刈屋武昭

『回帰分析の理論』

(一橋大学経済研究叢書 31)

岩波書店 1979. 3 ix+273 ページ

わが国の計量経済学界が，誇りとするに足ることの一つは，本来経済学部出身で，かつ数理統計学の最新の理論を身につけた研究者が，数は多くないが水準の高い研究を行っていることである。この本の著者はそのような研究者の中でも最も若い世代に属し，またこの本はそのような研究の最新の成果を代表するものとして，注目すべき力作である。

この本は回帰分析についての教科書ではなく，また高度の概説書といったものでもなく，むしろ回帰分析の理論をめぐって，著者自身の研究成果を中心としていくつかの問題について論じた理論的研究書である。従って決して読みやすくはないが，研究者にとっては興味深い多くの結果と，刺激的な問題をふくんだ有益な書物であるということができる。

内容は2部10章からなり，1部1〜6章は1変量回帰モデル，2部7〜10章は多変量回帰モデルを扱っている。

第1章「回帰係数の推定(I)」では，主として多次元正規分布の平均ベクトルの推定における James-Stein 推定量の理論の，正規線形回帰モデルへの拡張を論じている。James-Stein 推定量については，最近になってその応用上の有用性についても注目されるようになっており，回帰分析への応用も一つの重要な話題である。この本では Berger, Efron-Morris 等による新しい結果をとり入れて，理論的な側面がくわしく論ぜられている。しかし応用上の観点からすれば，問題を単に抽象的な「2次形式損失のもとでの回帰係数の推定」として扱うのではなく，むしろ与えられた説明変数の値のいくつかの組に対応する y の値の同時予測として考えた方が，その意味が理解しやすいのではないかと思われるが，どうであろうか。

第2章「回帰係数の推定(II)」は，第1章の結果の若干の拡張，およびいわゆる Ridge 推定量を論じている。

第3章「検定仮説の識別可能性」では，まず標題のような概念を定義した後，その誤差分散行列に関する仮説への応用を論じ，そうして OLSE と GLSE(一般最小2乗推定量)との選択の問題を考察している。ここでは OLSE の効率が小さくなるような誤差分散行列 $\boldsymbol{\Sigma}$ についての仮説が，一般に検出できないという重要な指摘がなされている。著者はここで相対効率があまり落ちないという意味で OLSE を用いてもよいような $\boldsymbol{\Sigma}$ の範囲について論じている。この章の問題は，実際的にも重要であるが，次のような形に拡張することができると思う。すなわち $\boldsymbol{\Sigma}$ について，何個かの母数をふくんだ仮説

$$\boldsymbol{\Sigma}=\boldsymbol{\Sigma}(\theta,\cdots,\theta_k)$$

を考え，標本から $\theta,\cdots,\theta_k$ を推定した上で，推定された $\boldsymbol{\Sigma}$ を用いて GLSE を適用する場合，推定の効率は($\boldsymbol{\Sigma}$ が正確には上記の仮定にあてはまらない場合もふくめて)どうなるか，またその OLSE との優劣はどうかである。かりに OLSE の効率が小さい場合に GLSE を適用しようとすれば，何らかの形で $\boldsymbol{\Sigma}$ を推定しなければならないことになるから，このような吟味は必要であると思う。

第4章「t 検定，F 検定のロバストネス」で示されている基本的な事実は，誤差分布についての正規性の仮定の下に導かれたこれらの検定が，より一般な spherical 分布(同時密度が $\|x\|$ のみの関数になるような分布)の下でも望ましい性質をそのまま保つことを示している。このことは数学的な命題としては勿論正しいが，応用上の観点からはその意義は疑問である。というのは現実に誤差分布が正規分布に従わない場合，それが spherical になることはほとんど期待されないからである。この章の標題から想像されることは，むしろ誤差項が互いに独立に正規分布でない分布に従う場合の議論である。そのとき，両側 t 検定，F 検定の仮説の下での大きさ(仮説を棄てる確率)は名目的な水準とあまり変らないこと，もし分布が正規分布より long-tailed であれば逆にそれは小さくなること，従って第1種の誤りに関してはロバストであること，しかし検出力に関しては効率が落ちること等についてのべておくことが必要ではなかったかと思われる。

第5章「系列相関の検定のロバストネス」では，第4章と同じ意味で，正規分布を前提にして導いた分散行列に関する検定が，spherical 分布の場合に拡張できることを示し，また正規性の仮定の下での，いろいろな基準にする最適検定方式の導出を行っている。特に通常の Durbin-Watson 統計量にもとづく検定が，一般に片側対立仮説に対しては局所最強力不偏検定になること，しかし両側仮説に対する局所最強力不変不偏検定はそれとは違った形になることを示しているのが注目される。

第6章「予備検定推定量」では，線形回帰式の係数についての線形仮説を考え，それが棄却されれば，ふつう

の OLSE によって係数を推定し，棄却されなければ仮説を線形制約としてその下で最小2乗推定を行うという，検定–推定の2段推定問題を論じている。このようにして得られる推定量は，ふつうの損失関数に関しては非許容的であることを指摘した後，他の観点も考慮していろいろな性質をのべている。なおこの問題が第1章で展開された James-Stein 推定量と密接な関係があることは，一応指摘されているけれども，James-Stein 推定量の一般化の観点から，この問題を扱い，更に多段階の検定–推定方式へとより一般的な場合を考慮することが，可能でもありまた実際上にも有用な意味を持ち得るのではなかろうか。

第2部に入って，第7章，第8章はいわゆる「Zellner モデル」における推定，検定の問題を扱っている。それはそれぞれ異る説明変数の組(一部に重複があってもよい)を持つ p 本の回帰式

$$\boldsymbol{y}_i=\underset{\sim}{\boldsymbol{\beta}}_i'\underset{\sim}{\boldsymbol{x}}_i+\boldsymbol{\varepsilon}_i \qquad i=1\cdots p$$

を同時に考える問題である。ここに $\boldsymbol{y}_i$ は第 i 番目の被説明変数，$\underset{\sim}{\boldsymbol{x}}_i$ はそれに対応する説明変数ベクトル，$\boldsymbol{\beta}_i$ は回帰係数ベクトル，$\boldsymbol{\varepsilon}_i$ は誤差項である。このモデルについて n 組のデータが与えられたとすれば，

$$\boldsymbol{y}_{ij}=\underset{\sim}{\boldsymbol{\beta}}_i'\underset{\sim}{\boldsymbol{x}}_{ij}+\boldsymbol{\varepsilon}_i, \qquad i=1\cdots p \quad j=1\cdots n$$

と表すことができる。この場合 $\boldsymbol{\beta}_i$ の推定において，他の式を同時に考慮することは一見無意味なように思われるが，そうではないことを指摘したのが Zellner である。すなわち ε_i が互いに独立でなく，かつその分散構造が知られていれば，GLSE を適用することによって，個々の式に OLSE を適用する場合よりも分散の小さい推定量が得られるのである。分散行列が未知ならば OLSE による推定の残差項から推定して，GLSE の式に推定値を代入すればよい。このような推定量の性質については，わが国の若い研究者によるものもふくめいくつかの研究があるが，第7章ではそれらが紹介されている。第8章ではこれと関連して ε_i が互いに独立であるという仮説の検定問題を扱い，最良不変検定の観点から検定統計量を導出している。ここでも片側局所最良検定は残差の相関係数を用いる直観的にわかりやすい形になるが，両側検定は複雑な形になることが示されているのが興味深い。

第9章「成長曲線モデルと相似性の検定」では p 種類のデータに関する n 時点の値 $\boldsymbol{y}_i(t)$ $i=1\cdots p, t=1\cdots n$ について想定された「成長曲線モデル」

$$\boldsymbol{y}_i(t)=\boldsymbol{\mu}_i(t)+\boldsymbol{\varepsilon}_{it}$$

について，$\mu_i(t)$ の「相似性」の仮説を検定する問題から，多変量分散分析 MANOVA，および一般多変量分散分析 GMANOVA の問題を導き出している。

そうして最後の章「一般多変量回帰モデルにおける検定」において，GMANOVA がくわしく論ぜられている。ここで考えられるモデルは，

$$\boldsymbol{Y}=\boldsymbol{X}_1\boldsymbol{B}\boldsymbol{X}_2+\boldsymbol{E}$$

$$\boldsymbol{Y}: NXp \quad \boldsymbol{X}_1: N\times m \quad \boldsymbol{X}_2: (p_1+p_2)\times p$$

$\boldsymbol{E}$ の列ベクトルは互いに独立に $N(0, \boldsymbol{\Sigma})$ に従う。

というもので，$\boldsymbol{B}$ が未知母数である。このモデルの下で，

$$\boldsymbol{R}_1\boldsymbol{B}\boldsymbol{R}_2=\boldsymbol{R}_0 \qquad \boldsymbol{R}_1: m_1\times m,\ \boldsymbol{R}_2: (p_1+p_2)\times p_2$$

という形の仮説を検定する問題を考える。不変性を考慮にすることによって，統計量の基準形を導出し，そこからいくつかの基準によって，最適な検定方式をいくつか導き出している。この章の内容はまだ独立な書物ではほとんど論ぜられていないもので，ここでまとまった形で理論的に説明されていることは極めて有益であると思う。

個々の部分については若干の注文ものべたが，全体として極めて印象的なことは，つねに推定や検定の方法が，何らかの最適性の基準と結びつけて数学的に論ぜられ，いいかげんな「思いつき」を数値計算で合理化するというような安易な方法はきびしく排除されていることである。そのことは著者の見識を表すとともに，この本の品格を高めている。

しかしながらいわずもがなのことをつけ加えれば，回帰分析においては種々の手法の厳密な最適性のほかに，数値的な性質の大体の傾向もまた大切であり，具体的な応用の場面ではそれをさけて通ることはできない。とくにこの本ではほとんどふれられていない，「変数選択」の問題「非正規性」の問題等の扱いにおいては，*ad hoc* な方法と，数値的な扱いを除くわけにはいかない。この本の著者によって，そのような側面をも考慮に入れた続篇が書かれることを希望したい。

しかしとにかく，このような理論的に高度な内容豊かな本が若い研究者によって書かれたことを，わが国の計量経済学界のために喜びたいと思う。 〔竹内 啓〕