動学的パネルデータモデル

奥井亮
京都大学経済研究所

大阪大学集中講義「データ科学特論 II」
平成 28 年 9 月 1 日

目次

概要

パネルデータを用いて、経済変数の動学構造を分析する手法を解説する。

動学パネルデータ分析で重要となる点

- 観測個体間の異質性に考慮する必要がある。
- 短い期間のデータしか得られない。

1. パネルデータ分析の一般的な解説
2. (定常) 動学パネルデータモデル
3. 非定常パネルデータモデルの推定

目次

パネルデータ

パネルデータとは、複数の観測個体を複数の時点に渡って観測したデータである。

$$y_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T.$$

- i: 観測個体 (個人、企業、国など)
- t: 時点 (年、４半期など)

例：y_{it} は個人 i の時点 t での所得。

パネルデータを用いることで、観測できない異質性を考慮した分析ができる。

データの例

対数所得のパネルデータの例

個人 \ 時点	$t = 1$	t=2	
$i = 1$	0.0327	-0.240	
$i = 2$	0.280	-0.822	
$i = 3$	0.197	0.264	

データの例、その二

他のパネルデータのファイルの例

個人	時点	対数所得
$i = 1$	$t = 1$	0.0327
$i = 1$	$t = 2$	-0.240
$\vdots$	$\vdots$	$\vdots$
$i = 2$	$t = 1$	0.280
$i = 2$	$t = 2$	-0.822
$\vdots$	$\vdots$	$\vdots$

固定効果モデル

まず基本的なパネルデータのモデルである、固定効果モデルを考える。

$$y_{it} = \beta' x_{it} + \eta_i + \epsilon_{it}$$

- $\{y_{it}, x'_{it}\}$, $t = 1, \ldots, T$, $i = 1, \ldots, N$: データ。
- β: 係数。これを推定したい。
- η_i: 固定効果。観測できない。
- ϵ_{it}: 誤差項。

固定効果モデルについて

- ここでは、N が大きく、個人間は、i.i.d. を仮定する。
- 固定効果については、何も仮定しない。
- このようなモデルを考える目的は、次の通りである。欠落変数によるバイアスが、時間を通じて一定なものによって引き起こされているときに、パネルデータがあればバイアスを回避できる。固定効果は、そうした「時間を通じて一定な」「観測できない」「x_{it} と相関しているかもしれないもの」を表現している。

欠落変数バイアス

欠落変数のバイアスをいかにして回避するかは計量経済学の中心的な課題である。

- 例えば、企業の売り上げと広告の関係を調べたいとしよう。売り上げと広告費は正の相関があるであろう。しかし、このことは、広告を増やせば売り上げが増えることを必ずしも意味しない。大企業であれば、売り上げも広告費も大きいことが予想されるため、売り上げと広告費の相関は、企業の大きさからくる見せかけの相関の可能性がある。
- この場合は、企業の大きさというものを考慮しない分析のため、バイアスが発生すると考える事ができる。
- こうしたバイアスを欠落変数のバイアスと呼ぶ。
- 統計学では、交絡因子という用語の方がよく使われているようである。

仮定

次を仮定する。

- $E(\epsilon_{it} | x_{i1}, \ldots, x_{iT}, \eta_i) = 0$
- $(x_{i1}, \ldots, x_{iT}, \epsilon_{i1}, \ldots, \epsilon_{iT})$, $i = 1, \ldots, N$ は個人間について、i.i.d. である。
- (x'_{it}, ϵ_{it}) は少なくとも 4 次までのモーメントを持つ。
- 多重共線性の問題は存在しない。

推定

係数 β の推定を考える。
なお、単に y_{it} を x_{it} に回帰しただけでは、一致性のある推定量を得る保障はないことを注意すること。
基本的な推定のアイデアは、次の通りである。

1. η_i を取り除く。
2. η_i を取り除いたデータに対して、回帰をかける。

固定効果の消去

固定効果 η_i を消去するために、まず、各個人の時間を通じた平均を考える。すると、時間を通じた平均は、次の方程式を満たす。

$$\bar{y}_i = \beta' \bar{x}_i + \eta_i + \bar{\epsilon}_i.$$

この式を元の式から引くことによって、η_i を消去する。

$$y_{it} - \bar{y}_i = \beta'(x_{it} - \bar{x}_i) + \epsilon_{it} - \bar{\epsilon}_i$$

つまり、平均からの乖離を考えると、それらは、固定効果に依存しない。
次の表記を導入する。$\ddot{y}_{it} = y_{it} - \bar{y}_i$。$\ddot{x}_{it}$ と $\ddot{\epsilon}_{it}$ も同様に定義する。

$$\ddot{y}_{it} = \beta' \ddot{x}_{it} + \ddot{\epsilon}_{it}$$

固定効果推定量

先の、平均からの乖離についてのモデルは、固定効果が入っていない。そのため、平均からの乖離のモデルを回帰すると、$N \to \infty$ で T が固定されているときに一致性のある推定量を得ることが出来る。
固定効果推定量:

$$\hat{\beta} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it} \ddot{x}_{it}' \right)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it} \ddot{y}_{it}$$

この推定量には、いくつかの名前がある。

1. Fixed Effects Estimator （固定効果推定量）
2. Within Group Estimator
3. Least Squares Dummy Variables (LSDV) estimator

LSDV

固定効果推定量には、観測個体を表すダミー変数を用いた別の解釈ならびに計算方法がある。この解釈にしたがって、固定効果推定量は最小 2 乗ダミー変数 (LSDV) 推定とも呼ばれる。
まずは、観測個体ごとのダミー変数を定義しよう。

$$D1_{it} = \begin{cases} 1 & (i = 1 \text{ のとき}) \\ 0 & (i \text{ が } 1 \text{ でないとき}) \end{cases}$$

同様にして、$D2_{it}, \ldots, DN_{it}$ という N 個のダミー変数を定義する。
ダミー変数を用いると固定効果モデルは

$$y_{it} = \alpha_1 D1_{it} + \cdots + \alpha_N DN_{it} + \beta' x_{it} + \epsilon_{it}$$

と書き直すことが出来る。このモデルを OLS で推定すると、固定効果推定量を得ることができる。

時間効果

観測個体間では変化しないが、時間を通じて変化する要素を、時間効果によって表現することができる。

$$y_{it} = \lambda_t + \beta' x_{it} + \epsilon_{it}$$

ただし、λ_t が時間効果である。
このモデルは、

$$(y_{it} - \bar{y}_t) = \beta'(x_{it} - \bar{x}_t) + \epsilon_{it} - \bar{\epsilon}_t$$

と横断面での平均を引くことによって、時間効果を除去した上で推定すると良い。
また時間ダミー ($T1_{it} = \mathbf{1}\{t = 1\}$ など) をモデルに含めて、LSDV のような推定をすることもできる。

個人効果と時間効果の入ったモデル

個人効果と時間効果の両方とも入ったモデルも考える事ができる。

$$y_{it} = \eta_i + \lambda_t + \beta' x_{it} + \epsilon_{it}$$

このモデルが最も一般的に使用されているモデルであろう。
個人効果と時間効果の両方とも、$y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}$ といった変換をすることで、消去することができる。
また個体ダミーと時間ダミーの両者を入れたモデルを推定する LSDV のような推定もできる。

漸近分布その一

固定効果推定量の漸近分布を求めよう。ここでは、N が無限に行く一方で、T は固定されているような、漸近理論を考える。

$$\begin{aligned}\sqrt{N}(\hat{\beta}-\beta) &= \left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}'\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\ddot{\epsilon}_{it} \\ &= \left(\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}'\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\epsilon_{it}\end{aligned}$$

ここでは、$\sum_{t=1}^{T}\ddot{x}_{it}\ddot{\epsilon}_{it}=\sum_{t=1}^{T}\ddot{x}_{it}\epsilon_{it}$ という性質を使用した。

漸近分布その二

通常の漸近理論を使って、

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}' \to_p E\left(\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}'\right)$$

ならびに、

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\epsilon_{it} \to_d N\left(0, E\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\epsilon_{it}\epsilon_{is}\ddot{x}_{it}\ddot{x}_{is}'\right)\right)$$

を示すことが出来る。

漸近理論その三

したがって、固定効果推定量の漸近分布は次の様になる。

$$\sqrt{N}(\hat{\beta} - \beta) \to_d N\left(0, V_{FE}\right).$$

ここで、漸近分散は、

$$V_{FE} = \left(E\left(\sum_{t=1}^{T} \ddot{x}_{it}\ddot{x}_{it}'\right)\right)^{-1} E\left(\sum_{t=1}^{T}\sum_{s=1}^{T} \epsilon_{it}\epsilon_{is}\ddot{x}_{it}\ddot{x}_{is}'\right)\left(E\left(\sum_{t=1}^{T} \ddot{x}_{it}\ddot{x}_{it}'\right)\right)^{-1}$$

である。

漸近分散の推定（準備編）

まず、

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}' \to_p E\left(\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}'\right)$$

ですので、$E\left(\sum_{t=1}^{T}\ddot{x}_{it}\ddot{x}_{it}'\right)$ は問題無く推定できる。
次に、

$$E\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\epsilon_{it}\epsilon_{is}\ddot{x}_{it}\ddot{x}_{is}'\right)$$

の推定を考える。

漸近分散の推定 (cluster-robust)

まず、残差項をとる。

$$\hat{\ddot{\epsilon}}_{it} = \ddot{y}_{it} - \hat{\beta}' \ddot{x}_{it}$$

そして、誤差項の代わりに、残差項を使って、漸近分散の推定をする。

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \hat{\ddot{\epsilon}}_{it} \hat{\ddot{\epsilon}}_{is} \ddot{x}_{it} \ddot{x}_{is}'.$$

この推定量は、一致性を持つ。

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \hat{\ddot{\epsilon}}_{it} \hat{\ddot{\epsilon}}_{is} \ddot{x}_{it} \ddot{x}_{is}' \to_p E\left(\sum_{t=1}^{T} \sum_{s=1}^{T} \epsilon_{it} \epsilon_{is} \ddot{x}_{it} \ddot{x}_{is}' \right).$$

Cluster-Robust 漸近分散推定量について

- この推定量は Cluster-Robust あるいは Heteroskedasticity and Autocorrelation Consistent と呼ばれる。(Arellano (1987))
- その名の通り、この推定量は、各個人の時系列に系列相関があっても、分散不均一でも、一致性を持つ。
- 多くの計量用パッケージで、この漸近分散を基にした、標準誤差や、検定統計量を計算することができる。
- 通常は、この漸近分散推定量を使うべきである。

なぜ、Cluster-Robust を使うべきか

- 不均一分散は、経済データの分析では、常に問題となる。
- また、系列相関も、近年、重要な問題であると指摘されている。(Bertrand, Duflo and Mullainathan (2004))
- また、たとえ、系列相関が問題ではなくとも、今のところ、既存の計量パッケージには、少し問題があり、系列相関がないという活かすことが難しい。

保育所の整備が母親の就業率を上げるか

Asai, Kambayashi and Yamaguchi (2015) の研究を紹介しよう。

- この論文は、都道府県別のパネルデータを用いて、保育所定員率と母親の就業率の関係を調べたものである。
- 保育所定員率と母親の就業率には正の相関がある。

$$\widehat{\text{母親就業率}}_{it} = \underset{(0.077)}{0.686} \quad \text{保育所定員率}_{it} \quad + \text{定数項}$$

$$\bar{R}^2 = 0.491, \quad N = 47, \quad T = 5$$

欠落変数のバイアスの疑い

しかし、この相関関係は、欠落変数のバイアスが含まれている可能性がある。

- 都道府県ごとに女性の就業や保育に対する文化が異なっており、正の相関はそれらの文化の違いを反映しているだけの可能性がある。
  $\longrightarrow$ 都道府県効果を入れる必要がある。
- 日本全体での女性の就業や保育への考え方が変わってきており、そうした時代の風潮の影響を両者とも反映したため、相関がでている可能性もある。
  $\longrightarrow$ 時間効果を入れる必要がある。

結果

個人効果と時間効果を考慮すると、係数は負になり、統計的にも有意でなくなる。

$$\widehat{\text{母親就業率}}_{it} = \underset{(0.110)}{-0.147} \quad \text{保育所定員率}_{it} \quad + \text{都道府県効果}$$
$$+ \text{年効果} + \text{他の説明変数}$$
$$\bar{R}^2 = 0.979, \quad N = 47, \quad T = 5$$

- 詳しくは、Asai, Kambayashi and Yamaguchi (2015), Asai, Kambayashi and Yanamuchi (2016), 朝井、神林、山口 (2015)、山口 (2016) を参照。
- 注: Asai, Kambayashi and Yamaguchi (2015) は都道府県の人口の違いを考慮した推定量の調整を行っているので、上の結果は単なる固定効果推定値とは異なる。

目次

概要

動学パネルデータモデルについて学習する。

1. パネル AR(1) モデル。
2. GMM 推定。
3. 固定効果推定のバイアス修正。
4. 異質性と状態異存。

パネル AR(1) モデル

動学パネルデータモデルの最も単純なモデルである、パネル AR(1) モデルを考える。

$$y_{it} = \alpha y_{i,t-1} + \mu_i + \epsilon_{it}$$

- α: この係数の推定が焦点になる。状態依存の度合いを表している。
- μ_i: 個人効果。観測できない異質性を表現している。
- ϵ_{it}: 誤差項。系列相関がなく、η_i とは独立を仮定する。$E(\epsilon_{it}) = 0$ かつ $Var(\epsilon_{it}) = \sigma^2$ としよう。

説明変数のあるパネル AR(1) モデル

動学パネルデータモデルは、説明変数を含むことも多い。

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \beta' X_{it} + \epsilon_{it}.$$

- X_{it}: 説明変数
- 興味のある母数は α あるいは β(の一部) である。

応用

動学パネルデータモデルは、経済学の多くの分野での実証分析で使用されている。

- 労働経済学 (例. Meghir and Pistaferri, 2004).
- 経済成長論 (例. Levine, Loayza and Beck, 2000).
- 産業組織論 (例. Nickell, 1996).
- 国際経済学 (例. Clerides, Lach and Tybout, 1998).

推定する母数

説明変数のないモデルを考えよう。

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \epsilon_{it}.$$

推定したい母数は α である。

- 他の母数は、α の推定ができれば、比較的簡単に推定できる。
- 説明変数のあるモデルでも、推定の基本は同じである。

推定上の問題

動学パネルデータモデルの推定には、個人効果をいかにして取り除くという問題と、個人効果除くことで発生する内生性の問題を考える必要がある。

- OLS の一種である固定効果推定量は T が無限に行かないと一致性はない。
  また T がそれほど大きくないならバイアスが残る。
- 動学パネルデータモデルの推定は、操作変数法によって行われることが多い。
  最も有名な推定量は Arellano-Bond 推定量であろう。

固定効果推定

固定効果推定量は、個人効果を除いたモデルに対する OLS 推定量である。
まず、「固定効果変換」を行い、個人効果を除去する。

$$
\begin{aligned}
&y_{it} - \frac{1}{T-1}\sum_{t=2}^{T} y_{it} \\
=&\alpha\left(y_{i,t-1} - \frac{1}{T-1}\sum_{t=1}^{T-1} y_{it}\right) + \left(\epsilon_{i,t} - \frac{1}{T-1}\sum_{t=2}^{T}\epsilon_{it}\right)
\end{aligned}
$$

固定効果推定量とは、このモデルの OLS 推定量である。

- T が一定の場合、一致性を持たない。モデルの設定上、固定効果変換後のモデルの説明変数と誤差項は相関を持つ。
- T が無限に行くなら、相関はなくなり、一致推定が可能になる。

一階差分推定量

固定効果の除去のために、他の変換を使用しても、やはり問題は生じる。
例えば、一階差分をとって、個人効果を除去しよう。

$$y_{it} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + (\epsilon_{it} - \epsilon_{i,t-1}).$$

一階差分をとったモデルに OLS をあてはめて α を推定する。

$$\hat{\alpha}_{FD} = \frac{\sum_{i=1}^{N} \sum_{t=3}^{T} (y_{i,t-1} - y_{i,t-2})(y_{it} - y_{i,t-1})}{\sum_{i=1}^{N} \sum_{t=3}^{T} (y_{i,t-1} - y_{i,t-2})^2}.$$

この推定量を一階差分 (FD) 推定量という。

一階差分推定量も一致性を持たない

しかし、一階差分推定量は $N \to \infty$ で T を固定したときには一致性をもたない。

$$y_{it} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + (\epsilon_{it} - \epsilon_{i,t-1}).$$

このモデルでは、説明変数と誤差項に相関がある。なぜなら、$y_{i,t-1}$ と $\epsilon_{i,t-1}$ は設定上相関があるのである。

- これらの他にも、個人効果を除く変数変換の方法はあるが、どの変数変換を行っても同じ問題が発生する。

操作変数法

T が小さくとも機能する方法が望ましいために、動学パネルデータモデルの推定には操作変数法が用いられる。
一回差分をとると、

$$y_{it} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + \epsilon_{it} - \epsilon_{i,t-1}$$

となる。このとき

$$\begin{aligned} &E(y_{i,t-2}(\epsilon_{it} - \epsilon_{i,t-1})) \\ =&E(y_{i,t-2}(y_{it} - y_{i,t-1} - \alpha(y_{i,t-1} - y_{i,t-2})) = 0 \end{aligned}$$

が成り立つ。
このモーメント条件を元にした操作変数推定量が、Anderson and Hsiao (1981) の推定量である。

Anderson-Hsiao 推定量

Anderson-Hsiao 推定量は $y_{i,t-2}$ を操作変数として使用する。

$$\hat{\alpha}_{AH} = \frac{\sum_{i=1}^{N}\sum_{t=3}^{T} y_{i,t-2}(y_{it} - y_{i,t-1})}{\sum_{i=1}^{N}\sum_{t=3}^{T} y_{i,t-2}(y_{i,t-1} - y_{i,t-2})}.$$

- $y_{i,t-2}$ と $\epsilon_{it} - \epsilon_{i,t-1}$ は無相関である。

GMM

現在最も良く使用されている推定量は、Arellano and Bond (1991) あるいは、Holtz-Eakin, Newey and Rosen (1988) による GMM 推定量である。
Anderson-Hsiao 推定量では、$y_{i,t-2}$ しか操作変数として使用しなかったが、他の全てのラグも操作変数として使用可能である。

$$
\begin{aligned}
&E(y_{i,t-s}(\epsilon_{it}-\epsilon_{i,t-1})) \\
=&E(y_{i,t-s}(y_{it}-y_{i,t-1}-\alpha(y_{i,t-1}-y_{i,t-2}))=0
\end{aligned}
$$

が $s \geq 2$ で成り立つ。
これら全てのモーメント条件を用いた GMM 推定量が Arellano-Bond 推定量である。

他の操作変数

他のラグ付きの被説明変数も操作変数として使用できる。
Anderson-Hsiao 推定量は $y_{i,t-2}$ のみを操作変数として使っている。

$$\Delta\epsilon_i = \begin{pmatrix} \epsilon_{i3} - \epsilon_{i2} \\ \dots \\ \epsilon_{iT} - \epsilon_{i,T-1} \end{pmatrix}$$

と

$$\mathbf{Z}_i = \begin{pmatrix} [y_{i1}] & 0 & \dots & 0 \\ 0 & [y_{i1}, y_{i2}] & \dots & 0 \\ \dots & & \dots & 0 \\ 0 & \dots & 0 & [y_{i1}, \dots, y_{i,T-2}] \end{pmatrix}.$$

と定義しよう。

差分モーメント条件

全ての利用可能なラグを操作変数として使用すると、
Arellano-Bond 推定量が得られる。
Arellano and Bond (1991) and Holtz-Eakin, Newey and Rosen (1988) は、次のモーメント条件に基づく GMM で α を推定することを提唱した。

$$E(\mathbf{Z}_i' \Delta \epsilon_i) = 0.$$

Arellano-Bond 推定量

次の目的関数を最小化することで推定量を得られる。

$$\min_a \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{Z}'(\Delta y_i - a\Delta y_{i,-1}) \right)' W \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{Z}'(\Delta y_i - a\Delta y_{i,-1}) \right)',$$

なお、W は重み付け行列である。次の重み付け行列を使用すると、(均一分散のもとでの) 効率的な GMM 推定量が得られる。

$$\hat{W}^{opt} = \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{Z}_i \Delta\hat{\epsilon}_i \Delta\hat{\epsilon}_i' \mathbf{Z}_i \right)^{-1}.$$

GMM 推定量

GMM 推定量の明示的な式は

$$
\begin{aligned}
\hat{\alpha}_{AB} \quad = \quad & \left\{ \left(\sum_{i=1}^{n} \Delta y_{i,-1}' \mathbf{Z}_i \right) W \left(\sum_{i=1}^{n} \mathbf{Z}_i' \Delta y_{i,-1} \right) \right\}^{-1} \\
& \times \left(\sum_{i=1}^{n} \Delta y_{i,-1}' \mathbf{Z}_i \right) W \left(\sum_{i=1}^{n} \mathbf{Z}_i' \Delta y_i \right)
\end{aligned}
$$

である。

- Arellano–Bond 推定量は動学パネルデータモデルにおける最も頻繁に使用される推定量である。
- この推定量は、(STATA や EViews などの) 多くの統計分析ソフトで簡単に計算できる。

説明変数を含むモデル

同様の方法で、説明変数を含む動学パネルデータモデルの推定を行う事ができる。

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \beta' X_{it} + \epsilon_{it}.$$

X_{it} は外生 (誤差項と相関がない) と仮定するが、その際に二つ種類の外生性を区別する必要がある。

- 強外生: $E(X_{it}\epsilon_{is}) = 0$ が全ての t と s について成り立つ。この場合には、$E(X_{it}(\epsilon_{is} - \epsilon_{i,s-1})) = 0$ というモーメント条件が全ての t と s について成り立つ。
- 先決: $E(X_{it}\epsilon_{is}) = 0$ が $t \leq s$ についてのみ成り立つ。この場合は、$E(X_{it}(\epsilon_{is} - \epsilon_{i,s-1})) = 0$ というモーメント条件が $t < s$ についてのみ成り立つ。

GMM 推定の問題点

残念ながら GMM 推定量は常に良い性質を持つとは限らない。
特に GMM 推定量は次の問題に注意する必要がある。

- 弱い操作変数の問題。
- 操作変数が多い問題。

これらの問題については、多くの研究があり、また多くの解決案も提唱されている。

弱い操作変数の問題

操作変数が「弱い」とは、その操作変数と説明変数との相関が弱いことを意味する。

- 動学パネルデータモデルでは、$y_{i,t-2}$ が操作変数で、$y_{i,t-1} - y_{i,t-2}$ が説明変数である。
- 弱い操作変数の問題は、α が 1 に近いときに発生する。
- つまり、単位根過程に近い場合には、$y_{i,t-1}$ と $y_{it} - y_{i,t-1}$ の相関が小さくなる。

弱く操作変数の問題、その二

まず、操作変数と説明変数の共分散をみよう。

$$\begin{aligned} E(y_{i,t-1}(y_{it} - y_{i,t-1})) &= E(y_{i,t-1}((\alpha-1)y_{i,t-1} + \mu_i + \epsilon_{it})) \\ &= -(1-\alpha)E(y_{i,t-1}^2) + E(y_{i,t-1}\mu_i). \end{aligned}$$

もし、y_{it} が定常であれば、

$$y_{it} = \frac{\mu_i}{1-\alpha} + \sum_{j=0}^{\infty} \alpha^j \epsilon_{i,t-j}$$

となる。

弱く操作変数の問題、その三

したがって、共分散は、

$$
\begin{aligned}
& -(1-\alpha)E(y_{i,t-1}^2) + E(y_{i,t-1}\mu_i) \\
= \ & -\frac{(1-\alpha)}{(1-\alpha)^2}E(\mu^2) + \frac{1}{1-\alpha}E(\mu^2) - (1-\alpha)E\left(\left(\sum_{j=0}^{\infty}\alpha^j\epsilon_{i,t-j}\right)^2\right) \\
= \ & -\frac{\sigma_\epsilon^2}{1+\alpha}
\end{aligned}
$$

である。
$y_{i,t-1}$ の分散は、

$$
\frac{\sigma_\mu^2}{(1-\alpha)^2} + \frac{\sigma_\epsilon^2}{1-\alpha^2}.
$$

弱い操作変数の問題、その四

$(y_{it} - y_{i,t-1})$ の分散は、

$$2\sigma_\epsilon^2 + \frac{2}{1+\alpha}\sigma_\mu^2$$

である。
従って、相関係数は、

$$-\frac{\sqrt{(1-\alpha)^2}}{1+\alpha} \frac{\sigma_\epsilon^2}{\sqrt{2\sigma_\epsilon^2 + \frac{2}{1+\alpha}\sigma_\mu^2}\sqrt{\sigma_\epsilon^2 + \frac{1-\alpha}{1+\alpha}\sigma_\mu^2}}$$

となり、$\alpha \to 1$ の時、相関係数は 0 に近づく。

レベルモーメント

Arellano and Bover (1995) によるレベルモーメント条件を用いた GMM もよく使われている。

- Blundell and Bond (1998) は、先のモーメント条件は α が 1 に近いと弱いモーメント条件となると、指摘した。
- モーメント条件が弱いとは、この場合、$y_{i,t-2}$ と $(y_{i,t-1} - y_{i,t-2})$ の相関が弱くなることである。$\alpha = 1$ で単位根の場合には、相関はなくなる。

レベルモーメント条件では、操作変数から個人効果を除く。

$$
\begin{aligned}
&E((y_{i,t-1} - y_{i,t-2})(\mu_i + \epsilon_{it})) \\
=&E((y_{i,t-1} - y_{i,t-2})(y_{it} - \alpha y_{i,t-1})) = 0
\end{aligned}
$$

- 両方のモーメント条件を使用した GMM 推定量が、システム GMM 推定量である。

レベルモーメント条件

システム GMM 推定量に使われる追加的なモーメント条件は、レベルモーメント条件と呼ばれる。
このモーメント条件が満たされることをみよう。
まず、$E(\epsilon_{it}(y_{i,t-1} - y_{i,t-2})) = 0$ であることは簡単に分かるだろう。次に、

$$E(\mu_i(y_{i,t-1} - y_{i,t-2})) = 0$$

を示す。これは、$E(\mu_i y_{it})$ は、$t \geq 4$ の場合に、定数になる。
もし、初期値が定常分布から生成されているなら、$t = 3$ でも成り立つ。

システム GMM 推定量のバイアス

- Blundell and Bond (1998) は、シミュレーションによってシステム GMM 推定量のバイアスは小さいことを示して。
- Bun and Kiviet (2006) は、解析的にバイアスの近似式を導出し、Blundell and Bond (1998) の結果は、彼らのシミュレーションデザインによることが多い可能性を指摘した。基本的には、システム GMM 推定量もバイアスがある。
- Hayakawa (2007) は、差分モーメント条件と、レベルモーメント条件は、逆方向のバイアスをもたらすことを示した。そのため、両方のモーメント条件を使用することで、バイアスを小さく押さえることができる可能性を指摘している。

非定常性

システム GMM 推定量は、定常性の仮定に大きく依存している。もし、定常性の仮定が満たされないなら、システム GMM 推定量は機能しない可能性がある。
特に問題となる非定常性は次の二つである。

1. 単位根: α が 1 に近い。
2. 初期条件: y_{i0} が定常分布から生成されていない。

文献

- Bun and Kleibergen (2013) によると、Arellano-Bond 推定量やシステム GMM 推定量の依拠する母数の識別条件は、α が 1 に近い場合には、非常に特殊なものとなる。
- Bun and Sarafidis (2015) は、動学パネルデータモデルの優れた概説論文である。特に、初期条件に関して詳しく議論している。
- Hayakawa (2009) と Hayakawa and Nagata (2016) は、α が 1 に近くとも、初期条件が定常分布から生成されていない場合には、Arellano-Bond 推定量は信頼のおける推定値をもたらすことを示した。

操作変数が多い問題

- モーメント条件が多い場合には、GMM 推定量は、バイアスが発生することが知られている。しかし、一方で、推定量の分散は、モーメント条件が多いと小さくなる。
- 動学パネルデータモデルの GMM 推定の場合には、モーメント条件の数は自然と多くなる。
- 例えば、$y_{i2} - y_{i1}$ に対しては y_{i1} が操作変数となる。しかし、$y_{t10} - y_{t9}$ に対しては、$(y_{i1}, \ldots, y_{i9})$ の 9 つの操作変数がある。
- 操作変数の数、あるいはモーメント条件の数の次数は T^2 となる。そのため、T がそれなりの大きいだけでも、モーメント条件の数は非常に大きくなる可能性がある。

モーメント条件の数を選ぶ

Okui (2009) は、動学パネルデータモデルの推定で、モーメント条件が多いことから来るバイアスについて研究した。

- バイアスは、それぞれの説明変数に対しての操作変数の数に依存する。通常の GMM 推定と異なり、全体のモーメント条件の数ではない。
- モーメント条件が増えると、推定量の分散は小さくなる。
- モーメント条件の数を選ぶ方法も提唱されている。

Hayakawa 推定量

Hayakawa (2009) は、モーメント条件を減らすことでバイアスを小さくし、なおかつ T が大きい場合には有効 (最小分散) な推定量を提唱した。
Hayakawa 推定量では、"Forward orthogonal deviation" を用いて、個人効果を除去している。

$$
\begin{aligned}
y_{it}^{*} &= \sqrt{\frac{T-t}{T-t+1}}\left(y_{it} - \frac{y_{i,t+1} + \cdots + y_{iT}}{T-t}\right), \\
x_{it}^{*} &= \sqrt{\frac{T-t}{T-t+1}}\left(y_{i,t-1} - \frac{y_{i,t} + \cdots + y_{i,T-1}}{T-t}\right).
\end{aligned}
$$

と定義する。

Hayakawa 推定量その二

すると、変換後の変数は、

$$y_{it}^* = \alpha x_{it}^* + \epsilon_{it}^*.$$

という関係を満たす。
Hayakawa 推定量は、この変換後のモデルに対する操作変数推定量である。説明変数の x_{it}^* に対して、次の “Backward orthogonal deviation” で構築した操作変数を用いる。

$$z_{it} = \sqrt{\frac{T-t}{T-t+1}} \left(y_{i,t-1} - \frac{y_{i,t-1} + \cdots + y_{i1}}{t-1} \right).$$

この推定量は、各説明変数に対して操作変数を一つしか使用していない。

Roodman の方法

Roodman (2009) は、モーメント条件の数を減らすために、複数のモーメント条件をまとめ上げる “collapse” という方法を提唱した。Roodman のモーメント条件は、

$$E(\sum_{t=s+2}^{T} y_{i,t-s}(\epsilon_{it} - \epsilon_{i,t-1})) = 0$$

が $s = 2, \ldots, T-2$ について成り立つというものである。

- この方法は、Stata のパッケージがあるため、広く使われている。
- 説明変数のあるモデルでも同じようにできる。
- ただし、この方法でモーメント条件を減らすことでどの程度効率性が損なわれるのかは、はっきりとした答えはない。

どの推定量を使うべきか？

パネル AR(1) モデルの推定法は色々あり、どの推定量がよいのかには、議論が未だに終わっていない。

- 現状では、Arellano-Bond 推定量が最もよく使われていよう。主要な計量ソフトでは実装されている。
- T がある程度あるなら、誤差修正済み固定効果推定量が望ましい。
- 個人的には Hayakawa (2009) の推定量が色々と良い性質を持っていて魅力的である。
- Arellano-Bond 推定量と、システム GMM 推定量のどちらが良いかは、議論がつきない。Blundell and Bond (1998)、Bun and Kiviet (2006)、Hayakawa (2007)、Bun and Kleibergen (2013)、Bun and Sarafidis (2013) などを参照のこと。

固定効果推定量

固定効果推定量は、T が固定のもとでは一致性を持たないが、バイアス修正をかけることによって、かなり推定の信頼性を高めることができる。
固定効果推定量について、おさらいをしよう。
まず、各個人ごとの平均を引くことで、個人効果を除去する。

$$y_{it} - \bar{y}_i = \alpha(y_{i,t-1} - \bar{y}_i) + \epsilon_{it} - \bar{\epsilon}_i.$$

そして、この変換後のモデルに OLS を適応すると固定効果推定量を得ることができる。
しかし、$N \to \infty$ で T が固定の場合は、$\hat{\alpha}_{FE}$ は一致性を持たない。

固定効果推定量のバイアス

$\hat{\alpha}_{FE}$ は、説明変数と誤差項が相関しているので、一致性を持たない。

$$E((y_{i,t-1} - \bar{y}_i)(\epsilon_{it} - \bar{\epsilon}_i)) \neq 0$$

y_{is} と ϵ_{it} は $s \geq t$ のとき相関している。

- しかし、説明変数と誤差項の相関は、T が大きくなると小さくなる。
  そのため、$\hat{\alpha}_{FE}$ は $T \to \infty$ は一致性をもつ。
- しかし、バイアスは、T が非常に大きくない限り無視できない。(Hahn and Kuersteiner (2002), Alvarez and Arellano (2003)).

バイアス修正

Hahn and Kuersteiner (2002) は簡単に計算できるバイアス修正済み推定量を提唱した。

$$\hat{\alpha}_{HK} = \frac{T+1}{T}\hat{\alpha}_{FE} + \frac{1}{T}$$

$N, T \to \infty$ かつ $N/T \to \kappa$(ただし $0 < \kappa < \infty$) のとき

$$\sqrt{NT}(\hat{\alpha}_{HK} - \alpha) \to_d N(0, 1 - \alpha^2)$$

となる。

- バイアス修正をしない場合は、漸近分布は、平均が 0 とならない。
- T が固定であるなら、この推定量はやはり一致性をもたない。

文献

固定効果推定量のバイアス修正については、以下の文献も重要である。

- Kiviet (1995).
- Bun and Carree (2005).

他のモデル

この講義では、AR(1) モデルのみを取り扱うが、同様の議論が他の動学パネルデータモデルでも成り立つ。

- Lee (2012): AR(p).
- Okui (2010): 自己共分散と自己相関の推定。
- Arellano (2003): 他の動学パネルデータモデルについて。
- Lee, Okui and Shintani (2016): AR(∞)。

ハーフパネルジャックナイフ

Dhaene and Jochmans (2015) によるハーフパネルジャックナイフ (HPJ) を使用して、バイアス修正を行う。
T が偶数の場合を考えよう。

1. まず、パネルデータを前半と後半の二つのパネルデータに分ける。($\{\{y_{it}\}_{t=1}^{T/2}\}_{i=1}^{N}$ と $\{\{y_{it}\}_{t=T/2+1}^{T}\}_{i=1}^{N}$)
2. $\hat{H}(1)$ と $\hat{H}(2)$ を、それぞれ、$\{\{y_{it}\}_{t=1}^{T/2}\}_{i=1}^{N}$ あるいは $\{\{y_{it}\}_{t=T/2+1}^{T}\}_{i=1}^{N}$ を使った推定量とする。
3. HPJ 推定量は、

$$\tilde{H}^{HPJ} = 2\hat{H} - \frac{1}{2}\left(\hat{H}(1) + \hat{H}(2)\right).$$

となる。この推定量は、バイアスの最大項を消すことができる。

HPJ バイアス修正の基本的な考え方

- まず $\hat{H}$ が次のように書けることを証明する。

$$\hat{H} = H + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right) + \frac{B}{T} + o_p\left(\frac{1}{T}\right),$$

ただし、B は定数である。

- すると、

$$\bar{H} = H + O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right) + \frac{2B}{T} + o_p\left(\frac{1}{T}\right).$$

となる。

- 次数 $1/T$ のバイアスは $\bar{H} - \hat{H}$ と推定できる。
- そのため、HPJ 推定量 $\hat{H}^H = \hat{H} - (\bar{H} - \hat{H})$ は、次数 $1/T$ のバイスは除去できる。

動学パネルデータ分析の必要性

なぜパネルデータを用いて動学構造を分析する必要があるのだろうか？

例えば、個人の所得は、系列相関が強い。

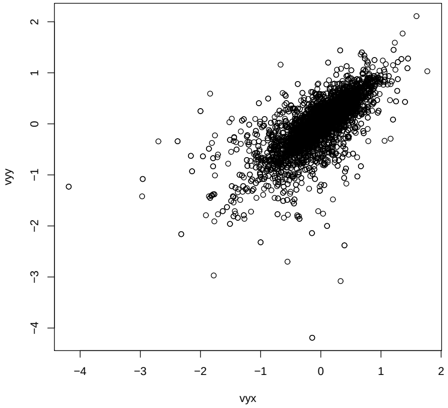
- 高所得の個人は、将来も高所得である傾向がある。

しかし、この発見に対しては二つの解釈が成り立つ。

- 観測できない異質性 (Unobserved heterogeneity)
- 状態依存 (State dependence)

実例

PSID での対数所得と、次の年の対数所得の関係



二つの解釈

1. 観測できない異質性: 所得の系列相関は観測できない個人ごとの永続的な違いによって生み出される。
   - 高所得の人は、生産性が高い。生産性の高い労働者は、生涯にわたって高い所得を得続けることができる。
   - この能力などの異質性が系列相関を生み出している。
2. 状態依存: 系列相関は、所得の動学構造がそもそも持続的であるからである。
   - たまたま、運良く高い所得をえると、なぜかそのまま高い所得を得ることができる。
   - この現象は、たとえば市場の不完全性などによって起こる。(賃金は下方硬直的であるというのは、昔からある議論である。)

二つの解釈を分けて考える必要性

これら二つの解釈を区別した上で、所得の動学構造を分析する必要がある。

1. 政策的含意が大きく異なる。
   1.1 観測できない異質性の場合
      - この場合に所得再分配がよいかどうかを判断するのは、経済学が苦手とする分野である。規範的な基準を持ち出す必要があるので。

   1.2 状態依存
      - 再分配政策は、経済学的な効率性の観点から支持することができる。つまり、リスクシェアリングの性質をもつ政策であればよい。また、所得の調整を速めるような政策も重要になる可能性がある。
2. 学問的に興味のある問題である。
3. 所得動学の特定化は、多くのマクロモデルで必要となる。

他の例

以上の問題意識は、他の多くの分析対象でも重要である。

- 国ごとに豊かさが違うのはなぜか？
  - 国ごとに定常状態での豊かさが異なるのか。
  - 経済成長が定常状態に移行するまで時間がかかるからなのか。
- 価格調整の速さ: 一物一価の法則を実証的に確かめるのは容易ではない。地域ごとに同じ物の値段が違っていても、それだけでは、何とも言えない。
  - 地域ごとの需要供給状況が異なるからかもしれない。
  - 従って、価格へのショックが起こった際にどれほどの速度で定常状態の価格に行くかを調べる必要が出てくる。

パネル AR(1) モデル

パネル AR(1) モデルは、この二つの解釈のそれぞれがどの程度の寄与をしているかを簡潔に分析するのに適している。
パネル AR(1) モデルは、

$$y_{it} = \mu_i + \alpha y_{i,t-1} + \epsilon_{it}.$$

であった。

- μ_i (個人効果) は観測されない異質性を表現している。
- α (AR(1) 係数) は状態異存の程度を表現している。

説明変数の含んだモデル

観測されない異質性と状態異存を区別することは、動学構造それ自体に興味がなく、ほかの説明変数の影響を調べたい場合でも重要である。
Asiedu and Lien (2011) は、民主主義の度合いが FDI に与える影響を分析した論文である。特に、その影響が、自然資源が豊富かどうかで変化するかを、検証した。
モデルは、

$$\begin{aligned} fdi_{it} =& \mu_i + \alpha fdi_{i,t-1} + \beta dem_{it} + \gamma dem_{it} \times nat_{it} \\ & + \text{他の変数} + \epsilon_{it} \end{aligned}$$

である。

- なぜ、このモデルに μ_i と $fdi_{i,t-1}$ を含めることが重要なのであろうか?

目次

パネル単位根検定

パネルデータ分析における単位根検定は、時系列分析の単位根検定を拡張したものと考えると良い。
パネル単位根検定は、多くの種類があるが、大まかに次の二つに分類される。(Hurlin and Mignon (2007) による分類)

- 第一世代
  観測個体間は独立であると仮定する。
- 第二世代
  横断面での相関を考慮した検定。パネル単位根検定は国別パネルなどのマクロパネルに対して使われることが多いため、横断面での相関は実証上重要な問題となる。

パネルデータを用いて単位根検定をする意義

パネルデータを用いることで、複数の観測個体全てに単位根があるかの全体像が提示できる。

- パネル単位根検定研究の初期の目的は、複数の時系列の情報を使うことで検出力を上げることであった。(時系列分析における単位根検定は、検出力が弱いことが知られている。)
- しかし、そもそも、複数の時系列をみるため、検定している問題が異なるのではないかという批判があった。

パネル単位根検定の主な論点は次の二つである。

- 如何にして、各観測個体の単位根検定を組み合わせるか。
- 観測個体間の相関をどのように考慮するか。

基本的な考え方

パネル単位根検定は、各時系列の単位根検定統計量 (あるいはその変換) の横断面での平均をとることで出来る。

- 時系列分析でみたように、単位根検定統計量の帰無仮説の下での漸近分布は非標準的な物になる。
- しかし、パネル単位根検定では、横断面での平均をとるので、通常の中心極限定理が効力を持ち、検定統計量の漸近分布は、正規分布になり、正規分布表を使用した検定ができる。
- なお、各観測個体の単位根検定統計量の漸近分布は平均が 0 でないため、パネル単位根検定統計量を構築する際には、平均が 0 になるようにする必要がある。

設定

データ: y_{it}、$t = 1, \ldots, T$、$i = 1, \ldots, N$
$T, N \to \infty$ の場合を考える。
時系列の場合と同じように、次のドリフト付きのモデルを考える。

$$\Delta y_{it} = \rho_i y_{i,t-1} + \mu_i + \epsilon_{it}$$

- ここでは、AR(1) モデルを考えている。一般的な AR モデルとその場合の Augmented Dickey-Fuller 検定に対応する方法は、後ほど簡単に議論する。
- トレンドの入ったモデルも後ほど議論する。
- ρ_i と μ_i は i に依存している。

帰無仮説：

$$\rho_i = 0, \quad \forall i.$$

LLC 検定

Levin, Lin and Chu (2002) の検定は、パネル単位根検定の嚆矢とも言える検定である。

$$\Delta y_{it} = \rho_i y_{i,t-1} + \mu_i + \epsilon_{it}$$

のモデルを、$\rho_i = \rho$ と全ての係数が同じとして ρ を推定し、t 統計量を作る。

$$t_{NT} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} (y_{i,t-1} - \bar{y}_{i,-1})(\Delta y_{it} - \bar{\Delta y_i})/\hat{\sigma}_i^2}{\sqrt{\hat{\sigma}^2 \sum_{i=1}^{N} \sum_{t=2}^{T} (y_{i,t-1} - \bar{y}_{i,-1})^2/\hat{\sigma}_i^2}}.$$

ただし、$\bar{y}_{i,-1} = \sum_{t=2}^{T} y_{it}/(T-1)$ かつ $\bar{\Delta y_i} = \sum_{t=2}^{T} \Delta y_{it}/(T-1)$ である。
なお、この検定統計量は、横断面での不均一分散に対応している。

LLC 検定、分散推定

$\hat{\sigma}_i^2$ は i 毎の回帰に基づいた誤差項の分散推定量である。

$$\hat{\sigma}_i^2 = \frac{1}{T-1}\sum_{t=2}^{T}((\Delta y_{it} - \bar{\Delta y_i}) - \hat{\rho}_i(y_{i,t-1} - \bar{y}_{i,-1}))^2$$

ただし、$\hat{\rho}_i$ は各 i 毎の ρ_i の推定量。
また $\hat{\sigma}^2$ は、次で定義される。

$$\hat{\sigma}^2 = \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}((\Delta y_{it} - \bar{\Delta y_i}) - \hat{\rho}(y_{i,t-1} - \bar{y}_{i,-1}))^2/\hat{\sigma}_i^2$$

ただし、

$$\hat{\rho} = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1} - \bar{y}_{i,-1})(\Delta y_{it} - \bar{\Delta y_i})/\hat{\sigma}_i^2}{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1} - \bar{y}_{i,-1})^2/\hat{\sigma}_i^2}.$$

LLC 検定、$T \to \infty$

上の t 検定統計量は、$T \to \infty$ のとき、

$$t_{NT} \to_d \frac{\sum_{i=1}^{N} s_i \int_0^1 \bar{W}_i(r) dW_i(r)}{\sqrt{\hat{\sigma}_N \sum_{i=1}^{N} s_i^2 \int_0^1 \bar{W}_i^2(r) dr}}$$

となる。ここで、W_i は標準ブラウン運動、
$\bar{W}_i(r) = W_i(r) - \int_0^1 W_i(s) ds$、$s_i$ は Δy_{it} の長期分散と、ϵ_{it} の分散の比である。
これをさらに $N \to \infty$ ととる。

- 分母は、定数に収束する。
- また、分子は、独立確率変数の和であるので、中心極限定理を適用できる。
- しかし、注意すべきは、$s_i \int_0^1 \bar{W}_i(r) dW_i(r)$ は平均が 0 でなく、$-s_i/2$ であることである。

LLC 検定、統計量

LLC 検定統計量は、

$$LLC = \frac{t_{NT}}{\sigma_T^*} - \frac{\mu_T^* T \sum_{i=1}^N \hat{s}_i}{\sigma_T^* \sqrt{\hat{\sigma}^2 \sum_{i=1}^N \sum_{t=2}^T (y_{i,t-1} - \bar{y}_{i,-1})^2 / \hat{\sigma}_i^2}}$$

である。

- μ_T^* は漸近的に $-1/2$ になる定数。LLC はシミュレーションで適切な値を求めている。
- σ_T^* は漸近的に $\sqrt{1/2}$ になる定数。LLC はシミュレーションで適切な値を求めている。
- $\hat{s}_i$ は s_i の推定量。

帰無仮説の元で、この検定統計量は、漸近的に $N(0,1)$ に従う。単位根の検定であるので、左側の片側検定 (つまり、大きく負であれば棄却する) になる。

IPS 検定

Im, Pesaran and Shin (2003) は、各観測個体の単位根検定統計量の平均による検定を提唱した。
各 i 毎の ρ_i に関する t 検定統計量は、

$$t_i = \frac{\sum_{t=2}^{T}(y_{i,t-1} - \bar{y}_{i,-1})(\Delta y_{it} - \bar{\Delta y_i})}{\sqrt{\hat{\sigma}_i^2 \sum_{t=2}^{T}(y_{i,t-1} - \bar{y}_{i,-1})^2}}.$$

である。帰無仮説の下で、$T \to \infty$ のとき、

$$t_i \to_d \frac{\int_0^1 \bar{W}_i(r)dW_i(r)}{\sqrt{\int_0^1 \bar{W}_i^2(r)dr}}$$

となる。

IPS 検定、検定統計量

IPS 検定統計量は、

$$IPS = \frac{\sqrt{N}(\sum_{i=1}^{N} t_i/N - \sum_{i=1}^{N} \mu(t_i)/N)}{\sqrt{\sum_{i=1}^{N} \sigma(t_i)^2/N}}$$

である。

- $\mu(t_i)$ は t_i の帰無仮説の下での平均である。IPS にシミュレーションに基づく数値がある。これは、定数やトレンド項がある状況では必要になる。
- $\sigma(t_i)$ は t_i の帰無仮説の下での分散である。IPS にシミュレーションに基づく数値がある。

帰無仮説の下で、$N, T \to \infty$ の時、$IPS \to_d N(0,1)$ になる。片側検定であり、IPS が大きく負の場合に帰無仮説を棄却する。

Combination 検定

Maddala-Wu (1999) と Choi (2001) は、検定統計量そのものでなく、p 値を組み合わせる検定法を提唱した。

- 各 i ごとに単位根検定を行い、p 値を計算する。p_i を i の p 値とする。そして、p_i の変換の平均をとる。
- Combination 検定の基本は、帰無仮説のもとで、p 値は、一様分布に従うことである。

Choi 検定

Choi (2001) は

1. p_i は Elliott, Rothenberg and Stock (1996) の Dickey-Fuller-GLS 検定の p 値をとる。
2. 検定統計量として、

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(p_i)$$

を使う。なお、Φ は標準正規分布関数である。なお、帰無仮説の元で、$Z \to_d N(0,1)$ である。

という方法を薦めている。

注意点

- パネル単位根検定の帰無仮説は
$$\rho_i = 0, \quad \forall i$$
である。したがって、帰無仮説が棄却された場合は、
ρ_i が 0 でない観測個体がある、ということになる。すべての観測個体が定常ということではない。
- ここでは、$T \to \infty$ のもとで理論的に正当化できる検定を考えた。T 固定の場合の検定統計量もある。有名なものでは、Harris and Tzavalis (1999) や Breitung (2000) など。

定常性検定

帰無仮説を定常過程とする検定もできる。こうした検定は、定常性検定と呼ばれる。
各観測個体での Kwiatkowski, Phillips, Schmidt and Shin (1992) の検定を組み合わせたものになる。

- Hadri (2000) は KPSS 検定統計量の横断面での平均に基づく検定を提唱した。
- Yin and Wu (2000) は KPSS 検定の p 値に基づく combination 検定を提唱した。combination 検定の考え方は、上で議論したものと同じなので、ここでは省略する。

Hadri 検定

まず KPSS 検定統計量を復習しよう。
各 i 毎の KPSS 検定統計量は、

$$KPSS_i = \frac{1}{T^2 \hat{\lambda}} \sum_{t=1}^{T} \left(\sum_{s=1}^{t} (y_{is} - \bar{y}_i) \right)^2$$

である。ただし、$\hat{\lambda}$ は y_{it} の長期分散の推定量である。
定常過程であれば、$T \to \infty$ のとき、

$$KPSS_i \to_d \int_0^1 (W_i(r) - rW_i(1))^2 dr$$

となる。

Hadri 検定、検定統計量

Hadri (2000) は次の検定統計量を提唱した。

$$Hadri = \frac{\sqrt{N}(\sum_{i=1}^{N} KPSS_i/N - 1/6)}{\sqrt{1/45}}.$$

- $1/6$ は $\int_0^1 (W_i(r) - rW_i(1))^2 dr$ の平均、$1/45$ はその分散である。

定常過程の帰無仮説の下で、$N, T \to \infty$ のとき、

$$Hadri \to_d N(0, 1)$$

となる。
Hadri 検定は、右側の片側検定である。つまり、検定統計量が正で大きい時に、定常過程の帰無仮説を棄却する。

拡張

これまで紹介してきた検定方法は、より一般的な状況に拡張することも可能である。

- 各 i ごとの ADF 検定や Phillips-Perron 検定を組み合わせることで、系列相関を考慮した検定もできる。なお、ADF でのラグの長さは、各 i ごとに異なっても良い。
- トレンド付きのモデルのための単位根検定を組み合わせることも可能である。
- ただし、これらの拡張を行うと、平均や分散の項が異なるため、横断面での平均をとって中心極限定理を適用する際の平均や分散の調整項が異なってくる。

横断面での相関

パネル単位根検定は、国別パネルなどのマクロパネルデータに対して使用されることが多い。
そのため、横断面での相関を考慮することが重要になる。
横断面での相関構造は、ファクター構造を用いてモデル化することが一般的である。

- Moon and Perron (2004) や Bai and Ng (2004) はファクターを推定することで、データから横断面での相関を取り除いて、単位根検定をすることを提唱した。
  この方法は、ファクターの推定法の解説が必要となるので、今回は議論しないことにする。しかし、よく使われてる方法である。
- Pesaran (2007) は、横断面での平均値をモデルに含めることでファクターの影響を取り除けることを示した。
- Breitung and Das (2005) は横断面の相関を考慮した GLS 推定に基づく検定を提唱した。ここでは省略する。

ファクター構造

次のモデルを考える。

$$\Delta y_{it} = \rho y_{i,t-1} + u_{it}$$

ただし、u_{it} はファクター構造を持っているとする。

$$u_{it} = \gamma_i f_t + \epsilon_{it}$$

f_t は全ての観測個体に共通だが時間を通じて変化する観測出来ない項である。

- 横断面での相関をファクター構造でモデル化している。つまり、u_{it} はどの i でも f_t を含むため、横断面で相関している。

CIPS 検定

Pesaran (2007) の方法は CPIS 検定と呼ばれ、Pesaran (2006) によって開発された common correlated effects (CCE) と呼ばれる手法を下にしている。次の横断面での平均を加えたモデルを推定する。

$$\Delta y_{it} = \rho y_{i,t-1} + \alpha_1 \Delta\bar{y}_t + \alpha_2 \bar{y}_{t-1} + v_{it}$$

ただし、$\Delta\bar{y}_t = \sum_{t=1}^{N} \Delta y_{it}/N$ かつ $\bar{y}_{t-1} = \sum_{t=1}^{N} y_{i,t-1}/N$ である。

- 非説明変数の横断面の平均 $\Delta\bar{y}_t$ をもモデルに加えているところが自明でないところである。

このモデルを推定することで、ρ の一致推定量を得ることができ、それに基づくの t 統計量を使用して、単位根検定ができる。

定常性検定

横断面での相関を考慮した定常性検定もある。

- Harris, Keybourne and McCabe (2005) や Bai and Ng (2005) などがある。

目次

終わりに

本日の講義では、パネルデータ分析の概略を紹介した。

1. 固定効果モデルと固定効果推定量。
   - 観測出来ないが時間を通じて変化しない要素を、固定効果によってモデル化し、固定効果推定量を用いて、そのような要素によってもたらされる欠落変数のバイアスを避けることができる。
2. (定常) 動学的パネルデータモデル
   - GMM によって推定することが通例である。
   - 観測出来ない異質性と状態異存
3. パネル単位根検定
   - 基本的には時系列分析の単位根検定を組み合わせる。
   - 横断面での相関を考慮する必要がある可能性。

文献

次に挙げる文献は、動学パネルデータモデルを学習する上で、有用な教科書である。

- Wooldridge, Jeffrey M., "Econometric Analysis of Cross Section and Panel Data," 2nd edition, (2010), the MIT press.
  動学パネルデータに限らず、ミクロデータ分析の計量経済学の手法についての基本書といえる。
- Arellano, Manuel, "Panel Data Econometrics," (2003), Oxford University Press.
  パネルデータに関しての専門的な教科書。
- 千木良、早川、山本「動学的パネルデータ分析」(2011), 知泉書館
  動学パネルデータモデルに関する詳細な専門書。

多くの統計分析ソフト (STATA, EViews, R, gretl など) では、動学パネルデータモデルを分析するための機能が実装されている。