計量経済学応用

9. 操作変数法

矢内　勇生

2019年5月20日

高知工科大学　経済・マネジメント学群

今日の目標

- 操作変数法の使い方を身につける
    - 操作変数法とは？
    - 操作変数法の具体的な実行方法は？

説明変数の内生性

- 回帰分析では、因果効果の推定にバイアスが生じることがある
  - 交絡変数を統制し損ねる（欠落変数バイアス）
  - 結果変数が説明変数に影響を与える（逆の因果関係）
  - 測定誤差
  - 交絡変数が測定・観察不能
- こういう状況下の説明変数を内生変数 (endogenous variable) と呼ぶ

例：教育が所得に及ぼす影響

- 結果変数：所得　Y_i
- 説明変数：教育（修学年数）D_i
- 交絡：能力　A_i
- 母回帰モデル（母集団における本当の関係）：

$$Y_i = \alpha + \beta D_i + \gamma A_i + \epsilon_i$$

- β を推定したい！
- **問題**：　A_i が観察できない
- 妥協案：交絡変数を除いた回帰分析 (**s**hort regression)

$$Y_i = \alpha_s + \beta_s D_i + \eta_i$$

Short Regression のバイアス

- 最小二乗推定量 $\hat{\beta}_s$ の期待値：

$$
\begin{aligned}
\mathrm{E}[\hat{\beta}_s] &= \frac{\mathrm{Cov}(Y, D)}{\mathrm{Var}(D)} \\
&= \frac{\mathrm{Cov}(\alpha + \beta D + \gamma A + \epsilon, D)}{\mathrm{Var}(D)} \\
&= \frac{\mathrm{Cov}(\beta D, D) + \mathrm{Cov}(\alpha + \gamma A + \epsilon, D)}{\mathrm{Var}(D)} \\
&= \beta + \gamma \frac{\mathrm{Cov}(A, D)}{\mathrm{Var}(D)}
\end{aligned}
$$

- $\gamma \frac{\mathrm{Cov}(A, D)}{\mathrm{Var}(D)}$ がバイアス
- $\gamma \neq 0$ かつ $\mathrm{Cov}(A, D) \neq 0$ であれば、バイアスが生じる
- $\frac{\mathrm{Cov}(A, D)}{\mathrm{Var}(D)}$ は、A を D に回帰したときの傾きパラメタ

操作変数法 (instrumental variable [IV] method)

- バイアスなしにβを推定するために、IV法を使う
- 操作変数 (instrument)：Z_i
- 操作変数は、以下の条件を満たす必要がある
  - 原因となる説明変数（内生変数）と相関がある
  - 結果変数には、**説明変数（内生変数）を通してのみ**影響を与える（除外制約 [exclusion restriction]）
- 操作変数法の公式（内生変数も操作変数も1つずつのとき）：

$$\beta_{\mathrm{IV}} = \frac{\mathrm{Cov}(Y, Z)}{\mathrm{Cov}(D, Z)}$$

操作変数を使った推定量

- 真のモデルを操作変数の公式に代入してみると

$$
\begin{aligned}
\beta_{\mathrm{IV}} &= \frac{\mathrm{Cov}(Y,Z)}{\mathrm{Cov}(D,Z)} = \frac{\mathrm{Cov}(\alpha+\beta D+\gamma A+\epsilon, Z)}{\mathrm{Cov}(D,Z)} \\
&= \frac{1}{\mathrm{Cov}(D,Z)}\left[\beta \mathrm{Cov}(D,Z)+\gamma \mathrm{Cov}(A,Z)+\mathrm{Cov}(\epsilon,Z)\right]
\end{aligned}
$$

- 除外制約があるので、$\mathrm{Cov}(A,Z)=0$ かつ $\mathrm{Cov}(\epsilon,Z)=0$
- 説明変数と操作変数には相関があるので $\mathrm{Cov}(D,Z)\neq 0$
- したがって、 $\beta_{\mathrm{IV}}=\beta$

操作変数法の推定量は2つの回帰係数の比

- 操作変数法で得られる推定量は、2つの回帰係数の比である

$$\beta_{\mathrm{IV}} = \frac{\mathrm{Cov}(Y, Z)}{\mathrm{Cov}(D, Z)} = \frac{\mathrm{Cov}(Y, Z)/\mathrm{Var}(\mathrm{Z})}{\mathrm{Cov}(D, Z)/\mathrm{Var}(Z)} = \frac{\rho}{\phi}$$

- ρ と ϕ は以下の回帰モデルで推定できる
  - 第1段階の回帰： $D_i = \alpha_1 + \phi Z_i + e_{1i}$
  - 誘導型回帰： $Y_i = \alpha_0 + \rho Z_i + e_{0i}$

2段階回帰分析法

- 2段階回帰 (two-stage least squares, 2SLS, TSLS)
- 第1段階の回帰

$$D_i = \alpha_1 + \ \phi Z_i + e_{1i}$$

- 第2段階の回帰

$$Y_i = \alpha_2 + \beta_{2\mathrm{SLS}} \hat{D}_i + e_{2i}$$

  - ただし、$\hat{D}_i$ は第1段階で得られた D_i の予測値

$$\hat{D}_i = \hat{\alpha}_1 + \hat{\phi} Z_i$$

2段階回帰の推定量と操作変数法の推定量

- 操作変数の数と内生変数の数が一致するとき、2段階回帰の推定量と操作変数法の推定量は一致する

$$
\begin{aligned}
\beta_{\mathrm{2SLS}} &= \frac{\mathrm{Cov}(Y, \hat{D})}{\mathrm{Var}(\hat{D})} \\
&= \frac{\mathrm{Cov}(Y, \alpha_1 + \phi Z)}{\mathrm{Var}(\alpha_1 + \phi Z)} \\
&= \frac{\phi \mathrm{Cov}(Y, Z)}{\phi^2 \mathrm{Var}(Z)} \\
&= \frac{\rho}{\phi} \\
&= \beta_{\mathrm{IV}}
\end{aligned}
$$

共変量を伴う操作変数法

- 2段階回帰を行うなら、両方の回帰に共変量（統制変数）を含める（共変量を X とする）
  - 第1段階： $D_i = \alpha_1 + \phi Z_i + \lambda_1 X_i + e_{1i}$
  - 第2段階： $Y_i = \alpha_2 + \beta_{2\mathrm{SLS}} \hat{D}_i + \lambda_2 X_i + e_{2i}$

Rで推定する場合

- 2つの回帰を別々には実行しない
  - ただし、第1段階だけは別の目的で実行する
- 2段階回帰を別々に実行すると、
  - 正しい（バイアスのない）回帰係数の推定値は得られる
  - **標準誤差が正しくない！**
- 操作変数法用に用意されたコマンドを使う

AER::ivreg()

- AER パッケージの `ivreg()` 関数を使う
- 使い方

```
fit <- ivreg(Y ~ D + X | X + Z, data = …)
```