経済予測とパラメーターの安定性

山　田　　　勇

1　はしがき

われわれが，たとえば消費函数とか生産函数とかといった経済関係式を打ち立ててこれらを分析し，それらの諸変数間の恒常的な関係[1]をえようとするのは，その際用いられた統計資料の関係する期間についての経済構造を分析するだけを問題とするためではない。もしもその期間だけの経済分析に終るならばそれは過去の単なる記述に止まるに過ぎないものとなってしまう。このような分析を行う目的は，むしろ過去の分析を通じて，将来の予測を行うことにあると見るのが妥当である[2]といわなければならない。

この考え方は何も経済学にかぎらず，他のあらゆる理論研究にも共通の事実であって，そのことはとくに物理学，天文学においていちじるしい。ただこれらの自然現象では，管理された実験が可能であって，何回も同じ条件のもとで繰返された実験資料が利用できる点が，経済現象の分析の場合といちじるしい対照をなしている。最近アメリカでは経済行動の調査に survey method が適用され，ある一定の条件のもとに繰返して同じ種類の質問を同質的な対象にしかけ，これによって自然科学的な管理実験を行おうとしている点は注目に値する。この場合は主として統計学でいわゆるクロス・セクションの研究に属するものといえよう。そこで，このようなクロス・セクションの資料が数期間に亘って同じ対象について行われ，時間的な変化の法則が見出されたならば，観察期間のつぎの時期以降の予測を行うのに，そのような時間的変化の法則を適用することは有効な方法となろう。1つの時期についてのいかに精密な調査も，理論的にはつぎの時期の予測には役立ちえない。

日本においてはアメリカにおけるような survey method はむしろ今後において期待しうるものであって，現在の段階では，利用しうる資料という観点からは，この問題は時期早尚といわざるをえない。

われわれが日本経済の分析を行う場合，いちばん多く利用しうる資料はクロス・セクションの資料よりは時系列の資料である。このことは外国の経済資料についてもあてはまる事実である。そこで，ここでは時系列資料を使って経済予測を行う場合，どのような点に注意しなければならないかということを考えてみよう。

2　統計予測の方法

それでは，時系列資料を使って統計予測を行うには従来どのような方法がとられてきたかを一べつしよう。これは大きく分けて2つの段階に分れる。1つは推定論(proplem of estimation)と呼ばれるものであって，モデルを確定して，そのなかに含まれるパラメーターの推定を行う問題である。このパラメーターの推定にあたっては，過去の統計資料を利用するのが常套手段である。もう1つは，予測論(problem of prediction)と呼ばれるものであって，推定の段階でえられたモデルとそのなかのパラメーターの推定値を使って将来値を予測する問題である。

いま一般的な線型確率的動学モデルをつぎの形で考えよう。

$$\sum_{j=1}^{n}\sum_{k=0}^{p}\alpha_{ijk}\, Y_{j,t-k}+\sum_{j=1}^{m}\beta_{ij}\, Z_{jt}+\eta_{it}=u_{it} \qquad (2\cdot1)$$

$$(i=1, 2, \cdots\cdots, n)$$

この式で Y は内生変数であって，n 個あり，Z は外生変数で m 個ある。さらに $Y_{j,t-1}, Y_{j,t-2}, \cdots\cdots, Y_{j,t-p}$ は先決変数で p 個あり，η は常数項，u は確率変数である。さらに α と β とは構造パラメーターをあらわす。この式の誘導形(reduced form)は

$$Y_{it}=\sum_{j=1}^{n}\sum_{k=1}^{p}\gamma_{ijk}\, Y_{j,t-k}+\sum_{j=1}^{m}\delta_{ij}\, Z_{jt}+\theta_{it}+v_{it} \qquad (2\cdot2)$$

$$(i=1, 2, \cdots\cdots, n)$$

となる。ここに γ, δ, θ は(2・1)式を誘導形(2・2)に変換することによってえられる構造パラメーターであり，また v は同様の変換によってえられる確率変数である。そこで過去の統計資料を使って γ, δ, θ の推定値 $\hat{\gamma}, \hat{\delta}, \hat{\theta}$ を推定し，これから α, β, η の推定値 $\hat{\alpha}, \hat{\beta}, \hat{\eta}$ をうることが

---

1) Trygve Haavelmo, "The Probability Approach in Econometrics", *Econometrica*, XII, Supplement. 1944, 山田勇訳編「計量経済学の確率的接近法」pp. 17—24.

2) Richard Ruggles, "Introduction", in *Long-range Economic Projection, Studies in Income and Wealth*. Vol 16. National Bureau of Economic Research, New York. 1954, p. 3.

できる。そこで $\hat{\gamma}, \hat{\delta}, \hat{\theta}$ を(2・2)式に代入して[3]

$$Y_{it}=\sum_{j=1}^{n}\sum_{k=1}^{p}\hat{\gamma}_{ijk}Y_{j,t-k}+\sum_{j=1}^{m}\hat{\delta}_{ij}Z_{jt}+\hat{\theta}_{it}\,(i=1,2,\cdots\cdots,n) \tag{2・3}$$

をうる。以上は誘導形法の場合である。もし単一方程式だけを問題にするならば(2・1)式から出発する代りに，直接(2・2)式から初めればよい。いずれにしろ，これまでの操作が推定の問題である。

以上の知識を基礎にして，つぎの予測方程式

$$Y_{i,t+\tau}=\sum_{j=1}^{n}\sum_{k=1}^{p}\hat{\gamma}_{ijk}Y_{j,t+\tau-k}+\sum_{j=1}^{m}\hat{\delta}_{ij}Z_{j,t+\tau}+\hat{\theta}_{it} \quad (2・4)$$

$$(\tau=1,2,\cdots\cdots,T)$$

によって $Y_{i,t+\tau}$ の予測を行うのがいわゆる予測の問題となるわけである[4]。

しかし，この際重要なことは(2・3)式の推定の段階における $\hat{\gamma}, \hat{\delta}, \hat{\theta}$ と予測の段階における(2・4)式のなかのそれらとが同じものであるかどうかの保証はないということであって，予測の問題の困難な点のいちばん重要なポイントの1つはここにある。

そこで，これらの2つの段階において，構造パラメーターの推定値が変化しないか，もし変化するとしても，これを変化しないものとして予測を行うためには，推定式(2・4)の γ, δ, θ の推定にあたって，これらが安定的であるような推定法を用いることが必要なこととなる。

従来，経済予測を行うにあたって，たとえば消費函数のなかの限界消費性向とか，産業連関分析の場合の投入係数は安定的なものでなければならないといわれてきたことは周知のことであろう。しかし，過去の統計資料を使って予測を行う場合，パラメーターの安定性は極めて重要な問題であり，このことは統計技術的にも充分考慮せらるべき問題なのであるが，従来の統計的研究では，この点余り注意が払われなかったといってよい。そこで，本稿では，とくにこの点に注目して経済予測の問題を考えることにしよう。

3　パラメーターの推定

(2・2)式の右辺には先決変数 $Y_{j,t-1},\cdots\cdots,Y_{j,t-p}$ と外生変数 Z が含まれているが，パラメーターの推定にあたっては，これらの両者を区別する必要はないから，これらを1つにして X であらわし，しかも i がただ1であるような単一方程式をとり上げよう。この場合は

$$Y=\sum_{i=1}^{N}a_iX_i+a_0+u \tag{3・1}$$

がえられる。ただし，$N=np,\ a_0=\theta$ である。そこで，a の推定の問題であるが，ここでは線型方程式を考えているから，いちばん簡単な推定法として，通常の最小二乗法を採用することにする[5]。いま(3・1)式の X, Y をそれらの算術平均 $\bar{X}, \bar{Y}$ からの偏差の形に変えれば

$$y=\sum_{i=1}^{N}a_ix_i+u \tag{3・2}$$

この式の x, y はそれぞれ $X-\bar{X},\ Y-\bar{Y}$ である。そこで，(3・2)式の u の分散 σ^2 を極小にするように最小二乗法を適用して a の推定値を求めた結果はつぎの如くである，

$$\hat{\boldsymbol{a}}=\sigma_y\boldsymbol{\sigma}^{-1}\boldsymbol{r}^{-1}\boldsymbol{r}_y$$

ここに $\hat{\boldsymbol{a}}$ は列ベクトル

$$\hat{\boldsymbol{a}}\equiv\{\hat{a}_1,\hat{a}_2,\cdots\cdots,\hat{a}_N\}$$

であり，$\boldsymbol{\sigma}$ は N 次の対角マトリックス

$$\boldsymbol{\sigma}\equiv\begin{pmatrix}\sigma_1 & 0\cdots\cdots 0\\ 0 & \sigma_2 \quad 0\\ & \cdots\cdots\cdots\\ 0 & 0\cdots\cdots\sigma_N\end{pmatrix}$$

をあらわし，$\boldsymbol{r}$ は N 次の正方マトリックス

$$\boldsymbol{r}\equiv\begin{pmatrix}1 & r_{12}\cdots\cdots r_{1N}\\ r_{12} & 1\ \cdots\cdots r_{2N}\\ & \cdots\cdots\cdots\cdots\\ r_{1N} & r_{2N}\cdots\cdots 1\end{pmatrix}$$

さらに $\boldsymbol{r}_y$ は列ベクトル

$$\boldsymbol{r}_y\equiv\{r_{y1},r_{y2},\cdots\cdots,r_{yN}\}$$

である。ここに $\sigma_y,\sigma_1,\sigma_2,\cdots\cdots,\sigma_N$ はそれぞれ $Y, X_1, X_2,\cdots\cdots,X_N$ の標準偏差，r_{ij} は X_i と X_j とのピアソン相関係数，r_{yi} は Y と X_i とのピアソン相関係数をあらわす。

(3・1)式のなかの常数項 a_0 の推定値 $\hat{a}_0$ は

$$\hat{a}_0=\bar{Y}-\hat{\boldsymbol{a}}'\boldsymbol{X} \tag{3・4}$$

から求められる。ここに $\bar{Y}$ はスカラーであり，Y の平均をあらわすが，$\hat{\boldsymbol{a}}'$ は列ベクトル $\hat{\boldsymbol{a}}$ の転置されたもの，すなわち行ベクトルを意味する，$\bar{\boldsymbol{X}}$ は列ベクトル

$$\bar{\boldsymbol{X}}\equiv\{\bar{X}_1,\bar{X}_2,\cdots\cdots,\bar{X}_N\}$$

である。

3)　v_{it} はその平均を考えて，これを零におく。

4)　L. R. Klein, *A Textbook of Econometrics*. Row, Peterson and Company, New York, 1953, p. 252.

5)　最小二乗法の仮定としては，通常つぎのものが挙げられる。(1) u_i の確率密度函数は u_i の分布に関して特別の仮定を設けないこと。(2) u_i は X_i に独立であること。(3) $E(u)=0$　(4) $E(u^2)=\sigma^2<\infty$　で，かつこれは X_i に独立であること。(5) u_i と u_j とはたがいに独立であること。本稿においてもこの仮定を採用する。

以上は通常の最小二乗推定値の求め方を述べたに過ぎない。

4　予測の問題

(3・3)式はつぎの式に変換できる。

$$\hat{\boldsymbol{a}}=\boldsymbol{S}_1^{-1}\boldsymbol{S}_2 \qquad (4\cdot1)$$

ここに

$$\boldsymbol{S}_1\equiv\begin{pmatrix}\Sigma x_1^2 & \Sigma x_1x_2\cdots\cdots\Sigma x_1x_N\\ \Sigma x_1x_2 & \Sigma x_2^2\cdots\cdots\ \Sigma x_2x_N\\ \cdots\cdots & \cdots\cdots\cdots\cdots\\ \Sigma x_1x_N & \Sigma x_2x_N\cdots\cdots\Sigma x_N^2\end{pmatrix}$$

$$\boldsymbol{S}_2\equiv\begin{pmatrix}\Sigma x_1y\\ \Sigma x_2y\\ \vdots\\ \Sigma x_Ny\end{pmatrix}$$

である。

ところで，$\hat{a}_i$ の分散（var $\hat{a}_i$ とあらわす）はつぎの如く求められることは周知のところであろう。

$$\text{var}\ \hat{a}_1=E[\hat{a}_i-a_i]^2=E(\hat{a}_i^2)-a_i^2 \qquad (4\cdot2)$$

上式において E は期待値をあらわす。つぎに $E(\hat{a}_i^2)$ を求めてみよう。いま，(4・1)式の右から $\hat{\boldsymbol{a}}'$ をかければ

$$\hat{\boldsymbol{a}}\hat{\boldsymbol{a}}'=\boldsymbol{S}_1^{-1}\boldsymbol{S}_2\boldsymbol{S}_2'\boldsymbol{S}_1^{-1} \qquad (4\cdot3)$$

両辺の期待値は

$$E(\hat{\boldsymbol{a}}\hat{\boldsymbol{a}}')=\boldsymbol{S}_1^{-1}E(\boldsymbol{S}_2\boldsymbol{S}_2')\boldsymbol{S}_1^{-1} \qquad (4\cdot4)$$

$E(\hat{\boldsymbol{a}}\hat{\boldsymbol{a}}')$ の対角要素が(4・2)式の右辺の $E(\hat{a}_i^2)$ を与えることはいうまでもない。そこで，(4・4)式からこれを求めた結果はつぎの如くである[6]。

$$\boldsymbol{A}=\boldsymbol{\alpha}+\sigma^2\boldsymbol{D}$$

上式において $\boldsymbol{A}$ は列ベクトル

$$\boldsymbol{A}\equiv\{E(\hat{a}_1^2),\ E(\hat{a}_2^2),\ \cdots\cdots,\ E(\hat{a}_N^2)\}$$

$\boldsymbol{D}$ は

$$\boldsymbol{S}^{-1}\begin{pmatrix}\Sigma x_1^2 & 0\cdots\cdots\cdots\cdots0\\ 0 & \Sigma x_2^2\quad\cdots\cdots0\\ \cdots\cdots & \cdots\cdots\cdots\cdots\\ 0 & 0\cdots\cdots\ \Sigma x_N^2\end{pmatrix}\boldsymbol{S}^{-1}$$

の対角要素 $D_1, D_2, \cdots\cdots, D_N$ を要素とする列ベクトル，すなわち

$$\boldsymbol{D}\equiv\{D_1, D_2, \cdots\cdots, D_N\}$$

をあらわす。また $\boldsymbol{\alpha}$ は列ベクトル

$$\boldsymbol{\alpha}\equiv\{a_1^2, a_2^2, \cdots\cdots, a_N^2\}$$

を示しており，σ^2 は u の分散である。いま(4・2)式の関係を(4・5)式に代入すれば，列ベクトル

$$\boldsymbol{B}\equiv\{\text{var}\ \hat{a}_1,\ \text{var}\ \hat{a}_2, \cdots\cdots,\ \text{var}\ \hat{a}_N\}$$

はつぎの式によって求められる。

$$\boldsymbol{B}=\sigma^2\boldsymbol{D} \qquad (4\cdot6)$$

しかし実際には σ^2 は測定できないから，その推定値 $\hat{\sigma}^2$ をもって代える。その結果は

$$\boldsymbol{B}=\hat{\sigma}^2\boldsymbol{D} \qquad (4\cdot7)$$

となり，これが最後に必要な分散の式である。

（1）そこでいま，(3・1)式のいちばん簡単な場合として

$$Y=a_0+a_1X_1+u \qquad (4\cdot8)$$

の場合の $\hat{a}_1$ の分散を(4・7)式から求めると

$$\text{var}\ \hat{a}_1=\frac{\hat{\sigma}^2}{(N-2)\hat{\sigma}_1^2} \qquad (4\cdot9)$$

がえられ，var $\hat{a}_0$ は(3・4)式から

$$\text{var}\ \hat{a}_0=\frac{(N-2)\hat{\sigma}_1^2+N\bar{X}_1^2}{N(N-2)\hat{\sigma}_1^2}\hat{\sigma}^2 \qquad (4\cdot10)$$

となる。上式のなかの $\hat{\sigma}_1^2$ は X_1 の標本分散である。予測にあたっては，var $\hat{a}_1$, var $\hat{a}_0$ が最小であることが必要であり，したがって，(4・9)式においては，標本の大きさ N が与えられたとすると，与えられた利用可能な資料のなかで $\hat{\sigma}_1^2$ の最大のものを(4・8)式に挿入することによって，a_1 の推定値の安定性をいちばん強く保証することとなる。(4・9)式の分子の $\hat{\sigma}_2$ は，与えられた資料が決定すれば，極小となっていることは最小二乗法の教えるところである。(4・10)式においては，$\bar{X}_1$ が小さくなればなるほど，a_0 の推定値の分散が小さくなることはいうまでもないが，X_1 の分散については，(4・10)式を $\hat{\sigma}_1^2$ について偏微分して，検査すればよい。すなわち

$$\frac{\partial}{\partial\sigma_1^2}\text{var}\ \hat{a}_0=-\frac{\bar{X}_1^2\hat{\sigma}^2}{(N-2)\hat{\sigma}_1^4}\leqq0 \qquad (4\cdot11)$$

となるから，a_0 の推定値の安定性は X_1 の標本分散が大きければ大きいほど大きくなることが知られる。したがって，(4・8)式のパラメーター，a_0, a_1 の安定性は，利用可能な資料のなかで，X_1 に該当する資料の標本分散が最大であり，その平均が最小のものを選ぶことが要望せられる。（N は通常の場合2よりも大であるから，$N-2$ はプラスである。）

この際注意すべきこととしては，以上のような条件を満たすものであれば，どんなものを持ってきてもよいということではなく，経済的に意味のある資料のなかから選択する必要のあることはいうまでもない。

（2）つぎに(3・1)式を

$$Y=a_0+a_1X_1+a_2X_2+u \qquad (4\cdot12)$$

とし，この場合の $\hat{a}_1, \hat{a}_2$ の分散を(4・7)式から求めると

6)　$E(u_i u_j)=0$,　$E(u)=0$,　U_i と X_i とは独立，$E(u^2)=\sigma^2<\infty$ でかつ X_i と独立の仮定を用いる。

$$\left.\begin{aligned}\text{var}\ \hat{a}_1&=\frac{1+\rho_{12}^2}{(N-3)(1-\rho_{12}^2)^2}\frac{\hat{\sigma}^2}{\hat{\sigma}_1^2}\\ \text{var}\ \hat{a}_2&=\frac{1+\rho_{12}^2}{(N-3)(1-\rho_{12}^2)^2}\frac{\hat{\sigma}^2}{\hat{\sigma}_2^2}\end{aligned}\right\}\quad(4\cdot13)$$

この式においても X, X_2の標本分散が大なるほど a_1, a_2 の推定値の安定性はいっそう大となることは容易に知られるところであろう。さらに X_1 と X_2 との標本相関係数 ρ_{12} については，これらに関して var $\hat{a}_1$, var $\hat{a}_2$ を偏微分することによって知ることができる。すなわち

$$\left.\begin{aligned}\frac{\partial}{\partial\rho_{12}}\text{var}\ \hat{a}_1&=\frac{2\rho_{12}(3+\rho_{12}^2)}{(N-3)(1-\rho_{12}^2)^2}\frac{\hat{\sigma}^2}{\hat{\sigma}_1^2}\\ \frac{\partial}{\partial\rho_{12}}\text{var}\ \hat{a}_2&=\frac{2\rho_{12}(3+\rho_{12}^2)}{(N-3)(1-\sigma_{12}^2)^2}\frac{\hat{\sigma}^2}{\hat{\sigma}_2^2}\end{aligned}\right\}\quad(4\cdot14)$$

ρ_{12}^2 は通常の場合１よりも小なる値であり，かつこの場合 N は３より大であり，さらに $\hat{\sigma}^2, \hat{\sigma}_2^2$ はノン・ネガチヴであるから，２式の左辺の符号は ρ_{12} によって決定される。したがって，もし X_1 と X_2 とが正相関の場合は，ρ_{12} の大(小)なるほど a_1, a_2 の推定値の安定性は小(大)さくなる。このことから，X_1 と X_2 とが正相関の場合には ρ_{12} のなるべく小さいものをとり，負相関の場合には ρ_{12} のなるべく大きいものをとることが望ましい。

a_0 の推定値の安定性については一般的なことは何もいえない。

(4•9)，(4•10)，(4•13)式を見ればわかるように，標本の大きさ N が大きければ大きいほど，パラメーターの安定性は大となることが知られよう。

5　例　題

以上の関係のうち，(4•8)式を仮設例によって説明しよう。

いま X_1 と Y とについてつぎの表が与えられたとしよう。

	Y	X_1	X_1'
	3	2	20
	5	4	21
	6	6	22
平均	4.7	4	21

この場合は，Y の説明変数として X_1 と X_1' とが経済的に意味のあるものであるが，X_1 の方は X_1' よりも平均が小である。またその標本分散は X_1' の方が小である。すなわち，X_1 の標本分散　$\hat{\sigma}_1^2=8$, X_1' の標本分散　$\hat{\sigma}_1'^2=2$　である。またこの際 u の標本分散　$\hat{\sigma}^2=0.17$　を利用して，(4•9)式を計算すれば，X_1 の場合は

$$\text{var}\ \hat{a}_1=\frac{0.17}{(3-2)\times 8}=0.02$$

X_1' の場合は

$$\text{var}\ \hat{a}_1'=\frac{0.17}{(3-2)\times 2}=0.085$$

となって，X_1' よりも X_1 を説明変数として用いた方が a_1 の推定値の安定性はいっそう大なることが知られ，このかぎりにおいては X_1 を用いる方が予測には都合のよいことがわかる。

つぎに a_0 の推定値の標本分散は，X_1 の場合は

$$\text{var}\ \hat{a}_0=\frac{(3-2)\times 8+3\times 16}{3\times(3-2)\times 8}\times 0.17=0.397$$

また X_1' の場合は

$$\text{var}\ \hat{a}_1'=\frac{(3-2)\times 2+3\times 441}{3\times(3-2)\times 2}\times 0.17=37.54$$

となり，この場合も a_1 の場合と同様のことが考えられる。