Specification Errors と 推定値の特性（II）

地主重夫

5. パラメーターに関する予備情報と Specificatin Errors

推定されるパラメーターに対する制約条件があらかじめ別の情報からわかっているときには，これを利用した方が好都合なことがある。ここでは2つの場合について考えてみる。

第1は推定されるパラメーターの値はわからないが，パラメーターの間の関係があらかじめわかっている場合であり，第2は推定されるパラメーターのうちいくつかの値が既知である。まず第1の場合からはじめよう。[1)][1] いまモデル

$$y = x\beta + e \qquad (75)$$

において，変数はすべて平均値からの偏差で表わされているものとする。ここで y は (n×1) 次の列ベクトル，x は (n×k) 次の行列，β は (k×1) 次の列ベクトル，e は (n×1) 次の列ベクトルである。さらに

$$E(e) = 0, \qquad E(e^2) = I\sigma^2$$

$$E(xe) = 0$$

1) 次の部分は〔1〕を若干展開したものである。

と仮定し，

$$T\beta=t \qquad (76)$$

なる関係式が既知であるとする。T は $(n\times k)$ 次の行列であり，その要素はすべて既知であるとする。また t は $(n\times 1)$ 次のベクトルでこの要素も既知であるとする。β は，制約条件式 (76) のもとで，(75) の推定分散を極小にするように推定されなければならない。もっとも (76) 式の β は真のパラメーターで未知であるから，これをその推定値でおきかえなければならない。いま λ を Lagrange 乗数とすれば，

$$\phi=(y-xb)'(y-xb)-\lambda'(Tb-t) \qquad (77)$$

上式を b に関して微分してこれを零とおき，

$$b'x'y=y'xb(=\text{scaler})$$

を考慮すれば

$$-2x'y+2x'xb_0-T'\lambda=0 \qquad (78)$$

ここで b_0 は (78) 式を満足する β の推定値である。(78) 式の左から $T(x'x)^{-1}$ をかけると，

$$-2\{T(x'x)^{-1}x'y-T(x'x)^{-1}x'xb_0\}$$
$$=T(x'x)^{-1}T^{-1}\lambda$$

上式を変形し，$b=(x'x)^{-1}xy$ を代入すれば，

$$\lambda=-2\{[T(x'x)^{-1}T']^{-1}T(x'x)^{-1}x'y$$
$$-[T(x'x)^{-1}T']^{-1}T(x'x)^{-1}x'xb_0]$$
$$=-2\{[T(x'x)^{-1}T']^{-1}[Tb-Tb_0]\} \qquad (79)$$

$Tb_0=t$ とし，上の λ を (78) 式に代入すれば，

$$b_0=b+(x'x)^{-1}T'[T(x'x)^{-1}T'](t-Tb) \qquad (80)$$

b は制約条件 (76) がない場合の β の最小二乗推定値であり，b_0 は制約条件 (76) のもとにおける β の最小二乗推定値であるから，(80) 式の右辺第 2 項は，b が制約条件 (76) を満足しない度合を表わしている。ところで (80) 式から，

$$b_0=(x'x)^{-1}x'y+(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}(t-Tb)$$
$$=(x'x)^{-1}x'(x\beta+e)+(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}$$
$$\times[t-T(x'x)^{-1}x(x\beta+e)]$$
$$=\beta+(x'x)^{-1}x'e+(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}[t-T\beta-T(x'x)^{-1}x'e]$$
$$=\beta+\{I-(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}T\}\{(x'x)^{-1}x'e\} \quad (81)$$

それゆえ，

$$(b_0-\beta)=\{I-(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}T\}\{(x'x)^{-1}x'e\}$$

$E('e)=0$ であるから $E(b_0-\beta)=0$ となり，b_0 は β の不偏推定値である。b_0 の標本分散の推定値は，

$$E(b_0-\beta)(b_0-\beta)'$$
$$=\sigma^2(x'x)^{-1}\{I-T'[T(x'x)^{-1}T'(x'x)^{-1}\} \quad (82)$$

ここで $\sigma^2(x'x)^{-1}$ は制約条件のない場合における β の最小二乗定推定値の分散である。$\sigma^2(x'x)^{-1}T'\{T(x'x)^{-1}T'(x'x)^{-1}\}$ は正の準定正値行列であるから，制約条件のもとにおける推定値 b_0 の分散は，制約条件のない場合における推定値 b の分散より大きくはない。したがって b_0 は β の最良推定値である。

しかし，制約条件（76）が定式上ただしくないとすればどうだろうか。いま，制約条件の正しい定式が

$$T\beta=\bar{t} \qquad \bar{t}\neq t \quad (83)$$

であるのに（76)を制約条件として β を推定したとしよう。いま

$$t=\bar{t}+\triangle t \quad (84)$$

としてこれを（80）式に代入すれば，

$$b_0-\beta=(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}\triangle t$$
$$+\{I-(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}T\}\{(x'x)^{-1}x'e\} \quad (85)$$

となる。(85）式の右辺第1項はもはや零にひとしくはないから，b_0 は β の不偏推定値ではない。また b_0 の分散は

$$E(b_0-\beta)(b_0-\beta)'=\sigma^2(x'x)^{-1}-(x'x)^{-1}T'[T(x'x)^{-1}T']^{-1}$$

$$\times\{\sigma^2 I-\triangle t\triangle t'[T(x'x)^{-1}T']^{-1}\}T(x'x)^{-1} \qquad (86)$$

ここで右辺第2項は，(82) 式における $\sigma^2(x'x)^{-1}T'\{T(x'x)^{-1}T'(x'x)^{-1}\}$ のように準定値形ではない。この符号は $\{\sigma^2 I-\triangle t\triangle t'[T(x'x)^{-1}T']^{-1}\}$ によってきまる。もし σ^2 が小さければこの値は負になる傾むきがあり，右辺第2項は正になるだろう。したがって，b^0 の分散は制約条件のない場合における最小二乗推定値より大きくなり，この結果，制約条件を利用しない方が有効性の高いパラメーターを推定できることになるだろう。その上，大標本では σ^2 が小さくなると同時に $(x'x)^{-1}$ は大きくなるであろうから，明らかに (86) から b_0 の分散はますます大きくなる。

以上の説明から，もし定式上正しいパラメーターへの制約条件が用いられるならば，パラメーターの不偏最良推定値がえられるが，もし定式上正しくないパラメーターの制約条件が利用されるならば，えられるパラメーターは制約条件を利用しない場合に比して有効性が低い。ただ，説明変数の間に強い相関関係がある場合，つまり線型重合が存在しているときには $(x'x)$ は大きくなるから $(x'x)^{-1}$ の要素の値は小さくなり，(86) から b_0 の分散は b の分散より小さくなって制約条件の利用がよい結果を生むという結果になるかもしれない。しかし，線型重合の存在しない場合には，パラメーターに関する予備的な情報，すなわち制約条件の利用には十分な注意が必要である。制約条件が既知なる場合におけるパラメーターの推定について簡単な例をあげて説明を加えてみよう。

いま次のような関係式を考える。

$$Y(t)=\alpha+\beta_1 X_1(t)+\beta_2 X_2(t)+e(t) \qquad (87)$$

さらに制約条件として

$$\beta_2=2\beta_1+1 \qquad (88)$$

前提仮説はみたされているものとし，(87) を平均値からの開離の形で表現すると，

$$y(t)=\beta_1 x_1(t)+\beta_2 x_2(t)+e(t) \qquad (87)'$$

(76) 式は,

$$[-2\ \ 1]\begin{bmatrix}\beta^1\\ \beta_2\end{bmatrix}=1$$

ここで

$$T\equiv[-2\ \ 1]\quad t\equiv 1 \tag{89}$$

x_1 と x_2 の単純相関係数を r_{12}, x_1 と x_2 との単純回帰係数を b_{12} のごとく表わすならば, (86) 式に対応する b_0 の分散・共分散行列は

$$\frac{\sigma^2}{(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_2}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & \dfrac{1}{\Sigma x_2^2}\end{pmatrix}$$

$$-\frac{[4\Sigma x_2^2(1+b_{12})+\Sigma x_1^2]^2\sigma^2-(\triangle t)^2-(1-r_{12}^2)n}{(1-r_{12}^2)n[4\Sigma x_2^2(1+b_{12})+\Sigma x_1^2]^2}$$

$$\times\begin{pmatrix}2\dfrac{b_{12}}{b_{21}}+3b_{12}+r_{12}^2 & -2b_{21}+r_{12}^2+b_{12}\\ -4b_{21}+b_{12}-2r_{12}^2 & \dfrac{b_{21}}{b_{12}}+4b_{12}+4r_{12}^2\end{pmatrix} \tag{90}$$

b_1 の分散は

$$\mathrm{var}(b_1)=\frac{\sigma^2}{(1-r_{12}^2)n}\frac{1}{\Sigma x_1^2}-\frac{[4\Sigma x_2^2(1+b_{12})+\Sigma x_1^2]\sigma^2}{(1-r_{12}^2)n[4\Sigma x_2^2(1+b_{12}+\Sigma x_1^2)]^2}$$

$$\times\left[2\frac{b_{12}}{b_{21}}+3_{21}+r_{12}^2\right]+\frac{(\triangle t)^2}{[4\Sigma x_2^2(1+b_{12}+\Sigma x_1^2]^2}\left[2\frac{b_{12}}{b_{21}}+3b_{21}+r_{12}^2\right] \tag{91}$$

また b_2 の分散は

$$\mathrm{var}(b_2)=\frac{\sigma^2}{(1-r_{12}^2)n}\frac{1}{\Sigma x_2^2}-\frac{\sigma^2}{(1-r_{12}^2)n[4\Sigma x_2^2(1+b_{12}+\Sigma x_1^2]^2}$$

$$\times\left[\frac{b^{21}}{b_{12}}+4b_{12}+4r_{12}^2\right]+\frac{(\triangle t)^2}{[4\Sigma x_2^2(1+b_{12}+\Sigma x_1^2]^2}\left[\frac{b_{21}}{b_{12}}+4b_{12}+4r_{12}^2\right] \tag{92}$$

もし, $r_{12}=0$ であるならば, 定式上の誤差 $\triangle t$ の存在いかんにかかわらず, 制約条件のある場合における推定値の分散は, これがない場合における推定値の分散と異ならず, 制約条件を利用しても推定値の有効性はなんら改善されない。しかし, $r_{12}\neq 0$ であり, しかも r_{12} の値が十分に大であるならば,

右辺第2項の分母にはいる $(1-r_{12}^2)$ のために第2項の絶対値は大きくなり，第3項の絶対値をこえる傾向があるだろう。このときには $var(b_2)$ は，制約条件のない場合の推定値の分散 $\sigma^2/(1-r_{12}^2)\Sigma x_2^2$ より小さくなり，制約条件の利用によってパラメーターの推定値はその有効性をますことになる。したがって，説明変数間に強い線型重合が存在することによって σ^2 および r_{12}^2 が大きくなるときには，制約条件を利用して推定されるパラメーターは有効性の高いものになる。線型重合のある場合に制約条件を用いた方がのぞましいのはこの理由によるのである。

第2に，あるパラメーターの値が既知の場合に移ろう。原理は前の第1の場合に全くひとしいから，ここでは簡単な例をあげて説明してみることにしよう。関係式は前の (87)′ と同じであるとする。

$$y(t)=\beta_1 x_1(t)+\beta_2 x_2(t)+e(t)$$

予備情報から[2)]

$$\beta_1=t \tag{89}$$

であるとする。ここで t はもちろん既知の定数である。前提仮説も第1の場合と同様であるとする。もちろん，われわれの知っているのは $\hat{b}_1$ であって β_1 ではない。これを考慮して (87)′ をかき直すと，

$$\begin{aligned} y(t)-\hat{b}_1 x_1(t) &= x_2(t)\beta_2+\{e+x_1(\hat{b}_1-\beta_1)\} \\ &= x_2\beta_2(t)+\hat{e} \end{aligned} \tag{90}$$

ここで

$$\hat{e}\equiv e+x_1(\hat{b}_1-\beta_1)$$

(90) 式では左辺が新らたな従属変数であり，説明変数は x_2 だけになり，簡単な2変数回帰方程式に変形されて計算手続きがきわめてたやすくなる。これは Stone[〔2〕] によって試みられたもので高い利用価値をもつが，ここではこの方法をとらず，第1の場合に採用した手続きを準用し方法の一貫性を保持したい。パラメーター b_0 の分散の推定値は (86) 式から，

2) 本例は〔1〕p. 333 からとられたものであるが結果は同一ではない。

$$\mathrm{var}(b_0)=\frac{\sigma^2}{(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_2}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & \dfrac{1}{\Sigma x_2^2}\end{pmatrix}$$

$$-\frac{\sigma^2-\triangle t^2(1-r_{12}^2)\dfrac{1}{\Sigma x_2^2}}{(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_2}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & -\dfrac{r_{12}^2}{\Sigma x_2^2}\end{pmatrix}$$

$$=\frac{\sigma^2}{(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_2}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & \dfrac{1}{\Sigma x_2^2}\end{pmatrix}-\frac{\sigma^2}{(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_2}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & \dfrac{r_{12}^2}{\Sigma x_2^2}\end{pmatrix}$$

$$+\frac{\triangle t^2(1-r_{12}^2)n}{\Sigma x_2^2(1-r_{12}^2)n}\begin{pmatrix}\dfrac{1}{\Sigma x_1^2} & -\dfrac{r_{12}^2}{\Sigma x_1x_1}\\ -\dfrac{r_{12}^2}{\Sigma x_1x_2} & \dfrac{r_{12}^2}{\Sigma x_2^2}\end{pmatrix}\tag{91}$$

(91) 式の右辺第 1 項は制約条件のない場合における b_0 の分散であり，第 2 項，第 3 項は制約条件を利用したことによる分散の追加分を表わしている。このうち，第 2 項は制約条件に定式上の誤差のない場合の分散の追加分であり，第 3 項は制約条件に定式上の誤差が存在することによる分散の追加分を示している。したがって，第 2 項と第 3 項の和が正ならば，制約条件を利用して推定された b_0 は，これを利用せずに推定された b よりも有効性の高いことを意味している。もし，r_{12}^2 が 1 ならば b_0 の分散は b の分散にひとしく，制約条件の利用によって推定値の有効性を増すことはないであろう。しかし r_{12}^2 が 1 にひとしくはないが十分に大きく，また定式上の誤差はさして大きくないのに観測標本数の増加によって σ^2 が大きくなると，(91) 式の右辺第 2 項の分母にある $(1-r_{12}^2)$ の値が大きくなるために，第 2 項の絶対値は第 3 項の絶対値をこえる傾むきがあり，制約条件を利用して推定される b_0 は b よりも有効な推定値になるだろう。$b_{0,1}$, $b_{0,2}$ の分散は (91) 式からそれぞれ次のとおりである。

$$\mathrm{var}(b_{0,1}) = \frac{\sigma^2}{(1-r_{12}^2)n}\frac{1}{\Sigma x_1^2} - \frac{\sigma^2}{(1-r_{12}^2)n}\frac{1}{\Sigma x_1^2} + \frac{\triangle t^2}{n\Sigma x_2^2 \Sigma x_1^2} = \frac{\triangle t^2}{\Sigma x_2^2 \Sigma x_1^2} \qquad (92)$$

この場合には，もし $\triangle t$ の値が小であり，しかも分母の値が大きく，小標本推定のときには，制約条件の利用によってパラメーターの有効性をかなり高めることができる。また $b_{0,2}$ の分散推定値は，

$$\mathrm{var}(b_{0,2}) = \frac{\sigma^2}{n(1-r_{12}^2)\Sigma x_2^2} - \frac{\sigma^2}{n(1-r_{12}^2)}\frac{r_{12}^2}{\Sigma x_2^2} + \frac{\triangle t^2 r_{12}^2}{\Sigma x_2^4} = \frac{\sigma^2(1-r_{12}^2)}{n(1-r_{12}^2)\Sigma x_2^2} + \frac{\triangle t^2 r_{12}^2}{\Sigma x_2^4}$$

$$= \frac{\sigma^2}{n\Sigma x_2^2} + \frac{\triangle t^2 r_{1.2}^2}{\Sigma x_2^4} \qquad (93)$$

$\sigma^2/n\Sigma x_2^2$ は $\beta_1 = t$ を原式に代入して通常の最小二乗法で推定した場合の $b_{0,2}$ の分散にひとしい。もし $r_{1.2}^2=0$ であるならば制約条件の利用によって推定値はなんら改善されない。しかし，制約条件を利用した場合と，利用しない場合とを比べて，推定値にどれほど改善がほどこされるかをみるためには，次の不等式を吟味しなければならない。

$$\frac{\sigma^2}{n(1-r_{12}^2)\Sigma x_2^2} \gtreqless \frac{\sigma^2}{n(1-r_{12}^2)\Sigma x_2^2} - \frac{\sigma^2}{n(1-r_{12}^2)}\frac{r_{12}^2}{\Sigma x_2^2} + \frac{\triangle t^2 r_{1.2}^2}{\Sigma x_2^4} \qquad (94)$$

この大いさは，(94) 式の右辺における第2項と第3項の和が正であるかどうかによってきまる。所与の r_{12}^2 に対し，σ^2 が $\triangle t^2$ に比してきわめて大きいならばこの和は負になるかもしれない。しかし r_{12}^2 が十分に大きいならば第2項の分母は非常に小さくなるため，この和が負になる可能性が強くなる。つぎに $r_{1,2}^2$ が 1，すなわち x_1 と x_2 が完全に相関しているときには b_2 に対して通常の最小二乗推定が不可能であったが，制約条件を利用すれば (93) 式から b_2 の推定が可能となる。それゆえ線型重合の存在する場合には制約条件がパラメーターの推定にとってきわめて便利である。

6. 結　　語

以上，Specification Errors（定式上の誤差）が，求めようとするパラメーターの推定値にいかなる影響を与えるかを種々の角度から検討した。そもそも計量経済学的分析における第1段階の要請は，解明しようとする問題を適確に定式化することで，ある財の需要の分析に当ってその財の価格や競争財の価格，可処分所得などはふつう説明変数として入るべきはずであると仮定されたときにこれを入れなかったり，あるいは誤った情報をモデルのなかに持ちこんだりしては，えられる結果に大きな誤差を生ずることを否めない。この定式化，モデル・ビルデングの段階をハーベルモー（T. Haavelmo）は "creative process" とよんでいるが[3]，まさしくこれは，エコノミストの創意と直観のさえを示す恰好の場所なのである。

参　考　文　献

〔1〕 Theil, H., *Economic Forecasts and policy*, 1958, pp. 331—333.

〔2〕 Stone, J. R., *Measurement of Consumers' Expenditures and Behaviour in the United Kingdom*, Vol. 1. pp. 303—305.

〔3〕 Haavelmo, T., *The Probability Approach in Econometrics*, 1944 p. 10.