制限つき回帰問題について

三枝義清

一、問　題

既に知られているように、線型回帰モデルの場合パラメーターに関する先験的な知識や外部情報を利用することにより、推定の精度を高めることができる。例えば、需要分析（特に時系列データによる時）で説明変数の間に高いmulticollinearityがある場合、その障害をパラメーターに関する知識を活用することにより、ある程度克服できる。Durbin〔2〕（〔　〕内の数字は末尾に掲載した参考文献の番号である。以下同様）の方法に従って横断面分析で得た所得弾性値を外部情報にして時系列データと組み合わすのも一つの試みだし、またFox〔3〕は、*Brookings Quarterly Econometric Model of the United States* において、Brandowのdemand matrixを基礎にした農産物の需要関数の計測を行なっているが、これは各農産物の所得ならびに価格弾性値に関する先験的な知識を積極的に利用した典型的な例とみられる。

パラメーターに関する知識が、等式のみからなる線型の制約式で与えられている場合には、Goldberger〔4〕などによって明らかにされているように、制約式が誤差を含むものであっても制約条件のない通常の回帰分析の方法を多少修正することによって処理できる。しかしパラメーターに関する有用な知識の中には、記号条件や、あるパラメーターが例えば0から1までの間の値をとるといった条件のように等号つきの一次式では表わせないものがある。例えばある生産関数の計測では投入量に関する一次の微係数が正、二次の微係数は負であるといった制約条件を付さねばならぬだろうし、また時間的に不変なマルコフ行列〔p_{ij}〕をマクロ的な時系列データを用いて推定する場合には、$p_{ij} \geqq 0,\ \sum_j p_{ij}=1$ といった条件を満たすようにそれぞれの p_{ij} の推定を行なわねばならない。横断面分析による所得弾性値を利用する場合でも点推定値としてでなく、ある幅をつけて利用する方がより実際的であろう。Judge and Takayama〔5〕が指摘しているように、線型の制約条件が不等号の条件式を含む場合の制限的回帰問題は、しからざる場合に比べてその取り扱いが著しく面倒になる。推定値の性質について制約条件が等号つきの条件式からなる場合に、対比できるような一般的な結果を導くことは難しい。

本稿では、制限つき回帰問題について知られている主なる成果を整理したあとで、ある一つのパラメーターに対して、$\beta_L \leqq \beta_1 \leqq \beta_u$ という制約条件が唯一つ与えられた場合の制限つき最小二乗推定値の性質の検討を行なったものである。

二、制約条件が一次式から成る場合の制限つき回帰推定

標本データ（X、Y）は次のような標準的な線型回帰モデルに従って生成されるものとする。

$$y=X\beta+u \quad \cdots\cdots(2.1)$$

$$E(u)=0,\quad E(uu')=\sigma^2 I \quad \cdots\cdots(2.2)$$

2・1式における変数yはT個の観測値からなる列ベクトル、Xは$T\times K$の行列で、その階数はTを越えぬものとする。Xの要素はすべて確定変数として扱う。変数uはT個の残差項からなる確率ベクトルで、いずれの残差項も期待値がゼロかつ同一の分散σ^2をもち、そしてお互いに無相関であると前提する。βは推定すべきK個のパラメーターからなる列ベクトル$(\beta_1, \beta_2, \cdots\cdots \beta_k)'$で、2・2式の$I$は$T\times T$の単位行列を表わすものとする。ベクトルおよび行列の転置は以下ダッシュ(')の記号で示すことにする。

(1) 通常の最小二乗法(O.L.S)による推定

βを推定するのに与えられた情報が標本データしか利用できない場合には2・3式に示すような最小二乗推定値bによってβを推定するわけだが、この場合のbの性質については多くのことが知られており、例えば期待値と分散—共分散行列は2・4、2・5式のようになる。分散—共分散行列は以下簡単に分散行列と略称し$V(b)$というように表わす。

$$b=(X'X)^{-1}X'y \quad \cdots\cdots(2.3)$$

$$E(b)=\beta \quad \cdots\cdots(2.4)$$

$$V(b)\equiv E(b-\beta)(b-\beta)'=\sigma^2(X'X)^{-1} \quad \cdots\cdots(2.5)$$

(2) 線型の制約条件のもとでの最小二乗推定

利用できる情報として標本データの外に、パラメーターβについて2・6式で表現されるような先験的な知識が

与えられたものとする。

$$R\beta=r \qquad \cdots\cdots(2.6)$$

2・6式におけるRは$J\times K$の行列、rは$J\times 1$の列ベクトルだが、Rもrもその要素はすべて既知とする。Rの階数は$J<K$としておく。2・6式を利用した時の制限つきの最小二乗推定値をb^*とすると、b^*は制約条件2・6のもとで残差平方和の$(y-X\beta)'(y-X\beta)$を最小にするようなβを選ぶことによって得られるものであるが、既に明らかにされているように（例えば Goldberger〔4〕によって）b^*は2・7式のように表わすことができる。なお通常の制限なしの最小二乗法（O.L.S）に対して制限つきの最小二乗法を以下 R.L.S と略記する。

$$b^*=b+(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r-Rb) \qquad \cdots\cdots(2.7)$$

2・7式のbは2・3式で与えられるO.L.Sによる推定値であり、b^*とbとの差は標本から推定されたRbの犯す誤差に依存することになる。b^*の標本分布については

$$E(b^*)=0 \qquad \cdots\cdots(2.8)$$

$$V(b^*)=V-VR'(RVR')^{-1}RV \qquad \cdots\cdots(2.9)$$

ただし　$V\equiv V(b)$

$V(b)$は positive definite な行列であるから、2・9式の右辺の第二項の行列（その階数はJで）は nonnegative definite である。従って$V(b)-V(b^*)$の対角線上の要素は非負となり、2・10式の関係が常に成り立つ。つまり、いずれのパラメーターβ_iについても2・6式の制約条件のもとでの制限つき推定値b_i^*の分散はいかなる場合でもO.L.Sによる推定値b_iの分散よりも大きくなることはない。

$$V(b_i^*) \leqq V(b_i) \qquad (i=1,\ 2,\ \ldots\ldots K) \qquad \cdots\cdots(2.10)$$

しかし残差平方和は当然のことであるが、O.L.Sによる場合に比べて増大する。すなわち

$$SSE(b) \equiv (y-Xb)'(y-Xb)$$

$$SSE(b^*) \equiv (y-Xb^*)'(y-Xb^*)$$

$$SSE(b^*)-SSE(b)=(r-Rb)'[R(X'X)^{-1}R'](r-Rb) \geqq 0$$

(3) 制約条件2・6式が誤差、特に specification error を含む場合

先験的に与えられた条件式2・6が誤差を含む場合を問題にしてみる。そのうちの一つのケースは、正しくは$\bar{r}=R\beta$ であるにも拘らず誤って $r=R\beta$ という制約条件を課して制限つきの最小二乗推定値b^*を求めた場合である。当然の結果として偏りをもった推定値になる。例えば時系列データによる需要分析に横断面分析で求めた所得弾性値がそのまま適用できるとは限らない。系統的な偏りをもつかも知れない。また回帰分析における説明変数の省略は誤って "$\beta_i=0,\ \beta_j=0$" といった制約条件を課すことに外ならない。制約条件が Specification error を含む場合の回帰問題は、既に Theil〔7〕や Toro-Vizcarrondo and Wallace〔9〕によって取り上げられている。

2・7式は制約条件 $r=R\beta$ の真偽の如何を問わずに成立する関係であるが、この制約条件が誤りである時のb^*の期待値は $\bar{r}=R\beta$ を真の条件式とすれば2・11式のようになる。

$$E(b^*)=\beta+(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}\Delta r \qquad \cdots\cdots(2.11)$$

ただし　$\Delta r=r-\bar{r}$

分散行列は(2)の場合と同じく2・9式で表わされるから、誤った制約条件のもとでの推定値b_i^*は偏りをもつが、

O.L.S による b_i に比べて b_i^* の分散が増大することはない。従って平均平方誤差（以下MSEと略記）で比較すると偏りの程度如何によっては制約条件が誤っていても、それを課すことによって b_i の分散よりも小さいMSEをもつ推定値 b_i^* を得ることができる。2・9′、2・11式を用いて $MSE(b^*)$ を求めると、

$$MSE(b^*)=E(b^*-\beta)(b^*-\beta)'=V(b^*)+E(b^*-\beta)E(b^*-\beta)' \quad \cdots\cdots(2.12)$$

$$MSE(b^*)=V-S^{-1}R'[RS^{-1}R']^{-1}\{\sigma^2 I-\Delta r\Delta r'[RS^{-1}R']^{-1}\}RS^{-1} \quad \cdots\cdots(2.13)$$

ただし　$S=(X'X)$, I は $K\times K$ の単位行列

単純な例として $\beta_1=\beta_1^*$ という唯一の条件が課せられた場合を取り上げてみる。これは $r=R\beta$ において

$$R=(1,\ 0,\cdots\cdots 0),\qquad r'=(\beta_1^*,\ 0,\cdots\cdots 0)$$

とおいたものである。従って

$$b_i^*=b_i+\frac{S^{i1}}{S^{11}}(\beta_1^*-b_1)\qquad (i=1,\ 2,\cdots\cdots K) \quad \cdots\cdots(2.14)$$

$$E(b_i^*)=\beta_i+\frac{S^{i1}}{S^{11}}(\beta_1^*-\beta_1)\qquad (i=1,\ 2,\cdots\cdots K) \quad \cdots\cdots(2.15)$$

ただし　$S^{-1}=(X'X)^{-1}=(S^{ij})$

一方2・13式の右辺における第二項の行列の i 行、j 列目の要素は2・16式のようになる。

$$\sigma^2\left\{1-\frac{(\beta_1^*-\beta_1)^2}{\sigma^2 S^{11}}\right\}\frac{S^{i1}S^{1j}}{S^{11}} \quad \cdots\cdots(2.16)$$

$V(b_i)=\sigma^2 S^{ii}$ であるから $MSE(b_i^*)$ は次のように書ける。

$$MSE(b_i^*)=V(b_i)\left[1-\frac{(S^{i1})^2}{S^{11}S^{ii}}\left\{1-\frac{(\beta_1^*-\beta_1)^2}{V(b_1)}\right\}\right] \cdots\cdots (2.17)$$

$(S^{i1})/\sqrt{S^{11}S^{ii}}$ はb_1とb_iの間の相関係数に外ならないから２・18式の関係を導出できる。

$$\frac{(\beta_1^*-\beta_1)^2}{V(b_1)}\leqq 1 \text{ であれば，その時に限り } MSE(b_i^*)\leqq V(b_i) \quad (i=1,\ 2,\cdots\cdots K) \cdots\cdots (2.18)$$

つまり $\beta_1^*=\beta_1$ が誤った情報であってもその誤差が b_1 の標準偏差を越えなければ，MSE 基準で考えて O. L. S による推定値よりも “better” な推定値を得ることができるわけである。

一般の２・13式については Toro-Vizcarrondo and Wallace 〔9〕によって，次の事柄が知られている。２・13式を変形して

$$V-MSE(b^*)=\sigma^2 S^{-1}R'\left(R^{-}S^{1}R'\right)^{-1}\left(RS^{-1}R'-\frac{\varDelta r\varDelta r'}{\sigma^2}\right)(RS^{-1}R')^{-1}RS^{-1} \cdots\cdots (2\ 19)$$

２・19式の右辺における行列 $\left(RS^{-1}R'-\frac{\varDelta r\varDelta r'}{\sigma^2}\right)$ が positive semi-definite であれば，その時に限り２・19式の左辺は positive semi-definite であるから，その場合には $V(b_i)\geqq MSE(b_i^*)$ $(i=1,\ 2,\cdots\cdots K)$ が成立する。$RS^{-1}R'-\frac{\varDelta r\varDelta r'}{\sigma^2}$ が positive semi-definite であるための必要かつ充分条件は Toro-Vizcarrondo and Wallace 〔9〕によれば２・20式で与えられる。

$$(R\beta-r)'(RS^{-1}R')^{-1}(R\beta-r)\leqq\sigma^2 \cdots\cdots (2.20)$$

$V-MSE(b^*)$ が positive semi-definite であることは如何なるベクトル $d\neq 0$ に対しても２・21式が成立することと同等である。

$$V(d'b) \geqq MSE(d'b^*) \quad \cdots\cdots \quad (2.21)$$

2・21式が成立する時 b^* は b よりも "better" な推定であると呼ぶことにすれば、2・20式は b^* が O.L.S による b よりも better な推定であるための必充条件になる。回帰分析においてある説明変数 x_i, x_j…… を除去するか否かの判定を MSE 基準に従って行なうものとすれば、この判定は $\beta_i=0$, $\beta_j=0$,…… といった制約条件のもとでの b^* が O.L.S による b に比べて better であるか否かを判定することと同じであるから、2・20式に従ってテストをすればよいことになる。

(4)　制約条件 2・6式が random error を含む場合

パラメーターに関する外部情報が標本データ（y、X）以外の別個の標本によって統計的な情報として与えられる場合であるが、Durbin〔2〕の扱った例が典型的である。

先ず簡単な場合について。例えばK個のパラメーターのうち、$\beta_1, \beta_2, \cdots \beta_G$ については 2・23、2・24式で示すような統計的情報が与えられていたものとしよう。これらのG個の パラメーターの組をベクトル $\beta_{1\cdot}$、残りの$K-G$個のパラメーターの組をベクトル $\beta_{2\cdot}$ で表わすことにして線型回帰モデル2・1式を2・22式のように表わしておく、従って X_1 は $T\times G$, X_2 は $T\times(K-G)$ の行列である。仮定2・24式によって外部情報として与えられる推定値$b_{1\cdot}^*$ は当然 O.L.S による $\beta_{1\cdot}$, $\beta_{2\cdot}$ の推定値である $b_{1\cdot}$ や $b_{2\cdot}$ とは無相関である。

$$y=X_1\beta_{1\cdot}+X_2\beta_{2\cdot}+u \quad \cdots\cdots \quad (2.22)$$

ただし　$\beta_{1\cdot}=(\beta_1, \beta_2, \cdots\cdots \beta_G)'$

$\beta_{2\cdot}=(\beta_G, \cdots\cdots \beta_K)$

$$b_{1\cdot}{}^* = \beta_{1\cdot} + v \quad \cdots\cdots \quad (2.23)$$

$$E(v) = 0,\ E(vu') = 0,\ E(vv') = \Psi \quad \cdots\cdots \quad (2.24)$$

$\beta_{1\cdot} = b_{1\cdot}{}^*$ という条件をおいた時の β の R.L S による推定値を $b^* = (b_1{}^*,\ b_2{}^*)$ とすれば、2・7式によって

$$b^* = b + S^{-1}R'(RS^{-1}R')^{-1}(b_{1\cdot}{}^* - b_{1\cdot}) \quad \cdots\cdots \quad (2.25)$$

ただし　$R = (I,\ O)$　ここで I は $G \times G$ の単位行列

従って $\beta_{2\cdot}$ の推定値 $b_{2\cdot}{}^*$ は

$$b_{2\cdot}{}^* = b_{2\cdot} - (X_2'X_2)^{-1}X_2'X_1(b_{1\cdot}{}^* - b_{1\cdot}) \quad \cdots\cdots \quad (2.26)$$

$b_{1\cdot}{}^*$ は $\beta_{1\cdot}$ の不偏推定値であるから $E(b_{2\cdot}{}^*) = \beta_2$　2・9式を利用して2・27式が得られる。この式は2・19式で $R = [I,\ O]$, $\Delta r \Delta r'$ の代わりに $\Psi = E(b_{1\cdot}{}^* - \beta_{1\cdot})(b_{1\cdot}{}^* - \beta_{1\cdot})'$ とおいたものに均しい。

$$V - V(b^*) = B\{V(b_{1\cdot}) - \Psi\}B' \quad \cdots\cdots \quad (2.27)$$

ただし　$B = \begin{bmatrix} I \\ -(X_2'X_2)^{-1}X_2'X_1 \end{bmatrix}$

従って

$$V(b_{2\cdot}{}^*) = V(b_{2\cdot}) - (X_2'X_2)^{-1}X_2'X_1\{V(b_{1\cdot}) - \Psi\}X_1'X_2(X_2'X_2)^{-1} \quad \cdots\cdots \quad (2.28)$$

2・27式より知れるように b^* が b よりも better であるためには、$V(b_{1\cdot}) - \Psi$ が positive semi-definite になっていなければならない。特に $\beta_{1\cdot}$ が単一のパラメーター β_1 から成る場合には、$V(b_i{}^*) \leqq V(b_i)$ $(i = 1,\ 2, \cdots K)$ となるためには $V(b_1{}^*) \leqq V(b_1)$ でなければならない。

推定式2・25式の中で $\beta_1.$ に関する推定値をみると、$b_1.^*$ という外部情報だけに頼って肝心の標本データ(y、X)を利用していない。標本データより得られる $b_1.$ を追加することにより、$b_1.^*$ よりも更に better な推定値を導くことができるはずである。Theil〔8〕はより一般的条件の下で外部情報の効果的利用を論じている。

β に関する補完的な外部情報(r、R)が次のような統計的モデルに従うものと前提する。

$$r=R\beta+v \qquad (2.29)$$

$$E(v)=0,\ E(vu')=0,\ E(vv')=\Psi \qquad (2.30)$$

2・2式の σ^2 および2・30式の Ψ が既知であったとすれば、β の線型(y、r に関して)不偏推定式の中で最良なもの(BLUE)は2・31式の b^{**} で与えられ、その分散は2・32式で表わされることが知られている。

$$b^{**}=\left(\frac{1}{\sigma^2}X'X+R'\Psi^{-1}R\right)^{-1}\left(\frac{1}{\sigma^2}X'y+R'\Psi^{-1}r\right) \qquad (2.31)$$

$$V(b^{**})=\left(\frac{1}{\sigma^2}X'X+R'\Psi^{-1}R\right)^{-1} \qquad (2.32)$$

外部情報(r、R)だけを用いて行なう β の BLUE を $\tilde{b}$ とすれば、その分散行列は2・34式で与えられるから

$$\tilde{b}=(R'\Psi R)^{-1}R'\Psi^{-1}r \qquad (2.33)$$

$$V(\tilde{b})=(R'\Psi^{-1}R)^{-1} \qquad (2.34)$$

b^{**} は次のように O.L.S による b と上述の $\tilde{b}$ の加重平均として表現することができる。

$$b^{**}=(W+\tilde{W})^{-1}(Wb+\tilde{W}\tilde{b}) \qquad (2.35)$$

ただし　$W=V(b),\ \tilde{W}=V(\tilde{b})$

Theil〔8〕によれば近似的ではあるが、2・36式に示すように b^{**} を2・7式や2・25式と比較可能な形で表わすことができる。

$$b^{**} \approx b + \frac{(X'X)}{\sigma^2} R'\Psi^{-1}(r - Rb) \quad \cdots\cdots (2.36)$$

簡単な例として外部情報（r、R）が前に述べた2・23、2・24式のモデルに従う場合を考えてみる。r、R は2・37式で与えられるから、この場合には $b^{**}-b$ は正確に（$r-Rb$）の一次結合になる。

$$R=(I,\ O) \quad r=(b_{1\cdot}^{*},\ 0)' \quad \cdots\cdots (2.37)$$

2・31式によって $\beta=(\beta_{1\cdot},\ \beta_{2\cdot})'$ の BLUE は次のようになる。ここでの X_1 および X_2 は2・22式における行列と同じものである。

$$b_{1\cdot}^{**}=b_{1\cdot}+(I+\Phi)^{-1}\Phi(b_{1\cdot}^{*}-b_{1\cdot}) \quad \cdots\cdots (2.38)$$

$$b_{2\cdot}^{**}=b_{2\cdot}-(X_2'X_2)^{-1}(X_2'X_1)(I+\Phi)^{-1}\Phi(b_{1\cdot}^{*}-b_{1\cdot})$$

ただし　$\Phi = V(b_{1\cdot})\Phi^{-1}$

分散を2・27式と比較可能な形で表わすと

$$V-V(b^{**})=B\{V(b_{1\cdot})-V(b_{1\cdot}^{**})\}B' \quad \cdots\cdots (2.39)$$

ただし　B は (2.27) の時と同じ

$$V(b_{1\cdot})-V(b_{1\cdot}^{**})=V(b_{1\cdot})\{\Phi+V(b_{1\cdot})\}^{-1}V(b_{1\cdot})$$

従って $b_{2\cdot}^{**}$ の分散は2・40式のようになる。当然のことであるが常に $V(b_i^{**}) \leqq V(b_i)$　$(i=1, 2, \cdots\cdots K)$.

$$V(b_{2\cdot})-V(b_{2\cdot}^{**})=(X_2'X_2)^{-1}X_2'X_1\{V(b_{1\cdot})-V(b_{1\cdot}^{**})\}X_1'X_2(X_2'X_2)^{-1} \quad \cdots (2.40)$$

特に $\beta_{1\cdot}$ が単一のパラメーター β_1 から成り、$E(v^2)=\sigma_v^2$ とすれば

$$b_1^*=\beta+v,\ E(v)=0,\ E(v^2)=\sigma_v^2$$

$$b_1^{**}=b_1+\frac{1}{1+\phi}(b_1^*-b_1) \qquad \cdots\cdots(2.41)$$

$$b_{2\cdot}^{**}=b_{2\cdot}-\frac{1}{1+\phi}(b_1^*-b_1)(X_2'X_2)^{-1}(X_2'X_1)$$

ただし　$\phi=\dfrac{\sigma_{v^2}}{V(b_1)}$

当然のことであるが σ_v^2 が減少するに従い b^{**} は制約条件が正確であった場合の R.L.S による推定値 b^* に近づく。一方 σ_v^2 が増大すれば O.L.S による推定値 b に近づいてゆく。各推定値 b_i^{**} の分散は2・17式に類似した式で与えられる。

$$\frac{V(b_i^{**})}{V(b_i)}=1-\frac{1}{1+\phi}\rho_{i1}^2 \qquad \cdots\cdots(2.42)$$

ただし　ρ_{i1} は O.L.S による b_1 と b_i の間の相関係数

三、二次計画法による制限回帰問題について

線型回帰モデル2・1式のあるパラメーター、例えば β_1 について3・1式のような型の先験的な情報が与えられた場合

$$\beta_L \leqq \beta_1 \leqq \beta_u \qquad \cdots\cdots(3.1)$$

ただし　β_L, β_u は既知とする

Theil〔8〕は3・1式を前節の(4)で述べた手続きに従って利用する方法を提唱している。例えば β_1 が0と1の間にあることが殆ど確実に知られていたとする。β_1 の推定値として0.5を採用した時の誤差を v とすれば v は殆ど確実に -0.5 から0.5までの値をとる。$E(v)=0$, かつ v が ± 2 シグマ の範囲外の値をとる確率が殆どゼロと仮定できるならば、上記の β_1 に関する知識は次のような統計的情報として定式化できよう。

$$0.5=\beta_1+v \quad \cdots\cdots (3.2)$$

ただし　$E(v)=0$, $E(v^2)=1/16$

3・2式を(4)の制約条件2・29と同様に扱うことによって β_1 についての制限つき推定値をうることができる。しかし(4)で述べた BLUE を得るには誤差項 v の期待値および分散についての妥当な知識を必要とする。これらの知識を欠く場合には、以下述べるように(3)の場合に準じた方法で処理しなければならないだろう。

線型回帰モデルに対して次のような型の線型の制約条件が追加されたものとしよう。ただし r、R は既知とする。

$$R\beta\leq r \quad \cdots\cdots (3.3)$$

2・6式が等号つきの条件式だけから成り立っているのに対して、3・3式は2・6式が一般化されて不等号つきの条件式も含む体系になっている。制約条件3・3式のもとで残差平方和の SSE を最小にするような制限つき最小二乗推定値 $\hat{\beta}$ で β を推定する問題を以下検討してみる。

$$SSE=(y-X\beta)'(y-X\beta)\equiv Q(\beta)$$
$$=y'y-2\beta'X'y+\beta'X'X\beta \quad \cdots\cdots (3.4)$$

通常の回帰問題では3・4式の行列 $X'X$ は positive definite か少なくとも positive semi-definite であるので、βを求める問題は制約条件 3・3のもとで目的関数の $-Q(\beta)$ を最大にするという典型的な二次計画問題として処理することができる。以下の議論では $X'X$ は positive definite としておく。$Q(\beta)$ はβについての凸関数であるから、$\beta=b$ のところで唯一の最小値をもつ。O.L.S による推定値bを本節では b^{o} と記すことにする。

もし $Rb^{o} \leqq r$ であれば b^{o} は当然求める最小解 $^{*}\beta$ になる。b^{o} が3・3式を満足せぬ場合には二次計画法で使われている計算手続きのいずれかに従ってβを求めればよいわけだが、当面の問題に対しては通常の条件つき極小化の手続きを反覆するところの Theil-Van de Panne method〔1〕によって検討するのが便利であろう。

3・3式の制約条件を次のようなJ個の条件式で表わし、個々の条件式を順次一番目、二番目……の条件式（不等号つきの）と呼ぶことにする。

$$R_1\beta\leqq r_1,\ R_2\beta\leqq r_2\cdots\cdots R_J\beta\leqq r_J \quad\cdots\cdots(3.5)$$

3・4式に属する条件式の部分集合を簡単に3・6式で表わそう。例えばj番目……l番目の条件式より成る集合の場合には $C=\{i,\ j,\cdots\cdots l\}$ で $R_c'=(R_i,\ R_j,\cdots\cdots R_l')$ となる。

$$R_c\beta\leqq r_c \quad\cdots\cdots(3.6)$$

ただし　R_c は $G\times K$ の行列，$0\leqq G\leqq K$

Theil-Van de Panne method によれば、最適解 $^{*}\beta$ は $R_c\beta=r_c$ というある等号つきの制約条件のもとで $Q(\beta)$ を最小にすることにより到達できる。

制約条件 $R_c\beta=r_c$ のもとでの $Q(\beta)$ の最小解は前節の(3)で述べた R.L.S による推定値の$^{*}b$に外ならないが、3・

5式のうちでどの制約条件が有効であったかを明記するために$b_{\hat{c}}$と表わすことにする。$b_{\hat{c}}$が求める最小解になるように部分集合$\hat{c}$をいかに見出だすかが Theil-Van de Panne method の骨子になっており、次のような手続きで$\hat{c}$を求めてゆく。b_ϕが3・3式を満足せぬ場合にはb_ϕが満足しなかった条件式のうち少なくとも一つの条件式が$\hat{c}$の中に含まれていなければならない(Boot〔1〕を参照)。b_ϕが満たさなかった条件式を一つずつ等号つきの条件式にしてR.L.Sによるβの推定値を求める。例えば$R_i b_\phi > r_i$であればb^iを求める。もしいずれのb^iも3・5式を満足せぬとすればそれぞれのb^iについて次の計算を続ける。b^iが満たさなかった条件式の中から一つずつ条件を取り出し(例えばh番目のものであれば)、$R_i\beta=r_i$, $R_h\beta=r_h$という制約条件のもとでR.L.Sによってb^{ih}を求める。このような手続きでb^cを生成してゆくわけだが、$c=\hat{c}$であるか否かの判定はb^cが制約条件3・3式を満足しかつ等号つきの制約条件$R_c\beta=r_c$に応ずるラグランジュ乗数ベクトルλが負となる、という二つの基準によって行なえばよい。

いずれの部分集合c(空集合ϕも含めて)が選ばれるにせよ、最適解$\hat{\beta}$は$R_c\beta=r_c$という条件のもとでの$Q(\beta)$の最小解であるから、前節の2・7式を用いて$\hat{\beta}$を次のように表現することができる。

$$\hat{\beta}=b+(X'X)R_c'[R_c(X'X)^{-1}R_c']^{-1}(r_c-R_c b) \quad \cdots\cdots(3.7)$$

いずれのR_cが最適解の中に現われるかは標本データ(y、X)の結果如何によるから、R_cの各要素は確率変数yに依存して変動する変数になる。

このために$\hat{\beta}$の期待値や分散行列の評価は以下述べるような単純な場合を除き一般には難しい。

単純な場合についての$\hat{\beta}$の標本分布について

ある一つのパラメーターβ_1について、3・8式のような制約条件が課せられている場合を検討してみよう。

$$\beta_L \leqq \beta_1 \leqq \beta_u \quad \cdots\cdots\cdots(3.8)$$

3・8式は $-\beta_1 \leqq -\beta_L,\ \beta_1 \leqq \beta_u$ であるから、一般式3・3で

$$R=\begin{pmatrix} -1, 0\cdots\cdots 0 \\ 1, 0\cdots\cdots 0 \end{pmatrix} \qquad r'=(-\beta_L,\ \beta_u)$$

とおいたものである。もし O.L.S で求めた $b\equiv b\phi$ が $b\phi<\beta_L$ であれば、求める最小解 $\hat{\beta}$ は $\hat{\beta}=b^1$ で与えられる。つまり $\beta_1=\beta_L$ という制約条件のもとでの R.L.S による b^* で与えられる。何故なら、b^1 は3・8式を満足し、かつ等号つきの条件式 $\beta_1=\beta_L$ に応ずるラグランジュ乗数 λ^1 を求めると

$$\lambda_1=\frac{b\phi-\beta_L}{S^{11}}<0 \quad \cdots\cdots\cdots(3.9)$$

ただし　S^{11} は $(X'X)^{-1}$ の1行列目の要素

$(X'X)$ は positive definite だから $S^{11}>0$

なお $\hat{\beta}=b^1$ とした時の SSE の増分は

$$Q(b^1)-Q(b\phi)=\frac{(\beta_L-b\phi)^2}{S^{11}} \quad \cdots\cdots\cdots(3.10)$$

$b\phi<\beta_L$ の時の最小解が $\hat{\beta}=\beta_L$ で与えられることは3・10式からも確かめることができる。一方 $b\phi>\beta_u$ であった場合には同じような議論で $\hat{\beta}=b^2$ で与えられることがいえる。

先ず最初に β_1 の推定値 $\hat{\beta}_1$ の標本分布を検討してみる。上述の結果をまとめると

$$\left.\begin{array}{lll} \beta_L\leqq b_1\leqq\beta_u & \text{であれば} & \hat{\beta}_1=b_1 \\ b_1<\beta_L & 〃 & \hat{\beta}_1=\beta_L \\ b_1>\beta_u & 〃 & \hat{\beta}_1=\beta_u \end{array}\right\} \quad \cdots\cdots\cdots(3.11)$$

ただし　b_1 は O.L.S による β_1 の推定値

ここで誤差項の u は規準型正規分布に従うという条件を線型回帰モデル2・1、2・2式に追加することにする。この場合には b は期待値 β、$\sigma^2(X'X)^{-1}$ なる分散行列をもつ K 変量の正規分布に従うから、b_1 が $b_1<\beta_L$ あるいは $b_1>\beta_u$ となる確率を次のように表わすことができる。

$$P\{b_1<\beta_L\}=H(Z_L)$$

$$P\{b_1>\beta_u\}=1-H(Z_u)=H(-Z_u)$$

ただし　$Z_L=\dfrac{\beta_L-\beta_1}{\sqrt{V(b_1)}},\ \ Z_u=\dfrac{\beta_u-\beta_1}{\sqrt{V(b_1)}},\ \ h(t)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}},\ \ H(Z)=\displaystyle\int_{-\infty}^{Z}h(t)\,dt$

3・11式より知れるように $\hat{\beta}_1$ の変域は境界値を含めて β_L までである。$\hat{\beta}_1$ が β_L, β_u となる確率はそれぞれ $H(Z_L)$, $H(Z_u)$ であるが、β_L から β_u までの区間についての確率分布は連続型であって、β_L から β_u までの範囲の切断正規分布 (truncated normal distribution) になる。この切断分布の一次と二次のモーメント（原点の周りの）は3・12、3・13式で与えられる。

$$一次のモーメント=\beta_1+\frac{\sqrt{V(b_1)}\{h(Z_L)-h(Z_u)\}}{H(Z_u)-H(Z_L)}\cdots\cdots\cdots\cdots(3.12)$$

$$二次のモーメント=\beta_1^2+V(b_1)+\frac{\sqrt{V(b_1)}\{(\beta_L+\beta_1)h(Z_L)-(\beta_u+\beta_1)h(Z_u)\}}{H(Z_u)-H(Z_L)}\cdots\cdots(3.13)$$

従って $\hat{\beta}_1$ の期待値は

$$E(\hat{\beta}_1)=\beta_L H(Z_L)-\beta_u H(-Z_u)+\beta_1\{H(Z_u)-H(Z_L)\}+\sqrt{V(b_1)}\{h(Z_L)-h(Z_u)\} \quad \cdots(3.14)$$

$\hat{\beta}_1$ は $\beta_1-\beta_L=\beta_u-\beta_1$ の場合を除き一般に偏りをもち、その程度は次式のように表わせる。

$$\frac{E(\hat{\beta}_1-\beta_1)}{V(b_1)}=Z_L H(Z_L)+Z_u H(-Z_u)+h(Z_L)-h(Z_u) \quad \cdots\cdots(3.15)$$

一方 $\hat{\beta}_1$ の平均平方誤差を求めると

$$\frac{E(\hat{\beta}_1-\beta_1)^2}{V(b_1)}=1+(Z_L^2-1)H(Z_L)+(Z_u^2-1)H(-Z_u)+Z_L h(Z_L)-Z_u h(Z_u) \quad \cdots\cdots(3.16)$$

3・16式の右辺に現われる $H(Z)$ と $h(Z)$ の間には3・17式で示す関係があるから

$$(Z^2-1)H(Z)+Zh(Z)=Z\int_{-\infty}^{Z}(Z-t)h(t)dt \quad \cdots\cdots(3.17)$$

$\beta_L\leqq\beta_1\leqq\beta_u$ という制約条件が正しければ次の3・18式が成立する。

$$\begin{aligned}&(Z_L^2-1)H(Z_L)+Z_L h(Z_L)\leqq 0\\&(Z_u^2-1)H(-Z_u)-Z_u h(Z_u)\leqq 0\end{aligned} \quad \cdots\cdots(3.18)$$

従って $\beta_L\leqq\beta_1\leqq\beta_u$ が正しければ常に $E(\hat{\beta}_1-\beta_1)^2\leqq V(b_1)$ となることが判かる。境界値の β_L あるいは β_u が β_1 に近づくにつれて $MSE(\hat{\beta}_1)$ は次第に減少する。3・16式による $MSE(\hat{\beta}_1)$ の評価には Zellner〔10〕の Table II が利用できる。いくつかの組み合わせ (Z_L, Z_u) について $MSE(\hat{\beta}_1)/V(b_1)$ の値を掲げたものが表1である。

表1　$E(\hat{\beta}_1-\beta_1)^2/V(b_1)$

Z_u \ Z_L	−3.0	−2.0	−1.0	−0.5	0	0.5	1.0
3.0	.9950	.9578	.7556	.5902	.4976	.6550	1.2396
2.0	.9578	.9204	.7182	.5528	.4602	.6176	1 2022
1 0	.7556	.7182	.5160	.2508	.4154	.4156	1 0000
0.5	.5902	.5528	.3506	.1852	.0926	.2500	0.4154

$\beta_L \leqq \beta_1 \leqq \beta_u$ という先験的情報が確実であり、かつ $V(b_1)$ に比べて β_1 の含まれる区間の幅が充分狭ければ、O.L.S による推定値よりも可成り better な (MSE 基準で考えて) 推定値を得ることが出来るわけである。不幸にして β_1 が指定した区間外にある場合には、$V(b_1)$ に比して $MSE(\hat{\beta})$ は増大しうるわけで、例えば $Z_L=1.0$ の場合を考えてみると表1の最右列のような結果になる。

$y=\beta_0+\beta_1x+u$, $\beta_L \leqq \beta_1 \leqq \beta_u$ という説明変数が一個よりなる線型回帰モデルについて、本節の初めで紹介した Theil の方法との比較を行なってみる。制約条件 $\beta_L \leqq \beta_1 \leqq \beta_u$ を3・19 式で規定される統計的情報に定式化できたとしよう。3・19 式では特に誤差項 v は $-\frac{(\beta_u-\beta_L)}{2}$ から $\frac{(\beta_u-\beta_L)}{2}$ までの範囲の一様分布に従うものと仮定してある。

$$r=\beta_1+v \qquad \cdots\cdots (3.19)$$

ただし $r=\frac{\beta_L+\beta_u}{2}$, $E(v^2)=\sigma_v^2=\frac{(\beta_u-\beta_L)}{2}$

2・41 式による β_1 の推定値 b_1^{**} の分散は次の如くなる。

$$\frac{V(b_1^{**})}{V(b_1)}=\frac{\phi}{1+\phi} \qquad \cdots\cdots (3.20)$$

ただし $\phi=\frac{V(b_1)}{V(b_1)}=\frac{(\beta_u-\beta_L)^2}{12V(b_1)}$

一方 $MSE(\hat{\beta}_1)$ を特に $\beta_1-\beta_u=\beta_L-\beta_u$ の場合について評価してみると、3・16式によ

って

$$\frac{MSE(\hat{\beta}_1)}{V(b_1)}=1+2(3\phi-1)H(-\sqrt{3\phi})-2\sqrt{3\phi}\,h(-\sqrt{3\phi}) \quad \cdots\cdots(3.21)$$

$\phi=0.5,\ 1.0,\ 2.0$ の場合について、$V(b_1^{**})$ と $MSE(\hat{\beta}_1)$ の比較を行なうと表2のような結果になる。

2・41式の b_1^{**} では σ^2 を既知としてあるので、実際には σ^2 の代わりに $S^2=\dfrac{SSE(b_1)}{T-2}$ を用いねばならない。σ^2 を S^2 で代用した場合には $V(b_1^{**})$ は σ^2 が既知であった場合に比べて増大する。小標本の場合の正確な分散式は Swamy and Mehta〔6〕によって導出されているが、表2の括孤内の数値は $T=4$ の場合の分散を掲げたものである。3・19式による定式化が妥当なものであれば、b_1^{**}に比べ3・11式による推定値 $\hat{\beta}_1$ の効率は可成り劣ることが判かる。

表 2

	0.5	1.0	2.0
$MSE(\hat{\beta}_1)/V(b_1)$	0.649	0.858	0.982
$V(b_1^{**})/V(b_1)$	⅓	0.5	⅔
	(0.386)	(0.562	(0.730)

もとに戻って β_1 以外のパラメーター β_i の推定値 $\hat{\beta}_i$ $(i=2,\ 3,\cdots\cdots K)$ の性質を検討してみる。$b_1<\beta_L$ となった場合の $\hat{\beta}_i$ は $\beta_1=\beta_L$ という制約条件のもとでの R.L.S による推定値 b_i^* に外ならないから、2・14式によって次のように表わせる。

$$b_1<\beta_L \quad \text{であれば} \quad \hat{\beta}_i=b_i+\frac{S^{i1}}{S^{11}}(\beta_L-b_1) \cdots\cdots\cdots\cdots(3.22)$$

同じようにして

$$b_1>\beta_u \quad \text{であれば} \quad \hat{\beta}_i=b_i+\frac{S^{i1}}{S^{11}}(\beta_u-b_1)$$

$$\beta_L\leqq b_1\leqq\beta_u \quad \text{であれば} \quad \hat{\beta}_i=b_i$$

ただし　$(X'X)^{-1}=S^{ij}$

$\hat{\beta}_i$ の条件つき期待値を求めるとそれぞれ

$$E(\hat{\beta}_i|b_1<\beta_L)=\beta_i+(\beta_L-\beta_1)\frac{S^{i1}}{S^{11}}$$

$$B(\hat{\beta}_i|b_1>\beta_u)=\beta_i+(\beta_u-\beta_1)\frac{S^{i1}}{S^{11}} \quad \cdots\cdots\cdots(3.23)$$

$$E(\hat{\beta}_i|\beta_L\leqq b_1\leqq\beta_u)=\beta_i+\frac{\sqrt{V(b_1)}\{h(Z_L)-h(Z_u)\}}{H(Z_u)-H(Z_L)}\cdot\frac{S^{i1}}{S^{11}}$$

従って

$$E(\hat{\beta}_i)=\beta_i+\rho_{i1}\sqrt{V(b_1)}\{Z_LH(Z_L)+Z_uH(-Z_u)+h(Z_L)-h(Z_u)\} \quad \cdots\cdots(3.24)$$

ただし，$\rho_{i1}=\dfrac{S^{i1}}{\sqrt{S^{ii}S^{11}}}=b_i$ と b_1 の間の相関係数

あるいは

$$\frac{E(\hat{\beta}_i-\beta_i)}{\sqrt{V(b_i)}}=\rho_{i1}\{Z_LH(Z_L)+Z_uH(-Z_u)+h(Z_L)-h(Z_u)\} \quad \cdots\cdots(3.25)$$

一方 $\hat{\beta}_i$ の条件つき分散はそれぞれ

$$V(\hat{\beta}_i|b_1<\beta_L)=V(\beta_i|b_1>\beta_u)=V(b_i)(1-\rho_{i1}^2)$$

$$V(\hat{\beta}_i|\beta_L\leqq b_1\leqq\beta_u)=V(b_i)\{1-\rho_{i1}^2(1-\theta)\} \quad \cdots\cdots(3.26)$$

ただし　$\theta=\dfrac{V(\beta_1|\beta_L\leqq b_1\leqq\beta_u)}{V(b_1)}$

以上の結果を用いて $\hat{\beta}_i$ の MSE を求めると、3・27式のように簡単な形にまとめることができる。

$$\frac{E(\hat{\beta}_i-\beta_i)^2}{V(b_i)}=1-\rho_{i1}{}^2\frac{\{V(b_1)-E(\hat{\beta}_1-\beta_1)^2\}}{V(b_1)} \qquad (i=1,\ 2,\cdots\cdots K)\cdots\cdots\cdots\cdots(3.27)$$

右辺の $E(\hat{\beta}_1-\beta_1)^2/V(b_1)$ は3・16式で支えられる。3・25、3・27式に見られる通り β_1 についての先験的情報3・8式を利用した場合には、ρ_{i1} が高いパラメーターの推定ほどその恩恵を受ける度合が大くなる。この関係は2・17、2・42式より知れるように b_i*, b_i** についても同じことである。

参 考 文 献

〔1〕 Boot, J. C. G., *Quadratic Programming*, North-Holland, 1964.

〔2〕 Durbin, J., "A note on regression when there is extraneous information about one of the coefficients", *Journal of the American Statistical Association*, Dec. 1953.

〔3〕 Fox, K. A., "A submodel of the agricultural sector", chapter 5 in *The Brookings Quarterly Econometric Model of the United States*, Edited by T. S. Duesenberry, G. Fromm, et al. Rand Mcnally & Company, 1965.

〔4〕 Goldberger, A. S., *Econometric Theory*, John wiley & sons, 1964.

〔5〕 Judge, G. G. and Takayama, T., "Inequality restrictions in regression analysis", *Journal of the American Statistical Association*, March 1966.

〔6〕 Swamy, P. A. V. B and Mehta, T. S., "On the Theil's mixed regression estimator", *Journal of the American Statistical Association*, March 1969.

〔7〕 Theil, H., *Economic Forecasts and Policy*, North-Holland, 1958.

〔8〕 Theil, H., and Goldberger, A S., "On pure and mixed statistical estimation in economics", *International Economic Review*, Janu 1961.

〔9〕 Toro-Vizcarrondo, C. and Wallace, T D., "Test of the mean square error criterion for restriction in linear regression", *Journal of the American Statistical Association*, June 1968.
〔10〕 Zellner, A., "Decision rules for economic forecasting", *Econometrica*, January-April 1963.

（研究員）