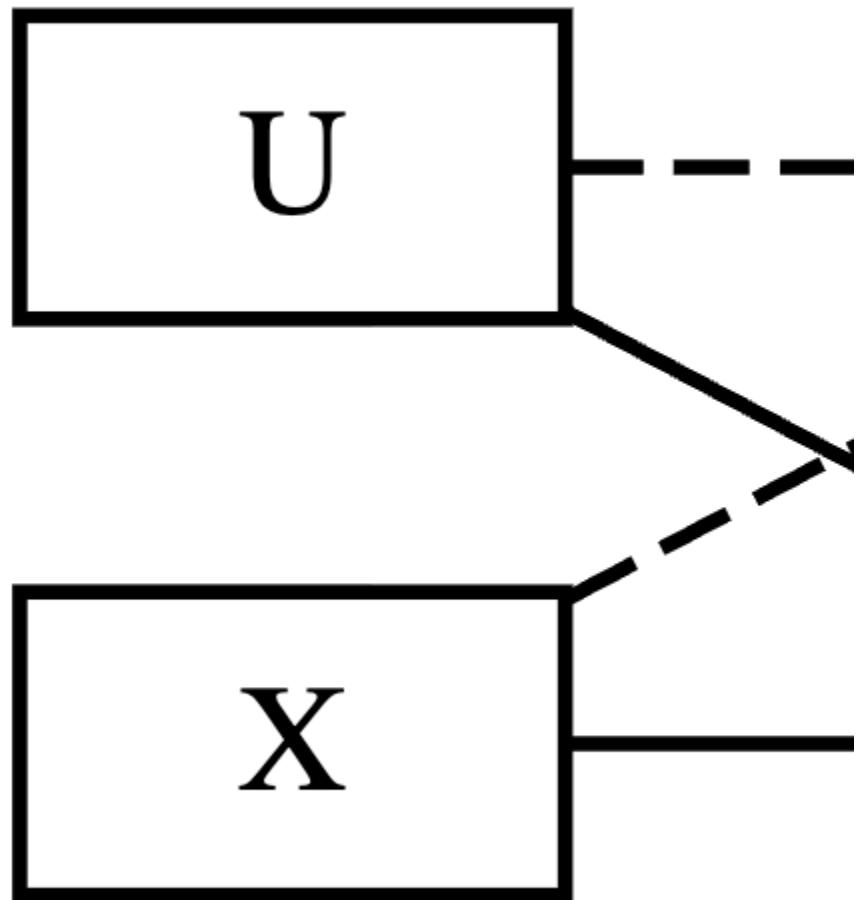
- D はコイントスで決める
 - $U_D = \text{コイントスの結果}$
- U_D は完全にランダムに決まっているので、 U は存在しない

2.12 DAG による表現



- 矢印 = 影響
- U 以外のその他要因は省略

2.13 DAG による表現: RCT



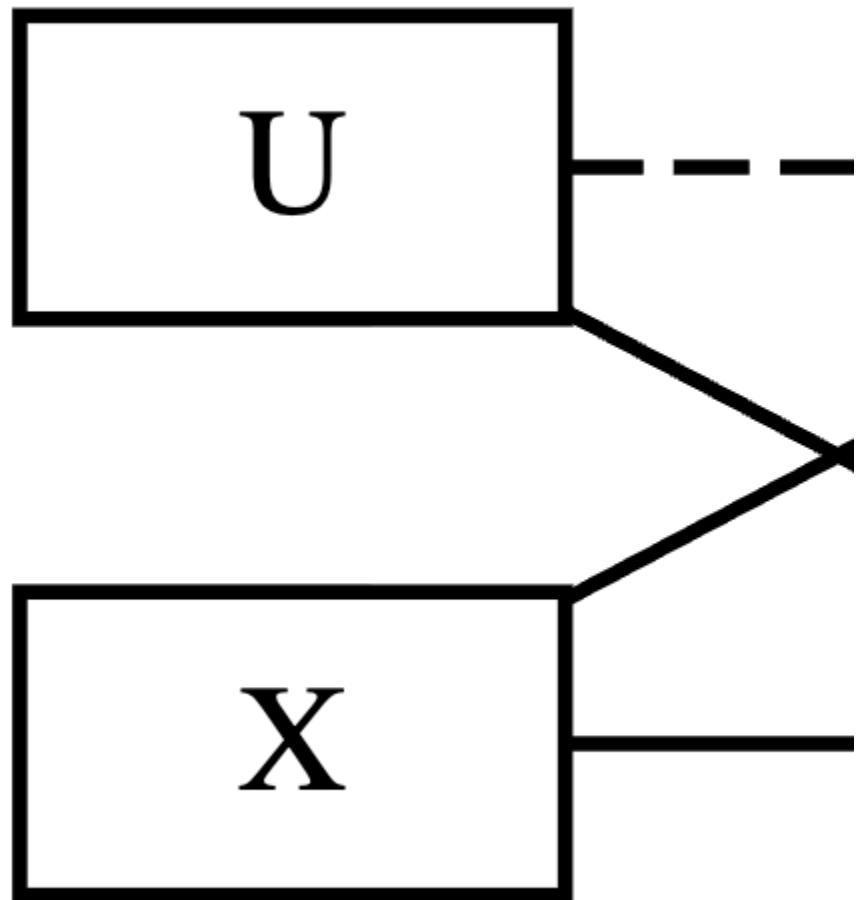
- 点線: 独立

2.14 条件付き独立による識別

- X が共通であれば、子供の数はランダムに決まっている

$$\rightarrow \beta_D = \tau_D$$

2.15 DAG による表現: RCT



2.16 推定方法

- ・ 残差回帰を活用するために、以下のように書き換える

$$Y - E[Y | X] = \beta_D \times (D - E[D | X]) + U_Y$$

- ・ 以下の手順で推定できる

1. $E[Y | X], E[D | X]$ を機械学習等で推定する
2. Y の予測誤差を D の予測誤差で回帰する

3 操作変数

3.1 構造モデル

- ・構造モデルは、

$$Y = \tau_D \times D + f_Y(X) + \underbrace{U_Y}_{E[U_Y|D,X]=0}$$

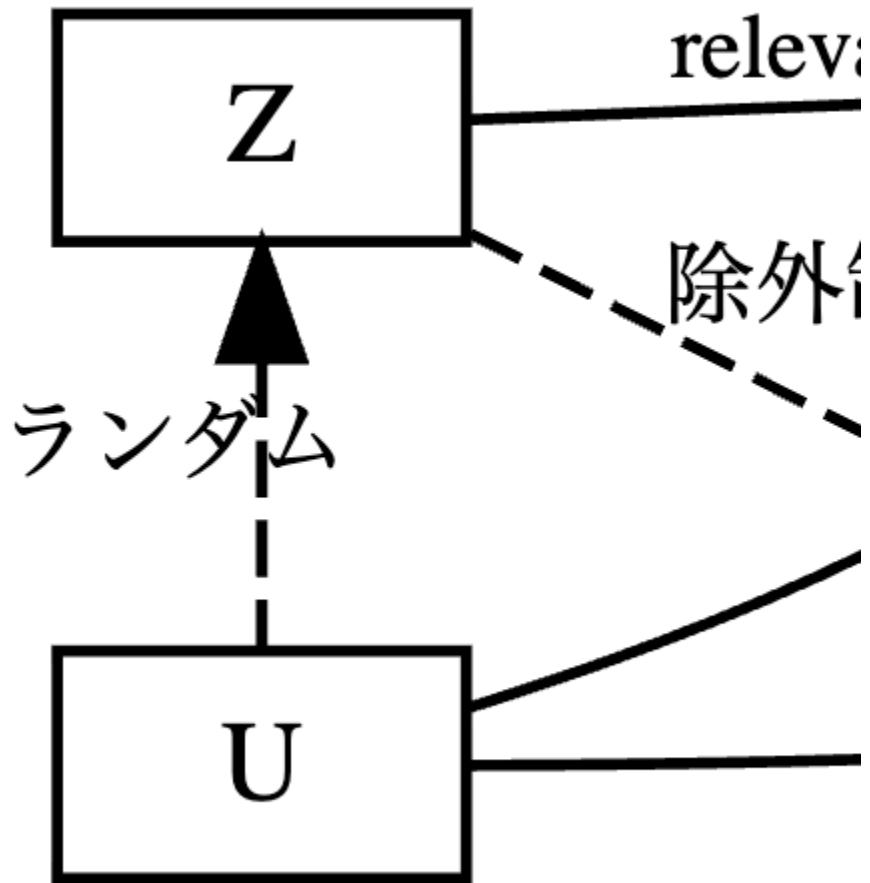
$$D = \tau_Z \times Z + f_D(X) + \underbrace{U_D}_{E[U_D|Z,X]=0}$$

$$Z = f_Z(X) + \underbrace{U_Z}_{E[U_Z|X]=0}$$

3.2 識別の仮定

- ・ U_Z が、 U_Y, U_D から独立
 - ▶ 除外制約 + ランダムな操作変数
- ・ $\tau_Z \neq 0$
 - ▶ relevance

3.3 DAG による表現



3.4 Two-stage least square

- ・以下を推定する
- ・ $Y = \beta_Y \times D + f_Y(X) + \underbrace{u_Y}_{E[u_Y|D,X]}$
- ・ $D = \beta_D \times Z + f_D(X) + \underbrace{u_D}_{E[u_D|Z,X]}$
- ・ $f_Y, f_D = X$ についての複雑な関数

3.5 残差回帰

- 以下のように書き換えられる

- $$Y - E[Y | X]$$
$$= \beta_Y \times (D - E[D | X]) + u_Y$$
- $$D - E[D | X]$$
$$= \beta_D \times (Z - E[Z | X]) + u_D$$

3.6 R での実装

```
E[Y|X]: sample fold 1/10
E[Y|X]: sample fold 2/10
E[Y|X]: sample fold 3/10
E[Y|X]: sample fold 4/10
E[Y|X]: sample fold 5/10
E[Y|X]: sample fold 6/10
E[Y|X]: sample fold 7/10
E[Y|X]: sample fold 8/10
E[Y|X]: sample fold 9/10
E[Y|X]: sample fold 10/10 -- Done!

E[Z1|X]: sample fold 1/10
E[Z1|X]: sample fold 2/10
E[Z1|X]: sample fold 3/10
E[Z1|X]: sample fold 4/10
E[Z1|X]: sample fold 5/10
E[Z1|X]: sample fold 6/10
E[Z1|X]: sample fold 7/10
E[Z1|X]: sample fold 8/10
E[Z1|X]: sample fold 9/10
E[Z1|X]: sample fold 10/10 -- Done!

E[D1|X]: sample fold 1/10
E[D1|X]: sample fold 2/10
E[D1|X]: sample fold 3/10
E[D1|X]: sample fold 4/10
E[D1|X]: sample fold 5/10
E[D1|X]: sample fold 6/10
E[D1|X]: sample fold 7/10
E[D1|X]: sample fold 8/10
E[D1|X]: sample fold 9/10
E[D1|X]: sample fold 10/10 -- Done!
```

```

PLIV estimation results:

, , nnls

            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.00247    0.1946  0.0127 9.90e-01
D_r          0.10109    0.0182  5.5624 2.66e-08

```

3.7 Reference

Bibliography

Griffen, A.S., Nakamuro, M. and Inui, T. (2015) “Fertility and maternal labor supply in Japan: Conflicting policy goals?,” Journal of the Japanese and International Economies, 38, pp. 52–72.

Wang, L., Richardson, T. and Robins, J. (2025) “Causal Inference: A Tale of Three Frameworks,” arXiv preprint arXiv:2511.21516 [Preprint].