多変量時系列変動要因分析モデル*

——不確実性への接近——

刈屋　武昭

§1　不確実性への接近法

本節では経済理論と情報の関係を2つの論点から議論し，第2節以降で述べる多変量時系列変動要因分析モデルの有効領域をさぐる。

経済情報の果す役割は，貿易の自由化，金融の自由化，等経済の自由化とともに一層重要になりつつある。その理由として，これら自由化により経済現象の不確実性が一層増加している点が挙げられよう。そしてその不確実性の中で経済主体が自らの最適化を行う場合，的確な経済情報の把握・分析が必要となり，その情報によって自らの行動を経済情勢に適応させていくことが必要となっている。このような情報をベースとした経済行動の変化は競争理論が仮定する主体間の「情報の完全性」の問題と深く関係する。すなわち現実の世界では情報は不完全であり，またその収集には費用を必要とすることから，情報の分布は資本力と関係して不均一になる傾向がある。そのため，情報収集および分析能力を保有する(情報寡占的)主体の行動もしくはその提供する情報に対して他の主体は自らの行動を適応させたり，全く逆の行動をすることで利益機会の可能性を確保したりする。このことが一時的にはゲーム理論のいう「結託」等の現象をみかけ上呈することにもなろうし，それによって不安定性を増加させるかもしれない。他方，金融市場等では資本力がかなりの程度直接的に影響力をもつから，各主体間の相手の行動を意識せざるを得なくなっており，情報量(とその分析能力)および資本力を媒介とした経済主体間の相互依存性は，不確実性の増加とともに一層深まっているようにみえる。とくに為替市場，株価市場，債券(先物)市場等ではこの側面が顕著といえよう。このことは，競争経済理論の仮定する「各主体の行動の独立性」と抵触する。すなわち，不確実性の増加は，経済理論が競争による効率性と経済厚生の最適化を主張するための2つの基本的前提を弱める，と推論する。この推論の正当性は，情報の不均一性，分析能力，資本力の不均一性といった変数をもとに簡単なモデルで確認することもできようが，情報独占があり，資本力が他を圧倒していると想定すれば，競争の自由化は，金融市場等では経済厚生の最大化が必ずしも達成されないことは自明であろう。上の議論で注意する点は，情報と資本力の差に基づいてゲーム理論的「結託」のような現象が生ずるとしてもそれはあくまでも結果的かつ一時的であり，常にそれは確率的に変動する。すなわちゲームの構造としても安定的でない。要約すれば，金融面での自由化の場合，競争原理による経済厚生の拡大を追求しても，その基本的前提としての情報の完全性の仮定が満たされないため，経済主体の行動は他の主体の行動を考慮することが必要となり，その結果全体の行動は一層不安定(Volatile)となる可能性が高い。この行動の不安定化は行動の確率的変動を促し，不確実性を一層拡大しているとみることができよう。このような不確実性の拡大は経済とくに金融構造の不安定化をもたらし，経済理論による現象への接近を難しくしている。為替市場が典型的な例であろう。また財市場でもオイルショックのような経済構造の急激な変化がある場合，多くの主体が情報力と資本力の不均一性をもとに利益機会を追求して不確実性を拡大したことは記

＊　本研究は，日本証券研究奨励財団による研究助成(「福祉制度と福祉政策の厚生経済学的評価と計測」代表鈴村興太郎)を受けた。記して感謝したい。

憶に古くない。なおこのような不確実性の拡大は，後に述べる人間の欲望に根ざす不確実性に依る部分が大きい。

このような不確実な現象に対する接近法として時系列論者は，特定の経済構造を前提にした経済理論およびその視点からの計量分析が非有効的であるという批判をした。彼等は複雑な経済現象に対して帰納的アプローチを主張し，経済構造をブラックボックス的に扱うことで不確実性の高まりつつある現象に時系列論的に接近した。その結果，複雑性のため理論化しにくい金利構造や金融市場の問題において多くの実証分析の方法と実証結果を生んだ。その代表的方法として多変量自己回帰(VAR)モデルによる因果分析が挙げられる。しかしパラメータ数およびそのモデルの構造と関係して少数の変数しか扱えないという限界もあり，金融市場のように不確実性を自己拡大する可能性がある現象に対しては一層の分析手法の発展が要請されている。

もう１つの論点を述べるため，Lucas(1975)が情報の伝播の遅れにより，経済変動を説明していることを思い起そう。彼のモデルの背後には，当然のことながら構造の不変性(それは経済主体の行動関数のパラメータの不変性，技術関係のパラメータの不変性によって保証される)を前提とし，情報の時間的に遅れを伴った伝播のプロセスにより各主体の不変な行動関数を通して決まる変数の値が時間的に変化し，それが基本的に経済変動を引き起す，としている。そこでは，各主体の行動方程式の変数の中には予想変数が含まれており，その予想形成方式としては合理的予想形成仮説が導入されている。もちろん合理的予想は政策変数を含むすべての変数およびパラメータの関数であるからそれらの変化に対して反応できる構造になっている。しかしその仮説を主張する多くの場合，政策変数およびその関数のパラメータの変化のみを問題とし，その政策の変化が予想の変化に反映し，人々はそれによってたとえば消費量を変更する，と想定する。その場合変更するのは消費量であって，消費行動を示す消費関係のパラメータ(たとえば消費性向)は不変であるとする。このパラメータの不変性については，次のように疑問を提起できる。まず合理的予想形成仮説ではモデルのすべてのパラメータを既知とせざるを得ない。さもないと，経済政策等の変化の影響が経済構造を通じて自らの最適化のための目的関数の値への影響を評価できないことになる。このことは，各主体は経済政策等の変化により他の主体の合理的予想値を知ることができることを意味し，その予想値を用いてその主体のとる行動をも評価できることになる。従って，各経済主体は他の主体の行動を知ることができるから，その情報を利用して自らの行動(パラメータ)を変更することで，より大きな厚生を獲得できる可能性が生ずる。これを具体的なモデルで示したのが Kariya(1983)であった。もちろんこれに対して，モデルに含まれるすべての主体が上の可能性をもつのであるから，結局人々は行動を変更しないのが得策と判断してこれまでと同じ方式で行動する，と主張することもできよう。しかし，それはモデルの中にあるメカニズムでなく，もう１つの仮定にすぎない。逆に，各主体はより大きな厚生(もしくは利益)を得る可能性がある場合，確率的(ランダム)な行動をとるとみることもできる。その結果は，パラメータはある構造をもった確率的モデルに従うと定式化するか，その構造は不安定であるため現象を確率的なゲーム論的モデルに帰すことも考えられる。いずれにしても，各主体の行動(パラメータ)は不確実性をもつこととなり，その場合その不確実性は「モデルは既知」とする完全情報的な合理的予想形成仮説に起因する。このことは，完全情報が不確実性を生じせしめるというパラドックスを意味している。このパラドックスは，実は人間のもつ基本的欲望に根ざした心理的不確実性と密接に関係し，その心理的不確実性が天候の変化や外的ショック，政治変革等，による不確実性を拡大したり，相互に反応したりする。それは，目の見えない人が四つ角で道路を渡ろうとしているとき，ある人が「今渡っても安全です」といったとき，目の見えない人は

(1) その人の言うことが信頼できるかという不確実性。

(2) 道路に急に新しい車が入って来ないかという不確実性。

という2つの不確実性に直面することの周知の例と対応している。合理的予想モデルが誤差項を通してモデル化している不確実性は，主として(2)のタイプの不確実性であり，モデルのパラメータを既知として合理的予想を採用し，完全情報の結果生ずる行動主体のパラメータに関する不確実性が(1)のタイプのものである。(1)のタイプの不確実性はモデル化が難しく，ゲーム的な世界の問題でもあろうが，理論的(数学的)な取扱いに関して十分把握されていない。以上を要約してみよう。将来の不確実性に対して期待変数を含むモデルの場合，モデルのパラメータを既知と仮定して合理的予想形成仮説を採用すると，その予想値の中に政策の変化を反映させることはできるが，その結果各主体がお互いにその予想値を知るという完全情報的な世界が発生し，そのことが逆に(1)のタイプの不確実性を生じさせる。このパラドックスを解決する1つの方法は，モデルのパラメータが未知として，合理的予想形成仮説を放棄することである。その場合，政策変数の変化等に対応して人々の予想値の変更を考慮する方法は，未知パラメータの予想値に基づいて主体毎に合理的予想形成の発想に基づいてモデルから形成しているとみてもよいであろうし，また別途政策変化を読み込んだ形でしているとみてもよい。それによって主体間の相互の行動に対して不確実性を導入し，行動の独立性を確保することができよう。

以上2つの議論は，ともに競争的均衡を前提としたときに情報を媒介として生ずる不確実性という点では共通しているが，前者は不完全情報のもとで自由競争を通して主として金融市場で生ずる不確実性であるのに対して，後者は完全情報であるため生ずる不確実性である。このような不確実性は人間行動の欲望に根ざした不確実性と関係し，現象的にはゲーム論的現象を呈することになろう。この論文では，このような不確実性の高い現象に対する実証分析法としてこれまでの計量モデル，時系列分析モデルを補う多変量時系列変動要因分析モデル (Multivariate Time Series Variance Component Model 以下 MTV モデル)を提案する。このMTV モデルは，形式的には主成分要因分析の時系列化(ダイナミック化)であり，第2節ではその理論的基礎を与える。それによって複雑な多変量時系列データの変動構造を把握もしくは近似し，構造解析モデルのみならず予測モデルとして利用できることになる。そのモデルの応用領域としては，上記のような不確実性を内在化する金融市場における変動要因分析，とくに為替や金利の変動要因分析や，株価変動要因分析，景気変動要因分析等，不確実性が高く変動構造が複雑で必ずしも計量経済学的分析が十分成功していない領域が考えられる。他方またベクトル自己回帰(VAR)モデルのもつ欠点を補うという意味で，VAR モデルの補完的モデルとしても重要である。さらにまた，主成分分析が景気指数，物価指数，福祉指数，貯蓄予測指数等，実際的な経済指標作成に大きな役割を果しているが，我々のMTV モデルはこれらの指標の時系列的変動にその理論的基礎を与えることになる。なおMTV モデルと密接に関連したモデルとしては，金利変動構造の解析を狙った Gewecke (1977)， Singleton (1980) および Gewecke and Singleton (1981) のダイナミック因子分析モデルがある。そこでは，一般性の追求と定常時系列分析的発想のみによる取扱いのため，こと応用性に関しては一層の発展が必要であろう。なお彼等のモデルは見方が因子分析的であるため，MTV モデルと異なるばかりか，因子分析のもつ統計的に未解決な論点(計量心理学者 vs 数理統計学者間の論点)を内包している。応用としては景気分析に MTV モデルを用いた刈屋(1985)がある。

§2 MTV モデルの理論的基礎

いま時点 t の p 次元確率ベクトルを

$$\boldsymbol{x}_t=(x_{1t},\cdots,x_{pt})' \quad (t=0,\pm1,\pm2,\cdots)$$

とし，各 x_{it} は共通変動要因

$$f_{jt}(j=1,\cdots,q)$$

をもつ確率モデル(過程)

(2.1) $$x_{it}=\mu_{it}+\alpha_{i1}f_{1t}+\cdots+\alpha_{ip}f_{pt}$$

$(i=1,\cdots,p)$，ただし

$$(2.2)\quad \begin{cases} \alpha_{1j}^2+\alpha_{2j}^2+\cdots+\alpha_{pj}^2=1 \\ \hfill (j=1,\cdots,p) \\ \alpha_{1j}\alpha_{1k}+\alpha_{2j}\alpha_{2k}+\cdots+\alpha_{pj}\alpha_{pk}=0 \\ \hfill (j\neq k) \end{cases}$$

に従うものとする。ベクトル表現では

$$(2.3)\quad \boldsymbol{x}_t=\boldsymbol{\mu}_t+\boldsymbol{\alpha}_1 f_{1t}+\cdots+\boldsymbol{\alpha}_p f_{pt}$$

$$(2.4)\quad \begin{cases} \boldsymbol{\alpha}_j'\boldsymbol{\alpha}_j=1,\ \boldsymbol{\alpha}_j'\boldsymbol{\alpha}_k=0 \quad (j\neq k) \\ \boldsymbol{\mu}_t=(\mu_{1t},\cdots,\mu_{pt})' \\ \boldsymbol{\alpha}_j=(\alpha_{1j},\cdots,\alpha_{pj})' \end{cases}$$

となる。(2.2)もしくは(2.4)の条件は，後に述べるモデルの識別性の条件である。ここで $\boldsymbol{\mu}_t$ は $\boldsymbol{x}_t$ の平均値

$$\boldsymbol{\mu}_t=E(\boldsymbol{x}_t)$$

で，時間に依存してもよい。しかし以下では，簡単化のため $\boldsymbol{\mu}_t$ を既知とし，$\boldsymbol{x}_t-\boldsymbol{\mu}_t$ を考えることで $\boldsymbol{\mu}_t=\boldsymbol{0}$ と仮定する。(2.4)をもつ(2.3)のモデルで次の仮定を満たすものを MTV モデル(多変量時系列変動要因分析モデル)という。

［仮定］ MTV モデルの仮定

(1) $F_j=\{f_{jt}:t=0,\pm1,\pm2,\cdots\}$ は平均0の2次定常過程に従う。とくに ARMA(p_j,q_j) を想定する。

(2) 任意の連続する n 個の $f_{jt}\in F_j(t=r+1,\cdots,r+n)$ に対して

$\boldsymbol{f}_j^r=(f_{j,r+1},\cdots,f_{j,r+n})':n\times1$

を作ると正値定符号の意味で次の不等式が成立する。

$\mathrm{Cov}(\boldsymbol{f}_1^r)>\mathrm{Cov}(\boldsymbol{f}_2^r)>\cdots>\mathrm{Cov}(\boldsymbol{f}_p^r)$

$\mathrm{Cov}(\boldsymbol{f}_j^r)$ は $\boldsymbol{f}_j^r$ の分散行列を示す。

(3) F_j と F_k は無相関である $(j\neq k)$：

$E(f_{jt}f_{ks})=0 \quad (j\neq k)$

このMTVモデルの性質を述べる前に，通常の主成分要因分析モデル(PCモデル)と因子分析モデルとの相異を述べておこう。PCモデルでは，(1)の代りにその特別な場合として

(1)′ F_j はホワイトノイズ系列

とする。従って時点 t と s の共通要因(主成分) f_{jt} と f_{js} は無相関$(t\neq s)$となり，実際の分析ではこれを確保するために階差をとるなどの変数変換が行われる(たとえば King(1966) をみよ)。他方，MTV モデルで F_j が定常時系列過程に従うことを許し，それによって空間的(クロスセクション的)相関変動構造のみならず，時間的相関変動構造をも把握しようとする。さらにPCモデルの場合，上記仮定(2)に対しては F_j のホワイトノイズ性から

(2)′ $\mathrm{Var}(f_{1t})>\mathrm{Var}(f_{2t})>\cdots>\mathrm{Var}(f_{pt})$

でよい。(2)および(2)′はそれぞれのモデルの識別性に関係する。仮定(3)はPCモデルでも仮定される。

他方因子分析の場合，(2.1)のモデルの代りに q 個の共通因子を想定して

$$(2.5)\quad x_{it}=\mu_{it}+\beta_{i1}f_{1t}+\cdots+\beta_{iq}f_{qt}+\varepsilon_{it}$$

の形で表現されるモデルを用いる。共通因子 f_{jt} については，$F_j=\{f_{jt}\}$ がホワイトノイズ系列で F_j と F_k は互いに無相関$(j\neq k)$を仮定するのはPCモデルと同じだが，

$$(2.6)\quad \mathrm{Var}(f_{1t})=\cdots=\mathrm{Var}(f_{qt})=1$$

を課して識別性を議論する。そのため十分成功しない。また特殊因子としての誤差項の系列

$$E_i=\{\varepsilon_{1t}:t=0,\pm1,\pm2,\cdots\}$$

はホワイトノイズ系列であり，E_i と $E_j(i\neq j)$ は互いに無相関，また E_i と F_j も互いに無相関と仮定する。(2.6)を仮定するための係数ベクトル

$$\boldsymbol{\beta}_j=(\beta_{1j},\cdots,\beta_{pj})':p\times1$$

は直交 $\boldsymbol{\beta}_j'\boldsymbol{\beta}_k=0\,(j\neq k)$ しているが，長さは1でない。因子分析の場合，$\boldsymbol{x}_t$ の分散共列は

$$(2.7)\quad \begin{cases} \mathrm{Cov}(\boldsymbol{x}_t)=\boldsymbol{BB}'+\boldsymbol{D} \\ \boldsymbol{B}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_q) \\ \boldsymbol{D}=\mathrm{diag}\{\gamma_1,\cdots,\gamma_p\} \end{cases}$$

と表現される。ただし $\gamma_i=\mathrm{Var}(\varepsilon_{it})$，$\mathrm{diag}\{a_1,\cdots,a_p\}$ は対角要 $a_1,\cdots,a_p$ をその順にもつ対角行列である。(2.7)の構造は任意の $q\times q$ の直交行列 $\boldsymbol{C}$ をとって $\boldsymbol{BC}$ でおきかえても変らない。従って係数 $\boldsymbol{\beta}_j$ が一意的に識別できないという問題をもち，このことに絡む議論が百出している。また与えられたデータから $\boldsymbol{B}$ と $\boldsymbol{D}$ を推定する推定法についても多くの議論がある。

さてMTVモデルの性質を調べよう。MTVモデルも必要ならば因子分析的に q 個の変動要因を

とって

$$(2.8)\quad \begin{cases} x_{it}=\alpha_{i1}f_{1t}+\cdots+\alpha_{iq}f_{qt}+\eta_{it} \\ \eta_{it}=\alpha_{iq+1}f_{q+1t}+\cdots+\alpha_{ip}f_{pt} \end{cases}$$

と書きかえて因子分析的に眺めることもできる($\boldsymbol{\mu}_t=\mathbf{0}$ の仮定を用いている)。しかしアプリオリに一定数の変動要因を必ずしも想定しないこと，また(2.8)の η_{it} と $\eta_{kt}(i\neq k)$ は相関をもつことが因子分析と異なっている。MTV モデルの場合，(2.3)から $\boldsymbol{x}_t$ の分散行列は

$$(2.9)\quad \begin{cases} \boldsymbol{\Sigma}=\sum_{i=1}^{p}\lambda_j\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j'=\boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}' \\ \boldsymbol{A}=[\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_p]:p\times p, \\ \boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\cdots,\lambda_p\} \end{cases}$$

となる。ここで λ_j は仮定(1)(2)から

$$(2.10)\quad \lambda_1>\cdots>\lambda_p,\quad \lambda_j=\mathrm{Var}(f_{jt})$$

となる。従って $\boldsymbol{A}$ は $\boldsymbol{\alpha}_j$ と f_{jt} の符号の同時変化($\boldsymbol{\alpha}_j\to-\boldsymbol{\alpha}_j, f_{jt}\to-f_{jt}$)を除いて一意的に決まり，$\boldsymbol{\Lambda}$ は一意的である。すなわち $\boldsymbol{A}$ は $\boldsymbol{\Sigma}$ の固有ベクトルの作る直交行列

$$(2.11)\quad \boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$$

であり，$\boldsymbol{\Lambda}$ は固有値行列である。注意すべき点は，(2.9)をみる限り PC モデルの場合と全く同様である。しかし MTV モデルの場合，$\boldsymbol{x}_t$ は時系列的な相関をもつため，(2.9)だけでは十分でない。いま $t=1,\cdots,n$ に対して第 i 変量 x_{it} の時系列的構造をみるため

$$(2.12)\quad \boldsymbol{y}_i=(x_{i1},\cdots,x_{in})':n\times 1$$

とおくと，$\boldsymbol{y}_i$ の分散行列は

$$(2.13)\quad \mathrm{Cov}(\boldsymbol{y}_i)=\sum_{j=1}^{p}\alpha_{ij}{}^2\boldsymbol{R}_j$$

となる。ここで

$$(2.14)\quad \begin{cases} \boldsymbol{f}_j=(f_{j1},\cdots,f_{jn})':n\times 1 \\ \boldsymbol{R}_j=\mathrm{Cov}(\boldsymbol{f}_j) \end{cases}$$

である。$\boldsymbol{R}_j$ は F_j の定常性の仮定から時点 t に依存せず，仮定(2)から正値定符号の意味で

$$(2.15)\quad \boldsymbol{R}_1>\boldsymbol{R}_2>\cdots>\boldsymbol{R}_p$$

となる。さらに $\{x_{it}\}$ の i と t についてすべての変動構造をみるため

$$(2.16)\quad \boldsymbol{y}=(\boldsymbol{y}_1',\cdots,\boldsymbol{y}_p'):np\times 1$$

とおくと，$\boldsymbol{y}$ の分散行列は

$$(2.17)\quad \boldsymbol{\Omega}=\sum_{j=1}^{p}\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j'\otimes\boldsymbol{R}_j$$

と表現される。ここで $\boldsymbol{A}\otimes\boldsymbol{B}$ は行列 $\boldsymbol{A}$ と $\boldsymbol{B}$ のクロネッカー積($a_{ij}\boldsymbol{B}$)を示す。PC モデルでは $\boldsymbol{R}_j=\lambda_j\boldsymbol{I}_n$ より，(2.17)の $\boldsymbol{\Omega}$ は

$$\boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}'\otimes\boldsymbol{I}_n=\boldsymbol{\Sigma}\otimes\boldsymbol{I}_n$$

となることを注意しておく。さて PC モデルを要因分析モデルと考える正当性は，Darroch (1965) および Okamoto and Kanazawa (1968) に与えられている。その結果は現在の MTV モデルで各 t の $\boldsymbol{x}_t$ に対しても適用可能である。それを我々の脈絡の中で述べる。この結果は後に拡張された形で証明される(定理3をみよ)。

定理1　(Darroch (1965))任意の t と任意の q を固定する。そして(2.9)の分散行列 $\boldsymbol{\Sigma}$ をもつ確率ベクトル $\boldsymbol{x}_t$(平均は一般性を失うことなく $\mathbf{0}$ と仮定)を $\boldsymbol{C}\boldsymbol{w}_t$ なる形の確率ベクトルで近似する問題を考える。ただし $\boldsymbol{C}$ は $\boldsymbol{C}'\boldsymbol{C}=\boldsymbol{I}_q$ を満たす $p\times q$ の定数行列，$\boldsymbol{w}_t$ は

$$\boldsymbol{w}_t=(w_{1t},\cdots,w_{qt})':q\times 1$$

の任意の確率ベクトルである。このとき

$$(2.18)\quad E\|\boldsymbol{x}_t-\boldsymbol{C}\boldsymbol{w}_t\|^2$$

を最小にする $\boldsymbol{C}\boldsymbol{w}_t$ は

$$\boldsymbol{C}\boldsymbol{w}_t=\boldsymbol{A}_q\boldsymbol{f}^t$$

に限る。最小値は $\sum_{j=q+1}^{p}\lambda_j$ である。ただし

$$(2.19)\quad \begin{cases} \boldsymbol{A}_q=[\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_q],\ \boldsymbol{f}^t=(f_{1t},\cdots,f_{qt})' \\ f_{jt}=\boldsymbol{\alpha}_j'\boldsymbol{x}_t \end{cases}$$

この定理により，各 t に対して $\boldsymbol{x}_t$ の変動を任意の q に対して平均2乗誤差(2.18)の基準で最もよく近似するモデルは MTV モデルであり，共通変動要因 f_{jt} は $\boldsymbol{\Sigma}$ の主成分 $\boldsymbol{\alpha}_j'\boldsymbol{x}_t$ で与えられる。またこの結果は，$t=1,\cdots,n$ を同時に扱った場合に直接的に拡張できる。実際(2.16)の $\boldsymbol{y}$ を用いると，モデルは

$$(2.20)\quad \boldsymbol{y}=(\boldsymbol{A}\otimes\boldsymbol{I})\boldsymbol{f},\quad \boldsymbol{f}=(\boldsymbol{f}_1',\cdots,\boldsymbol{f}_p')':np\times 1$$

と表現される。

定理2　任意の q を固定し，(2.17)の分散行列

$\boldsymbol{\Omega}$ をもつ確率ベクトル $\boldsymbol{y}$ を

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w},\ \boldsymbol{w}=(\boldsymbol{w}_1',\cdots,\boldsymbol{w}_q')':nq\times1$$

で近似する問題を考える。ただし $\boldsymbol{C}$ は $\boldsymbol{C}'\boldsymbol{C}=\boldsymbol{I}_q$ を満たす $p\times q$ の定数行列, $\boldsymbol{w}$ は $\boldsymbol{w}_i:n\times1$ と $\boldsymbol{w}_j:n\times1$ が無相関 $(i\neq j)$ である任意の $nq\times1$ の確率ベクトルとする。このとき

$$(2.21)\qquad E\|\boldsymbol{y}-(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}\|^2$$

を最小にする $(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}$ は

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}=(\boldsymbol{A}_q\otimes\boldsymbol{I})\boldsymbol{f}^*$$
$$\boldsymbol{f}^*=(\boldsymbol{f}^{1\prime},\cdots,\boldsymbol{f}^{n\prime})$$
$$=(f_{11},\cdots,f_{q1};\cdots;f_{1n},\cdots,f_{qn})'$$

に限る。ただし $\boldsymbol{f}^t$ は(2.9)で与えられる。

証明: (2.21)の平均2乗誤差は, 定理1の $\boldsymbol{x}_t$ と $\boldsymbol{w}_t$ を用いると

$$\sum_{t=1}^{n}E\|\boldsymbol{x}_t-\boldsymbol{C}\boldsymbol{w}_t\|^2$$

と書ける。従って定理1から結果がでる。

この定理も以下の定理3に含まれる。

(2.21)の平均2乗誤差は

$$\mathrm{tr}\,E(\boldsymbol{X}-\boldsymbol{C}\boldsymbol{W})(\boldsymbol{X}-\boldsymbol{C}\boldsymbol{W})'$$

と表現されることに注意せよ。ただし

$$\boldsymbol{X}=[\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n]=[\boldsymbol{y}_1,\cdots,\boldsymbol{y}_p]':p\times n$$
$$\boldsymbol{W}=[\boldsymbol{w}_1,\cdots,\boldsymbol{w}_q]:q\times n$$

である。

上記定理1と2では, (2.10)の仮定のみで成立し(2.15)の仮定を用いていない。それは最小にすべく目的関数(2.18)もしくは(2.21)が, $\boldsymbol{x}_t$ の分散行列 $\boldsymbol{\Sigma}$ のみに依存し, $\boldsymbol{R}_j$ に含まれるパラメータに依存していないからである。定理2では $\boldsymbol{\Sigma}$ の固有値が時間に依存しないことが用いられている。その意味で(2.21)をモデルの最適性の基準とする限り, f_{jt} の分散が一定という仮定が必要となろう。もちろん MTV モデルの仮定では定常性を f_{jt} に課しているのでそれは満たされている。他方, 次の定理は定理1と2を含むより強い結論で, 時系列的変動構造をもつ MTV モデルの最適性を示すものである。そこでは(2.15)の仮定が用いられる。

定理3 q を任意に固定する。(2.17)の分散行列 $\boldsymbol{\Omega}$ をもつ(2.16)の $\boldsymbol{y}:np\times1$ を

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}$$

で近似する問題, すなわち各 $\boldsymbol{y}_i:n\times1$ を $\sum_{j=1}^{q}c_{ij}\boldsymbol{w}_j$ で近似する問題 $(i=1,\cdots,p)$ を考える。ただし $\boldsymbol{C}=(c_{ij})$ は $\boldsymbol{C}'\boldsymbol{C}=\boldsymbol{I}_q$ を満たす任意の $p\times q$ の定数行列, $\boldsymbol{w}_j:n\times1$ は平均 $\boldsymbol{0}$, 分散行列 $\mathrm{Cov}(\boldsymbol{w}_j)=\boldsymbol{E}_j$ をもち, $\boldsymbol{w}=(\boldsymbol{w}_1',\cdots,\boldsymbol{w}_q')'$

$$(2.22)\qquad \begin{cases}\boldsymbol{E}_1>\boldsymbol{E}_2>\cdots>\boldsymbol{E}_q\\ \mathrm{Cov}(\boldsymbol{w}_j,\boldsymbol{w}_k)=\boldsymbol{0}\quad(j\neq k)\end{cases}$$

を満たす任意の確率ベクトルである。このとき $n\times n$ の平均2乗誤差行列

$$(2.23)\qquad \boldsymbol{\Phi}\equiv\sum_{i=1}^{p}E\left[\boldsymbol{y}_i-\sum_{j=1}^{q}c_{ij}\boldsymbol{w}_j\right]\times\left[\boldsymbol{y}_i-\sum_{j=1}^{q}c_{ij}\boldsymbol{w}_j\right]'$$

を非負値定符号の意味で最小にするのは

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}=(\boldsymbol{A}_q\otimes\boldsymbol{I})\boldsymbol{f}^*$$
$$\boldsymbol{f}^*=(f_{11},\cdots,f_{q1};\cdots;f_{1n},\cdots,f_{qn})':qn\times1$$
$$\boldsymbol{A}_q=[\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_q],\ f_{jt}=\boldsymbol{\alpha}_j'\boldsymbol{x}_t$$

であり, その場合に限る。最小値行列は

$$\sum_{j=q+1}^{p}\boldsymbol{R}_j$$

で与えられる。

証明 まず(2.16)の $\boldsymbol{y}$ を $(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}$ で近似するときの平均2乗誤差行列

$$(2.24)\qquad \boldsymbol{\Psi}=E[\boldsymbol{y}-(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}][\boldsymbol{y}-(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}]'$$

を考察する。また

$$(2.25)\qquad \mathrm{Cov}(\boldsymbol{f})=\boldsymbol{R}=\begin{pmatrix}\boldsymbol{R}_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{R}_p\end{pmatrix}:np\times np$$

$$(2.26)\qquad \mathrm{Cov}(\boldsymbol{w})=\boldsymbol{E}=\begin{pmatrix}\boldsymbol{E}_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{E}_q\end{pmatrix}:nq\times nq$$

とおく。このとき

$$(2.27)\qquad \mathrm{Cov}\left(\begin{pmatrix}\boldsymbol{f}\\ \boldsymbol{w}\end{pmatrix}\right)=\begin{pmatrix}\boldsymbol{R} & \boldsymbol{B}\\ \boldsymbol{B}' & \boldsymbol{E}\end{pmatrix}:\ n(p+q)\times n(p+q)$$

が非負値定符号であることから, $\boldsymbol{E}$ と $\boldsymbol{B}$ は

$$(2.28)\qquad \boldsymbol{R}-\boldsymbol{B}\boldsymbol{E}^{-1}\boldsymbol{B}'\geq\boldsymbol{0}$$

を満たさねばならない。これが $\boldsymbol{E}$ と $\boldsymbol{B}$ に関する

制約である。さて(2.24)の行列 $\boldsymbol{\Psi}$ は(2.27)を用いると

$$\boldsymbol{\Psi}=(\boldsymbol{A}\otimes\boldsymbol{I})\boldsymbol{R}(\boldsymbol{A}'\otimes\boldsymbol{I})-(\boldsymbol{A}\otimes\boldsymbol{I})\boldsymbol{B}(\boldsymbol{C}'\otimes\boldsymbol{I})-(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{B}'(\boldsymbol{A}'\otimes\boldsymbol{I})+(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{E}(\boldsymbol{C}'\otimes\boldsymbol{I})$$

と評価される。従って

$$\begin{aligned}\boldsymbol{\Psi}_1&\equiv(\boldsymbol{A}'\otimes\boldsymbol{I})\boldsymbol{\Psi}(\boldsymbol{A}\otimes\boldsymbol{I})\\&=\boldsymbol{R}-\boldsymbol{B}\boldsymbol{E}^{-1}\boldsymbol{B}'\\&\quad+[(\tilde{\boldsymbol{C}}\otimes\boldsymbol{I})-\boldsymbol{B}\boldsymbol{E}^{-1}]\\&\quad\times\boldsymbol{E}[(\boldsymbol{C}\otimes\boldsymbol{I})-\boldsymbol{B}\boldsymbol{E}^{-1}]'\\&\geq\boldsymbol{R}-\boldsymbol{B}\boldsymbol{E}^{-1}\boldsymbol{B}'\end{aligned}\tag{2.29}$$

となる。ただし

$$\boldsymbol{C}=\boldsymbol{A}'\boldsymbol{C},\ \ \boldsymbol{C}'\boldsymbol{C}=\boldsymbol{I}_q\tag{2.30}$$

である。次に(2.23)の目的関数 $\boldsymbol{\Phi}$ は任意の $\boldsymbol{a}:n\times1$ に対して

$$\begin{aligned}\phi=\boldsymbol{a}'\boldsymbol{\Phi}\boldsymbol{a}&=\mathrm{tr}[\boldsymbol{I}\otimes\boldsymbol{a}']\boldsymbol{\Psi}[\boldsymbol{I}\otimes\boldsymbol{a}]\\&=\mathrm{tr}[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}']\boldsymbol{\Psi}_1[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}],\ \ \tilde{\boldsymbol{a}}=\boldsymbol{A}\boldsymbol{a}\end{aligned}$$

となることに注意すると，(2.29)から

$$\phi\geq\mathrm{tr}[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}'](\boldsymbol{R}-\boldsymbol{B}\boldsymbol{E}^{-1}\boldsymbol{B}')[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}]$$

となる。等号は

$$\boldsymbol{B}\boldsymbol{E}^{-1}=\boldsymbol{C}\otimes\boldsymbol{I}\tag{2.31}$$

の場合に限る。そのとき最小値は

$$\phi_1=\mathrm{tr}[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}'][\boldsymbol{R}-(\tilde{\boldsymbol{C}}\otimes\boldsymbol{I})\boldsymbol{E}(\tilde{\boldsymbol{C}}\otimes\boldsymbol{I})]\times[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}]\tag{2.32}$$

で与えられる。さらに(2.28)(2.30)(2.32)から

$$[\tilde{\boldsymbol{C}}'\otimes\boldsymbol{I}]\boldsymbol{R}[\tilde{\boldsymbol{C}}\otimes\boldsymbol{I}]-\boldsymbol{E}\geq0\tag{2.33}$$

が得られる。従って

$$\begin{aligned}\phi_1&\geq\mathrm{tr}[\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}']\\&\quad\times[\boldsymbol{R}-(\boldsymbol{M}\otimes\boldsymbol{I})\boldsymbol{R}(\boldsymbol{M}\otimes\boldsymbol{I})][\boldsymbol{I}\otimes\tilde{\boldsymbol{a}}]\\&=\sum_{i=1}^{p}\tilde{\boldsymbol{a}}'\boldsymbol{R}_j\tilde{\boldsymbol{a}}-\sum_{i=1}^{p}\sum_{j=1}^{p}m_{ij}{}^2\tilde{\boldsymbol{a}}'\boldsymbol{R}_j\tilde{\boldsymbol{a}}\\&=\mathrm{tr}\,\boldsymbol{D}[\boldsymbol{I}-\boldsymbol{M}]=\phi_2\end{aligned}\tag{2.34}$$

となる。ただし

$$\begin{cases}\boldsymbol{M}=\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}',\ \ \boldsymbol{M}^2=\boldsymbol{M},\ \ \mathrm{rank}(\boldsymbol{M})=q\\\boldsymbol{D}=\mathrm{diag}\{d_1,\cdots,d_p\},\ \ d_j=\tilde{\boldsymbol{a}}'\boldsymbol{R}_j\tilde{\boldsymbol{a}}\end{cases}\tag{2.35}$$

である。(2.34)で等号が成立するのは，(2.33)，(2.34)より

$$(\boldsymbol{M}\otimes\boldsymbol{I})\boldsymbol{R}(\boldsymbol{M}\otimes\boldsymbol{I})-(\tilde{\boldsymbol{C}}\otimes\boldsymbol{I})\boldsymbol{E}(\tilde{\boldsymbol{C}}'\otimes\boldsymbol{I})=\boldsymbol{0}$$

すなわち $\tilde{\boldsymbol{C}}'\tilde{\boldsymbol{C}}=\boldsymbol{I}$ より

$$\boldsymbol{E}=(\tilde{\boldsymbol{C}}'\otimes\boldsymbol{I})\boldsymbol{R}(\tilde{\boldsymbol{C}}\otimes\boldsymbol{I})\tag{2.36}$$

の場合に限る。ここで(2.15)を用いると

$$d_1>\cdots>d_p\quad(\tilde{\boldsymbol{a}}\neq\boldsymbol{0}\text{ のとき})\tag{2.37}$$

であることに注意する。(2.35)の $\boldsymbol{M}$ は巾等だから直交行列 $\boldsymbol{P}=(p_{ij})$ を用いて

$$\boldsymbol{M}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}',\ \ \boldsymbol{J}=\mathrm{diag}\{1,\cdots,1,0,\cdots,0\}$$

と表現されるから，(2.34)の ϕ_2 は

$$\phi_2=\sum_{i=1}^{p}d_i-\sum_{i=1}^{p}d_i\sum_{j=1}^{q}p_{ij}{}^2$$

となる。さらに(2.37)と

$$\sum_{i=1}^{p}\sum_{j=1}^{q}p_{ij}{}^2=q$$

から

$$\sum_{j=1}^{q}p_{ij}{}^2=1\quad(i=1,\cdots,q)$$

$$\sum_{j=1}^{q}p_{ij}{}^2=0\quad(i=q+1,\cdots,p)$$

のとき，またそのときに限り ϕ_2 は最小となり，その最小値は $\sum_{i=q+1}^{p}d_i$ となる。そのとき $\boldsymbol{P}$ は

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{P}_1&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{P}_2\end{pmatrix}$$

の形をとる。その結果(2.35)より

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{I}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}\end{pmatrix},\ \ \tilde{\boldsymbol{C}}=\begin{pmatrix}\boldsymbol{T}\\\boldsymbol{0}\end{pmatrix},\ \ \boldsymbol{T}'\boldsymbol{T}=\boldsymbol{I}_q\tag{2.38}$$

を得る。この $\tilde{\boldsymbol{C}}$ が(2.26)の $\boldsymbol{E}$ に対して(2.36)を満たすためには

$$\boldsymbol{G}=\boldsymbol{T}'\boldsymbol{D}_1\boldsymbol{T}$$

$$\boldsymbol{G}=\mathrm{diag}\{g_1,\cdots,g_q\},\ \ \boldsymbol{D}_1=\mathrm{diag}\{d_1,\cdots,d_q\}$$

が必要である。ただし

$$\boldsymbol{g}_i=\tilde{\boldsymbol{a}}'\boldsymbol{E}_i\tilde{\boldsymbol{a}},\ \ g_1>\cdots>g_q>0$$

である。これと(2.37)から

$$\boldsymbol{T}=\mathrm{diag}\{\pm1,\cdots,\pm1\}$$

が結論される(複合任意)。これと(2.38)，(2.30)より

$$\boldsymbol{C}=[\pm\boldsymbol{\alpha}_1,\pm\boldsymbol{\alpha}_2,\cdots,\pm\boldsymbol{\alpha}_q]\tag{2.39}$$

となる。ただし複合は任意である。この $\boldsymbol{C}$ を用いると(2.31)と(2.36)より

$$\boldsymbol{E}=\begin{pmatrix}\boldsymbol{R}_1&&\boldsymbol{0}\\&\ddots&\\\boldsymbol{0}&&\boldsymbol{R}_q\end{pmatrix},\ \ \boldsymbol{B}=\begin{pmatrix}\pm\boldsymbol{R}_1&&\\&\ddots&\\&&\pm\boldsymbol{R}_q\\&\boldsymbol{0}&\end{pmatrix}$$

が決まる。この $\boldsymbol{E}$ と $\boldsymbol{B}$ に対して

$$\mathrm{Cov}(\boldsymbol{w})=\boldsymbol{E},\ \ \mathrm{Cov}(\boldsymbol{f},\boldsymbol{w})=\boldsymbol{B}\tag{2.40}$$

を満たす任意の確率ベクトル $\boldsymbol{w}$ を用いて

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}$$

を作れば，それは(2.23)の $\boldsymbol{\Phi}$ を非負値定符号の意味で最小にする。もちろん $\boldsymbol{C}$ は(2.39)で与えられる。$\boldsymbol{w}$ が一意的に決まることを示すため，

$$\boldsymbol{w}^*=\boldsymbol{B}'\boldsymbol{R}^{-1}\boldsymbol{f}$$

とおく。このとき(2.40)を満たす任意の $\boldsymbol{w}$ に対して

$$E(\boldsymbol{w}-\boldsymbol{w}^*)(\boldsymbol{w}-\boldsymbol{w}^*)'$$
$$=\boldsymbol{E}-\boldsymbol{B}'\boldsymbol{R}^{-1}\boldsymbol{B}-\boldsymbol{B}'\boldsymbol{R}^{-1}\boldsymbol{B}+\boldsymbol{B}'\boldsymbol{R}^{-1}\boldsymbol{B}=\boldsymbol{0}$$

を得る。従って $\boldsymbol{w}=\boldsymbol{w}^*$ を得，そのとき

$$(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}^*=(\boldsymbol{A}_q\otimes\boldsymbol{I})\boldsymbol{f}^*$$

を一意的に得る。

定理2と定理3の両方とも全変量 $\boldsymbol{y}$ を $(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}$ の形の確率ベクトルで近似する問題を考察しているが，その近似の評価基準は定理3の方がより強いものとなっている。実際定理2の評価基準は，定理3の $\boldsymbol{\Phi}$ を用いて

$$\mathrm{tr}\,\boldsymbol{\Phi}=E\|\boldsymbol{y}-(\boldsymbol{C}\otimes\boldsymbol{I})\boldsymbol{w}\|^2$$

と表現される。従って定理3は定理2を含んでいる。これに対して定理1が定理3に含まれることをみるには，定理3で $n=1$ の場合を考えて定常性の仮定により1を t とみなせばよい。そのとき

$$\boldsymbol{\Phi}=\sum_{i=1}^{p}E\Bigl(y_{it}-\sum_{j=1}^{p}c_{ij}w_{jt}\Bigr)^2=E\|\boldsymbol{x}_t-\boldsymbol{C}\boldsymbol{w}_t\|^2$$

となる。

定理3の評価基準 $\boldsymbol{\Phi}$ は，$\boldsymbol{x}_t$ の時系列的変動をも考慮しており，その意味で定理3はMTVモデルが変動要因分析モデルとしての最適性をもっていることを示している。すなわち，MTVモデルでは空間的のみならず時系列的にも相関関係をもって変動する p 個の変量 $\boldsymbol{x}_t$ の背後に互いに無相関に時系列変動する q 個の変動要因があると想定する。このとき定理3は最初の q 個の共通変動要因をとれば，現象をもっともよく近似できることを示している。この点に関して次の点を指摘しておく。通常の因果関係を前提とした回帰分析で，たとえば消費の変動を所得の変動のみで説明する場合，消費変動のうち所得の変動がその変動要因として最も支配的であるとみているにすぎず，誤差項の中に他の変動要因があることを否定していない。その意味では，MTVモデルでは変動要因 f_{jt} は未知でありそれを $\boldsymbol{x}_t$ の変動原因とみることは必ずしもできないが，構造として似ている。共通変動要因の数 q としていくつとればよいかは，$(\boldsymbol{A}_q\otimes\boldsymbol{I})\boldsymbol{f}^*$ が $\boldsymbol{y}$ の全体の変動をどの程度近似しているかという通常の主成分要因分析の視点から累積寄与率に基づいて行うことができる。実際 $\boldsymbol{y}$ の変動をその分散行列 $\boldsymbol{\Omega}$ の対角和(分散和)

$$\mathrm{tr}\,\boldsymbol{\Omega}=\sum_{j=1}^{p}(\mathrm{tr}\,\boldsymbol{\alpha}_j\boldsymbol{\alpha}_j')(\mathrm{tr}\,\boldsymbol{R}_j)=n\sum_{j=1}^{p}\lambda_j=n\,\mathrm{tr}\,\boldsymbol{\Sigma}$$

とすれば，$\boldsymbol{f}_j=(f_{j1},\cdots,f_{jn})'$ の変動は分散行列 $\boldsymbol{R}_j$ の対角和(分散和)$n\lambda_j$ となる。それゆえ累積寄与率は次式で与えられる。

$$\eta_q=\sum_{j=1}^{q}\lambda_j\Big/\sum_{j=1}^{p}\lambda_j$$

若干のコメントをする。$p=1$ の場合もしくは第 i 要素 x_{it} のみの時間的変動は，f_{jt} の時系列変動に対応し，各 f_{jt} は定常時系列過程に従うのであるから，その1次結合としての x_{it} も定常時系列過程に従う。その際，たとえば f_{1t} と f_{2t} で x_{it} の変動の95%を説明できるとすれば(平均を考慮して)

$$x_{it}-\mu_{it}=\alpha_{i1}f_{1t}+\alpha_{i2}f_{2t}+\varepsilon_{it}$$

と書ける。f_{jt} がARMAモデルに従うとすれば $x_{it}-\mu_{it}$ もARMAモデルに従う。その場合，誤差とみなす ε_{it} もARMAモデルであるため，その中に若干の情報を残すことになる。

次に識別性の問題について述べる。まず $\boldsymbol{y}$ の2つの分散行列の $\boldsymbol{\Omega}=\sum_{i=1}^{p}\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i'\otimes\boldsymbol{R}_i$ と $\boldsymbol{\Omega}^*=\sum_{i=1}^{p}\boldsymbol{\alpha}_i^*\boldsymbol{\alpha}_i^{*\prime}\otimes\boldsymbol{R}_i^*$ に対して $\boldsymbol{\Omega}=\boldsymbol{\Omega}^*$ ならば

$$\boldsymbol{\alpha}_i=\boldsymbol{\alpha}_i^*,\quad \boldsymbol{R}_i=\boldsymbol{R}_i^*\,(i=1,\cdots,p)$$

であることが容易に示される。次に $\boldsymbol{y}$ の2つの表現 $(\boldsymbol{A}\otimes\boldsymbol{I})\boldsymbol{f}$ と $(\boldsymbol{A}^*\otimes\boldsymbol{I})\boldsymbol{f}^*$ に対して

$$(\boldsymbol{A}\otimes\boldsymbol{I})\boldsymbol{f}=(\boldsymbol{A}^*\otimes\boldsymbol{I})\boldsymbol{f}^*$$

が成立するとする。このとき条件(2.11)より $\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$，従って

$$\boldsymbol{f}=(\boldsymbol{A}'\boldsymbol{A}^*\otimes\boldsymbol{I})\boldsymbol{f}^*\equiv(\boldsymbol{B}\otimes\boldsymbol{I})\boldsymbol{f}^*$$

となる。ここで両辺の分散行列を求めると

$$\boldsymbol{R}=(\boldsymbol{B}\otimes\boldsymbol{I})\boldsymbol{R}^*(\boldsymbol{B}'\otimes\boldsymbol{I})$$

となり，$\boldsymbol{R}$ と $\boldsymbol{R}^*$ が(2.25)の形をしていることを用いて両辺に左と右から $[\boldsymbol{I}\otimes\boldsymbol{\alpha}']$ と $[\boldsymbol{I}\otimes\boldsymbol{\alpha}]$ をか

けると

$$D = BD^*B'$$

を得る(定理3の証明をみよ)。ここで $\boldsymbol{a}$ は $n\times 1$ の任意のベクトル，$\boldsymbol{D}$ と $\boldsymbol{D}^*$ はそれぞれ要素 $d_i = \boldsymbol{a}'\boldsymbol{R}_i\boldsymbol{a}$, $d_i^* = \boldsymbol{a}'\boldsymbol{R}_i^*\boldsymbol{a}_i$ をもつ $n\times n$ の対角行列である。従って $d_1 > \cdots > d_p > 0$, $d_1^* > \cdots > d_p^* > 0$ より $\boldsymbol{B}=\boldsymbol{I}$ すなわち $\boldsymbol{A}=\boldsymbol{A}^*$ が結論される。

3番目に，上の結果は景気指標等主成分分析を用いた経済指標の時系列変動にモデル論的基礎を与える。実際，主成分分析モデルでは $\boldsymbol{x}_t$ は時系列的に互いに無相関と仮定する。従ってその前提のもとでは，各主成分 $f_{jt} = \boldsymbol{\alpha}_j'\boldsymbol{x}_t$ も時系列的に無相関な筈である。しかし実際の経済指標は時系列的に相関をもって変動している。上の結果はこのギャップを埋めている。実際，主成分 f_{jt} の定常的変動が許される。これと関連して，[仮定] で仮定した $\{f_{jt}\}$ の定常性を弱めることを考えよう。モデル(2.1)の不確実な現象近似モデルとしての最適性を定理1に求めるとすれば，定常性の仮定を次のように弱めることができる。まず λ_{jt} を f_{jt} の分散とし，時間に依存してもよいとする。このとき，

$$\boldsymbol{\Sigma}_t = \mathrm{Var}(\boldsymbol{x}_t) = \boldsymbol{A}\boldsymbol{\Lambda}_t\boldsymbol{A}'$$

$$\boldsymbol{\Lambda}_t = \mathrm{diag}\{\lambda_{1t}, \cdots, \lambda_{pt}\}$$

となる。定理1はこの $\boldsymbol{\Sigma}_t$ に対してもそのまま成立する。従って少なくともモデルとしては f_{jt} の分散は時間に依存してよい。データからパラメータの推定可能性をみるため

$$E(\boldsymbol{X}'\boldsymbol{X}) = \sum_{t=1}^{n} E(\boldsymbol{x}_t\boldsymbol{x}_t') = \boldsymbol{A}\sum_{t=1}^{n}\boldsymbol{\Lambda}_t\boldsymbol{A}'$$

に注目する。この関係からモーメント法の発想に立って $\boldsymbol{A}$ の推定値として $\boldsymbol{X}'\boldsymbol{X}$ の固有ベクトルの作る直交行列をとることができる。そのとき $\boldsymbol{X}'\boldsymbol{X}$ の固有値 d_j は $\Sigma\lambda_{jt}$ の推定値とみることができる。従って $f_{jt} = \boldsymbol{\alpha}_j'\boldsymbol{x}_t$ の分散が λ_{jt} であることから，x_{jt} が時間の簡単な関数である場合，それを f_{jt} の推定値に基づいて推定できる。たとえば $f_{1t} = \boldsymbol{\alpha}_1'\boldsymbol{x}_t$ が時間の1次式で近似できるようなトレンド的変動をする場合，$\hat{f}_{1t} = \hat{\boldsymbol{\alpha}}_1'\boldsymbol{x}_t$ の分散

$$\frac{1}{n}d_1 = \frac{1}{n}\sum(\hat{f}_{1t})^2 \simeq \frac{a}{n}\sum(t-\bar{t})^2 = O(n)$$

となる。ただし $\boldsymbol{x}_t$ の平均を $\boldsymbol{0}$ と仮定していることに注意せよ。それゆえ，共通変動要因としてトレンド的変動が含まれていてもよい。

§3 MTV モデルの利用法

MTV モデルは共通変動要因分析モデルであるから，共通変動要因が何であるかを発見することを狙う。そのため共通変動要因 f_{jt} を定常時系列モデル

$$f_{jt} \sim \mathrm{ARMA}(p_j, q_j)$$

でモデル化し，その構造とその要因への大きさを示す係数パラメータ $\alpha_{ij}(i=1, \cdots, p)$ から要因の解釈を行う。この解釈の方法は PC モデルの場合と同様だが，経済学的知識等に基づいても主観性が入りやすい(なお主成分々析の経済分析的視点からの一般的解説については刈屋(1978)をみよ)。以下この解釈の助けとなるばかりでなく，現象の構造を把握する代表的な分析法をみてみよう。

[変動要因分析] いま§2で述べた累積寄与率等に基づいて q 個の変動要因がとられると，各 x_{it} の変動(分散)は

(3.1) $\mathrm{Var}(x_{it}) \simeq \alpha_{i1}^2\lambda_1 + \cdots + \alpha_{iq}^2\lambda_q$

と表現される。実際の応用では(2.1)の x_{it} の代りにそれを基準化した変数

(3.2) $z_{it} = (x_{it} - \mu_{it})/\sqrt{\sigma_{ii}}$

を用いることが考えられるが，その場合 $\mathrm{Var}(z_{it}) = 1$ となるため(3.1)は

(3.3) $1 \simeq \alpha_{i1}^2\lambda_1 + \cdots + \alpha_{iq}^2\lambda_q$

となり，通常の PC モデルのように因子負荷量の2乗 $\alpha_{ij}^2\lambda_j$ が第 j 共通変動要因に起因する変動の大きさになっている。以下説明を簡単にするため(3.2)の z_{it} について述べる。

[空間的相関分析] 次に時点 t を固定したときの z_{it} と z_{jt} のクロスセクションの相関係数は，q 個の変動要因に起因する大きさとして

(3.4) $\mathrm{Correl}(z_{it}, z_{jt}) \simeq \alpha_{i1}\alpha_{j1}\lambda_1 + \cdots + \alpha_{iq}\alpha_{jq}\lambda_q$

と表現される。これも PC モデルの場合と同様である。

[時間的相関分析] MTV モデルの特徴は，z_{it} と z_{jt-l} の相関構造が共通変動要因 f_{kt} の時系列的変動としての分散行列 $\boldsymbol{R}_k(k=1, \cdots, q)$ を用いて

$$(3.5)\quad \mathrm{Correl}(z_{it}, z_{jt-l}) \simeq \alpha_{i1}\alpha_{j1}\mathrm{Cov}(f_{1t}, f_{1t-l}) + \alpha_{i2}\alpha_{j2}\mathrm{Cov}(f_{2t}, f_{2t-l}) + \cdots + \alpha_{iq}\alpha_{jq}\mathrm{Cov}(f_{qt}, f_{qt-l}) = \alpha_{i1}\alpha_{j1}\gamma_1(l) + \cdots + \alpha_{iq}\alpha_{jq}\gamma_q(l)$$

と表現される点にある。ここで $\gamma_k(l)$ は f_{kt} の第 l 階自己共分散である。すなわち与えられた変量 z_{it} と z_{jt} の空間的相関構造が(3.4)で把握されるのに対して，その時間的相関構造が(3.5)で把握される。自己共分散 $\gamma_k(l)$ は自己相関係数 $\rho_k(l)$ を用いて

$$\gamma_k(l) = \lambda_k \rho_k(l)$$

と書けることに注意せよ。データから α_{ij} と $\gamma_k(l)$ が推定されれば，各 l に対して(3.5)の推定値が得られる。f_{kt} に ARMA(p_k, q_k) を仮定する場合，そのモデルから $\gamma_k(l)$ は得られる。ここで注意をすべき点は，(3.5)の表現からわかるように z_{it} と z_{jt-l} の相関係数は z_{it-l} と z_{jt} の相関係数に等しい，ということである。その意味でこのモデルを用いて直接的に因果関係をみることはできないが，変数の選択の段階で z_{it} に加えて z_{it-1} 等を加えて分析することでそれをみることができよう。

［予測］ 各変動要因 f_{kt} に対して ARMA(p_k, q_k) モデルを想定することで，l 期先の値 f_{kt+l} を t 期迄の f_{kt} の値で予測できる。従ってモデル(2.1)に基づいて p 個の変数 $z_{it}(i=1, \cdots, p)$ のすべての l 期先の値 z_{it+l} が予測可能となる。この予測可能性の背後では，当然のことながら z_{it} の係数の時間的安定性と f_{kt} の ARMA(p_k, q_k) モデルの安定性を前提にする。いま特定の変数 x_{1t}(たとえば為替レート)の予測をしたい場合，その変数の変動が一定割合(たとえば95%)把握される変動要因の数を選択する。その変動要因 f_{kt} の選択は，必ずしもその分散が大きい順に選択する必要はなく，$\alpha_{1k}f_{kt}$ の分散 $\alpha_{1k}{}^2\lambda_k$ でみて大きい順に f_{kt} を選択することが考えられる。それによって x_{1t} の予測能力を上げることができよう。

§4 MTV モデルの推定とモデル選択

第2節の議論によりモデル(2.1)の係数ベクトル $\boldsymbol{\alpha}_i$ は，$\boldsymbol{x}_t$ の分散行列 $\boldsymbol{\Sigma}$ の固有ベクトルであるから，$\boldsymbol{\alpha}_i$ の推定値は対応する標本分散行列の固有ベクトルとすることができる。今簡単化のため $\boldsymbol{\mu}_t$ を既知とすると，それは

$$(4.1)\quad \boldsymbol{S} = \frac{1}{n}(\boldsymbol{X}-\boldsymbol{M})(\boldsymbol{X}-\boldsymbol{M})' = \frac{1}{n}\sum_{t=1}^{n}(\boldsymbol{x}_t-\boldsymbol{\mu}_t)(\boldsymbol{x}_t-\boldsymbol{\mu}_t)'$$

$$\boldsymbol{M} = (\boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_n),\quad \boldsymbol{X} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$$

と表現される。この $\boldsymbol{S}$ の期待値が(2.1)モデルのもとで $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}'$ となることをすでにみた。従ってモーメント法的発想から $\boldsymbol{A}$ の推定値を $\boldsymbol{S}$ を対角化する直交行列で推定できる。他方，定理2より $\boldsymbol{V} = \boldsymbol{X}-\boldsymbol{M}$ の変動を平均2乗誤差の意味で最も近似する式は

$$\boldsymbol{A}_q\boldsymbol{A}_q{}'\boldsymbol{V}$$

で与えられる。そこでその平均2乗誤差を構成する

$$(4.2)\quad \delta = \|\boldsymbol{V}-\boldsymbol{A}_q\boldsymbol{A}_q{}'\boldsymbol{V}\|^2 = \mathrm{tr}(\boldsymbol{V}-\boldsymbol{A}_q\boldsymbol{A}_q{}'\boldsymbol{V})(\boldsymbol{V}-\boldsymbol{A}_q\boldsymbol{A}_q{}'\boldsymbol{V})'$$

を最小にする $\boldsymbol{A}_q$ の推定量を求めてみよう。(4.1)の $\boldsymbol{S}$ を用いて

$$\boldsymbol{S} = \boldsymbol{H}\boldsymbol{D}\boldsymbol{H}',\quad \boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_p$$

$$\boldsymbol{D} = \mathrm{diag}\{d_1, \cdots, d_p\}\quad (d_1 > \cdots > d_p)$$

と書くと

$$\frac{1}{n}\delta = \mathrm{tr}\,\boldsymbol{D} - \mathrm{tr}\,\boldsymbol{A}_q{}'\boldsymbol{H}\boldsymbol{D}\boldsymbol{H}'\boldsymbol{A}_q = \mathrm{tr}\,\boldsymbol{D} - \mathrm{tr}\,\boldsymbol{B}'\boldsymbol{D}\boldsymbol{B},\quad \boldsymbol{B} = \boldsymbol{H}'\boldsymbol{A}_q$$

とかける。$\boldsymbol{B}'\boldsymbol{B} = \boldsymbol{I}_q$ に注意せよ。このとき定理3の証明または Darroch (1965) と同様に議論すると

$$\boldsymbol{B} = \begin{pmatrix} \pm 1 & & 0 \\ & \ddots & \\ & & \pm 1 \\ & 0 & \end{pmatrix}\quad (\text{複合任意})$$

となる。従って

$$\hat{\boldsymbol{A}}_q = (\pm\boldsymbol{h}_1, \cdots, \pm\boldsymbol{h}_q)$$

となる。これは $\boldsymbol{S}$ に基づく固有値 $d_1 > \cdots > d_q$ に対応する固有ベクトルに他ならない。そのとき λ_k の推定値は d_k となる。これらの推定値を用いると，第3節で述べた空間的相関係数や時系列的相関係数の推定値が得られる。また f_{kt} の推定値は

$$\hat{f}_{kt} = \hat{\boldsymbol{\alpha}}_k{}'\boldsymbol{x}_t$$

であるので，これに基づいて ARMA(p_k, q_k) をモデル化でき，それに基づいて予測できる。

トレンドの問題もしくは μ_t の推定の問題について述べる。モデルが well-defined であるために f_{jt} に定常性を課した。しかし経済時系列データの場合，各変数には異なるトレンドが存在する場合が多い。従って $\boldsymbol{\mu}_t$ の推定のときにこれらのトレンドを何らかの方法で除去しておくことが望ましい。しかしたとえば階差法による場合，モデル(2.1)の表現する変動はこれら階差をとられた変数の変動となり，もとの変数の変動のモデル化とは異なる。他方，溝口・刈屋(1983)で議論したように，現象記述的立場からトレンドを含めた時系列分析をするという立場もありえよう。

MTV モデルの応用としては，為替変動要因分析とその予測([13])，景気変動要因分析とその予測([4])，株価変動要因分析とその予測，等不確実性の強い現象の予測モデルとして利用できる。なおシステマティックな計算機プログラムとしては日経 NEEDS の藤原俊朗氏が開発したものがある。

(一橋大学経済研究所)

参考文献

[1] 刈屋武昭(1983)「新しい経済指標作成の方法」『経済研究』34巻，40-49。

[2] 刈屋武昭(1978)「要因の発見と指標化」溝口・刈屋編『統計学』青林書院新社。

[3] 溝口・刈屋(1983)『経済時系列分析入門』日本経済新聞社。

[4] 刈屋武昭(1985)「景気動向指標採用系列の変動要因分析」未発表。

[5] Darroch, J. N. (1965), "An Optimal Property of Principal Components," *Ann. Math. Statist.*, 36, 1579-1582.

[6] Gewecke, J. F. (1977), "The Dynamic Factor Analysis of Economic Time Series Models," *Latent Variables in Socio-Economic Models*, North-Holland.

[7] Gewecke, J. F., and Singleton, K. J. (1981), "Maximum Likelihood "Confirmatory" Factor Analysis of Economic Time Series," *International Economic Review*, 22, 37-54.

[8] Kariya, T. (1983), "Optimal Rational Expectations," *Hitotsubashi Journal of Economics*, 24, 101-108.

[9] King, B. F. (1966), "Market and Industry Factors in Stock Price Behavior," *Journal of Business*, 28, 139-190.

[10] Lucas, R. E. Jr. (1975), "An Equilibrium Model of the Business Cycle," *Journal of Political Economy*, 83.

[11] Okamoto, M., and Kanazawa, M. (1968), "Minimization of Eigenvalues of a Matrix and Optimality of Principal Components," *Ann. Math. Statist.*, 39, 859-863.

[12] Singleton, K. J. (1980), "A Latent Time Series Model of the Cyclical Behavior of Interest Rates," *International Economic Review*, 21.

[13] 日本銀行調査統計局計量分析係(1985)「為替レートの変動分析と予測」未発表。