日本銀行金融研究所「金融研究」第11巻第2号（平成4年7月）

Johansenの共和分検定について

川　崎　能　典

1.　目的・構成

マクロ経済データの多くが強いトレンド成分を持つことは古くから知られた問題である。これらのトレンドはこれまで伝統的に計量分析の一つの障壁として捉えられ、しばしばtime trendに回帰してその残差間で分析を行うとか、階差をとって分析を行うといった対応がとられてきた。しかし、こうした対応をとった場合には、トレンドについての情報を一切捨ててしまっている危険性があり、事実そうした分析ではモデルの当てはまりが芳しくないケースが多い。

一方で、複数のマクロデータを比較・観察すると、極めて似通った動きが認められる場合や、あるいはある変数の全く逆のパターンを示す変数が観察される、という場合が多い。このような場合には、各変数が背後に抱えているトレンド成分は共通のものであると考えられる。近年実証分析で頻繁に用いられるようになった共和分（cointegration）によるアプローチは、こうした共通のトレンド成分の情報を取り入れてモデルの説明能力を向上させようとするやり方である。

本論文で扱うJohansenの尤度比検定（likelihood ratio test、以下JohansenのLR検定と略記する）とは、非定常変数による多変量自己回帰モデル（vector autoregression model、以下VARと略記する）から出発して、変数間に何通りの共和分の関係が認められるかを検定するものである。それはまた誤差修正モデル（error correction model、以下ECMと略記する）の最尤推定法という側面も合わせ持ち、少なくとも米国では実証分析

本論文は、筆者が東京大学大学院在籍中に日本銀行金融研究所から委嘱された研究活動の一環としてとりまとめたものである。本論文の作成に当たっては、國友直人（東京大学）、矢島美寛（東京大学）、山本　拓（一橋大学）の各氏より有益なコメントを頂いた。なお、本論文のデータ解析に当たっては、加藤健吾（日本銀行）の協力を得た。

に欠かせないツールとなっている。わが国においても今後その使用頻度は増していくと考えられる。

本論文の目的は、従来の分析法が抱えていた問題と照らし併せてJohansen のLR検定を概説、その意義・問題点を整理することにある。また、Johansen のLR検定が3変数以上のモデルで真価を発揮することから、5つの主要な時系列から成る多変量モデルによる解析例を示す。

本論文の構成は以下のとおりである。

2.では、共和分とはどういう現象を指すのかを定義する。そして、従来共和分の有無を検定するために使われることの多かったEngle-Granger の2段階法について概観し、それとの相違点を対比する形で、Johansen のLR検定の特徴を簡単に紹介する。3.ではVAR と ECM の対応関係を示す。4.ではJohansen のLR検定を簡単に紹介する。5.では5変数VARモデルによるデータ解析と応用分析も試みる。6.では、ECMの最尤推定法というJohansen検定の解釈、統計的考え方、実用上の問題点の3つの論点からJohansen のLR検定の意義について考察を加える。

なお、Johansen のLR検定の統計理論的枠組みの詳細については補論1で、また、この検定を構成する際に用いる統計的漸近理論のうち重要な役割を果たすブラウン運動については補論2で説明する。

2. 共和分とEngle-Grangerの2段階法

Nelson and Plosser [1982] は、モデルの当てはまりという観点からみて、マクロ経済データの多くは決定論的なトレンド（deterministic trend）を持つというよりは、その動きがランダム・ウォーク的（すなわち非定常）であることを指摘した。[1] こうした指摘は、ほぼ同時期に研究の進んだ時系列データの非定常性の検定方法と相俟って、各種経済データが非定常、とくに単位根過程（unit root process）であるか否かの議論が盛んに行われるようになった。[2]

ところで、これらの変数をプロットしてみると往々にして似通った動きが認められ、1つ1つが「独立」[3] な非定常過程に従っているとは必ずしも考えにくい。このような非定常な変数間の従属性を把握する概念が共和分（cointegration）である。

共和分とはどういう現象を指すのか、その最も簡単な例が第1図に示されている。x_t、y_t はそれぞれ1階の階差で定常性を確保できる変数であると仮定し、$I(1)$変数と呼ぶ。[4] このとき $z_t=y_t-kx_t$ なる線形結合につい

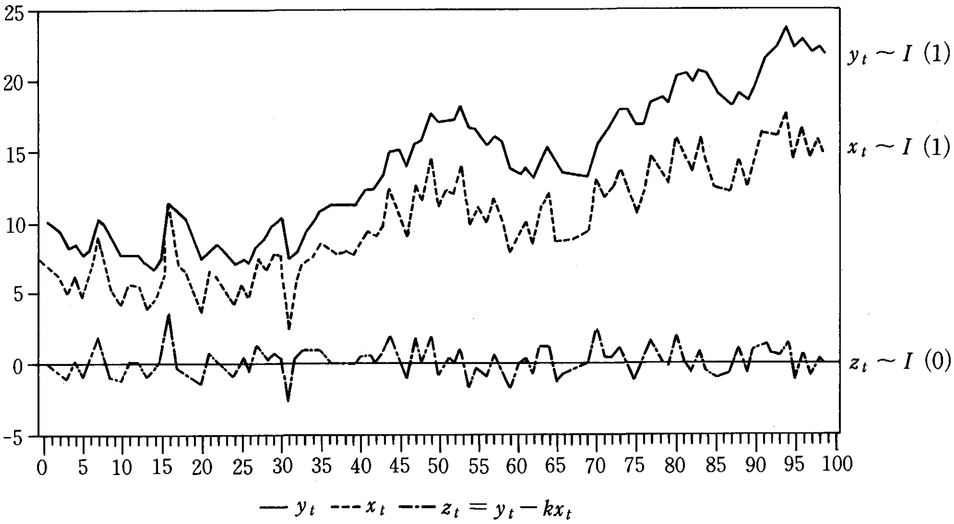
1）本論文において、ある時系列が定常であるとは、平均・分散が時間によらず一定で、自己共分散が時差のみに依存することを指す。

2）最も単純なケースとしてAR(1)過程 $y_t=ay_{t-1}+\epsilon_t$ を考えよう。$By_t=y_{t-1}$ というラグ作用素 B を用いて書き直すと $(1-aB)y_t=\epsilon_t$ となるが、$\{y_t\}$ が定常過程であるためには方程式 $1-aB=0$ が複素平面上の単位円周外に根を持たなければならない。これに対し、1階のランダム・ウォーク過程では $a=1$ であり、これを指して単位根過程という。単位根過程に関する手法のサーベイやそれらをマクロ経済学データへ応用する際の問題点等についてはCampbell and Perron [1992]、Diebold and Nerlove [1990] 等を参照されたい。

3）ここでいう独立は確率変数の独立ではなく、複数の非定常変数が線形独立の関係にあることを指す。以後この意味で用いるときは引用符をつける。

4）一般に、d 階の階差をとることで定常性が満たされる非定常変数を $I(d)$ 変数と呼ぶ。これはintegrated

第１図　共和分関係にある２変数



（出所）吉田[1989]

て、z_t が定常となるような k が存在すれば、x_t と y_t とは共和分関係にあるという。

［定義１］x_t、y_t は $I(1)$変数と仮定する。x_t、y_t が共和分関係にあるとは、ある定数 k が存在し、$z_t \overset{\text{def}}{=} x_t - ky_t$ が定常となるときである。これを $z_t \sim I(0)$と書く。[5] またこのときベクトル（１，$-k$）を共和分ベクトルという。

さて、共和分の有無を統計的に検証する手法として従来よく用いられてきたのが、Engle and Granger [1987] の２段階法（two-step method）である。

２変数モデルを例にとろう。前提として、２変数 x_t、y_t が $I(1)$変数とみなせることを検定で確認しておかなければならない。その上で、これら２変数の間に共和分が存在するかもしれないことを念頭に置きつつ、次のような回帰を行う。

$$y_t = c + \beta x_t + u_t$$

次に、残差の推定値 $\{\hat{u}_t\}$ が定常であるか否かを ADF（Augmented Dickey Fuller）テストで検定する。すなわち、

$$\Delta \hat{u}_t = -\rho \hat{u}_{t-1} + \sum_{i=1}^{k} a_i \Delta \hat{u}_{t-i} + e_t$$

を計測し、$\rho = 0$ を帰無仮説とする検定を実行する（対立仮説は $\rho > 0$）。その結果 $\{\hat{u}_t\}$ が定常（$\rho > 0$）と結論されれば、x_t と y_t の間は共和分関係にあることになる。[6]

さて、２段階法は必ずしもその分析対象が

---

of order d の略で、いまの例では x_t、y_t は $I(1)$変数、階差をとらずとも元々定常である変数は $I(0)$変数である。

5）記号～は、統計学においては通常「次の分布に従う」という意味で用いられる。$I(0)$等は分布を表すものではないが、「定常過程に従う」ことをこう表現するのが慣例となっている。

6）ADF テストの帰無仮説は「$\{\hat{u}_t\}$ は非定常」であるから、２段階法での帰無仮説は「共和分が存在しない」になる。

2変数モデルである場合に限らず適用できるが、その場合の重要な前提条件は「共和分が何個存在しているかが既知」ということである。

後述のGrangerの表現定理（Granger's representation theorem、Engle and Granger［1987］）によれば、一般にp変量VARで各変数が$I(1)$という仮定の下では、共和分の個数は1個以上$p-1$個まで存在しうる。したがって、2変数であれば共和分は存在しても1個であるが、3変数以上の$I(1)$変数を扱う際には、まず何通りの共和分が存在するのか決めなければ分析を開始することはできない。

本論文の目的は、複数の共和分関係の検出手法であるJohansenのLR検定について、いくつかの角度から検討を加えることである。したがって、3変数以上の一般的な多変量時系列を念頭に置くことになるので、その準備として共和分の定義を一般化しておこう。

［定義2］ 各成分が$I(1)$であるベクトル$\mathbf{x}_t$が共和分の関係にあるとは、ある定数ベクトルβが存在して$\mathbf{z}_t=\beta'\mathbf{x}_t$の各成分が$I(0)$となることである（これをまた$\mathbf{z}_t \sim I(0)$と書く）。また、このときの$\beta$を共和分ベクトル（cointegration vector）という。

3. VARとECM

時系列解析の一手法であるVARは、現在では経済データの分析手法としては極めてポピュラーなものとなっている。しかし、非定常変数をVARによって分析する際に留意すべき点として、時系列データを定常化することの妥当性という問題がある。この問題に触れつつ、VARとECMを関係づけるのが本節の目的である。

そもそも時系列解析の理論は、主に定常性を有するデータに対する解析法として発展を遂げてきた。そして非定常変数については階差をとることで定常化したとみなし、定常時系列の議論を適用する手法がとられてきた（Box and Jenkins［1976］）。しかし、非定常な変数間に共和分が存在する場合には、この手法は適切さを欠く。

いま2変数のVARモデル（ラグは2期）を考えよう。x_t、y_tはともに$I(1)$変数であると仮定する。

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix}$$

ここで$(\mu_x, \mu_y)'$、$(\epsilon_t, \eta_t)'$はそれぞれ定数項および誤差項を表す。[7] 誤差項は独立同一正規分布に従うと仮定する。

さて、定常時系列の議論に持ち込むため、この式の左辺がそれぞれx_t、y_tの階差系列Δx_t、Δy_tとなるように変形してみよう。表記の簡略化のため上の式の係数行列をそれぞれΠ_1、Π_2と書くことにすれば、直ちに次の表現を得る。

$$\begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \Gamma \begin{pmatrix} \Delta x_{t-1} \\ \Delta y_{t-1} \end{pmatrix} - \Pi \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \tag{1}$$

ここでΓもΠも2行2列行列であり、

7） プライム（$'$）はここでは行列やベクトルの転置を表す。

$$\Gamma \overset{\text{def}}{=} -\mathrm{I} + \Pi_1$$
$$\Pi \overset{\text{def}}{=} \mathrm{I} - \Pi_1 - \Pi_2$$

で定義される（Ｉは単位行列)。

ところで、仮定より Δx_t、Δy_t は定常であるから、(1)式右辺は定常であることが必要である。右辺第３項は定数、右辺第１、第４項は定常であるから、右辺第２項もまた定常でなければならない。[8] 一方、第２項中の２つの成分（x_{t-2}, y_{t-2}）はともに非定常変数であるため、この項の各成分が定常であるためには(i)Πがゼロ行列、(ii)Πと$(x_{t-2}, y_{t-2})'$を掛け合わせたベクトルの各成分は定常、のどちらかである必要がある。

(i)は、従来の「階差をとってからVARを当てはめる」方法が妥当性を持つケースである。このとき x_t、y_t は「１次独立」な非定常変数であり、先に定義した言葉を用いれば共和分が存在しないケースである。(ii)では x_{t-2}、y_{t-2} の線形結合は定常性を満たしており、このケースを指して「共和分関係にある」と定義したのである。(ii)の場合、２×２行列 Π は、２×１行列と１×２行列の積として次のように分解できることが知られている（Granger の表現定理、Engle and Granger [1987])。

$$\Pi \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} (\beta_1,\ \beta_2) \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix}$$
$$= \begin{pmatrix} \alpha_1 \cdot EC_{t-2} \\ \alpha_2 \cdot EC_{t-2} \end{pmatrix}$$

ここで $EC_{t-2} \overset{\text{def}}{=} \beta_1 x_{t-2} + \beta_2 y_{t-2}$ である。これを(1)式に代入して得られる(2)式は、いわゆる誤差修正モデルとしても解釈可能である。

$$\begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \Gamma \begin{pmatrix} \Delta x_{t-1} \\ \Delta y_{t-1} \end{pmatrix} - \begin{pmatrix} \alpha_1 \cdot EC_{t-2} \\ \alpha_2 \cdot EC_{t-2} \end{pmatrix} + \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \qquad (2)$$

線形結合 $EC_t \overset{\text{def}}{=} \beta_1 x_t + \beta_2 y_t$ が定常であるとみなせる場合、ベクトル（β_1, β_2）を定義２に従って共和分ベクトルといい、EC_{t-2} を「誤差修正項」(error correction term) と呼ぶ。ベクトル（α_1, α_2)′は送出ベクトル（loading vector）と呼ばれるが、これは、「誤差修正項」EC_{t-2} をそれぞれ α_1 倍、α_2 倍して各々第１方程式・第２方程式の説明変数としてシステムに送り出す役割を負っているからである。

さて、(2)式の解釈としてよく用いられる「誤差修正メカニズム」を概説すれば以下のとおりである。

いまの例（２変数）の続きで考えよう。まず現象として観測されることは、x_t、y_t の２変数は非定常であるが、その線形結合が（検定の結果）定常とみなされる、すなわち $\beta_1 x_t + \beta_2 y_t = EC_t \sim I(0)$、$t = 1, \cdots, T$ である（EC_t の平均は一般性を失うことなく、ある定数 c と仮定することができる)。

実際には各期において値 $\{EC_t\}$ が観測されるのだが、十分長い期間を通じて、平均的には $\beta_1 x_t + \beta_2 y_t = c$ という関係が２変数間に成り立っていることが想像される。このようにして共和分から抽出した決定論的関係のことを、本論文では一貫して相互依存関係と呼ぶことにする。ECM の文脈ではこの相互依存関係を、時間に依らず成り立つ均衡状態という意味で、「長期均衡関係」（long-run

8） $x_t \sim I(d_x)$, $y_t \sim I(d_y) \Rightarrow x_t + y_t \sim I(\max(d_x, d_y))$。詳しくは Granger [1983] 参照。

equilibrium relationship）と呼ぶ。

ECMの視点に立つ場合、この相互依存関係を経由することにより、共和分という現象に対して次のような解釈を行う。「相互依存関係とは経済システムでの一種の均衡状態であり、実際にはゆらぎ$\{EC_t\}$が、「長期均衡」からの乖離（＝error）として発生している。」ここでさらに、「経済システム内には、過去に生じた乖離を「長期均衡」に向けて修正するメカニズム（error correction mechanism）が存在する」と考えて、$\{EC_t\}$の過去の値をモデルの説明変数に加えるのである。

以上の説明から分かるように、ECMとは観測者の立場を表明している呼び名である。例に即して確認すれば、今期のx_t、y_tの変動（Δx_t, Δy_t）は、過去の変動（Δx_{t-1}, Δx_{t-2}; Δy_{t-1}, Δy_{t-2}）によって説明されるだけでなく、2変数間に成り立っている相互依存関係（あるいは「長期均衡関係」）から、過去においてどれだけの乖離（＝ereor）が発生していたかによっても説明される、とする立場である。なお、この解釈の問題点については6.で述べる。

4.　Johansenの共和分検定

(1)　JohansenのLR検定

前節に述べたように、3変数以上のVARモデルから出発してECMを考える際の最大の難関は、行列Πのランク、すなわち共和分の個数[9]をいかに決定するかである。3変数モデルであれば、Πのランクは1かもしれないし2かもしれない（つまり非定常な3変数を結び付ける共和分ベクトルが1つしかないかもしれないし2通りあるかもしれない）。Johansen [1988]の尤度比検定とは、Πのランクがいくつであるかを検定で決め、同時に未知パラメターの最尤推定量を求める方法である。

Johansen [1988]の方法は次のようなk次のVARから出発する。

$$X_t = \Pi_1 X_{t-1} + \cdots + \Pi_k X_{t-k} + \epsilon_t, \quad t = 1, \cdots, T \tag{3}$$

ここで、X_tはp変量の確率変数ベクトルであり、誤差項ϵ_tは平均0分散Λ（$p \times p$）の独立同一正規分布に従うとする。

X_tの各要素が全て$I(1)$であれば、次の$p \times p$行列

$$\Pi = \mathrm{I}_p - \Pi_1 - \cdots - \Pi_k$$

のランクはpより小さくなければならない。これは、次のようにVARのECM表現を考えれば直ちに判ることである。

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} - \Pi X_{t-k} + \epsilon_t \tag{4}$$

ここで、$\Gamma_i = -\mathrm{I} + \Pi_1 + \cdots + \Pi_i$（$i = 1, \cdots, k-1$）である。前節での2変数の例と全く同様、$X_t$の各成分が$I(1)$という仮定の下では、(4)式の$\Pi$はフルランクではあり得ない。そこで、$X_t$の要素が共和分関係にある（換言すれば$0 < r = \mathrm{rank}\ \Pi < p$）という仮説は、$\Pi$が$\Pi = \alpha\beta'$と分解可能であることに対応する。ここで$\alpha$、$\beta$は$p \times r$行列である（以上、Grangerの表現定理）。

(4)式をECMとみなしたときの行列α、β'の持つ意味も全く前節と平行である。このときβ'のr個の行ベクトルは、X_tに対する共

9）　これを共和分階数（rank of cointegration）という。

和分ベクトルになっている。つまり、線形結合 $\beta' X_t$ が $I(0)$であり、これは共和分の関係そのものである。行列 α の第 i 行ベクトル（$i=1,\cdots,p$）は、r 個の共和分関係（あるいは誤差修正項）の、第 i 方程式における重みである。

(4)式において、Π の最尤推定は次のような手順で行われる。最初の目標は、$\Gamma_1,\cdots,\Gamma_{k-1}$ を撹乱母数[10]として扱い、Π（あるいは β）に関する集約対数尤度を構成することである。まず ΔX_t と X_{t-k} を $\Delta X_{t-1},\cdots,\Delta X_{t-k+1}$ に回帰し、それぞれ残差ベクトル R_{0t} と R_{kt} を得る。この2本の補助回帰から、残差に関する積率行列を

$$S_{ij} \overset{\text{def}}{=} T^{-1}\sum_{t=1}^{T} R_{it}R'_{jt}, \quad i,\ j=0,\ k$$

で定義する。このとき集約対数尤度は、次の回帰モデルに対応している。

$$R_{0t}=-\alpha\beta' R_{kt}+u_t \tag{5}$$

β を一旦固定して考えれば、(5)式は R_{0t} を $-\beta' R_{kt}$ に回帰して、β 所与の形で $\hat{\alpha}$、$\hat{\Lambda}$ を求めることができる。すなわち、

$$\hat{\alpha}(\beta)=-S_{0k}\beta(\beta' S_{kk}\beta)^{-1}$$
$$\hat{\Lambda}(\beta)=S_{00}-S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}$$

を得る。

Johansen [1988] は、尤度関数の最大化は $|\hat{\Lambda}(\beta)|$ の β に関する最小化と同値であることを示した。解 $\hat{\beta}$ は固有値問題

$$|\lambda S_{kk}-S_{k0}S_{00}^{-1}S_{0k}|=0 \tag{6}$$

の解である p 個の固有値 $\hat{\lambda}_1>\cdots>\hat{\lambda}_p>0$ に対応する固有ベクトルを $\hat{V}' S_{kk}\hat{V}=I$ に従って基準化した $\hat{V}=[\hat{v}_1,\cdots,\hat{v}_p]$ より得られる。ここまでで未知パラメターの推定は終わっており、次の段階では $\hat{v}_1$ から $\hat{v}_p$ までのうち、何個までが共和分ベクトルなのかを検定する。もし共和分階数が r なら、$\hat{\lambda}_1$ から $\hat{\lambda}_r$ までの対応する r 個のベクトルを並べて $\hat{\beta}$ とする。

共和分ベクトルの数 r（換言すれば行列 Π のランク）を検定するに当たって、Johansen はこの p 個の固有値に基づく方法を提唱した。これは高々 r 個の共和分ベクトルが存在するという仮説を、残りの $p-r$ 個の固有値 $\lambda_{r+1},\cdots,\lambda_p$ がゼロという仮説検定で置き換えるというものである。

Johansen の記法に従い、「高々 r 個の共和分ベクトルが存在する」という仮説を $H_2(r)$ と書こう。$H_2(r)$ 対 H_1（無制約）[11]の検定は、LR 統計量

$$-T\sum_{i=r+1}^{p}\ln(1-\hat{\lambda}_i) \tag{7}$$

に基づいて行われる。$0\leq r<p$ であるから、帰無仮説の下、r の値として可能性があるのは $r=0,1,\cdots,p-1$ であり、それぞれの r の値に対して個別に LR 検定が行われる。

もしちょうど r 個の共和分関係が存在しているのであれば、$r+1$ 個から p 個までの固有ベクトルに基づいて説明変数を追加しても、モデルの尤度はほとんど増加しないはずである。そこで、対立仮説 H_1 の下で推定された最も冗長（redundant）なモデルと、「r

10) 撹乱母数（nuisance parameter）とは、統計的推測を行うに当たってさしあたり関心のない未知母数を指す。

11) 対立仮説 H_1 の下では共和分ベクトルが何個存在するかを考慮しない、という意味での無制約である。

個の cointegration vector が存在する」という帰無仮説の下でのモデルとで、尤度比検定を考える。これは結局、$\lambda_{r+1}=\cdots=\lambda_p=0$ を $\lambda_i \geq 0\ \forall i$ に対して行う検定に帰着する。

Johansen [1988, Theorem 3] によれば、LR 統計量(7)の極限分布は、

$$tr\left\{\int_0^1 dBB'\left[\int_0^1 BB'du\right]\int_0^1 BdB'\right\} \tag{8}$$

である。ここで $B(u)$ は $p-r$ 次元ブラウン運動であり、分散行列は I_{p-r} である。ここまで述べてきた検定方式を、極限分布の形からトレース検定と呼ぶ。

さらに、$H_2(r)$ 対 $H_2(r+1)$ $(r=0,\cdots,p-1)$ の検定も同様に構成できる。これは共和分を追加的にもう１個考慮するモデルの冗長性の検定であり、トレース検定の場合から容易に類推されるように、この際の検定統計量は $-T\ln(1-\hat{\lambda}_{r+1})$ である。その極限分布は上の行列の最大固有値であることから、この検定は最大固有値検定と呼ばれる。

Johansen は、検定で常識的に用いられるパーセント点に対し、$p-r=1, 2, 3, 4, 5$ のケースでシミュレーションを行って、極限分布の数表を作成している（Johansen [1988] の Table 1)。極限分布は共和分の個数にしか依存しないことに注意しよう。また、利用可能な数表の都合上 p は高々５までである。

(2)　定数項の分解

Johansen LR 検定の大枠は(1)に述べたとおりであるが、実際の分析に当たっては定数項を含むモデルを考える必要があろう。しかし、これまで提案されてきた非定常性の検定と同様、Johansen の LR 検定も定数項を含むか否かで検定統計量の極限分布が微妙に異なってくる。

モデルに定数項が含まれるか否か自体重要な場合分けの節目であるが、モデルの解釈を考慮すれば、推定された定数項はドリフト項である場合（すなわち線形トレンド）と、誤差修正項の平均を構成する場合という２通りの可能性があることに注意が必要である。

この点は、次のように明示的に表される。いまドリフト項（ないし線形トレンド）を表すベクトルを μ_0 と書き、誤差修正項 $\beta' X_{t-k}$ の平均ベクトルを β'_0 と書こう。このとき(4)式に対応する表現として

$$\Delta X_t-\mu_0=\sum_{i=1}^{k-1}\Gamma_i(\Delta X_{t-i}-\mu_0)$$
$$-\alpha[\beta' X_{t-k}-\beta'_0]+\epsilon_t \tag{9}$$

を得る。これより直ちに、ECM 型で推定される定数ベクトル（例えば μ）は

$$\mu \overset{\text{def}}{=} (I-\sum_{i=1}^{k-1}\Gamma_i)\ \mu_0+\alpha\beta'_0$$

である。

Johansen [1989] はこの点を考慮し、推定された μ を、ドリフトを表すベクトルと誤差修正項の平均ベクトルに分解する式を与えている。式の詳細については補論1の (A-20)、(A-21)式を参照されたい。

この際の検定統計量は、前節に述べた２つの検定方式で帰無仮説の下

$$\left\{\int_0^1 (dU)F'\left[\int_0^1 FF'dt\right]^{-1}\int_0^1 F(dU)'\right\} \tag{10}$$

のトレースと最大固有値とにそれぞれ弱収束する。ただし、ここで、$U(t)=\{U_1(t),\cdots,U_{p-r}(t)\}$ は $p-r$ 次元ブラウン運動、$F(t)=\{F_1(t),\cdots,F_{p-r}(t)\}$ は各要素が

$$F_i(t)=U_i(t)-\int_0^1 U_i(s)ds,$$

$$(i=1,\cdots,p-r-1)$$

$$F_i(t)=t-\frac{1}{2},\ (i=p-r)$$

で定義される$p-r$次元の確率過程である。数表はJohansen and Juselius [1990] の Table A1に与えられている。

5. データ解析例

(1) 分析の対象と計測結果

本節では、JohansenのLR検定を現実の経済データに適用した計測を試みた。具体的には、物価に関連した主要経済変数として分析対象となることの多い3変数（マネーサプライ、名目GNP、物価）に、近似注目されている資産ストックの2変数（土地、株式）を加えた5変数について、共和分関係を計測することとした。[12)]

分析にはM2+CD（平残値、季節調整済)、名目GNP（季節調整済)、市街地価格指数（全国平均)、東証一部株価時価総額（季節調整済)、物価デフレータのいずれも対数値を使用した。Johansenの方法により共和分関係の検証に入る前に、これら5つの変数の非定常性を予め確認しておく必要性がある。ADFテスト（Dickey and Fuller [1981]）とPhillips-Perronテスト（Phillips and Perron [1988]）を用いて検定したところ、いずれの変数においても非定常であるとの結果が得られた。[13)] また、出発点であるVARモデルのラグについては、ラグ8期のモデルに対し、

第1表 JohansenのLR検定の結果

帰無仮説	$r\leq 0$	$r\leq 1$	$r\leq 2$	$r\leq 3$	$r\leq 4$
トレース検定	145.83**	92.84**	47.26**	19.49*	3.14
(5%有意点)	(68.91)	(47.18)	(29.51)	(15.20)	(3.96)
最大固有値検定	52.99**	45.57**	27.77**	16.36*	3.14
(5%有意点)	(33.18)	(27.17)	(20.78)	(14.04)	(3.96)

	マネー	GNP	土　地	株　価	物　価
推定された共和分ベクトル（行ごとに）	−42.38	46.57	6.03	5.70	−32.62
	−61.60	75.91	− 3.46	8.22	−26.48
	−44.31	10.82	26.10	7.52	12.55
	− 7.46	− 0.43	14.79	− 3.19	0.02

(注) 1. *は有意水準5%で棄却、**は有意水準1%で棄却を表す。
2. サンプルは1956/2～1990/2、半期データ、ラグは6期を選択。

12) 変数の選択に当たっては、必ずしも特定の経済理論を前提とした訳ではない。ここでの目的は特定の経済理論の検証ではなく、むしろJohansenのLR検定を現実のデータに当てはめることで、どのような計測結果やインプリケーションが得られるか調べることにある。

13) マネーサプライと地価については$I(1)$、$I(2)$双方の可能性が認められたが、ここでは$I(1)$として扱っている。

より小さいラグを持つ制約的なモデルの妥当性を検定するラグランジュ乗数検定を行い、ラグ値6を選択した。

第1表はJohansenのLR検定の実行結果である。2つの検定の帰無仮説・対立仮説のとり方は前節で扱ったが、表に則して改めて説明すれば、トレース検定とは、

• $r \leq k$ vs. $r \leq 5$ $(k=0, \cdots, 4)$

を、最大固有値検定とは、

• $r \leq k$ vs. $r \leq k+1$ $(k=0, \cdots, 4)$

をkに関して逐次的に実行して、共和分の個数を決定する方法である。

トレース・最大固有値検定のどちらにおいても、共和分の個数が0～3個であるとの帰無仮説（$r=0$, $r \leq 1$, $r \leq 2$, $r \leq 3$）は全て5％有意水準で棄却されている一方、4個以下であるとの帰無仮説（$r \leq 4$）は受容されており、これら5変数の間に共和分が4個存在することを示唆している。

これは、マネー、名目GNP、地価、株式時価総額、物価が、個別にみればランダム・ウォーク的な挙動を示すものの、5変数まとめてみたときに5次元空間を全くランダムに動いているのではなしに、むしろその動きが1次元的であることを示している。ECM流にいえば、これら5変数は「乖離」の方向や大きさは様々であるにしても、「均衡部分空間」というべき1本の「均衡」直線の周りを変動している、と解釈される。

なお、マネーサプライ指標として、M2＋CD平残値の代わりに末残値を使用した場合、地価変数として市街地価格（全国平均）の代わりに市街地価格（6大都市平均）を使用した場合も参考までに計測したが、いずれも共和分の個数は4個と推定されており、マネー、名目GNP、地価、株式時価総額、物価の5変数を使う限り、ほぼ同様な結論が得られた。ただし、株式時価総額の代わりに日経平均株価を使用した場合、もはや共和分は4個存在しないとの結果が得られている。[14)]

さて、推定されたモデルがどの程度現実の動きを説明できるかを調べるため、各変数の実績値とECMによるインサンプル予測値を比較してみた。レベルでみた場合、全ての変数において両者はほとんど一致しており、レベル面での当てはまりはほぼ完全であった。そこでより細かな当てはまりを確かめるため、前期比（つまり対数値の差分）での実績値と予測値をプロットしてみた（第2図～第6図)。

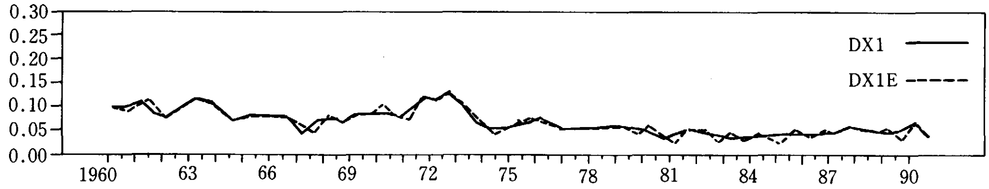
そもそも変動幅の大きい株式時価総額については、やや両者が乖離しているケースも見受けられるが、大きな動きはほぼカバーしているといえよう。その他の4変数については推定値が現実の動きをよく追跡しているといえる。

(2) 制約の検定：応用例

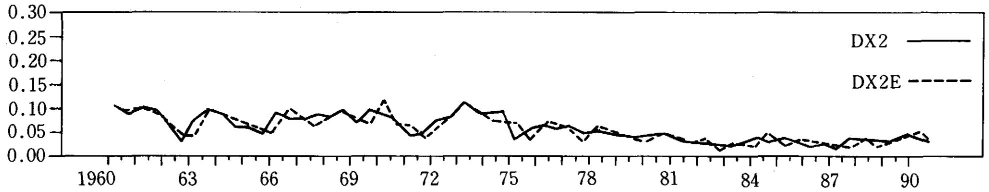
ここでは各変数に対する誤差修正項の効き具合いを統計的に検定することを考えよう。

14) これらの結果は、あるいは何らかの資産選択関係（関数）を反映しているのかもしれない。しかしながら、前述のように、ここでの目的は特定の経済理論の検証ではなく、また共和分の検定は対象となったデータの特性（前述のとおりマネーサプライと地価に関しては$I(2)$の可能性が認められる）にも左右されうるので、これ以上この問題には深入りしない。

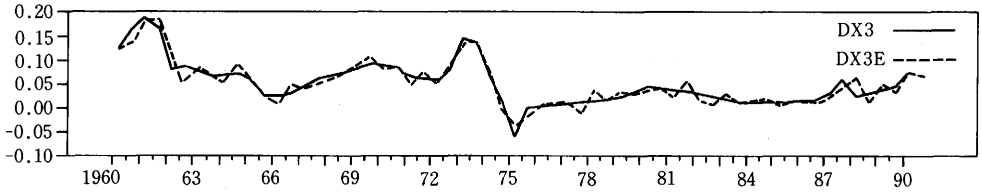
第２図　マネーサプライ（対数値の１階階差）



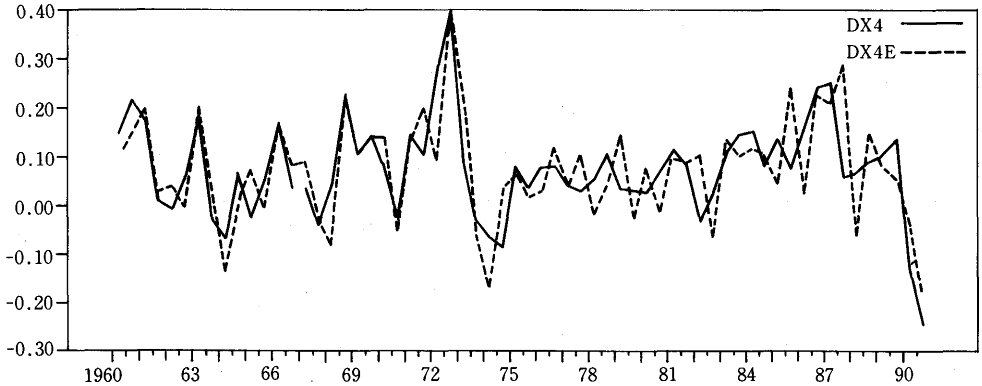
第３図　名目 GNP（対数値の１階階差）



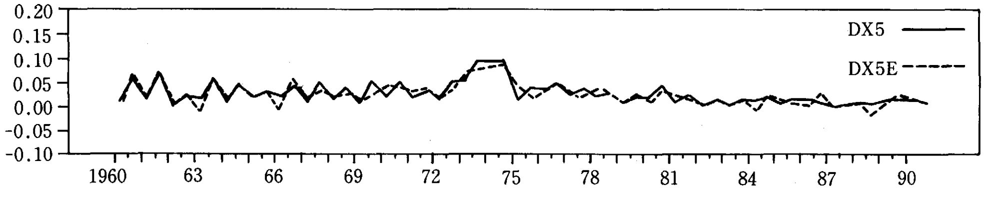
第４図　地　価（対数値の１階階差）



第５図　株　式（対数値の１階階差）



第６図　デフレータ（対数値の１階階差）



いま我々のモデルには、4個の誤差修正項が存在する。これらは行列 α の各行（送出ベクトル）によって重みづけられて、当該変数の現在の変動にフィードバック作用をもたらす。

もし行列 α のある行に関して「成分が全てゼロ」という仮説が受容されれば、その重みによってフィードバックを受けているはずのある変数に「誤差修正項の効果があるとは認められない」ということになろう。

具体的な方法は以下のとおりである。例えば、いま「第1変数が誤差修正項の影響を受けない」との仮説を考え、行列 α の第1行が全て0という制約を与えればよい。これはJohansenの記法に従えば

$$H_4: \alpha = A\psi, \ \Pi = \alpha\beta' = A\psi\beta'$$

と表現される。ただし、ここで、

$$A = \begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}, \ \psi = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix}$$

である。

この H_4 の下で、H_2 の場合と同様にして得られる固有値問題の解を

$$\hat{\lambda}_{4.1} > \cdots > \hat{\lambda}_{4.p}$$

とする。一方 H_2 の下での固有値を従来どおり

$$\hat{\lambda}_1 > \cdots > \hat{\lambda}_p$$

と書くことにすると、H_4 対 H_2 の検定統計量は

$$T \sum_{i=1}^{r} \ln \left\{ \frac{1 - \hat{\lambda}_{4.i}}{1 - \hat{\lambda}_i} \right\}$$

である。$A: p \times m$, $\psi: m \times r$ のときに、これが漸近的には自由度 $r(p-m)$ の χ^2 分布に従うことを利用して、制約の当否を検定する。

我々のケースでは、マネーから物価に至るまで、全ての場合でこのゼロ制約は有意水準1％で棄却された。つまり、どの変数にも誤差修正項は必要であり、毎期毎期安定的に「均衡関係」の近くに値が実現しているとみなせる変数は1つもなかったことになる。

なお、本論文では行わなかったが、共和分ベクトル β に対する制約の検定も可能である。これは推定された共和分関係について、より単純な形を考えてよいかどうかの検定に用いられる。例えば、2つの変数が共和分関係にあるときに、その係数同士が符号だけ反対で大きさは等しい、という仮説である。

(3) 手数項の扱いと補助的検定

今回のデータ解析では、パッケージRATS Ver. 3.1に予め用意されてあるプログラムを、本論文の目的に応じて修正の上で利用した。定数項の取扱いが重要であることは4.に述べたが、RATSに用意されているプログラムは、Johansen [1989] の公式によりECMで推定された定数項をドリフトと誤差修正項の平均とに分解する。したがって、算出された検定統計量はJohansen and Juselius [1990] のTable 1の有意点と比較すべきである、という点に注意する必要がある。

大部分の応用分析はこのTable 1を用いるケースであると思われる。しかし、ここではさらに、特殊な状況ではあるが形式上考えられる2つのケースについて、今回の分析がその当該ケースではなかったことを述べておく。特殊な2つのケースとは、第1に推定さ

れた定数項がゼロベクトルではないが全て誤差修正項の平均になっている場合であり、第2に推定された定数項がゼロベクトルである場合（線形トレンドが存在しない場合）である。

さて、「共和分が r 個で定数項が全て誤差修正項の平均」という仮説を Johansen にならって $H_2^*(r)$ と呼ぼう。第2の場合については、$H_2^*(r)$ 対 $H_2(r)$ の検定を実施する。この検定については補論1を参照されたい。我々のケースでは H_2^* は棄却され、推定された定数項は誤差修正項の平均を構成するとともに、部分的にはドリフト項であることが確認された。[15)]

H_2^* が棄却されたことは、必ずしも $\mu=0$ の棄却を意味しないので、念のため第2のケースも考慮しておく。この場合は、推定された定数項の t-統計量で有意性を検定すればよい。Johansen の方法は誤差修正モデルの推定を行うので、定数項について通常の回帰分析の枠組みでの検定は有効である。この結果、定数項の成分全てがゼロであるという仮説は棄却された。

6. いくつかの考察 — 結びに代えて

以上 Johansen の LR 検定の概要および一般によく用いられるその解釈について、例題的な解析を織りまぜながら解説してきた。そこで本節では、(1)ECM という解釈法、(2)統計的手法の側面、(3)実用上の問題点、の3つの角度から Johansen の LR 検定について考察を加え、結びとしたい。

（ECM という解釈と Johansen の LR 検定）

まず第1は、Johansen の LR 検定の結果として得られる相互依存関係の解釈に伴う問題である。

あるベクトル値時系列 x_t に対しスカラーベクトル β が存在し、$\beta' x_t=\epsilon_t \sim I(0)$ となる、ということが共和分の定義であった。単純に事実だけを述べれば、共和分とは「確率的トレンドの間の多重共線性（畠中[1991], p.212)」であり、共和分が何通り認められるかということは VAR におけるランク落ちの程度を示すものである。ECM では、この共和分関係において「平均的」に成り立っていると思われる、$\beta' x_t=c$ なる相互依存関係を取り出し、これを1つの均衡状態と思うことで「誤差修正」という解釈に到達した。この「誤差修正」という解釈は、モデルの形から生じるストレートな解釈であるとはいえ、いくつかの問題を含んでいる。

まず第1は、「誤差」に先立つ、非定常変数間の相互依存関係の解釈である。3.に述べたように ECM の文脈では、共和分から決定論的な共線関係を取り出して、それを変数間に本来成り立つべき安定的な関係であると想定して、「長期均衡関係」と解釈する。

だが、「均衡」にしても「長期」にしても、収束の概念を惹起しやすいので注意を要する。ここで「長期均衡関係」と呼ばれている

15) 推定された4通りの共和分関係と誤差修正項の平均を合わせて考えることで、「均衡」直線ともいうべきものを算出することが可能である。原理的には順次代入・消去を繰り返すだけであるが、ここでは次のような1次元の部分空間が得られる。

$$\frac{X_1-6.83}{4.91}=\frac{X_2-1.58}{3.72}=\frac{X_3-13.49}{4.18}=\frac{X_4-18.15}{7.40}=\frac{X_5}{1.00}$$

ものは、データを長期にわたって利用できる状況下で初めて観察可能な「平均的」関係を抽出したものである。ECMを用いたこれまでの応用分析をみる限りでは、3.に述べた紋切り型の解釈にとどまっており、背後にある「均衡」概念の脆弱性を問題にすることはほとんどないように思われる。そもそも(4)式における$\mathbf{\Pi} X_{t-k}$を「誤差修正項」と呼ぶことが適当かどうかも議論の余地があろう。現在の変動が、1期前ではなくk期前の「均衡」からの誤差の影響を受ける、と考えることが自然なことであろうか。1期前から誤差の影響を受ける形にECMを書き直すことは可能である。ラグ2期のVARを例にとれば

$$
\begin{aligned}
& X_t = \mathbf{\Pi}_1 X_{t-1} + \mathbf{\Pi}_2 X_{t-2} + \epsilon_t \\
\Rightarrow\ & \Delta X_t = (-\mathrm{I} + \mathbf{\Pi}_1)\Delta X_{t-1} \\
& \qquad - (\mathrm{I} - \mathbf{\Pi}_1 - \mathbf{\Pi}_2) X_{t-2} + \epsilon_t \\
& \qquad = -\mathbf{\Pi}_2 \Delta X_{t-1} - (\mathrm{I} - \mathbf{\Pi}_1 - \mathbf{\Pi}_2) X_{t-1} \\
& \qquad + \epsilon_t
\end{aligned}
$$

であるが、このような変形は、予測に貢献する可能性のあった直近k個の誤差修正項のうち$k-1$個までをインサンプルでの説明に費やすことになる。また、上の事実に加えて行列$\mathbf{\Pi}$の分解は基準化なしには不定であることから、ここでいうECMは、二重の意味で一意性を欠いている。

Johansen のLR検定は、経済データ分析の立場からすればECMの最尤推定法として位置づけることもできるが、これまで述べたようにECMという解釈の援用に無理が感じられる点も存在する。よってここでは、ECMとは切り離して以下のようにJohansenのLR検定を評価しておこう。Johansenの方法での根本的な関心事は、見せかけの回帰(spurious regression)ないし多重共線性の問題を避け、かつデータ変換に伴う情報の損失も避けて、非定常変数によるVARモデルを推定することである。モデルは1階階差データのVARに、さらに$\mathbf{\Pi} X_{t-k}$という変数間の1次従属性を情報として持つ項が加わっており、その分だけ当てはまりの向上が期待される。

(統計的手法の側面)

次に、JohansenのLR検定を統計的手法の側面から考察してみよう。

本論文で再三述べているように、Johansenの方法は、複数の共和分関係を検出できる点が最大の特色である。一言でいえば、JohansenのLR検定はGrangerの表現定理の枠組みをフルに生かすべくして考えられた、拡張方向としては極めて自然な手法である。

手法の大枠はJohansen [1988] で示されたが、さらにJohansen [1989] では、ECMの解釈、検定統計量の極限分布の2点にモデルの定数項が果たす意義が明らかにされ、さらにJohansen and Juselius [1989] において共和分ベクトル、送出ベクトルに関する制約の検定方法も提示され、これによって統計的手法としての実用性は完備されたと評価してよかろう。

Johansenのやり方で巧妙と思われるのは、検定を構成する際に、仮説の下での共和分ベクトルの推定量の分布を求めることを回避している点である。回帰理論に持ち込んで直接共和分ベクトルを相手にすれば、仮に共和分が存在する場合には、推定量の一致性は得られているものの、その分布は全く未知である(Stock [1987])。また共和分が存在しない場合であっても、推定量はゼロの周りに非正規で分布することが知られており(Phillips [1984]、Granger and Newbold [1974])、いず

れにせよこの種の議論に基づく検定の構成は、解析的に大きな困難を伴う。Johansen は ECM という定常性を満たすモデルを経由することでこの種の標準的な仮定の枠からはみだした回帰分析を回避し、正準相関分析の問題として共和分階数の検定を固有値問題に帰着させ、その副産物である固有ベクトルによって共和分ベクトルを推定している。

（実用上の問題点）

次に、Johansen の LR 検定を実行する際に生じる問題点を指摘しておこう。１つは Johansen の LR 検定の前提条件に関することであり、もう１つは検定それ自体に伴う問題点である。

Johansen の LR 検定を実行するには、全変数が $I(1)$であることが前提である。しかし、経済変数が真に $I(1)$プロセスであるかどうかは議論の分かれるところであろう。これまで提案されてきた確率過程の非定常性についての検定法の多くは、帰無仮説に非定常をとるので、Johansen の LR 検定には有利な形式といえるかもしれない。しかし、多くの実証論文に指摘されているように、構造変化を考慮に入れれば、必ずしもサンプル期間全てにわたって確率的トレンド（stochastic trend）を持つとは結論できないようである（例えば Takeuchi [1991] を参照)。Johansen の手法自体の問題ではないが、経済変数をどういったプロセスで特徴づけたらよいかは、難しい問題である。さらに、単位根を単根しか考慮しないことや、常に VAR というシステム的アプローチをとらねばならないことなども制約的に感じられるかもしれない。とくに VAR という定式化がデータの記述に不適当な場合には、Johansen 流の ECM を考えることは誤った結論を導く可能性が大きいといえよう。

このように前提条件にいくつかの問題を抱えつつも、Johansen の LR 検定は米国の実証分析の分野においてはすでに共和分検定の決定版ともいうべき信頼を寄せられているようである。しかしながら、漸近理論に基づく検定である以上、有限小標本でのパフォーマンスが問題となろう。この点を論じたものとして Podivinsky [1990] がある。彼は標本数が 50と100の２つの場合でシミュレーションを行い、Johansen [1988] の作成した有意点について調べている。これによると Johansen のトレース検定では、仮説 $r=0$（つまり共和分が存在しない）は概して棄却されやすく、$r\leq 1$, $r\leq 2$,…は受容されやすい傾向にあり、小標本においては検定のサイズが合わないことが指摘されている。

有限小標本でのサイズの不一致については、実証分析における Johansen の LR 検定の使用頻度を考えれば、今後さらに多くの研究がなされてしかるべきであると思われる。

補論 1.　Johansen の LR 検定の詳細

（手法の概要）

$X_t=(x_{1t},\cdots,x_{pt})'$ は p 変量ベクトルで、要素の全てが $I(1)$変数であるとする。このとき次のようなベクトル自己回帰モデルを考えよう。

$$X_t=\Pi_1 X_{t-1}+\cdots+\Pi_k X_{t-k}+\mu+\epsilon_t \quad \text{(A-1)}$$

ここで $X_t,\cdots,X_{t-k}$, μ, ϵ_t は $p\times 1$ ベクトル、$\Pi_1,\cdots,\Pi_k$ は $p\times p$ 行列を表すものとし、ラグ k は既知、ないし実際の分析においては残差がホワイトノイズであるとみなせるよう十分に大きくとるものとする。(A-1)式を変形して次の表現を得る。

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} - \Pi X_{t-k} + \mu + \epsilon_t \quad \text{(A-2)}$$

ただし、ここで

$$\Gamma_i \overset{\text{def}}{=} -\mathrm{I} + \Pi_1 + \cdots + \Pi_i, \ (i = 1, \cdots, k-1)$$
$$\Pi \overset{\text{def}}{=} \mathrm{I} - \Pi_1 - \cdots - \Pi_k$$

である。

Π の階数によって、プロセスは以下の3通りに分類される。

1. rank $\Pi = p$ (full rank)：X_t の成分は全て定常
2. rank $\Pi = r$ ($0 < r < p$)：r 個の共和分ベクトルが存在する
3. rank $\Pi = 0$：X_t の成分は「独立」な I(1)変数

システム内に r 個の共和分が存在すれば、Granger の表現定理から $\Pi = \alpha\beta'$ となる $p \times r$ 行列 α, β が存在し、(A-2)式は

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} - \alpha\beta' X_{t-k} + \mu + \epsilon_t \quad \text{(A-3)}$$

となる。$\beta' X_{t-k}$ は本論に解説したような意味での'error'であり、β'の行ベクトルを共和分ベクトル (cointegration vector) と呼ぶ。本論でも述べたように、α の行ベクトルは'error'をシステムに送り出すという意味で送出ベクトル (loading vector) と呼ばれている。

Johansen の LR 検定とは、行列 Π のランクを逐次的に検定することで共和分の個数を定める方法であり、以下の2つが考案されている。いずれも帰無仮説は「$H_2(r)$：r 個の共和分ベクトルが存在する」、すなわち rank $\Pi \leq r$ ($r = 0, \cdots, p-1$) であるが、対立仮説のとり方が異なっている。

1. 対立仮説としては共和分ベクトルの個数を考慮せず、H_1：rank $\Pi \leq p$ をとる。H_2、H_1 それぞれの下で推定したモデルの尤度比が検定統計量である。この検定方法をトレース検定 (trace test) という。
2. 対立仮説として H_2 ($r+1$)：rank $\Pi \leq r+1$ をとる。$H_2(r)$、$H_2(r+1)$ それぞれの下で推定したモデルの尤度比が検定統計量である。この検定方法を最大固有値検定 (maximal eigenvalue test) という。

これら2つの統計量の小標本分布は明らかでないが漸近分布は求められており、検定の名前は統計量の極限分布に由来する。以下では検定統計量を導くことを目標に、まず各仮説の下での尤度を求める。

（仮説 H_1 の下での尤度）

$$H_1: \Delta X_t = \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1} \Delta X_{t-k+1} - \Pi X_{t-k} + \mu + \epsilon_t \quad \text{(A-4)}$$

本節ではトレース検定の統計量の構成のため、まず対立仮説の下での尤度を示す。(A-4)式で Π のランクに何ら制約を設けずに（つまり rank $\Pi \leq p$ だと思って）推定したときの尤度を $L(H_1)$ としよう。ここで簡略化のため

$$Z_{0t} \overset{\text{def}}{=} \Delta X_t$$
$$Z_{1t} \overset{\text{def}}{=} (\Delta X'_{t-1}, \cdots, \Delta X'_{t-k+1}, 1)'$$
$$Z_{kt} \overset{\text{def}}{=} X_{t-k}$$

とし、さらに Z_{1t} に対応するパラメターをひとまとめにした $p \times \{p(k-1)+1\}$ 行列を Γ とする。これで(A-4)式は一気に簡明に表されて、

$$Z_{0t} = \Gamma Z_{1t} + \Pi Z_{kt} + \epsilon_t, \quad (t = 1, \cdots, T) \quad \text{(A-5)}$$

となる。

Π を固定すれば、Γ の最小二乗推定量は ϵ_t が正規分布という仮定の下、$Z_{0t}-\Pi Z_{kt}$ を Z_{1t} に回帰することで得られ、次の正規方程式を得る。

$$\sum_{t=1}^{T} Z_{0t}Z'_{1t}=\Gamma\sum_{t=1}^{T} Z_{1t}Z'_{1t}+\Pi\sum_{t=1}^{T} Z_{kt}Z'_{1t} \tag{A-6}$$

ここで簡略化のため積率行列を

$$M_{ij}\overset{\text{def}}{=}\frac{1}{T}\sum_{t=1}^{T} Z_{it}Z'_{jt}, \qquad (i,j=0,1,k) \tag{A-7}$$

と書く。これより正規方程式(A-6)は次のように書き直すことができる。

$$M_{01}=\Gamma M_{11}+\Pi M_{k1}$$

これを Γ について解けば

$$\hat{\Gamma}=M_{01}M_{11}^{-1}-\Pi M_{k1}M_{11}^{-1} \tag{A-8}$$

次に、ΔX_t を $\Delta X_{t-1}, \Delta X_{t-2}, \cdots, \Delta X_{t-k+1}$ に回帰したときの残差ベクトルを R_{0t}、X_{t-k} を ΔX_{t-1}, $\Delta X_{t-2}, \cdots$, ΔX_{t-k+1} に回帰したときの残差ベクトルを R_{kt} と定義しよう。これより

$$\hat{Z}_{0t}=\left(\sum_{t=1}^{T} Z_{0t}Z'_{1t}\right)\left(\sum_{t=1}^{T} Z_{1t}Z'_{1t}\right)^{-1}\cdot Z_{1t}$$
$$=M_{01}M_{11}^{-1}Z_{1t}$$

となり

$$R_{0t}=Z_{0t}-M_{01}M_{11}^{-1}Z_{1t} \tag{A-9}$$

同様に

$$R_{kt}=Z_{kt}-M_{k1}M_{11}^{-1}Z_{1t} \tag{A-10}$$

以上の議論から、Π を固定したとき、最終的に(A-6)式は

$$Z_{0t}-\hat{\Gamma}Z_{1t}-\Pi Z_{kt}=R_{0t}-\Pi R_{kt}$$

と表され、誤差項独立同一正規分布の仮定の下、Π、Λ の尤度関数は

$$|\Lambda|^{-T/2}\exp\left\{-\sum_{t=1}^{T}(R_{0t}-\Pi R_{kt})'\Lambda^{-1}\times(R_{0t}-\Pi R_{kt})/2\right\}$$

となる。ここでさらに $S_{ij}=(1/T)\Sigma_{t=1}^{T} R_{it}R'_{jt}\,(i,j=0,k)$ とおくと、

$$S_{ij}=M_{ij}-M_{i1}M_{11}^{-1}M_{1j}\ (i,j=0,k)$$

これを用いて Π、Λ の最尤推定量は

$$\hat{\Pi}=S_{0k}S_{kk}^{-1} \tag{A-11}$$

$$\hat{\Lambda}=S_{00}-S_{0k}S_{kk}^{-1}S_{k0} \tag{A-12}$$

と明示的に書ける。$\hat{\Gamma}$ については、$\hat{\Pi}$ を(A-8)式に代入すればよい。またこのときの尤度(の $-2/T$ 乗)は

$$L_{\max}^{-2/T}(H_1)=|\hat{\Lambda}|$$

である。インデックスに H_1 とあるように、ここまでは共和分の有無を考慮せずに(A-4)式を推定したときの推定量、およびモデルの尤度を求めた。

(仮説 H_2 の下での尤度)

検定統計量の構成には、仮説 H_2 の下での尤度 $L(H_2)$ がさらに必要である。再述すれば、トレース検定における帰無仮説は H_2：$\Pi=\alpha\beta'$、すなわち「$\mathrm{rank}\,\Pi\leq r$ $(r=0,\cdots,p-1)$ であり $p\times r$ 行列 α, β が存在して $\Pi=\alpha\beta'$」である。最大固有値検定での対立仮説 $H_2(r+1)$ の下での尤度は、本小節の議論から直ちに導かれる。前節の議論と全く同様に、今度は β を固定した上で $\hat{\alpha}$、$\hat{\Lambda}$ を β の関数として表す。

R_{0t} を $\beta' R_{kt}$ に回帰することにより、

$$\hat{\alpha}(\beta)=S_{0k}\beta(\beta' S_{kk}\beta)^{-1} \quad \text{(A-13)}$$

$$\begin{aligned}\hat{\Lambda}&=S_{00}-S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}\\&=S_{00}-\hat{\alpha}(\beta)(\beta' S_{kk}\beta)\hat{\alpha}(\beta)'\end{aligned} \quad \text{(A-14)}$$

を得る。一方尤度については、

$$\begin{aligned}L_{\max}^{-2/T}(\beta)&=|\Lambda(\beta)|\\&=|S_{00}-\hat{\alpha}(\beta)(\beta' S_{kk}\beta)\hat{\alpha}(\beta)'|\\&=\frac{|S_{00}||\beta'(S_{kk}-S_{k0}S_{00}^{-1}S_{0k})\beta|}{|\beta' S_{kk}\beta|}\end{aligned}$$

いま右辺の最小化が尤度最大化と同値であるから、ここで2次形式の比の最小化の議論を適用すればよい。この問題では、方程式

$$|\lambda S_{kk}-S_{k0}S_{00}^{-1}S_{0k}|=0 \quad \text{(A-15)}$$

を解けばよい。

(A-15)式を解いて得られる p 個の解（固有値）$\hat{\lambda}_1>\cdots>\hat{\lambda}_p$ の各々に対し固有ベクトルを求め、それを並べた $p\times p$ 行列 $\hat{V}=(\hat{v}_1, \cdots, \hat{v}_p)$ を、$\hat{V}' S_{kk}\hat{V}=I$ という規準を満たすように選ぶ（そうしなければ各 $\hat{v}$ については方向だけが決まって大きさが不定である）。

帰無仮説 (H_2) の下で共和分は r 個であり、$\hat{V}$ の列ベクトルのうち、大きい順にとった固有値 $\hat{\lambda}_1, \cdots, \hat{\lambda}_r$ に対応する固有ベクトルを並べて得られる $p\times r$ 行列が行列 β の最尤推定量である;

$$\hat{\beta}=(\hat{v}_1, \cdots, \hat{v}_r)$$

そしてこのときの尤度（の $-2/T$ 乗）は

$$L_{\max}^{-2/T}(H_2)=|S_{00}|\prod_{i=1}^{r}(1-\hat{\lambda}_i) \quad \text{(A-16)}$$

となる。

（検定統計量の構成）

前小節までの議論に基づき、トレース検定、最大固有値検定の統計量を導こう。

$$\begin{aligned}L_{\max}^{-2/T}(H_1)&=|\hat{\Lambda}|\\&=|S_{00}|\prod_{i=1}^{p}(1-\hat{\lambda}_i)\end{aligned}$$

であったから、H_2 対 H_1 の尤度比検定統計量は

$$-2\ln(Q;H_2|H_1)=-T\sum_{i=r+1}^{p}(1-\hat{\lambda}_i) \quad \text{(A-17)}$$

となる。同様に $H_2(r)$ 対 H_2 $(r+1)$ の尤度比検定統計量は

$$\begin{aligned}&-2\ln(Q;H_2(r)|H_2(r+1))=\\&-T\ln(1-\hat{\lambda}_{r+1})\end{aligned} \quad \text{(A-18)}$$

となる。

（統計量の極限分布と有意点）

Johansen [1989] の定理4.1によれば、統計量 (A-17) と (A-18) は、$T\to\infty$ のときに次の行列

$$\left\{\int_0^1 (dU)F'\left[\int_0^1 FF'dt\right]^{-1}\int_0^1 F(dU)'\right\} \quad \text{(A-19)}$$

のトレースと最大固有値とにそれぞれ弱収束する。ただし、ここで、$U(t)=\{U_1(t), \cdots, U_{p-r}(t)\}$ は $p-r$ 次元ブラウン運動、$F(t)=\{F_1(t), \cdots, F_{p-r}(t)\}$ は各要素が

$$F_i(t)=U_i(t)-\int_0^1 U_i(s)ds, \quad (i=1, \cdots, p-r-1)$$

$$F_i(t)=t-\frac{1}{2}, \quad (i=p-r)$$

で定義される $p-r$ 次元の確率過程である。

極限分布がブラウン運動に依存しているため、実際に Johansen の LR 検定を行うときにはシミュレーションによる Johansen and

Juselius [1990] の表を用いる(数表は区間[0, 1]上のブラウン運動を非常に細かいステップのランダム・ウォークで近似して作成されている)。暗に線形トレンドが存在すると思って検定を行った場合、用いるべき表は Johansen and Juseius [1990] の Table A1 である。

(定数項の分解)

(A-3)式において推定された μ は、実はシステムのドリフト項であるのか、それとも誤差修正項の平均であるのか分からない。この点については4.(2)ですでに述べたので、ここでは(9)式でいう μ_0 と β_0 の推定法を結果だけ紹介する。

Johansen [1989] は、定数項 μ の有する意味を明示的に考察した上で次の分解公式を導き出した。未知母数の推定が一旦済めば、

$$\mu_0=\beta_\perp[\alpha'_\perp\Pi_1(1)\beta_\perp]^{-1}\alpha'_\perp\mu \qquad \text{(A-20)}$$

によって μ_0 が推定される。ここで $\beta'_\perp$ と $\alpha_\perp$ は $p\times(p-r)$ 行列であり、以下の諸式を満たす。

$$\beta'_\perp\beta=0$$

$$\alpha'_\perp\alpha=0$$

$$\alpha'_\perp\Lambda\alpha_\perp=I$$

$$\alpha_\perp\alpha'_\perp=\Lambda^{-1}[I-\alpha(\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1}]$$

$$\Pi_1(1)=I-\sum_{i=1}^{k-1}\Gamma_i-(k-2)\Pi$$

このとき β'_0 は、

$$\beta'_0=(\alpha'\alpha)^{-1}\alpha'\left\{I+\left[\sum_{i=1}^{k-1}\Gamma_i-I\right]\right.$$

$$\left.\times\beta_\perp[\alpha'_\perp\Pi_1(1)\beta_\perp]^{-1}\alpha'_\perp\right\}\mu \qquad \text{(A-21)}$$

と表される。

実際の分析に当たっては、この定数項の分解は大変重要である。なぜなら、複数の「長期均衡関係」から「均衡部分空間」を抽出したとき、それが空間の原点を必ず通るという保証はない。2つの変数に共和分の関係が認められるとき、それが原点を通る直線になることは稀で、むしろ切片が存在する方が自然であろう。

逆に、予備検定等によって非定常性が確認されている変数であっても、データのプロットをみる限りでは線形トレンドが存在しそうにない、と判断できる場合もあろう。この場合には、「μ として推定したものが実は全て μ_1 である」という仮説を考えるのが適当であろう。それが Johansen and Juselius [1990] でいうところの仮説 H_2^* である。

(仮説 H_2^*)

$$H_2^*:\Pi=\alpha\beta',\ \mu=\alpha\beta'_0$$

この仮説の後半の意味するところは、モデルの特定化の段階でトレンド項として入れてあったものが、実は送出ベクトルの張る空間に含まれている、ということである。

いま誤差修正項とトレンド項を取り出してみると、帰無仮説 H_2^* の下、新たに $\beta^*\overset{\text{def}}{=}(\beta',\ \beta'_0)'$, $X^*_{t-k}\overset{\text{def}}{=}(X'_{t-k},\ 1)'$ によって

$$\alpha\beta'X_{t-k}+\mu=\alpha\beta'X_{t-k}+\alpha\beta'_0$$

$$=\alpha\beta^{*\prime}X^*_{t-k}$$

と表すことができる。このとき当初線形トレンドとして特定化した μ は誤差修正項に吸収されてしまい、モデルに線形トレンドは存在しない。

ところで、Johansen and Juselius [1990] の例では、予めデータをみた段階で H_2^*、H_2 のどちらを考えるべきなのかを決定し、H_2^* 対

H_2の検定は、その*a priori*な選択が妥当であることの補強作業の意味合いが強いように思われる。実際のところ経済データに限っていえば、モデルH_2^*に基づいて分析しなければならないことは稀であろう。マネー、GNPといった重要な変数は明らかに線形トレンドを持っており、このような変数を選択する限りはμがベクトルとしてゼロというケースは考えにくい。

検定$H_2^*(r)$対$H_2(r)$に用いる統計量を構成するには、まず$X_{t-k}^{*\prime}=Z_{kt}^{*\prime}$とし、それに応じて$S_{ij}$を$S_{ij}^*$に置き換えて定義し直した固有方程式を解く。その解を$\hat{\lambda}_i^*$と記すと、検定統計量は

$$-T\sum_{i=r+1}^{p}\ln\left\{\frac{1-\hat{\lambda}_i^*}{1-\hat{\lambda}_i}\right\}$$

であり、これは漸近的に$\chi^2(p-r)$に従う。

ところで、仮説H_2^*の下で共和分の個数を検定するときには、統計量の極限分布が（仮説H_2の下での分布とは）異なってくることに注意しなければならない。行列(A-19)でFを

$$F_i(u)=U_i(u),\quad (i=1,\cdots,p-r)$$
$$F_i(u)=1,\qquad (i=p-r+1)$$

で定義し直した$(p-r+1)\times(p-r+1)$行列のトレース、最大固有値が検定統計量の漸近分布になる。シミュレーションによる分布の有意点はJohansen and Juselius [1990]のTable 3に与えられている。

本論文の分析がそうであるように、経済分析であれば大概の場合「定数項の分解」によってμが2つに分かれるのが自然であって、補強的に用いる以外に検定H_2^*を考えなければならないことは稀であろう。ただ、Johansen and Juselius [1990]には、H_2^*を考えるのが妥当な状況としてデンマークの経済データを用いた4変数モデルによる分析例が挙げられている。

補論2.　ブラウン運動と不変原理について

Johansenの結論は、それまでに数多く行われてきた非定常時系列モデルの推定論の上に成り立っている。それらの結論に共通していることだが、JohansenのLR検定も最終的にはブラウン運動に依存する極限分布を利用する。[16)]

本節の目的は、1次元の場合に限定してランダム・ウォーク過程の極限として直観的にブラウン運動を理解することである。この種の極限定理は「不変原理（invariance principles)」と呼ばれ、1次元の場合は「Donskerの不変原理」と呼ばれる。一方JohansenのLR検定は多変量自己回帰モデルに基づくものであるから、多変量版の不変原理が必要である。これについては、誤差項にどの程度の相関や分散不均一性まで許してよいかを示しつつ、Phillips and Durlauf [1986]にTheorem 2.1として与えられているので参

16)　ここでは標準ブラウン運動を単にブラウン運動と呼ぶことにする。確率過程$\{B(t)、t\geq 0\}$が標準ブラウン運動に従うとは、(1)$B(0)=0$、(2)$0\leq t_1<t_2<\cdots<t_n$に対し、$B(t_2)-B(t_1), \cdots, B(t_n)-B(t_{n-1})$は独立、(3)$B(t)-B(s)$（但し$t\geq s$）は、平均0、分散$t-s$の正規分布に従う、という3つの条件を満たすときをいう。

照されたい。[17] このときの収束先は多次元のブラウン運動となる。

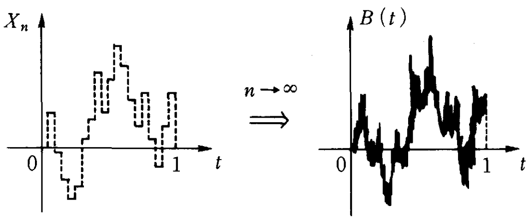
(Donsker の不変原理)

$\{\xi_n\}$を同一分布に従う独立確率変数列で、$E(\xi_1)=0$, $0<\sigma^2=Var(\xi_1)<\infty$ とし、$S_n=\xi_1+\cdots+\xi_n (n\geq 1)$, $S_0=0$ とおく。$[0,1]$区間上の連続関数空間 $C=C[0,1]$ に値をとる確率変数 $X_n (n\geq 1)$ を、

$$X_n(t,\ \omega)=\frac{1}{\sigma\sqrt{n}}S_{[nt]}(\omega)+(nt-[nt])\times\frac{1}{\sigma\sqrt{n}}\xi_{[nt]+1}(\omega),\quad 0\leq t\leq 1 \quad \text{(A-22)}$$

と定める。ただし、$[nt]$ は nt の整数部分を表す。1階のランダム・ウォーク・プロセスを無限 MA 表現することを考えれば、S_i のような部分和の極限を考える必要性は明らかである。

Donsker の不変原理によれば、$n\to\infty$ のとき、X_n の確率分布は(1次元の、つまり $C[0,1]$ 上の)ブラウン運動に弱収束する。



これを

$$X_n(t)\Rightarrow B(t)\quad \text{as}\quad n\to\infty \quad \text{(A-23)}$$

と書く (Billingsley[1968], §10)。

$X_n(t)$は、$i=1,\cdots,n$ に対し $X_n(i/n,\omega)=S_i/\sigma\sqrt{n}$ であり、ω を固定すれば各区間 $[\frac{i-1}{n},\frac{i}{n}]$ において線形である(第2項に注目)。Donsker の不変原理とは、$[0,1]$ 区間上のブラウン運動がランダム・ウォーク過程の極限として得られることを示している。

$X_n(t)$は, $[0,1]$ 区間で連続になるように構成されているのだが、非定常時系列の推測理論への応用では、次の形の random function に対する不変原理を直接考えることが多い。

$$X'_n(t,\omega)=\frac{1}{\sigma\sqrt{n}}S_{[nt]}(\omega),\ 0\leq t\leq 1 \quad \text{(A-24)}$$

X_n が折れ線の連続関数であるのに対し、X'_n は ω を固定したときには $[0,1]$ 上の階段関数(つまり不連続関数)になっている(上の図を参照)。このとき X'_n の確率分布も($C[0,1]$ 上の)ブラウン運動に収束する。すなわち、

$$X'_n(t)\Rightarrow B(t)\quad \text{as}\quad n\to\infty \quad \text{(A-25)}$$

である[18] (Billingsley[1968], §16)。

ここまで述べた確率論での結果を利用し、

17) 従属確率変数列に対する不変原理は、Billingsley[1968]の後半に詳しく論じられているが、その後の確率論研究者の成果も踏まえつつ、Phillips[1987]、Phillips and Durlauf[1986]は、一貫して誤差項に弱い相関や分散不均一性の存在を許している。しかし、Johansen の LR 検定では誤差項に正規分布が仮定されているので、本論文ではこれらの点については触れない。

18) 実は(35)式は、空間 C における収束より微妙な問題を多く含んでいるが、空間 D $[0,1]$ における位相の入れ方、確率測度の収束については、Billingsley[1968]§14−§16を参照されたい。

規準化された部分和 $X'_n(t)$ についての連続汎(はん)関数の収束先として、非定常時系列モデルにおける各種統計量の極限分布が求められる。理論的帰結については、Phillips [1986]、Phillips and Durlauf [1986] 等を参照されたい。

以上

[東京大学大学院経済学研究科
(現統計数理研究所助手)]

【参考文献】

畠中道雄、『計量経済学の方法』、創文社、1991年

吉田知生、「通貨需要関数の安定性をめぐって—ECM による計測」、『金融研究』第8巻第3号、日本銀行金融研究所、1989年10月

Billingsley, P., *Convergence of Probability Measures*, Wiley, 1968.

Box, G. E. P., and G. M. Jenkins, *Time Series Analysis: Forecasting and Control, 2nd ed.*, Holden-Day, 1976.

Campbell, J. Y., and P. Perron, "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots," *NBER Macroeconomic Annual*, 1992, pp. 141-219.

Dickey, D. A., and W. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica* 49, 1981, pp. 1057-72.

Diebold, F. X., and M. Nerlove, "Unit Roots in Economic Time Series: A Selective Survey," in T. B. Fomby, and G. F. Rhodes, eds., *Advances in Econometrics*, Vol. 8, Co-Integration, Spurious Regressions, and Unit Roots, Greenwich: JAI Press, 1990, pp. 3-69.

Engle, R., and C. W. J. Granger, "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica* 55, 1987, pp. 251-76.

Granger, C. W. J., "Co-Integrated Variables and Error-Correcting Models," University of California San Diego Discussion Paper, 1983.

———, and P. Newbold, "Spurious Regressions in Econometrics," *Journal of Econometrics* 2, 1974, pp. 111-20.

Johansen, S., "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control* 12, 1988, pp. 231-54.

———, "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," mimeo, Institute of Mathematical Statistics, University of Copenhagen, 1989.

———, and K. Juselius, "Maximum Likelihood Estimation and Inference on Cointegration — with Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52, 1990, pp. 169-210.

Nelson, C. R., and C. I. Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics* 10, 1982, pp. 139-62.

Phillips, P. C. B., "Time Series Regression with a Unit Root, " *Econometrica* 55, 1987, pp. 277-301.

———, "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics* 33, 1986, pp. 314-40.

———, and S. N. Durlauf, "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies* 53, 1986, pp. 473-95.

———, and P. Perron, "Testing for Unit Root in Time Series Regression," *Biometrika* 75, 1988, pp. 335-46.

Podivinsky, J. M., "Testing Misspecified Cointegrating Relationships," Working Paper No. 19/90, Department of Econometrics, Faculty of Economics Commerce & Management, Monash University, 1990.

Stock, J., "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica* 55, 1987, pp. 1035-56.

Takeuchi, Y., "Trends and Structural Changes in Macroeconomic Time Series," *Journal of Japan Statistical Society* 21, 1991, pp. 13-25.