

2

2010 年度 学部 エコノメトリックス 第 23 回講義メモ

2011 年 1 月 11 日

§ 4.4 パネルデータの分析 (続)

ランダム効果モデルの GLS 推定

GLS 推定を行うために、まず回帰モデル

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

における誤差項 $\mathbf{u}$ の分散共分散行列 ($n \times n$)

$$\boldsymbol{\Omega} \equiv \mathbf{E}(\mathbf{u}\mathbf{u}')$$

の構造を知っておく必要がある。

3

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{pmatrix}, \quad \text{ここで } \mathbf{u}_i = \begin{pmatrix} u_{i1} \\ \vdots \\ u_{iT} \end{pmatrix} = \begin{pmatrix} v_i + \varepsilon_{i1} \\ \vdots \\ v_i + \varepsilon_{iT} \end{pmatrix}$$

であるから、 $v_i \perp \varepsilon_{it}$ かつ $v_i \sim (0, \sigma_v^2)$, $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$ という仮定より

$$\boldsymbol{\Sigma} \equiv \mathbf{E}(\mathbf{u}_i \mathbf{u}_i') = \sigma_v^2 \mathbf{1}\mathbf{1}' + \sigma_\varepsilon^2 \mathbf{I}_T$$

という $T \times T$ の行列を用いると、

$$\boldsymbol{\Omega} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Sigma} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{O} \\ \mathbf{O} & \cdots & \mathbf{O} & \boldsymbol{\Sigma} \end{pmatrix} \quad (4.4.3)$$

4

という構造になっていることがわかる。

次に、 $\boldsymbol{\Omega}^{-1/2}$ を求めることにしよう。(4.4.3) 式より、 $\boldsymbol{\Sigma}^{-1/2}$ が求まれば、自動的に求まることが分かる。 $\mathbf{P}_1 \equiv \mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1} = 1/T\mathbf{1}\mathbf{1}'$ であることを用いると、

$$\boldsymbol{\Sigma} = T\sigma_v^2\mathbf{P}_1 + \sigma_\varepsilon^2\mathbf{I}_T$$

となる。さらに、適当なスカラー λ を用いて、

$$\boldsymbol{\Sigma}^{-1/2} = (1/\sigma_\varepsilon)(\mathbf{I}_T - \lambda\mathbf{P}_1)$$

と表現できるものとする、

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_\varepsilon^2}(\mathbf{I}_T + (\lambda^2 - 2\lambda)\mathbf{P}_1)$$

となるので、

$$\mathbf{I}_T = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}$$

5

$$= \frac{T\sigma_v^2}{\sigma_\varepsilon^2}\mathbf{P}_1 + \mathbf{I}_T + \frac{T\sigma_v^2}{\sigma_\varepsilon^2}(\lambda^2 - 2\lambda)\mathbf{P}_1 + (\lambda^2 - 2\lambda)\mathbf{P}_1$$

これを整理すると、

$$\begin{aligned} \left(1 + \frac{T\sigma_v^2}{\sigma_\varepsilon^2}\right)(\lambda^2 - 2\lambda + 1) &= 1 \\ (\lambda - 1)^2 &= \left(1 + \frac{T\sigma_v^2}{\sigma_\varepsilon^2}\right)^{-1} \\ \lambda &= 1 \pm \left(1 + \frac{T\sigma_v^2}{\sigma_\varepsilon^2}\right)^{-1/2} \end{aligned}$$

となり、

$$\boldsymbol{\Omega}^{-1/2} = \frac{1}{\sigma_\varepsilon}(\mathbf{I}_n - \lambda\mathbf{P}_D)$$

6

が求められた。よって GLS 推定量は

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'(\mathbf{I}_n - \lambda\mathbf{P}_D)^2\mathbf{X})^{-1}\mathbf{X}'(\mathbf{I}_n - \lambda\mathbf{P}_D)^2\mathbf{y} \quad (4.4.4)$$

であり、その分散共分散行列は

$$\text{Var}(\hat{\boldsymbol{\beta}}_{GLS}) = \sigma_\varepsilon^2(\mathbf{X}'(\mathbf{I}_n - \lambda\mathbf{P}_D)^2\mathbf{X})^{-1}$$

となる。

このように、GLS が実行可能 (feasible) であるためには、 σ_v^2 と σ_ε^2 の値が分かっている必要がある。(λ だけではダメな事に注意。)(4.4.4) 式より、 $\lambda = 0$ の時には、GLS 推定量と OLS 推定量は同じになり、また $\lambda = 1$ の時には、GLS 推定量は 'within estimator' と同じになることが分かる。

7

 Ω の推定

σ_v^2 と σ_ε^2 が既知でない場合、GLS が実行可能となるためには、これらを推定する必要がある。

まず、(4.4.1) 式の OLS 残差二乗和を $n - k - m$ で割って $\hat{\sigma}_\varepsilon^2$ を得る。次に、(4.4.2) 式の OLS 残差二乗和を $(m - k)T$ で割ると、 $\sigma_v^2 + \widehat{\sigma_\varepsilon^2}/T$ が求まるので、

$$\hat{\sigma}_v^2 = \sigma_v^2 + \widehat{\sigma_\varepsilon^2}/T - \frac{1}{T}\hat{\sigma}_\varepsilon^2$$

として求める。

8

ランダム効果モデルの推定上の注意

ランダム効果モデルを適用する場合、確率変数である 'individual effect' v_i と説明変数 $\mathbf{x}_{it}$ とが無相関である必要がある。もし、相関がある場合には説明変数と攪乱項（誤差項）の間に相関があることになり、GLS 推定量は不偏性、一致性を持たなくなる。

相関が存在する場合でも、 $\hat{\beta}_{LSDV}$ は不偏推定量（但し効率性はもたない）であるから、この性質を用いた検定が提案されている。

Hausman 検定は、specification test と呼ばれるものの一つである。 v_i と $\mathbf{x}_{it}$ とが無相関であるという帰無仮説 H_0 の下では、 $\hat{\beta}_{LSDV}$ も $\hat{\beta}_{GLS}$ もどちらも不偏である。

$$\Psi \equiv \text{Var}(\hat{\beta}_{LSDV} - \hat{\beta}_{GLS}) = \text{Var}(\hat{\beta}_{LSDV}) - \text{Var}(\hat{\beta}_{GLS})$$

9

として、次のように検定統計量を構成する。

$$W = (\hat{\beta}_{LSDV} - \hat{\beta}_{GLS})' \Psi^{-1} (\hat{\beta}_{LSDV} - \hat{\beta}_{GLS})$$

W は帰無仮説 H_0 の下で自由度 $(k - 1)$ の χ^2 分布に従うので、この性質を用いて検定を行う。

勿論、Hausman 検定によって帰無仮説 H_0 が棄却された場合には、GLS 推定量は一致性がないと判断されるので、用いてはならない。

Stata11 によるパネルデータ分析の例

```
. xtreg epwr temp, fe i(area)
```

```
Fixed-effects (within) regression      Number of obs   =      12
Group variable: area                  Number of groups =       3

R-sq:  within = 0.2071                Obs per group:  min =       4
      between = 0.0011                  avg   =      4.0
      overall  = 0.0000                  max   =       4

                                F(1, 8)      =      2.09
                                Prob > F       =     0.1863
```

epwr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
temp	83.91575	58.04681	1.45	0.186	-49.94043	217.7719
_cons	2704.637	1875.632	1.44	0.187	-1620.579	7029.853
sigma_u	2813.4397					
sigma_e	232.78508					
rho	.99320058	(fraction of variance due to u_i)				

```
F test that all u_i=0:      F(2, 8) =    582.31      Prob > F = 0.0000
```

```
. estimates store fixed
```

```
. xtreg epwr temp, re i(area)
```

```
Random-effects GLS regression      Number of obs   =      12
Group variable: area              Number of groups =       3

R-sq:  within = 0.2071                Obs per group:  min =       4
      between = 0.0011                  avg   =      4.0
      overall  = 0.0000                  max   =       4

                                Wald chi2(1)   =      2.31
                                Prob > chi2     =     0.1282
```

epwr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
temp	83.76955	55.05976	1.52	0.128	-24.1456	191.6847
_cons	2709.358	2807.825	0.96	0.335	-2793.877	8212.594
sigma_u	3964.5577					
sigma_e	232.78508					
rho	.99656421	(fraction of variance due to u_i)				

```
. estimates store random
```

```
. hausman fixed random
```

	---- Coefficients ----		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) fixed	(B) random		
temp	83.91575	83.76955	.1462058	18.38083

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(1) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
= 0.00
Prob>chi2 = 0.9937

```
. xtreg epwr temp, be i(area)
```

```
Between regression (regression on group means) Number of obs   =      12
Group variable: area                          Number of groups =       3

R-sq:  within = 0.2071                Obs per group:  min =       4
      between = 0.0011                  avg   =      4.0
      overall  = 0.0000                  max   =       4

                                F(1,1)        =      0.00
                                Prob > F       =     0.9785
```

epwr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
temp	-61.87946	1832.103	-0.03	0.979	-23340.95	23217.19
_cons	7412.608	59205.95	0.13	0.921	-744870.4	759695.6

```
. list
```

	year	temp	area	epwr	_est_f~d	_est_r~m
1.	2004	31	1	8585	1	1
2.	2005	31.8	1	8254	1	1
3.	2006	31.1	1	8193	1	1
4.	2007	33	1	9036	1	1
5.	2004	33.7	2	3213	1	1
6.	2005	31.3	2	3025	1	1
7.	2006	30.3	2	2947	1	1
8.	2007	29.2	2	3030	1	1
9.	2004	33.3	3	4848	1	1
10.	2005	33.2	3	4506	1	1
11.	2006	35	3	4699	1	1
12.	2007	34.6	3	4637	1	1