第 11 回　操作変数法（8）

村澤 康友

2022 年 12 月 22 日

今日のポイント

1. 説明変数と誤差項が無相関であることが OLS 推定量の一致性の必要十分条件．説明変数の欠落や内生性は OLS 推定量に偏りをもたらす．
2. 説明変数と相関があり，誤差項と無相関の変数を操作変数（IV）という．線形モデルの一致推定には係数の数だけ IV が必要．IV を用いる推定手法を操作変数（IV）法という．
3. 各説明変数を全ての IV に回帰して回帰予測を求め，それに被説明変数を回帰する手法を 2 段階最小 2 乗法（2SLS）という．IV 法は 2SLS で実行する．

目次

1　OLS 推定量の偏り

1.1　OLS 推定量の一致性（p. 191）

$((y_1, x_1), \dots, (y_n, x_n))$ を無作為標本とする．簡単化のため定数項のないモデルで考える．y_i の x_i 上への定数項のない線形モデルは

$$\begin{aligned} y_i &= \beta x_i + u_i \\ \mathrm{E}(u_i) &= 0 \end{aligned}$$

β の OLS 推定量を b_n とすると

$$b_n = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

$\mathrm{E}\left(x_i^2\right) > 0$（$x_i$ は 0 以外の値を取り得る）とする．

定理 1.

$$\mathrm{E}(x_i u_i) = 0 \Longleftrightarrow \operatorname*{plim}_{n\to\infty} b_n = \beta$$

証明．b_n の式に $y_i = \beta x_i + u_i$ を代入すると

$$\begin{aligned} b_n &= \frac{\sum_{i=1}^n x_i(\beta x_i + u_i)}{\sum_{i=1}^n x_i^2} \\ &= \beta + \frac{\sum_{i=1}^n x_i u_i}{\sum_{i=1}^n x_i^2} \\ &= \beta + \frac{(1/n)\sum_{i=1}^n x_i u_i}{(1/n)\sum_{i=1}^n x_i^2} \end{aligned}$$

大数の法則より

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^n x_i^2 = \mathrm{E}\left(x_i^2\right)$$

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^n x_i u_i = \mathrm{E}(x_i u_i)$$

スルツキーの定理より

$$\operatorname*{plim}_{n\to\infty} b_n = \beta + \frac{\mathrm{E}(x_i u_i)}{\mathrm{E}\left(x_i^2\right)}$$

したがって一致性の必要十分条件は $\mathrm{E}(x_i u_i) = 0$. □

注 1. $\mathrm{E}(u_i) = 0$ なので $\mathrm{E}(x_i u_i) = \mathrm{cov}(x_i, u_i)$.

注 2. 回帰モデルなら繰り返し期待値の法則より $\mathrm{E}(u_i|x_i) = 0 \Longrightarrow \mathrm{E}(x_i u_i) = 0$. すなわち一致性の必要十分条件を満たす.

1.2 欠落変数バイアス（p. 139）

(Y, X, Z) を確率ベクトルとする．Y の (X, Z) 上への重回帰モデルは

$$\begin{aligned} Y &= \alpha + \beta X + \gamma Z + U \\ \mathrm{E}(U|X, Z) &= 0 \end{aligned}$$

説明変数から Z が欠落すると

$$Y = \alpha + \beta X + V$$

ただし $V := \gamma Z + U$.

定理 2.

$$\mathrm{E}(XV) = \gamma\,\mathrm{E}(XZ)$$

証明.

$$\begin{aligned} \mathrm{E}(XV) &= \mathrm{E}(X(\gamma Z + U)) \\ &= \gamma\,\mathrm{E}(XZ) + \mathrm{E}(XU) \end{aligned}$$

繰り返し期待値の法則より第 2 項は

$$\begin{aligned} \mathrm{E}(XU) &= \mathrm{E}(\mathrm{E}(XU|X, Z)) \\ &= \mathrm{E}(X\,\mathrm{E}(U|X, Z)) \\ &= 0 \end{aligned}$$

□

注 3. したがって $\gamma = 0$ または $\mathrm{E}(XZ) = 0$ でない限り OLS 推定量は偏りをもつ.

定義 1. 説明変数の欠落によって生じる OLS 推定量の偏りを**欠落変数バイアス**という.

1.3 内生性バイアス（p. 191）

確率ベクトル (Y_1, Y_2, X) は次の連立方程式を満たす.

$$\begin{aligned} Y_1 &= -\gamma_{1,2} Y_2 + \beta X + U_1 \\ Y_2 &= -\gamma_{2,1} Y_1 + \beta X + U_2 \\ \mathrm{E}\left(\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} | X\right) &= \mathbf{0} \\ \mathrm{var}\left(\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} | X\right) &= \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2} \end{bmatrix} \end{aligned}$$

第 1 式の OLS 推定を考える.

定理 3.

$$\mathrm{E}(Y_2 U_1) = \frac{-\gamma_{2,1}\sigma_{1,1} + \sigma_{1,2}}{1 - \gamma_{2,1}\gamma_{1,2}}$$

証明. 繰り返し期待値の法則より $\mathrm{E}(U_1|X) = 0 \Longrightarrow \mathrm{E}(XU_1) = 0$ なので

$$\begin{aligned} \mathrm{E}(Y_2 U_1) &= \mathrm{E}((-\gamma_{2,1} Y_1 + \beta X + U_2) U_1) \\ &= -\gamma_{2,1}\,\mathrm{E}(Y_1 U_1) + \sigma_{1,2} \\ &= -\gamma_{2,1}\,\mathrm{E}((-\gamma_{1,2} Y_2 + \beta X + U_1) U_1) + \sigma_{1,2} \\ &= \gamma_{2,1}\gamma_{1,2}\,\mathrm{E}(Y_2 U_1) - \gamma_{2,1}\sigma_{1,1} + \sigma_{1,2} \end{aligned}$$

これを $\mathrm{E}(Y_2 U_1)$ について解けばよい. □

注 4. したがって $\gamma_{2,1} = 0$ かつ $\sigma_{1,2} = 0$ でない限り OLS 推定量は偏りをもつ.

定義 2. システム（連立方程式）の外部で決定される変数を**外生変数**という.

定義 3. 外生変数を所与としてシステムの内部で決定される変数を**内生変数**という.

定義 4. 説明変数に内生変数があることで生じる OLS 推定量の偏りを**内生性バイアス**という.

2 操作変数（IV）法

2.1 操作変数（IV）（p. 192）

(Y, X, Z) を確率ベクトルとする．簡単化のため定数項のないモデルで考える．Y の X 上への定数項のない線形モデルは

$$\begin{aligned} Y &= \beta X + U \\ \mathrm{E}(U) &= 0 \end{aligned}$$

定義 5. 線形モデルの説明変数と相関があり，誤差項と相関がない変数を**操作変数**（*Instrumental Variable, IV*）という．

注 5. $\mathrm{E}(ZX) \neq 0$ で $\mathrm{E}(ZU) = 0$ なら Z は β の推定の IV.

注 6. 回帰モデルなら繰り返し期待値の法則より $\mathrm{E}(U|X) = 0 \implies \mathrm{E}(XU) = 0$. また X が 0 以外の値を取り得るなら $\mathrm{E}(XX) = \mathrm{E}\left(X^2\right) \neq 0$. したがって X が IV となる.

定義 6. 操作変数を用いる推定手法を**操作変数**（*IV*）**法**という.

定理 4.

$$\beta = \frac{\mathrm{E}(ZY)}{\mathrm{E}(ZX)}$$

証明. $U := Y - \beta X$ より

$$\begin{aligned}\mathrm{E}(ZU) &= \mathrm{E}(Z(Y - \beta X)) \\ &= \mathrm{E}(ZY) - \beta\,\mathrm{E}(ZX)\end{aligned}$$

左辺$=0$ より結果が得られる. □

注 7. この式に MM 法を適用して β を推定する.

2.2 識別（p. 195）

定義 7. 母数の一致推定量が存在するなら母数は**識別可能**という.

注 8. 線形モデルの係数の識別には推定する係数の数だけ IV が必要.

2.3 IV 推定量（p. 194）

$((y_1, x_1, z_1), \ldots, (y_n, x_n, z_n))$ を無作為標本とする．簡単化のため定数項のないモデルで考える．y_i の x_i 上への定数項のない線形モデルは

$$\begin{aligned} y_i &= \beta x_i + u_i \\ \mathrm{E}(u_i) &= 0 \end{aligned}$$

定義 8. β の *IV* **推定量**は

$$b_n := \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i}$$

注 9. IV を用いた MM 法と解釈できる.

注 10. $z_i = x_i$ なら IV 推定量$=$ OLS 推定量.

定理 5.

$$\operatorname*{plim}_{n\to\infty} b_n = \beta$$

証明. 復習テスト. □

定理 6.

$$\sqrt{n}(b_n - \beta) \xrightarrow{d} \mathrm{N}\left(0, \frac{\mathrm{var}(z_i u_i)}{\mathrm{E}(z_i x_i)^2}\right)$$

証明. b_n の式に $y_i = \beta x_i + u_i$ を代入すると

$$\begin{aligned} b_n &= \frac{\sum_{i=1}^n z_i(\beta x_i + u_i)}{\sum_{i=1}^n z_i x_i} \\ &= \beta + \frac{\sum_{i=1}^n z_i u_i}{\sum_{i=1}^n z_i x_i} \end{aligned}$$

式変形すると

$$\sqrt{n}(b_n - \beta) = \frac{(1/\sqrt{n})\sum_{i=1}^n z_i u_i}{(1/n)\sum_{i=1}^n z_i x_i}$$

大数の法則より

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^n z_i x_i = \mathrm{E}(z_i x_i)$$

$\mathrm{E}(z_i u_i) = 0$ なので中心極限定理より

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n z_i u_i \xrightarrow{d} \mathrm{N}(0, \mathrm{var}(z_i u_i))$$

スルツキーの定理とクラーメルの定理より

$$\frac{(1/\sqrt{n})\sum_{i=1}^n z_i u_i}{(1/n)\sum_{i=1}^n z_i x_i} \xrightarrow{d} \mathrm{N}\left(0, \frac{\mathrm{var}(z_i u_i)}{\mathrm{E}(z_i x_i)^2}\right)$$

□

注 11. $\mathrm{var}(z_i u_i)$ は White の推定量で推定する.

系 1. $\mathrm{var}(u_i|z_i) = \sigma^2$ なら

$$\sqrt{n}(b_n - \beta) \xrightarrow{d} \mathrm{N}\left(0, \frac{\sigma^2}{\mathrm{E}(z_i x_i)^2/\,\mathrm{E}(z_i^2)}\right)$$

証明. $\mathrm{E}(z_i u_i) = 0$ より

$$\begin{aligned}\mathrm{var}(z_i u_i) &= \mathrm{E}\left((z_i u_i)^2\right) \\ &= \mathrm{E}\left(z_i^2 u_i^2\right)\end{aligned}$$

繰り返し期待値の法則より

$$\begin{aligned}\mathrm{E}\left(z_i^2 u_i^2\right) &= \mathrm{E}\left(\mathrm{E}\left(z_i^2 u_i^2 | z_i\right)\right) \\ &= \mathrm{E}\left(z_i^2 \mathrm{E}\left(u_i^2 | z_i\right)\right) \\ &= \mathrm{E}\left(z_i^2 \operatorname{var}(u_i | z_i)\right) \\ &= \sigma^2 \mathrm{E}\left(z_i^2\right)\end{aligned}$$

これを前定理の結果に代入すればよい． □

2.4 弱い IV（p. 202）

定義 9. $\mathrm{E}(ZX) \approx 0$ で $\mathrm{E}(ZU) = 0$ なら Z は**弱い** ***IV*** という．

注 12. $\beta = \mathrm{E}(ZY)/\mathrm{E}(ZX)$ より $\mathrm{E}(ZX) \approx 0$ だと推定値が不安定になる．また前定理より IV 推定量の漸近分散は $\operatorname{var}(ZU)/\mathrm{E}(ZX)^2$．したがって $\mathrm{E}(ZX) \approx 0$ だと推定の精度が低い．

3 2 段階最小 2 乗法（2SLS）

3.1 2 段階最小 2 乗法（2SLS）（p. 205）

(Y, X, Z) を確率ベクトルとする．Y の X 上への線形モデルは

$$\begin{aligned} Y &= \alpha + \beta X + U \\ \mathrm{E}(U) &= 0 \end{aligned}$$

Z を β の推定の IV とする．X を Z に回帰した回帰予測を $\hat{X}$ とすると，$\hat{X}$ は Z の線形変換なので $\mathrm{E}(ZU) = 0 \Longrightarrow \mathrm{E}\left(\hat{X}U\right) = 0$．すなわち $\hat{X}$ も IV．

定義 10. 各説明変数を全ての IV に回帰して回帰予測を求め，それに被説明変数を回帰する手法を ***2* 段階最小 *2* 乗法**（*2-Stage Least Squares, 2SLS*）という．

注 13. IV 法は 2SLS で実行する．「IV の数＞係数の数」でも 2SLS なら全ての IV を使える．

注 14. 本当に 2 段階で実行すると正しい標準誤差が得られない．実際は（一般化した）MM 法で実行する．

3.2 構造形と誘導形（pp. 196, 206）

定義 11. 変数間の理論的な関係を表した連立方程式を**構造形**という．

注 15. 説明変数に内生変数がある式は 2SLS で推定する．

定義 12. 内生変数について構造形を解いた式を**誘導形**という．

注 16. 誘導形の説明変数は外生変数（＝ IV）のみなので，2SLS の第 1 段階で使う．

4 今日のキーワード

欠落変数バイアス，外生変数，内生変数，内生性バイアス，操作変数（IV），操作変数（IV）法，識別可能，IV 推定量，弱い IV，2 段階最小 2 乗法（2SLS），構造形, 誘導形

5 次回までの準備

復習 教科書第 8 章，復習テスト 11
予習 教科書第 9 章