R による実証分析 第 11, 12 章

前川 大空 *

2025 年 3 月 14 日

第 11 章 操作変数法

■p.240 (11.3) **式の定式化**　一本目の式では, $\gamma_0 \neq 0$ で関連性の条件付けが厳密には必要なはず. 気持ち的な定義でも, 代わりに説明したい変数の『決定要因』であるため.

■p.241 Z **は内生変数** X **の" 決定要因"**　関連性って因果関係ほどの強い意味を持っているの？『因果関係を媒介する』程度でただの**相関**では？例えば, 弾力性推定の文脈では IV は『供給ショックに繋がる要因で, 価格変化を通じてのみ (需要ショックではなく！), 需要量に影響を与える』. この例では『内生変数である価格の決定要因』であるから, 因果関係といえそうだ. ただ, 問題は果たしてこれが一般的にも因果関係を意味するのか.

■p.241, **注 1 線形仮定は不要**　当然. だが多くの講義資料は入門のため, 教育的配慮からか線形性を仮定して説明をしてくれる. この教育的配慮は他の部分でも見て取れて, たしかこの教科書の欠落変数バイアスに関する議論の部分でも本来不要な線形性の仮定を置いて説明を簡単化していた.

■p.243 **出席率の操作変数選択**　**相関するだけの要因は IV として不適切か？**例えば, 定期価格 (通学時間が決定要素の一つ, だが出席率にはあまり因果効果がなさそう.) を IV として用いることは IV の定義からは出来ないのだろうか？通学時間↑ → 定期価格↑, 出席率↓で定期=出席に負の相関関係 (×因果) があるが, 定期価格は学業成績には影響を与えなさそう. 外生性と関連性の数学的な定義は満たせそうに思えるが . . ..

■p.244 (11.6) **式, 二本目**　$X_{2,i}\gamma_2$ は**外生変数も入れて**二段階目回帰するのは, 測定誤差が $X_{2,i}$ と相関することを防ぐため, とあるが. これは何故相関するの？普通に欠落変数バイアス？

■p.247 (11.7) **式, の実行**　γ が未知であるため不可能 ← 一段階目回帰の推定値を使えば分かるのでは

■p.249 11.3 **節とそれ以前の議論の違い**　内生性が無くなって因果効果の測定が可能, とは言葉ではいうものの, 実際にそれが可能なことを確認する為には明示的に Rubin の因果モデルを通じて確認する必要がある.

■p.249 11.3 **節とそれ以前の議論の違い**　もしかしてマッチング法が処置群に対しての ATET のみを考えていたのと関連しているのでは？条件付き独立を仮定して Rubin の因果モデルの枠組みで計算を進めると β で推定出来るのは ATET になる, みたいな (**未確認**). 処置群に属しているか, 操作変数の状態がどうかに関わらず, **人のタイプ**で条件付けていることに注意しよう.

* 一橋大学経済学部 4 年, 五年一貫専修コース公共経済プログラム

■p.253 Quiz1: **データの解釈と関数形**　分析や結果において, 回帰分析の係数が「1% 変化した時に〜」といった形で解釈されている. 線形単回帰にある変形を加えた為に, 解釈の仕方が通常のそれ（1 単位変化した時に〜）とは違う記述になっているのだ. さて, その**変形とは具体的になんだろう.** また, **第一段階と第二段階でそれぞれ用いられている functional form はなんだろうか.**

■p.254 Quiz2:(11.8) **式の導出過程**　2〜3 行目の式変形では, 離散確率変数の場合の繰り返し期待値の法則, $E[\cdot] = E[E[\cdot|Z]] = \Sigma_Z E[\cdot|Z=z] \cdot P(Z=z)$ を, 今回 2 値変数（ダミー変数）なので書き出している. さて, **何故 2〜3 行目の式変形は成立するのだろう？**具体的には, なぜ $2 \times 2 = 4$ 項あるはずなのに, 変形後は 3 項しか書き出されていないのだろうか？

■p.259 **結果の解釈, 一般化**　(2) は D の影響を受ける, Rubin の因果モデルでいうところの Complier 的立ち位置の人々! 今回に限らず Complier に注目するのが IV method では重要そうだ.

第 12 章 パネルデータ回帰分析

■p.267 **個人効果, 時点効果の識別の問題 (本題！)**　この問題の本質はダミー変数の罠, つまり過剰なダミー変数による完全な**多重共線性**にある. これを解決するために, 他のダミー変数で説明できる（＝線形従属である）ダミーをモデルから落としたり（若しくは標準化したり）する必要があることは知られているが .... **Q.** ここで落としている変数の数は 1 つ. 文中では $\delta_{t,1}$ を除外している. **何故個人効果ダミーは落とさない？ A. 個人効果に定数項が纏められているため.** ちゃんと考えないと意外と分からない話. 理解には連立一次方程式システムの理解が必要となるため, **Appendix** にてこの問題は詳述する.

■p.268 **タイポ**　$\hat{\phi}_i = \sum_{i=1}^n I_{i,j}\hat{\phi}_j$ ではなく $\hat{\phi}_i = \sum_{\mathbf{j}=1}^n I_{i,j}\hat{\phi}_j$ だろう. i じゃなくて j. 次の計算でも用いる.

■p.268 **OLSE と within 推定量の一致**　少し詰まったので行間を埋めておく.

$$
\begin{aligned}
&MSE = \frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(Y_{i,t} - X_{i,t}\beta_1 - \sum_{j=1}^n I_{i,j}\phi_j)^2 \\
&\text{FOC regarding } \phi_i : \frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T 2(Y_{i,t} - X_{i,t}\beta_1 - \sum_{l=1}^n I_{i,l}\phi_l)(-I_{i,j}) = 0 \\
&\frac{1}{nT}\sum_{t=1}^T 2(Y_{j,t} - X_{j,t}\beta_1 - \sum_{l=1}^n I_{j,l}\phi_l)(-1) = 0 \ (\because I_{i,j} = 1 \text{ if } i = j,\ 0 \text{ o.w.}) \\
&\frac{1}{T}\sum_{t=1}^T(Y_{j,t} - X_{j,t}\beta_1 - \sum_{l=1}^n I_{j,l}\phi_l) = 0,\ \frac{1}{T}\sum_{t=1}^T Y_{j,t} - \frac{1}{T}\sum_{t=1}^T X_{j,t}\beta_1 - \frac{1}{T}\sum_{t=1}^T\phi_j = 0 \\
&\hat{\phi}_j = \bar{Y}_j - \hat{\beta}_1\bar{X}_j \text{ (at the optimum)}
\end{aligned}
$$

同様の計算は (12.8) 式の導出でも用いられる.

■p.270 **『個体効果パラメータは解釈が難しい』**　何故？定数項をまとめているからか, T が限られているために漸近理論を適合できないからか.

■p.270 Quiz3: **within 推定量と OLSE が一致**　ではわざわざ **within 推定量を用いるメリットとは何だろう？**

■p.270 **二方向固定効果モデルの推定法**　どうやって推定されているのか明示されていない．普通に重回帰？

■p.275 **仮定** 1　駆け込み需要．動学的最適化問題を考える, マクロ計量？では重要な仮定っぽくて面白い．

■p.276 **図 12.3 の説明**　**Q.** 図で言うとどの部分がそれぞれの仮定に対応しているのか？ **A.** **仮定 1:** $E[Y_{i,1}|D=1]$ の点が平行トレンドに合致する点にあること．つまり, $E[Y_{i,1}|D=1]=E[Y_{i,1}(0)|D=1]$. **仮定 2:** (0) の四点が平行四辺形状に位置していること．

■p.278 **『実務上は (12.7) 式の重回帰が便利』**　じゃあ DID の意味ってなんなんだ, 分かりやすさ？ただの交差項がある二方向固定効果モデルにすぎんだろう．

■p.279 **重回帰モデルのパラメータ解釈と記法**　β_1, β_2 はそれぞれ順に $0 \to 1$, $1 \to 2$ 期間でのトリートメントによる傾きへの影響．$\beta_1 = 0$ なら $0 \to 1$ 期で処置群/対照群間での差異が変わらず**平行トレンドが成立**することを意味する．あと, ページ下部の重回帰モデルについての疑問: $\delta_{t,1}$ でもいいよなこれ．$\Delta Trend$ と表記を対応 (先頭の被説明変数と t を合致) させてるんだろうが．

■p.281 **『真の $ATET_2$ はゼロ』**　当たり前ではない．元々の $ATET_2$ の定義より以下の計算により確認可能．

$$
\begin{aligned}
&ATET_2 = E[Y_{i,2}(1) - Y_{i,2}(0)|D_i = 1], Y_{i,t}(d) = \phi_i + tD_i + \epsilon_i \\
&\therefore Y_{i,t}(1)|D_i = 1 = Y_{t,2}(0)|D_i = 1 \ (\because \text{ Y is independent of d}) \\
&\text{Thus, } ATET_2 = E[0|D_i = 1] = 0
\end{aligned}
$$

Appendices

付録 A　ダミー変数の罠とその解決策

重回帰分析の仮定の一つとして,『説明変数間に完全な多重共線性がない』というものがある．これは, **パラメータの識別 (推定値が一意に定まるか) に関する問題**であり, 一般的な解決策は, ダミー変数の内一つを落とすことだろう．また, 時系列モデルでは定数項を落とす場合もみられる．ではこの解決策の正統性とは？

多重共線性 (Multicollinearity) の問題点

多重共線性とは, ある変数ベクトルが他変数ベクトルの線形結合で表せることを意味する．即ち:

$$\text{There exists } \alpha \neq \mathbf{0}, \text{such that } \mathbf{0} = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_{K-1} X_{K-1} + \alpha_K X_K$$

と線形従属な場合には, 変数ベクトルは完全な多重共線性を持つ．そもそも, 何故多重共線性が識別の問題になるかを考えてみよう．

OLSE の行列表記の公式を念頭に置いた説明

OLSE に使用される行列 $\mathbf{X^T X}$ に逆行列が存在するのは列フルランク行列のみであるため．$\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X^T X})$ で, 列ベクトルが線形独立でない⇔正則ではない⇔逆行列が存在しない．

もう少し丁寧な説明, ダミー変数を念頭に置いて

OLS 推定値は $\hat{Y}_i = \sum_{k=1}^{K} \hat{\beta}_k X_{i,k}, i \in \{1, \dots, n\}$ である. 行列表記をすると:

$$\underbrace{\mathbf{X}}_{n\times K} \cdot \underbrace{\hat{\beta}}_{K\times 1} = (\underbrace{X_1}_{n\times 1}, \underbrace{X_2}_{n\times 1}, \dots, \underbrace{X_K}_{n\times 1}) \cdot \underbrace{\hat{\beta}}_{K\times 1} = \underbrace{\hat{\mathbf{Y}}}_{n\times 1}$$

と, 連立一次方程式を行列表記した形になる. 我々の目標は $\hat{\beta}$ を推定することにあり, この推定には『連立方程式の解 $\hat{\beta}$ が**一意**』と同値の条件が $\mathbf{X}$ に課される必要がある. これ即ち:

$$\text{解が一意} \Leftrightarrow \mathbf{X}\text{ が列フルランク行列} \overset{\text{def}}{\Leftrightarrow} \operatorname{rank}(\mathbf{X}) = K$$

である. [*1] 最後の階数条件が多重共線性の制約として課されるのはこの同値関係のためである. 一般の重回帰モデルでは, X_1 は定数項ベクトルであり, すべての成分が 1 となる. 簡単のため, 残りの $K-1$ 個の変数を, $K-1$ 個の属性を持つダミー変数 $D_{i,2} \sim D_{i,K}$ として議論を進める. ダミー変数はその定義より, すべてを回帰モデルに入れ込んでしまえば, $\sum_I D_i = 1$ となる. 明示的には, 推定値は以下のようになる:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 D_{i,2} + \hat{\beta}_3 D_{i,3} + \cdots + \hat{\beta}_K D_{i,K}, \text{ where } \sum_{k=2}^{K} D_{i,k} = 1, \forall i \in \{1, \dots, n\}$$

再度行列表記に直せば:

$$\mathbf{D}\hat{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{1} & D_2 & D_3 & \dots & D_K \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_K \end{bmatrix} = \begin{bmatrix} 1 & D_{1,2} & D_{1,3} & \dots & D_{1,K} \\ 1 & D_{2,2} & D_{2,3} & \dots & D_{2,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & D_{n,2} & D_{n,3} & \dots & D_{n,K} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_K \end{bmatrix} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix} = \hat{\mathbf{Y}}$$

ダミー変数の性質に注意して, サンプルを属性順に並び変えてやると:

$$\mathbf{D}' = \begin{bmatrix} \mathbf{1} & D'_2 & D'_3 & D'_4 & \dots & D'_K \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}}_{\text{行ごとに, 2〜K 列のうち一度だけ 1 をとる}}$$

ここで, 明らかに

$$\mathbf{1} = D'_2 + D'_3 + \cdots + D'_K \Leftrightarrow \mathbf{1} - D'_2 - D'_3 - \cdots - D'_K = \mathbf{0} \Rightarrow \mathbf{1}, D'_2, D'_3, \dots, D'_K \text{は線形従属}$$

[*1] 分からない場合は線形代数 2 を復習してもよい. 不定方程式とは**解が一意に定まらず, 無限個存在する**ようなものだった. この場合は**行列が列フルランクではない**ことをおさえておきさえすれば十分だ.

解の自由度は明らかに 1 で, 列の並び替えでは rank が保存されるため, [*2]

$$\mathrm{rank}(\mathbf{D}') = \mathrm{rank}(\mathbf{D}) = K - 1 < K = \#\mathbf{D} \text{ の列}$$

$\mathbf{D}$ が列フルランクでない $\Rightarrow$ $\mathbf{D}\hat{\boldsymbol{\beta}} = \hat{\mathbf{Y}}$ の解 $\hat{\boldsymbol{\beta}}$ が無数に存在するため, 最適な $\hat{\boldsymbol{\beta}}$ は見つからず, **識別不可能**.

解決策

$\mathrm{rank}(\mathbf{D}') = \mathrm{rank}(\mathbf{D}) = K - 1$ から分かるように, **落とすべき変数は一つだけ**である．ではどの変数を落とせば多重共線性は解消されるのだろうか.

ダミー変数の内一つを落とす

基準となるカテゴリを落としてやればよい．D_2 を落として議論を進めても一般性を失わない.

$$\mathbf{D}'' = \begin{bmatrix} \mathbf{1} & D'_3 & D'_4 & \dots & D'_K \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \xrightarrow{\text{基本変形}} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

明らかに $D'_3, D'_4, \dots, D'_K$ では $\mathbf{1}$ を表現できず, より正確には基本変形により得た Reduced row echelon form が列フルランク行列であることから, 多重共線性は解消されている．また, 残ったダミーのパラメータは, **基準となったカテゴリ 2**($D_{i,2} = 1$ の個人 i) **との属性効果の差異**として解釈が可能である.

定数項を落とす

同様の議論によって定数項を落とすことでも多重共線性は解消される．しかし, パラメータの解釈が困難となることに注意を払う必要がある．例えば, カテゴリ 2 の個人について, $\hat{y}_i = \hat{\beta}_2 D_{i,2} = \hat{\beta}_2$ である．すなわち, OLS 推定量 $\hat{\beta}_k$ は『カテゴリ k での被説明変数の推定値』を示す．以上から, **属性による因果効果を主眼とする場合, ダミー変数の内基準となるものを落とす方が適切**だと結論づけられる.

パネルデータ回帰での対応策

同様の議論はパネルデータモデルにも行えて, 教科書の例ならば, **個人効果は後者の方法で, 時点効果は前者の方法で多重共線性を排除している**ことが分かる．[*3] 以上が先述の解答に与えられた答えの論拠である.

[*2] 連立方程式を書き直したものだと理解していれば, 方程式の順番を並び替えても何らそのシステムに変化がないことは明白だろう.

[*3] 時点効果の対応策としてもう一つ挙げられているのは, 係数を一つ標準化してやることである．これは不定方程式の構造を考えてやれば前者の方法と同じであることが分かる．自由度 1 の不定方程式の解は $\hat{\beta}_k = f_k(c), c \in \mathbb{R}$ で表せた．標準化は $\hat{\beta}_k = 0$ となるように c を固定, つまり無数の解から一つを選び取る形で行われる．ダミーを落とすのは方程式の解を一意に定めるためで, また係数を常に 0 に固定するのとおなじ意味であることからも分かるだろう.

付録 B　図表の作成練習

各種ネット資料を参考にして, 13 章の補足で作成されているような図表を LaTeX 環境で再現する.

■練習 1: 基本的な環境の使用例　下の表では, ページ内での図表の位置を指定する table 環境, table 環境でのタイトル付けのための caption, 真ん中に据えるための center 環境, そして図表作成のための tabular 環境を利用している. [1] を参考とした. [*4]

表 1　記述統計表の再現 その 1

Unique	Missing Pct.	Mean	SD	Min	Median	Max	Histogram
wage	3631	0	1.65	0.53	-3.58	1.67	4.05
married	2	0	0.44	0.50	0.00	0.00	1.00
union	2	0	0.24	0.43	0.00	0.00	1.00
health	2	0	0.02	0.13	0.00	0.00	1.00
school	13	0	11.77	1.75	3.00	12.00	16.00
exper	19	0	6.51	2.83	0.00	6.00	18.00
black	2	0	0.12	0.32	0.00	0.00	1.00

■練習 2: 脚注と表の分割　threeparttable 環境を利用して, 脚注を表に記載することが出来, 図を分割している. ここでは各 rule のコマンドによって表に黒線を引いている. threeparttable 環境の中では, tabular 環境で表を書いた後に, tablenotes 環境で脚注を記載している. [2] には他の環境でも脚注をつける方法や, それが上手くいく方法などが詳しく記載されている. これで, 第 13 章の表 13.3 を再現することが出来た.

参考文献

[1] 高谷 遼 (2025), 【LaTeX】表の作成, Takatani Note, `https://takataninote.com/tex/table.html`, 閲覧日 2025/03/13

[2] Yarakashi Kikohshi(2024), 表の直下に脚注を付けたい, `https://qiita.com/Yarakashi_Kikohshi/items/31593aa58798470292c7`, 閲覧日 2025/03/13

[*4] HP 等なので正式なものではないが, 一応情報整理のために参考文献の項に置いておく.

表 2　回帰結果の再現 その 2

	(1)	(2)	(3)
切片	0.034	0.040	
	(0.062)	(0.084)	
married	0.112***	0.084***	0.051*
	(0.016)	(0.015)	(0.021)
union	0.184***	0.182***	0.083***
	(0.017)	(0.017)	(0.022)
health	-0.052	-0.033	-0.009
	(0.057)	(0.054)	(0.048)
school	0.103***	0.084***	
	(0.004)	(0.005)	
exper	0.050***	0.043***	
	(0.003)	(0.003)	
black	-0.145***	-0.127***	
	(0.023)	(0.022)	
Industry	No	Yes	Yes
Occupation	No	Yes	Yes
Individual FE	No	No	Yes
Year FE	No	No	Yes
Num.Obs.	4360	4360	4360

$p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
備考: 対数化賃金を被説明変数とした線形回帰モデルの OLS 推定値. 丸括弧内は標準誤差.