

# 共和分分析

一橋大学大学院経済学研究科 田中 勝人

## 1. はじめに

本稿のメイン・テーマである共和分 (cointegration) は、和分過程と呼ばれる非定常過程に従う時系列間の特殊な関係を表す概念である。和分過程とは、階差を取れば定常過程となるような非定常過程である。このとき、共和分とは、このような和分過程に従う時系列のある線形結合が定常過程になることをいう。

和分過程の線形結合は、一般には和分過程であるので、共和分は例外的な現象である。経済学的には、線形結合は長期的な均衡関係からの乖離を表しており、共和分は、その乖離が安定的で定常的であることを意味する。また、統計学的には、共和分は、線形結合が非定常から定常に退化した 1 次従属関係であるから、確率的な多重共線と解釈することができる。

共和分の例として、消費  $C_t$  と所得  $Y_t$  の時系列の間で回帰関係

$$C_t = \alpha + \beta Y_t + u_t$$

を考えよう。これは、経済学では消費関数と呼ばれる確率モデルである。ここで、消費と所得がともに和分過程であるとしよう。このとき、誤差項  $u_t$  が定常となるような  $\alpha$  と  $\beta$  が存在すれば、消費と所得は共和分関係をもつことになり、上の回帰式は意味があることになる。他方、どのように  $\alpha$  と  $\beta$  を選んでも、誤差項  $u_t$  が非定常のままであれば、上の回帰式は無意味である。共和分とは、簡単にいえば、和分過程の変数間で意味のある回帰が存在する場合である、ということができよう。

一般に、和分過程が  $d$  回の差分ではじめて定常となるならば、 $I(d)$  過程という。特に、 $I(1)$  過程に従う時系列を**単位根系列** (unit root series) という。単位根系列は、和分過程の中で最も頻繁に観察される時系列である。したがって、典型的な共和分分析とは、単位根系列間の回帰分析である、ということができる。

単位根系列の概念は、計量経済学の分野において、非定常性の源泉であるトレンドとの関連で 1970 年代後半に生み出されたものである。それ以前は、トレンドといえば、時間の関数で表されるような**確定的トレンド**だけを意味していた。しかし、単位根系列の登場により、単位根系列に固有の**確率的トレンド**が脚光を浴びることになった。

非定常時系列が、確定的トレンドをもつのか、あるいは確率的トレンドをもつのかを区別することは重要である。確定的トレンドをもつ時系列を分析するには、時間を説明変数とする回帰を考えればよい。他方、確率的トレンドをもつ時系列に時間のトレンドをあてはめても無意味であるが、統計的にはもっともらしい見せかけの回帰（spurious regression）が得られる。しかし、結局のところ、それは誤った結論を導くことになる。

確定的トレンドと確率的トレンドの違いは、経済政策へのインプリケーションにおいても重要な違いをもたらす。前者においては、時系列の将来に強い影響を及ぼすのは時間的な確定要因のみである。それに対して、後者においては、時系列に内在する確率的要因の影響がいつまでも持続して、その変動を不確定なものにする。トレンドが確定的か確率的かを統計的に調べるには、**単位根検定**（unit root test）を行う必要があり、そのために、さまざまな検定方式が提案されている。

共和分分析は、単位根検定において確率的トレンドを含む単位根系列とみなされた複数の時系列間の回帰関係を分析するものであり、1980 年代後半から始まった。そして、単位根系列に関する統計的推測が通常とは異なる理論に従い、実際的な観点からも異なる意味合いをもつことが明らかになった。その結果、単位根検定と共和分分析は、理論および実証両面において、計量経済学におけるホットなトピックとなった。そのインパクトは、1950 年代に盛んに研究された同時方程式モデルの推定問題を凌ぐほどに強く、共和分分析は計量経済学の新たな重要分野として受け入れられることとなった。

共和分分析がもたらした代表的な貢献としては、すでに述べたことも含めて次のことが挙げられる。

- (i) 経済時系列の非定常性として、確率的トレンドの概念を導入した。
- (ii) 実際の時系列が確率的トレンドを含むかどうかを調べるために、単位根検

定を生み出した。

- (iii) 確率的トレンドを含む時系列の統計的性質は、定常的な場合とは非常に異なることを理論的に明らかにした。
- (iv) 共和分関係がない場合でも、変数間の回帰はもっともらしく見え、見せかけの回帰となりうる理由を解明した。
- (v) 共和分関係がない場合には、階差変換した定常系列による分析は正当化される。しかし、共和分関係がある場合には、階差変数のほかに、共和分を反映したレベル変数を使わなければ特定化の誤りをもたらすことを示した。
- (vi) 和分過程に従う任意の複数時系列間に、1次独立な均衡関係はいくつ存在するか、すなわち、1次独立な共和分関係の最大数はいくつであるかを推測する方式を確立した。

以下では、単位根検定を出発点として共和分分析に至る非定常的な世界に読者を誘い、統計学的な観点から上記のことを説明しつつ、この世界を概観したい。そして、同時に、残された課題を検討して、今後の研究に資することとしたい。

## 2. 和分と共和分

この章は本稿の導入部であり、問題意識と諸概念を説明する。具体的には、まず、経済時系列に特有な非定常性を例示し、そのような時系列を分析するための基本的なモデルとして、ランダム・ウォークを取り上げる。さらに、その拡張である和分過程を導入する。そして、非定常な和分過程の世界が定常的な場合とどのように違うのかについて、統計量の分布の観点から概説する。最後に、本稿のメイン・テーマである非定常時系列間の回帰関係である共和分の明確な定義と概念について述べる。

### 2.1 経済時系列の非定常性

経済時系列は非定常的なものが多い。例えば、GDP（国内総生産）などのように、生産に関連する時系列は、確定的なトレンドに沿って推移し、トレンドからの乖離も時間的に一定でないのが普通である。

しかし、経済時系列がもたらす非定常性は、必ずしも確定的なトレンドを伴

うわけではない。例として、TOPIX (東証株価指数) と為替レート of データを取り上げよう。図 2-1 には、これらの月次データ 114 個 (1992/1 - 2001/6 の各月平均) の対数値系列が示されている。また、その自己相関係数が図 2-2 に示されている。

図 2-1                      図 2-2

図 2-1 の対数値系列を見るかぎり、どちらの系列も一定値のまわりで変動し、ばらつきも時間とともに変化しているようには思えない。しかし、図 2-2 に示してあるように、自己相関係数は減衰の程度がゆるやかであり、定常性からの乖離がうかがえる。

実際、これらの系列に AR (Autoregressive: 自己回帰) モデルをあてはめると、AIC (赤池情報量選択規準) により、次の結果が得られる。

$$\text{TOPIX: } y_t = 0.31 + 1.22 y_{t-1} - 0.26 y_{t-2}$$

$$\begin{aligned} \text{為替レート: } y_t = & 0.10 + 1.35 y_{t-1} - 0.44 y_{t-2} + 0.22 y_{t-3} \\ & - 0.38 y_{t-4} + 0.13 y_{t-5} + 0.08 y_{t-6} \end{aligned}$$

一般に、AR( $p$ ) モデルは

$$(1) \quad y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

のように表現される。ここで、i.i.d.( $0, \sigma^2$ ) は、平均 0、分散  $\sigma^2$  の独立同一分布を意味する。このとき、特性方程式

$$\phi(x) = 1 - \phi_1 x - \dots - \phi_p x^p$$

において、 $\phi(1) = 0$  ならば、AR( $p$ ) モデルは**単位根** (unit root) をもつといい、非定常となる。上で推定された AR(2) および AR(6) モデルも、 $\phi(1)$  の値が 0 に近いことから、ほぼ単位根をもつことがわかる。

対象とする時系列が、実際に単位根をもつかどうかは、単位根検定により統計的に判断すべきことである。また、単位根検定により、TOPIX と為替レートが単位根系列とみなされた場合、これらの系列間の関係の有無を判断することも必要になる。これは、単位根系列間の回帰分析、すなわち、共和分分析にほかならない。

しかし、単位根の世界では、通常の場合とは異なり、相関係数をはじめとするさまざまな検定統計量が特有の分布に従うなど、定常的な世界では起こりえない現象が起きるために、分析の際に注意すべき事項がいくつかある。以下では、このことを念頭において、単位根の世界へ分け入ることにしたい。

## 2.2 ランダム・ウォークと I(1) 系列

単位根をもつ場合の最も単純なモデルは、**ランダム・ウォーク**

$$(2) \quad x_t = x_{t-1} + \varepsilon_t = x_0 + \varepsilon_1 + \cdots + \varepsilon_t, \quad (t = 1, \dots, T)$$

である。ここで、誤差項の系列  $\{\varepsilon_t\}$  は、i.i.d.  $(0, \sigma^2)$  に従う。また、初期値  $x_0$  は確率的であっても非確率的でもよい。ただし、確率的な場合には、 $\{\varepsilon_t\}$  とは独立であると仮定される。

ランダム・ウォークは、異時点間で独立な誤差項が同じウェイトで集積されたものであり、過去からの影響は時点にかかわらず同等である。このようなランダム・ウォークの期待値は、各  $t$  について  $E(x_t) = E(x_0)$  であるが、**自己共分散**は

$$\begin{aligned} (3) \quad \text{Cov}(x_s, x_t) &= \text{Cov}(x_0 + \varepsilon_1 + \cdots + \varepsilon_s, x_0 + \varepsilon_1 + \cdots + \varepsilon_t) \\ &= V(x_0) + E[(\varepsilon_1 + \cdots + \varepsilon_s)(\varepsilon_1 + \cdots + \varepsilon_t)] \\ &= V(x_0) + \sigma^2 \min(s, t) \end{aligned}$$

となる。これより、 $x_t$  の分散は、 $V(x_t) = V(x_0) + \sigma^2 t$  となり、時間とともに比例的に増大して行くので、ランダム・ウォークは非定常過程である。

ところで、(2) 式のランダム・ウォーク  $\{x_t\}$  に対して、1 回の階差を適用すると、

$$(4) \quad (1 - L)x_t = \Delta x_t = \varepsilon_t, \quad (t = 1, \dots, T)$$

を得る。ここで、 $L$  は**ラグ・オペレータ**と呼ばれ、 $Lx_t = x_{t-1}$ 、 $L^j x_t = x_{t-j}$  のように、時点を戻す作用素である。他方、 $\Delta = 1 - L$  は**階差オペレータ**と呼ばれる作用素である。

ランダム・ウォークの階差系列は i.i.d. であるが、それを拡張して、従属的な定常過程を考えることもできる。すなわち、原系列  $\{y_t\}$  の階差系列  $\{u_t\}$  が

$$(5) \quad u_t = \Delta y_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} = \alpha(L) \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

と表される場合を考えよう。ここで、 $\{\alpha_j\}$  は定数の係数列で、

$$(6) \quad \sum_{j=0}^{\infty} j|\alpha_j| < \infty, \quad \alpha_0 = 1, \quad \sum_{j=0}^{\infty} \alpha_j \neq 0$$

をみたすものとする。また、 $\alpha(L)$  は、**ラグ多項式**

$$\alpha(L) = 1 + \alpha_1 L + \alpha_2 L^2 + \cdots = \sum_{j=0}^{\infty} \alpha_j L^j$$

である。(6)の最初の条件は、 $\{u_t\}$ の定常性を保証する条件であるが、そのためだけであれば、係数の2乗和が収束すること、すなわち、 $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ で十分である。しかし、ここでは、2.4節で展開する議論のためにより強い条件が課されている。他方、(6)の最後の条件は、 $\alpha(1) \neq 0$ と同値であり、 $\alpha(L)$ が因数 $\Delta$ をもたない( $\alpha(x) = 0$ が単位根をもたない)ことを意味する。これは、(5)のモデルが退化することを排除する条件である。

以上より、 $u_t = \Delta y_t$ は平均0、自己共分散

$$(7) \quad \gamma_h = \text{Cov}(u_t, u_{t-h}) = \sigma^2 \sum_{j=0}^{\infty} \alpha_j \alpha_{j+h}, \quad (h \geq 0)$$

の定常過程となる。一般に、(5)のように $\text{MA}(\infty)$ で表され、しかも、 $\alpha(1) \neq 0$ (MA部分に単位根がない)であるような定常確率過程 $\{u_t\}$ は定常な**線形過程**と呼ばれる。

線形過程の例としては、Box-Jenkins (1976)の $\text{ARMA}(p, q)$ 過程がある。ARMAは、Autoregressive Moving Averageの省略形であり、 $\text{ARMA}(p, q)$ 過程は、

$$(8) \quad \phi(L)u_t = \theta(L)\varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

で定義される。ここで、

$$\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p, \quad \theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$$

である。階差系列 $\{\Delta y_t\}$ が $\text{ARMA}(p, q)$ 過程に従うとき、原系列 $\{y_t\}$ は $\text{ARIMA}(p, 1, q)$ 過程に従うという。ARIMAは、Autoregressive Integrated Moving Averageの省略形である。ランダム・ウォークは、 $\text{ARIMA}(0, 1, 0)$ である。

(8) の ARMA( $p, q$ ) 過程が定常となる条件は、AR 部分の特性方程式  $\phi(x) = 0$  の根が、すべて単位円外にあることである。このとき、MA( $\infty$ ) 表現

$$(9) \quad u_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} = \alpha(L) \varepsilon_t$$

が可能であり、係数列  $\{\alpha_j\}$  は指数的に減少することが知られているので、(6) の最初の条件は自動的にみたされる。

他方、MA 部分の特性方程式  $\theta(x) = 0$  の根がすべて単位円外にあるとき、ARMA( $p, q$ ) 過程は**反転可能**であるという。反転可能な場合には、(8) の  $u_t$  は AR( $\infty$ ) 表現が可能となる。

1 回の階差系列が定常な線形過程となるならば、原系列は I(1) 過程に従う。ARMA( $p, 1, q$ ) 過程は、I(1) の特別な場合であり、ランダム・ウォークは最も単純な I(1) である。

これに対して、(5) で表される確率過程で、MA 部分に単位根をもたない定常過程は I(0) であるという。定常かつ反転可能な ARMA( $p, q$ ) 過程は I(0) の特別な場合である。なお、MA 部分に単位根を 1 つもつような反転不可能な定常過程は I(-1) であるという。

**図 2-3** には、TOPIX と為替レート対数値の階差系列が示されている。対数値系列が単位根をもつ場合には、この階差系列は I(0) であると考えられる。単位根をもたないならば、階差系列は I(-1) となる。この点の実証については、あとで行うことにする。

**図 2-3**

なお、I(1) 系列を拡張して、任意の非負整数  $d$  に対して  $d$  回の階差ではじめて定常となる I( $d$ ) 系列を考えることができる。実際、I(2) 系列は現実に観察される。当然ながら、 $d$  が大きいほど非定常性は強くなる。しかし、現実の経済時系列では  $d$  が 2 以上の場合は例外的であり、以下の議論を複雑にするので、ここでは I(1) 系列だけを扱うことにする。

## 2.3 原系列と階差系列

Box-Jenkins (1976) が提唱した ARIMA( $p, d, q$ ) モデルによる分析は、3 段階の手続きとフィードバックを経て適切なモデルを決めようというものであ

る．すなわち，まず，階差の次数  $d$  の値だけでなく，AR 部分の次数  $p$  や，MA 部分の次数  $q$  の値を特定化 (identification) した上で，パラメータを推定 (estimation) し，モデルのよさを診断 (diagnostic checking) する．そして，診断の結果が悪ければ，フィードバックして，特定化，推定，診断を繰り返す．最終的に適切なモデルを選択する．このようなモデル・ビルディングの発想は，きわめて明快，単純であり，1970 年代におけるコンピュータの進展という追い風を受けてパッケージ化され，多くの学問分野で広く使われるようになった．

ところで，Box-Jenkins の方法は，次のことを仮定している．

- (i) 非定常とみなせる系列は，AR 部分のラグ多項式が単位根をもつ．すなわち，非定常性は，確定的トレンドではなく，系列そのものが生み出す確率的トレンドに起因する．
- (ii) 階差変換しても，原系列に関する情報は失われない．
- (iii) 複数の時系列を扱う場合，系列相互間の関係は，原系列と階差変換後の系列で不変である．

これに対して，原系列そのものを分析すべきである，という立場から次のような批判が寄せられた．

- (iv) 非定常とみなせる系列において，非定常性の要因であるトレンドが確定的か確率的かは自明なことではなく，統計的に判断すべき問題である．また，トレンドが確率的であっても，本当に単位根をもつかどうかは検定すべき問題である．
- (v) 階差変換は単位根があることを前提にしているもので，そうでない場合は，階差変換後の系列を分析することは，原系列に関して誤った情報をもたらす．
- (vi) 複数の時系列を扱う場合には，たとえ，それぞれの時系列が単位根をもっていたとしても，階差変換後の系列を分析することは情報の損失をもたらす場合がある．それは，本稿のテーマである共和分の関係が成立する場合である．

以上の理由により，非定常時系列を分析する場合には，階差変換後の系列ではなく，まず，原系列そのものを分析するのが望ましい．このような立場から，次節では， $I(1)$  系列を中心にして，その統計的な性質，特に漸近的な性質を述べることにしたい．

## 2.4 和分過程の統計量の分布

原系列  $\{y_t\}$  のデータ生成過程が  $I(1)$ ，すなわち，

$$(10) \quad \Delta y_t = u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j} = \alpha(L) \varepsilon_t, \quad y_0 = 0, \quad (t = 1, \dots, T)$$

であるとする．初期値の条件  $y_0 = 0$  は，任意の値（ただし， $\sqrt{T}$  のオーダーよりも小）としても以下の議論は変わらないが，煩雑さを避けるために仮定した．

誤差項  $\{u_t\}$  に対して，**長期分散**

$$\sigma_L^2 = \lim_{T \rightarrow \infty} \frac{1}{T} V \left( \sum_{t=1}^T u_t \right) = \sigma^2 \left( \sum_{j=0}^{\infty} \alpha_j \right)^2 = \sum_{j=-\infty}^{\infty} \gamma_u(j)$$

を定義する．ここで， $\gamma_u(j)$  は  $\{u_t\}$  のラグ  $j$  の自己共分散である．なお， $\sigma_L^2$  が長期分散と呼ばれる由来は， $\{u_t\}$  の**スペクトラム**を  $f_u(\lambda)$  とするとき，

$$f_u(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_u(j) e^{-ij\lambda} = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} \alpha_j e^{-ij\lambda} \right|^2, \quad (i = \sqrt{-1})$$

であるから， $\sigma_L^2 = 2\pi f_u(0)$  は，周波数  $0$ （＝最長周期）におけるスペクトラムの値  $f_u(0)$  に比例することによる．

長期分散に対応して， $\{u_t\}$  の分散  $\gamma_u(0)$  は，**短期分散**と呼ばれる．誤差項の系列  $\{u_t\}$  が i.i.d. の場合には，長期分散と短期分散は一致する．一般には， $\{u_t\}$  のラグ  $j$  の自己相関を  $\rho_u(j)$  とすれば，

$$\frac{\sigma_L^2}{\gamma_u(0)} = \sum_{j=-\infty}^{\infty} \rho_u(j) = 1 + 2 \sum_{j=1}^{\infty} \rho_u(j)$$

となるので，正の相関が強い場合は長期分散の方が短期分散よりも大きくなる傾向がある．例えば， $\{u_t\}$  が，係数  $\phi$  の  $AR(1)$  に従う場合には，上の比は  $(1 + \phi)/(1 - \phi)$  となり，この値は  $\phi$  が正の場合に大きくなる．

さて、(10) の生成過程に従う  $\{y_t\}$  の漸近的な挙動を調べるために、連続時間で定義された確率過程

$$(11) \quad Y_T(r) = \frac{1}{\sqrt{T}\sigma_L} y_{[Tr]} = \frac{1}{\sqrt{T}\sigma_L} \sum_{t=1}^{[Tr]} u_t, \quad (0 \leq r \leq 1)$$

を考えよう。ここで、 $[Tr]$  は、 $Tr$  を越えない最大整数を表す。  $Y_T(r)$  は、関数空間  $D[0, 1]$  (区間  $[0, 1]$  で定義される右連続で左極限をもつような実数値関数全体の空間) 上の確率過程であり、線形過程  $\{u_t\}$  の部分和過程 (partial sum process) と呼ばれる。図 2-4 に図示されているように、 $Y_T(r)$  はジャンプをもつ階段関数である。なお、 $D[0, 1]$  は、連続性や完備性などを扱うために Skorohod 位相が与えられ、距離空間となる (Billingsley(1968, Chap. 4))。

図 2-4

部分和過程の定義域は連続な集合である。ただし、そのパスは不連続である。これに対して、定義域もパスも連続な確率過程として、標準ブラウン運動 (standard Brownian motion) を定義する。それは、 $[0, 1]$  上で定義され、次の条件をみたす確率過程  $\{W(r)\}$  である。

- (i)  $P(W(0) = 0) = 1$
- (ii) 任意の時点  $0 = r_0 < r_1 < \dots < r_n \leq 1$  に対して、増分  $W(r_i) - W(r_{i-1})$ ,  $(i = 1, \dots, n)$  は互いに独立に  $N(0, r_i - r_{i-1})$  に従う。

図 2-5 には、標準ブラウン運動が図示されている。この図を描くために、次の展開表現を利用した (Chan-Wei (1988))。

$$(12) \quad W(r) = \sum_{n=1}^{\infty} \frac{\sqrt{2} \sin\left(n - \frac{1}{2}\right) \pi r}{\left(n - \frac{1}{2}\right) \pi} Z_n, \quad \{Z_n\} \sim \text{i.i.d.} N(0, 1)$$

ブラウン運動のパスは、かなりジグザグな形状であるが、その理由は微分不可能なことによる。実際、平均変化率

$$\frac{W(r+h) - W(r)}{h} = \frac{W(r+h) - W(r)}{\sqrt{h}} \frac{1}{\sqrt{h}}$$

において、 $h \rightarrow 0$  のとき、右辺の第 1 因数は  $N(0,1)$  に従う確率変数であるが、第 2 因数が発散、したがって全体として発散してしまうことになる。

図 2-5

図 2-5 の標準ブラウン運動は，図 2-4 の部分和過程とは乖離があるが， $T \rightarrow \infty$  のとき，次の定理が成立する (Billingsley (1968, p. 137), Phillips (1987)) .

**定理 1.** (11) で定義された部分和過程  $\{Y_T(r)\}$  は， $T \rightarrow \infty$  のとき，連続なパスをもつ標準ブラウン運動  $\{W(r)\}$  に分布収束する．この分布収束を，

$$Y_T \Rightarrow W$$

と表す．

この定理は， $[0, 1]$  上で定義された確率過程  $\{Y_T(r)\}$  のパス全体の分布収束を与えるものである．時間  $r$  を固定して考えると， $Y_T(r) \rightarrow N(0, r)$  のように，通常の変数に対する CLT (中心極限定理) となる．これに対して，定理 1 は関数空間上の確率過程に関する定理であるので，**FCLT** (Functional Central Limit Theorem: 汎関数中心極限定理) と呼ばれている．

定理 1 が示唆することの第 1 の点は，I(1) 系列の分布がブラウン運動の分布で近似できる，ということである．さらに，より強い性質として，I(1) 系列のパスに関してもブラウン運動による近似が可能である．すなわち， $\{\varepsilon_t\}$  が  $p$  ( $> 2$ ) 次の絶対モーメントをもつならば，

$$\sup_{0 \leq r \leq 1} T^\delta \left| \frac{1}{\sqrt{T} \sigma_L} y_{[Tr]} - W(r) \right| \rightarrow 0 \quad \left( 0 < \delta = \frac{1}{2} - \frac{1}{p} < \frac{1}{2} \right)$$

が確率 1 で成り立つようなブラウン運動を構成することができる (Csörgő-Horváth (1993)) .

第 2 の点は，ランダム・ウォークと I(1) 系列は，長期分散の値を除けば，漸近的には同じものと考えられる，ということである．このことを示すために，線形過程  $u_t = \alpha(L) \varepsilon_t$  において， $\alpha(L)$  を次のように変形する．

$$\alpha(L) = \alpha(1) + \alpha(L) - \alpha(1) = \alpha(1) + (1 - L) \tilde{\alpha}(L)$$

ここで，

$$\tilde{\alpha}(L) = \sum_{j=0}^{\infty} \tilde{\alpha}_j L^j, \quad \tilde{\alpha}_j = - \sum_{l=j+1}^{\infty} \alpha_l$$

ラグ多項式  $\alpha(L)$  を上のように分解することを **B-N 分解** (Beveridge-Nelson decomposition) という (Phillips-Solo (1992)) . B-N 分解を適用することにより,

$$(13) \quad u_t = \alpha(L) \varepsilon_t = [\alpha(1) + (1-L)\tilde{\alpha}(L)] \varepsilon_t = \alpha(1) \varepsilon_t + \tilde{u}_t - \tilde{u}_{t-1}$$

を得る. ただし,  $\tilde{u}_t = \tilde{\alpha}(L)\varepsilon_t$  である. ここで, (6) の最初の条件から,

$$\sum_{j=0}^{\infty} |\tilde{\alpha}_j| = \sum_{j=0}^{\infty} \left| - \sum_{l=j+1}^{\infty} \alpha_l \right| \leq \sum_{j=0}^{\infty} \sum_{l=j+1}^{\infty} |\alpha_l| = \sum_{l=1}^{\infty} \sum_{j=0}^{l-1} |\alpha_l| = \sum_{l=1}^{\infty} l |\alpha_l| < \infty$$

を得るから,  $\{\tilde{u}_t\}$  は定常な線形過程となる.

以上のことから,

$$(14) \quad y_t = \sum_{j=1}^t u_j = \alpha(1) \sum_{j=1}^t \varepsilon_j + \tilde{u}_t - \tilde{u}_0$$

となり, I(1) 系列  $\{y_t\}$  は, 最右辺の第 1 項, すなわち, ランダム・ウォークの  $\alpha(1)$  倍にほぼ等しくなる. ここで, 定数  $\alpha(1)$  は長期分散に関連したパラメータである.

和分過程の統計量の漸近分布を扱うためには, FCLT とともに, 次の **CMT** (Continuous Mapping Theorem: 連続写像定理) も重要である (Billingsley (1968, p. 30)).

**定理 2.** 関数空間  $D[0, 1]$  上の確率過程  $\{Y_T(r)\}$  に対して,  $T \rightarrow \infty$  のとき,  $Y_T \Rightarrow W$  とする. このとき,  $D[0, 1]$  で定義された任意の連続関数  $h(y)$  に対して,  $T \rightarrow \infty$  のとき, 次のことが成り立つ.

$$h(Y_T) \Rightarrow h(W)$$

CMT の応用として,  $\{y_t\}$  が (10) の I(1) 過程にしたがうとき,  $\sum_{t=1}^T y_t^2$  の漸近分布を考えてみよう. まず, (11) を使うことにより,

$$\begin{aligned} S_T &= \frac{1}{T^2 \sigma_L^2} \sum_{t=1}^T y_t^2 = \frac{1}{T} \sum_{t=1}^T Y_T^2 \left( \frac{t}{T} \right) \\ &= \sum_{t=1}^T \int_{(t-1)/T}^{t/T} Y_T^2 \left( \frac{t}{T} \right) dr = \int_0^1 Y_T^2(r) dr \end{aligned}$$

を得る．ここで，

$$h(y) = \int_0^1 y^2(r) dr, \quad y \in D[0, 1]$$

は， $y$  の連続関数であるから，FCLT と CMT を使うことにより，次の結果を得る．

$$(15) \quad S_T = h(Y_T) \quad \Rightarrow \quad h(W) = \int_0^1 W^2(r) dr \equiv S$$

なお，極限確率変数  $S$  は，分布の意味で等しい次の表現も可能である．

$$(16) \quad S = \sum_{n=1}^{\infty} \frac{Z_n^2}{\left(n - \frac{1}{2}\right)^2 \pi^2} = \int_0^1 \int_0^1 (1 - \max(r, s)) dW(r) dW(s)$$

最初の等号は (12) の関係を使って示すことができる．このことから，極限分布の特性関数は，

$$\begin{aligned} \phi(\theta) &= E \left[ \exp \left\{ i\theta \int_0^1 W^2(r) dr \right\} \right] = E \left[ \exp \left\{ i\theta \sum_{n=1}^{\infty} \frac{Z_n^2}{\left(n - \frac{1}{2}\right)^2 \pi^2} \right\} \right] \\ &= \prod_{n=1}^{\infty} \left( 1 - \frac{2i\theta}{\left(n - \frac{1}{2}\right)^2 \pi^2} \right)^{-1/2} = \left( \cos \sqrt{2i\theta} \right)^{-1/2} \end{aligned}$$

となる．(16) の最後の等号は，

$$W(r) = \int_0^r dW(s)$$

を (15) の  $W(r)$  に代入して，積分順序を交換すれば得られる．分布の同等性は，特性関数が一致することにより示すこともできるが，ここでは省略する (Tanaka (1996, Chap. 5) を参照されたい) ．

## 2.5 和分から共和分へ

今までは，単一の  $I(1)$  系列を考えてきたが，ここでは2つの  $I(1)$  系列  $\{x(t)\}$  と  $\{y(t)\}$  を考えよう．これらの系列は，次のような2次元の  $I(1)$  過程に従うものとする．

$$(17) \quad \Delta z_t = \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} = C(L) \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, I_2)$$

ここで， $I_2$  は2次元の単位行列を表す．

他方,  $C_j$  は係数行列で,

$$(18) \sum_{j=0}^{\infty} j \|C_j\| < \infty, \quad C(1) = \sum_{j=0}^{\infty} C_j = \begin{pmatrix} A'_1 \\ A'_2 \end{pmatrix} \neq \begin{pmatrix} 0' \\ 0' \end{pmatrix}$$

をみたすものとする.  $\|C_j\|$  は, 行列  $C_j$  のノルムを表す. また,  $A'_i$  は, 行列  $C(1)$  の第  $i$  行ベクトルである. なお,  $C(1)$  に関する最後の条件は,  $C(L)$  の各要素が行ごとに因数  $1-L$  をもつことを排除するものである.

以上の設定のもとで, ラグ多項式  $C(L)$  の B-N 分解により, (17) 式は

$$(19) \Delta z_t = [C(1) + \Delta \tilde{C}(L)] \varepsilon_t, \quad \tilde{C}(L) = \sum_{j=0}^{\infty} \tilde{C}_j L^j, \quad \tilde{C}_j = - \sum_{l=j+1}^{\infty} C_l$$

と表現される. これより,

$$(20) z_t = z_0 + C(1) \sum_{j=1}^t \varepsilon_j + \tilde{C}(L) \varepsilon_t - \tilde{C}(L) \varepsilon_0$$

が得られる. ここで,  $\tilde{C}(L) \varepsilon_t$  は, (18) の最初の条件より定常過程となる.

このとき, 2次元の確率過程に対する次の FCLT が成り立つ (Chan-Wei (1988)).

$$Z_T \Rightarrow C(1)W, \quad Z_T(r) = \frac{1}{\sqrt{T}} z_{[Tr]}$$

今の場合,  $W = \{W(r)\}$  は 2次元の標準ブラウン運動である. すなわち, 互いに独立な 2つの標準ブラウン運動を成分にもつ確率過程である. そして, CMT を使うことにより,  $\{z_t\}$  の 2次モーメントに関して次の分布収束が成立する.

$$(21) \frac{1}{T^2} \sum_{t=1}^T z_t z_t' \implies C(1) \int_0^1 W(r) W'(r) dr C'(1) = C(1) H C'(1)$$

ここで,  $2 \times 2$  行列

$$H = \int_0^1 W(r) W'(r) dr$$

は, 確率 1 で正値定符号である.

他方, 行列  $C(1)$  は, 行列式  $D = |C(1)|$  が 0 のとき正則でなくなり,  $\{z_t\}$  の漸近的な 2次モーメント行列  $C(1) H C'(1)$  はランクが 1 に落ちることになる ( $C(1) \neq 0$  であるから, 0 に落ちることはない). したがって, 2つの非定常な系列  $\{x_t\}$  と  $\{y_t\}$  の間に確率的な線形関係が存在することになる.

以上のことは、次のようにも解釈することができる。今、(19)の最初の式の両辺に左から行ベクトル  $\alpha' = (-\beta, 1)$  をかけると、

$$(22) \quad \Delta(y_t - \beta x_t) = (A_2' - \beta A_1') \varepsilon_t + \Delta g'(L) \varepsilon_t$$

を得る。ただし、 $g'(L) = \alpha' \tilde{C}(L)$  である。ここで、 $C(1)$  が正則ならば、 $A_1'$  と  $A_2'$  は 1 次独立であるから、上の右辺第 1 項は 0 ベクトルとはならず、 $y_t - \beta x_t$  は  $I(1)$  のままである。

しかし、 $C(1)$  が正則でない場合は、 $\beta = A_2' A_1 / A_1' A_1$  とすれば、(22) の右辺第 1 項が 0 ベクトルとなり、両辺の因数  $\Delta$  がキャンセルされる。したがって、次の関係

$$(23) \quad y_t = \beta x_t + g'(L) \varepsilon_t$$

が得られる。すなわち、2 つの  $I(1)$  系列の 1 次結合  $\alpha' z_t = y_t - \beta x_t$  が定常系列となる。

このように、2 つの  $I(1)$  系列の 1 次結合が定常系列となるとき、これらの  $I(1)$  系列は**共和分**の関係にあるといい、係数ベクトル  $\alpha$  は**共和分ベクトル**と呼ばれる。共和分が成立するための条件は、長期分散に関連した行列  $C(1)$  が正則でないことである。 $I(1)$  系列間に成立する (23) の回帰を**共和分回帰**という。

共和分は、任意の複数個の  $I(1)$  系列に対しても定義される。すなわち、 $\{z_t\}$  が  $m$  次元の  $I(1)$  系列であるとき、1 次結合  $\alpha' z_t$  が定常ならば、 $z_t$  の各成分は共和分の関係にあるという。そのための条件は、行列  $C(1)$  が正則でないことである。特に、 $C(1)$  のランクが  $r$  ( $0 < r < m$ ) ならば、1 次独立な共和分ベクトルは  $m - r$  個存在する。この場合の値  $m - r$  を**共和分ランク**という。

### 3. 単位根検定

原系列が  $I(1)$  であるかどうかは先験的に決められることではなく、データから検証する必要がある。そのための検定は、本稿の冒頭でも述べたように、単位根検定と呼ばれる。この章では、単位根検定に使われるモデルを説明した上で、さまざまな単位根検定の方法について考える。

#### 3.1 検定問題の定式化と検定方式

単位根検定に使われる基本的なデータ生成過程は、AR タイプの

$$(24) \quad (1 - \rho L)(y_t - \alpha - \beta t) = \varepsilon_t, \quad y_0 = 0, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

であり、これは、

$$(25) \quad y_t = \alpha(1 - \rho) + \beta\rho + \beta(1 - \rho)t + \rho y_{t-1} + \varepsilon_t$$

のように書き換えることができる。

ここで、確定的トレンド  $\alpha + \beta t$  のパラメータを制約することにより、次の3つのモデルを考えることができる。

$$\text{モデル A: } y_t = \rho y_{t-1} + \varepsilon_t$$

$$\text{モデル B: } y_t = a + \rho y_{t-1} + \varepsilon_t$$

$$\text{モデル C: } y_t = a + bt + \rho y_{t-1} + \varepsilon_t$$

モデル A は確定的トレンドが全く存在しない場合 ( $\alpha = \beta = 0$ )、モデル B は定数項のみが存在する場合 ( $\beta = 0$ )、モデル C は制約のない一般的な場合である。

これらのいずれのモデルに対しても、検定問題は

$$(26) \quad H_0: \rho = 1 \quad \text{vs.} \quad H_1: \rho < 1$$

である。これは、「単位根あり」の帰無仮説を、「単位根なし」の対立仮説に対して検定するものである。ここで注意すべきことは、確定的トレンドを考慮したモデル C においても、 $H_0$  のもとでは、(25) 式からわかるように、 $b = 0$  となる。ただし、定数項  $a$  は任意である。すなわち、確率的トレンドが存在する  $H_0$  のもとでは確定的トレンドは存在しない。他方、確率的トレンドが存在しない  $H_1$  のもとでは  $b \neq 0$  となるから、確定的トレンドが存在する、という二分法が成り立っている。

もちろん、最初から線形トレンドはないとするモデル A と B では、確率的トレンドの有無だけが問題となる。ただし、モデル B の場合、 $H_0$  のもとでは定数項  $a$  は 0 に制約されることがわかる。

以上のモデルに対して、Dickey-Fuller (1979) は、2 種類の検定方式を提案した。1 つは、 $\rho$  の LSE (最小 2 乗推定量)  $\hat{\rho}$  に基づく係数検定、もう 1 つは  $\hat{\rho} - 1$  の  $t$  値に基づく  $t$  検定である。どちらの検定も、統計量の値が小さいときに  $H_0$  を棄却する。これらは **DF 検定** と総称されている。

ところで、検定を実行するためには帰無分布を導出する必要があるが、DF 検定統計量の帰無分布は、正規分布や  $t$  分布などの既知の分布には従わない。特に、標本サイズ  $T$  が大きくなるときは、前章で説明した標準ブラウン運動  $\{W(r)\}$  を含む分布に従う。

例えば、係数検定の統計量  $T(\hat{\rho} - 1)$  の帰無分布は次のように導出できる。モデル A の場合には、同等な次のモデル

$$(27) \quad \Delta y_t = \delta y_{t-1} + \varepsilon_t, \quad \delta = \rho - 1, \quad y_0 = 0, \quad (t = 1, \dots, T)$$

に変形する。 $\delta$  の LSE を  $\hat{\delta}$  とすれば、

$$\hat{\delta} = \hat{\rho} - 1 = \frac{\sum_{t=2}^T y_{t-1} \Delta y_t}{\sum_{t=2}^T y_{t-1}^2}$$

であり、次の分布収束が成り立つことが示される (Phillips (1987))。

$$A: \quad T(\hat{\rho} - 1) \Rightarrow \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W^2(r) dr} = \frac{U_A}{V_A}$$

ここで、極限確率変数の分子  $U_A$  は、**伊藤積分**による表現であり、

$$U_A = \int_0^1 W(r) dW(r) = \frac{1}{2} (W^2(1) - 1)$$

が成り立つ。

モデル B の場合は、 $H_0$  のもとでは  $a = 0$  であることに注意しながら、次の結果を得ることができる。なお、 $a = 0$  でなければ、 $\hat{\rho}$  は正規性に従うことが示される。

$$B: \quad T(\hat{\rho} - 1) \Rightarrow \frac{\begin{vmatrix} 1 & W(1) \\ \int_0^1 W(r) dr & \int_0^1 W(r) dW(r) \end{vmatrix}}{\begin{vmatrix} 1 & \int_0^1 W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 W^2(r) dr \end{vmatrix}} = \frac{U_B}{V_B}$$

ここで、

$$U_B = \int_0^1 W_{(2)}(r) dW(r), \quad V_B = \int_0^1 W_{(2)}^2(r) dr$$

$$W_{(2)}(r) = W(r) - \int_0^1 W(s) ds$$

である。  $W_{(2)}(r)$  は、区間  $[0, 1]$  上で  $W(r)$  を定数に回帰したあとの残差過程であり、**平均調整されたブラウン運動** (demeaned Brownian motion) と呼ばれる。

モデル C の場合は、  $H_0$  のもとでは  $b = 0$  となることに注意しながら、次の結果を得ることができる。なお、  $b = 0$  でなければ、  $\hat{\rho}$  は正規性に従うことが示される。

$$C: \quad T(\hat{\rho} - 1) \Rightarrow \frac{\begin{vmatrix} 1 & 1/2 & W(1) \\ 1/2 & 1/3 & \int_0^1 r dW(r) \\ \int_0^1 W(r) dr & \int_0^1 r W(r) dr & \int_0^1 W(r) dW(r) \end{vmatrix}}{\begin{vmatrix} 1 & 1/2 & \int_0^1 W(r) dr \\ 1/2 & 1/3 & \int_0^1 r W(r) dr \\ \int_0^1 W(r) dr & \int_0^1 r W(r) dr & \int_0^1 W^2(r) dr \end{vmatrix}} = \frac{U_C}{V_C}$$

ここで、

$$U_C = \int_0^1 W_{(3)}(r) dW(r), \quad V_C = \int_0^1 W_{(3)}^2(r) dr$$

$$W_{(3)}(r) = W_{(2)}(r) - 12 \left( r - \frac{1}{2} \right) \int_0^1 \left( s - \frac{1}{2} \right) W(s) ds$$

である。  $W_{(3)}(r) = W(r) - \hat{\kappa} - \hat{\lambda}r$  は、区間  $[0, 1]$  上で  $W(r)$  を直線  $\kappa + \lambda r$  に回帰したあとの残差過程であり、**トレンド調整されたブラウン運動** (detrended Brownian motion) と呼ばれる。ここで、

$$\hat{\kappa} = \int_0^1 (4 - 6r) W(r) dr, \quad \hat{\lambda} = \int_0^1 (12r - 6) W(r) dr$$

である。実際、  $\hat{\kappa}$  と  $\hat{\lambda}$  は、正規方程式

$$\int_0^1 W(r) dr = \hat{\kappa} \int_0^1 dr + \hat{\lambda} \int_0^1 r dr$$

$$\int_0^1 r W(r) dr = \hat{\kappa} \int_0^1 r dr + \hat{\lambda} \int_0^1 r^2 dr$$

の解である。

**図 3-1** には、3つのモデル A, B, C における  $T(\hat{\rho} - 1)$  の極限帰無分布の密度関数が描かれている。正規分布とは異なり、全体的に左にゆがんだ形状を

していることが見てとれる。また、これらの分布の分位点やモーメントなどの特性値を表 3-1 に示しておいた。密度関数や分位点などの計算方法については、Nabeya-Tanaka (1990a), Tanaka (1996, Chap. 6) を参照されたい。なお、有限標本における  $T(\hat{\rho}-1)$  の帰無分布は、シミュレーションに基づく結果が Fuller (1996) に数表化されている。

図 3-1      表 3-1

他方、 $\hat{\rho}-1$  の  $t$  値  $t_{\hat{\rho}-1}$  は、 $\hat{\delta}$  の  $t$  値  $t_{\hat{\delta}}$  に等しく、次の形をしている。

$$t_{\hat{\rho}-1} = \frac{\hat{\rho}-1}{s.e.(\hat{\rho})} = \frac{\hat{\delta}}{s.e.(\hat{\delta})} = t_{\hat{\delta}}$$

ここで、 $s.e.(\hat{\rho})$  は回帰分析とのアナロジーで、 $\hat{\rho}$  を回帰係数の LSE とみなしたときの標準誤差を与えるものである。モデル A では、

$$s.e.(\hat{\rho}) = \frac{\hat{\sigma}}{\sqrt{\sum_{t=2}^T y_{t-1}^2}}, \quad \hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=2}^T (y_t - \hat{\rho} y_{t-1})^2$$

であり、モデル B, C についても同様に定義される。そして、 $t$  統計量に関して、次の分布収束が成り立つ。

$$\text{モデル M (M=A, B, C): } t_{\hat{\rho}-1} = t_{\hat{\delta}} \Rightarrow \frac{U_M}{\sqrt{V_M}}$$

これら 3 つのモデルに対して、 $t$  統計量の極限帰無分布を正確に求めることは困難である。ただし、モーメントは正確に計算することが可能である。表 3-2 には、Fuller (1996) においてシミュレーションで計算された分位点、および Nabeya (1999) において数値積分で計算されたモーメントが示されている。 $T(\hat{\rho}-1)$  の分布とは異なり、歪度が正となっていることから、少し右にゆがんだ形状となっていることがわかる。有限標本における  $t$  統計量の帰無分布も、Fuller (1996) で報告されている。

表 3-2

### 3.2 I(1) 性の検定

前節で考えた単位根検定は、AR(1) の係数が 1 かどうかの検定であり、一般の I(1) 性を検定するためには制約的である。ここでは、I(1) 性に関する 2 通りの検定を考える。

まず、第 1 の方法は、データ生成過程として、

$$(28) \quad y_t = \alpha + \beta t + \eta_t, \quad \phi(L) \eta_t = \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

を考える。ここで、 $\phi(L)$  は  $p$  次のラグ多項式であり、 $\phi(x) = 0$  の根の絶対値は、すべて 1 以上で、絶対値が 1 となるものは高々 1 個で、その根は 1 に限られるものとする。

このとき、B-N 分解の考え方を使って、 $\phi(L)$  は、

$$(29) \quad \begin{aligned} \phi(L) &= 1 - \phi_1 L - \dots - \phi_p L^p = \phi(1) L + \Delta \psi(L) \\ &= \phi(1) L + (1 - L) \left( 1 - \psi_1 L - \dots - \psi_{p-1} L^{p-1} \right) \end{aligned}$$

のように表現することができる。ここで、 $\psi(x) = 0$  の根の絶対値はすべて 1 より大きく、

$$\psi_j = - \sum_{i=j+1}^p \phi_i, \quad (j = 1, \dots, p-1)$$

である。

さらに、(29) を使って、(28) は次のように表現することができる。

$$(30) \quad \Delta y_t = a + b t + \delta y_{t-1} + \psi_1 \Delta y_{t-1} + \dots + \psi_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

ここで、 $\delta = -\phi(1)$  である。また、

$$a = \phi(1) \alpha + \beta \sum_{j=1}^p j \phi_j, \quad b = \beta \phi(1)$$

である。

前置きが長くなったが、I(1) 性の検定には (30) のモデルを使うことになる。このモデルは、 $p = 1$  の場合、前節で説明した単位根モデルに帰着することを確かめられたい。そこで、前と同様にして、確定的トレンド  $\alpha + \beta t$  の有無によって、モデルとして次の 3 つを考えることができる。

$$\text{モデル A: } \Delta y_t = \delta y_{t-1} + \psi_1 \Delta y_{t-1} + \dots + \psi_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

$$\text{モデル B: } \Delta y_t = a + \delta y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

$$\text{モデル C: } \Delta y_t = a + bt + \delta y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

このとき、I(1) 性の検定問題は、対立仮説を片側として

$$H_0 : \delta = 0 \quad \text{vs.} \quad H_1 : \delta < 0$$

とすることができる。  $H_0$  のもとでは、  $\phi(1) = 0$  であるから、(28), (29) より

$$\Delta \psi(L)(y_t - \alpha - \beta t) = \varepsilon_t$$

となり、  $\{y_t\}$  は I(1) となる。他方、  $H_1$  のもとでは、  $\phi(1) > 0$  であり、このことは、仮定より、  $\phi(x) = 0$  の根の絶対値がすべて 1 より大きいことと同等となるから、  $\{y_t - \alpha - \beta t\}$  が I(0) となる。

以上の設定のもとで、係数  $\delta$  の LSE  $\hat{\delta}$  に基づく検定と、  $\hat{\delta}$  の  $t$  値に基づく検定を考えることができる。これらは **ADF 検定** (Augmented DF test) と総称される。

まず、係数検定は、  $T\hat{\delta}$  が小さいときに  $H_0$  を棄却するものである。ただし、その極限帰無分布は、

$$\text{モデル M (M=A, B, C): } T\hat{\delta} \Rightarrow \psi(1) \times \frac{U_M}{V_M}$$

となることが示される (Fuller (1996, Chap. 10)). ここで、  $U_M/V_M$  は前節で考察したモデルの場合の極限である。今の場合、極限確率変数が未知パラメータ  $\psi(1) = 1 - \psi_1 - \cdots - \psi_{p-1}$  に依存するので、その一致推定量  $\hat{\psi}(1)$  を求めて、検定統計量として  $T\hat{\delta}/\hat{\psi}(1)$  を使えば、前節の結果をそのまま使うことができる。一致推定量としては、

$$\hat{\psi}(1) = 1 - \hat{\psi}_1 - \cdots - \hat{\psi}_{p-1}$$

を使うことができる。ここで、  $\hat{\psi}_i$  は、各モデルから得られる  $\psi_i$  の LSE である。

他方、  $t$  値に基づく検定は、統計量として  $t_{\hat{\delta}}/s.e.(\hat{\delta})$  を使えば、前と同様の帰無分布が導出される (Fuller (1996)). したがって、ADF 検定としては  $t$  値に基づく検定を使うことが多い。

**図 2-1** の TOPIX と為替レートの対数値系列では、モデル C を使った検定で次の結果を得た。

	$p$ の値	$T \hat{\delta} / \hat{\psi}(1)$	$t$ 値
TOPIX	2	-6.8	-2.65
為替レート	6	-14.7	-1.92

係数検定の 10% 点は、表 3-1 から  $-18.245$ 、 $t$  検定の 10% 点は、表 3-2 から  $-3.13$  であるから、いずれの時系列においても、10% の有意水準で、単位根の帰無仮説は受容されることがわかる。

I(1) 性を検定する第 2 の方法は、モデルとして

$$(31) \quad (1 - \rho L)(y_t - \alpha - \beta t) = u_t = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

を考えるものである。ここで、誤差項  $\{u_t\}$  は定常な線形過程である。

確定的トレンドの有無に依存して 3 つのモデルが得られるが、ここでは  $\alpha = \beta = 0$  の最も単純なモデル

$$(32) \quad \Delta y_t = \delta y_{t-1} + u_t, \quad y_0 = 0, \quad (t = 1, \dots, T)$$

を考えよう。

このとき、 $\delta$  の LSE  $\hat{\delta}$  の極限帰無分布は、次のようになることが示される (Phillips (1987))。

$$(33) \quad T \hat{\delta} = \frac{\sum_{t=2}^T y_{t-1} \Delta y_t / T}{\sum_{t=2}^T y_{t-1}^2 / T^2} \Rightarrow \frac{\int_0^1 W(r) dW(r) + \lambda}{\int_0^1 W^2(r) dr}$$

ここで、

$$\lambda = \frac{1}{2} \left( 1 - \frac{\sigma_S^2}{\sigma_L^2} \right), \quad \sigma_S^2 = \sigma^2 \sum_{j=0}^{\infty} \alpha_j^2, \quad \sigma_L^2 = \sigma^2 \left( \sum_{j=0}^{\infty} \alpha_j \right)^2$$

である。 $\sigma_S^2$  と  $\sigma_L^2$  は、それぞれ、 $\{u_t\}$  の短期、長期分散である。(33) の結果は、次の分布収束と CMT (連続写像定理) から導出することができる。

$$\begin{aligned} \frac{1}{T^2} \sum_{t=2}^T y_{t-1}^2 &\Rightarrow \sigma_L^2 \int_0^1 W^2(r) dr \\ \frac{1}{T} \sum_{t=2}^T y_{t-1} \Delta y_t &= \frac{-1}{2T} \left[ \sum_{t=2}^T (y_t - y_{t-1})^2 - \sum_{t=2}^T y_t^2 + \sum_{t=2}^T y_{t-1}^2 \right] \\ &= \frac{1}{2T} \left[ y_T^2 - \sum_{t=2}^T (y_t - y_{t-1})^2 \right] \\ &\Rightarrow \frac{1}{2} \left( \sigma_L^2 W^2(1) - \sigma_S^2 \right) \\ &\Rightarrow \sigma_L^2 \int_0^1 W(r) dW(r) + \frac{1}{2} \left( \sigma_L^2 - \sigma_S^2 \right) \end{aligned}$$

もちろん、誤差項  $\{u_t\}$  が i.i.d. ならば、短期分散  $\sigma_S^2$  と長期分散  $\sigma_L^2$  は等しくなり、 $\lambda = 0$  となるので、上の結果は前の単位根検定の結果に帰着する。

しかし、(33) の結果は、 $T(\hat{\rho} - 1)$  の帰無分布が未知のパラメータ  $\sigma_S^2$  や  $\sigma_L^2$  を含むから、このままでは検定統計量として使うことはできない。そこで、これらのパラメータを一致推定量  $\hat{\sigma}_S^2$  と  $\hat{\sigma}_L^2$  で置き換えることにする。一致推定量としては、

$$\begin{aligned}\hat{\sigma}_S^2 &= \frac{1}{T} \sum_{t=1}^T (y_t - y_{t-1})^2 \\ \hat{\sigma}_L^2 &= \hat{\sigma}_S^2 + \frac{2}{T} \sum_{i=1}^l \left(1 - \frac{i}{l+1}\right) \sum_{j=i+1}^T (y_j - y_{j-1})(y_{j-i} - y_{j-i-1})\end{aligned}$$

を使うことができる。ただし、 $l$  は  $T$  に依存するラグ打ち切り数である。このとき、 $H_0$  のもとで、次の分布収束が成り立つことが示される (Phillips (1987))。

$$\begin{aligned}(34) \quad Z_\rho &= T(\hat{\rho} - 1) - \frac{1}{2}(\hat{\sigma}_L^2 - \hat{\sigma}_S^2) \left/ \left( \frac{1}{T^2} \sum_{t=2}^T y_{t-1}^2 \right) \right. \\ &\Rightarrow \frac{\int_0^1 W(r) dW(r)}{\int_0^1 W^2(r) dr}\end{aligned}$$

$T(\hat{\rho} - 1)$  から  $Z_\rho$  への変換は、短期分散と長期分散だけの推定が追加的に必要となるもので、**ノンパラメトリック修正**と呼ばれる。

同様に、 $t$  統計量に対しては、 $H_0$  のもとで次の分布収束が成り立つ。

$$t_\delta = \frac{\hat{\delta}}{\hat{\sigma}_S \sqrt{\sum_{t=2}^T y_{t-1}^2}} \Rightarrow \frac{\sigma_L}{\sigma_S} \frac{\int_0^1 W(r) dW(r) + \lambda}{\sqrt{\int_0^1 W^2(r) dr}}$$

このことより、ノンパラメトリック修正に基づいて次の統計量を  $t$  統計量として使うことができる。

$$\begin{aligned}(35) \quad Z_t &= \frac{\hat{\sigma}_S}{\hat{\sigma}_L} t_\delta - \frac{1}{2}(\hat{\sigma}_L^2 - \hat{\sigma}_S^2) \left/ \left( \hat{\sigma}_L \sqrt{T^{-2} \sum_{t=2}^T y_{t-1}^2} \right) \right. \\ &\Rightarrow \frac{\int_0^1 W(r) dW(r)}{\sqrt{\int_0^1 W^2(r) dr}}\end{aligned}$$

以上、I(1) 性を検定する 2 つの方式を説明した。ノンパラメトリック修正による方法は一般的であるが、有限標本のもとの分布は、極限分布との乖離

が大きいことが知られている。モデルに関する先験的な情報がある場合には、ADF 検定を使う方が望ましい。

### 3.3 局所対立仮説のもとでの検出力

前節までに紹介した単位根検定は、通常よく使われているものであるが、そのよさを評価するには、検出力を考慮する必要がある。さらに、何らかの最適性をもつ検定と比較することが望ましい。

次のモデルを考えよう。

$$(36) \quad y_t = \alpha + \beta t + \eta_t, \quad \eta_t = \rho \eta_{t-1} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.} N(0, \sigma^2)$$

ここで、誤差項に正規性を仮定したのは最適な検定を導出するためであり、漸近的な議論のためには不要である。また、確定的トレンド  $\alpha + \beta t$  の有無により、3つのモデルが得られるが、ここでは何も制約のない場合を考えよう。

このモデルに対して、さまざまな検定方式の検出力を調べるために次の検定問題を考える。

$$H_0 : \rho = 1 \quad \text{vs.} \quad H_1 : \rho = 1 - \frac{c}{T}$$

ここで、 $c$  は正の任意定数である。 $H_1$  は  $H_0$  から  $1/T$  のオーダーの距離にあり、漸近的には  $H_0$  に収束する。このような仮説を**局所対立仮説**という。局所対立仮説のもとでの統計量の分布を**局所対立分布**といい、そのもとでの検出力を**局所検出力**という。以下では、さまざまな検定方式の局所検出力を導出する。そのために、検定統計量の局所対立分布を求めてみよう。

検定統計量を与えるために、モデル (36) から得られる2種類の残差  $\hat{\eta}_t = y_t - \hat{\alpha} - \hat{\beta}t$  と  $\tilde{\eta}_t = y_t - \tilde{\alpha} - \tilde{\beta}t$  を定義しよう。ここで、 $\hat{\alpha}$  と  $\hat{\beta}$  は OLS 推定量 (通常 LSE)、 $\tilde{\alpha}$  と  $\tilde{\beta}$  は  $H_0$  のもとでの誤差項の相関を考慮した GLS 推定量 (一般化 LSE) である。したがって、 $\hat{\eta}$  は OLS 残差、 $\tilde{\eta}$  は GLS 残差である。

このとき、次の3つの検定統計量  $R_1, R_2, R_3$  を考えよう。

$$R_1 = T(\hat{\rho} - 1), \quad \hat{\rho} = \frac{\sum_{t=2}^T \hat{\eta}_{t-1} \hat{\eta}_t}{\sum_{t=2}^T \hat{\eta}_{t-1}^2}$$

$$R_2 = \frac{1}{T} \frac{\sum_{t=1}^T \hat{\eta}_t^2}{\sum_{t=2}^T (\hat{\eta}_t - \hat{\eta}_{t-1})^2}, \quad R_3 = \frac{1}{T} \frac{\sum_{t=1}^T \tilde{\eta}_t^2}{\sum_{t=1}^T (\tilde{\eta}_t - \tilde{\eta}_{t-1})^2}$$

$R_1$  は、3.1 節の DF 係数統計量と漸近的に同等である。 $R_2$  は、DW (Durbin-Watson) 統計量 (の逆数) である。これら 2 つが OLS 残差に基づくのに対して、 $R_3$  は GLS 残差を使った統計量である。 $H_0$  は、これらの値が小さいときに棄却される。なお、 $R_3$  に基づく検定は、適当な変換群のもとで LBIU (Locally Best Invariant and Unbiased: 局所最良不変不偏) である (Nabeya-Tanaka (1990b), Tanaka (1996, Chap. 9) 参照)。

これら 3 つの統計量の局所対立分布は、漸近的に次のようになることが示される。

$$R_1 \Rightarrow \frac{\int_0^1 Z(r) dY(r)}{\int_0^1 Z^2(r) dr}, \quad R_2 \Rightarrow \int_0^1 Z^2(r) dr,$$

$$R_3 \Rightarrow \int_0^1 [Y(r) - r Y(1)]^2 dr$$

ただし、

$$Y(r) = e^{-cr} \int_0^r e^{cs} dW(s) = W(r) - c e^{-cr} \int_0^r e^{cs} W(s) ds$$

$$Z(r) = Y(r) - \int_0^1 Y(s) ds - 12 \left(r - \frac{1}{2}\right) \int_0^1 \left(s - \frac{1}{2}\right) Y(s) ds$$

である。 $\{Y(r)\}$  は **O-U 過程** (Ornstein-Uhlenbeck process) と呼ばれる連続な確率過程であり、 $c = 0$  のとき標準ブラウン運動に帰着する。他方、 $\{Z(r)\}$  は、**トレンド調整された O-U 過程** である。これは、3.1 節で述べたトレンド調整されたブラウン運動の拡張である。

これらの極限分布を評価することにより、それぞれの検定の漸近的な局所検出力を求めることができる。ここでは、さらに、相対的なよさを評価するために、最強力検定と比較してみよう。モデル (36) において、対立仮説の値が  $\rho_1 = 1 - (\theta/T)$  と与えられた場合、 $\tilde{\eta}^{(1)}$  を  $\rho_1$  のもとで計算される GLS 残差、 $\tilde{\eta}^{(0)}$  を  $\rho = 1$  のもとで計算される GLS 残差とする。このとき、統計量

$$M_T(\theta) = \frac{\sum_{t=1}^T (\tilde{\eta}_t^{(1)} - \rho_1 \tilde{\eta}_{t-1}^{(1)})^2 - \sum_{t=1}^T (\tilde{\eta}_t^{(0)} - \tilde{\eta}_{t-1}^{(0)})^2}{\sum_{t=1}^T (\tilde{\eta}_t^{(0)} - \tilde{\eta}_{t-1}^{(0)})^2 / T}$$

が小さいときに  $H_0$  を棄却する検定は最強力不変である。

統計量  $M_T(\theta)$  は、 $\rho$  の真の値が  $\rho = 1 - (c/T)$  ならば、 $T \rightarrow \infty$  のとき、 $M_T(\theta) \Rightarrow M(c, \theta)$  となる (Elliott-Rothemberg-Stock (1996), Tanaka (1996, Chap. 9))。ただし、

$$M(c, \theta) = \theta^2 \left[ \int_0^1 Y^2(r) dr - \frac{2(\theta + 1)}{\delta} Y(1) \int_0^1 r Y(r) dr + \frac{\theta + 1}{3\delta} Y^2(1) - \frac{\theta^2}{\delta} \left( \int_0^1 r Y(r) dr \right)^2 \right] - \theta$$

であり、 $\delta = (\theta^2 + 3\theta + 3)/3$  である。

局所対立仮説  $H_1 : \rho = 1 - (c/T)$  のもとでの最強力検定の検出力を計算するためには、まず、 $M(0, c)$  の分布の  $100\alpha\%$  点  $x_\alpha$  を求め、次に  $P(M(c, c) < x_\alpha)$  を計算する。これが、有意水準  $\alpha$  における漸近的な検出力となる。そして、各  $c$  の値に対してこのように計算された検出力の全体を**検出力の包絡線**という。実行可能ないかなる検定も、包絡線を上回るような検出力をもたらすことは不可能であり、包絡線は到達可能な検出力の上限を表す。

図 3-2 には、5% 有意水準のもとでの  $R_1, R_2, R_3$  に基づく検定の検出力、および検出力の包絡線が示されている (同時に、MA 部分の単位根検定に関連した検出力も図示されているが、これらについては 3.4.2 節で説明する)。最も頻繁に使われる DF 検定 ( $R_1$ ) は、DW 検定 ( $R_2$ ) や LBIU 検定 ( $R_3$ ) よりも劣ることが見てとれる。例えば、 $c = 20$  のときの検出力は、次の通りである。

$$R_1 : 0.707 < R_2 : 0.746 < R_3 : 0.765$$

これらは漸近的な局所検出力であるが、有限標本における検出力の近似としても使うことができる。例えば、 $T = 200$  の場合、これらの検出力は  $\rho = 1 - (c/T) = 1 - (20/200) = 0.9$  における検出力の近似値である。なお、この中でベストな LBIU 検定も、 $H_0$  から離れるにしたがって検出力の振る舞いが悪くなっている。この振る舞いを改良するような検定を考案することも可能であるが、ここでは省略する (詳しくは、Tanaka (1996, Chap. 9) を参照のこと)。

図 3-2

### 3.4 さまざまな拡張

前節で述べてきた単位根検定あるいは I(1) 性の検定は、さまざまな方向に拡張されている。その代表的なトピックは次の通りである。

- (i) 季節性を含む時系列の単位根検定
- (ii) MA 部分の単位根検定
- (iii) 定常性の検定
- (iv) フラクショナルな和分次数の検定
- (v) 構造変化を含む時系列の単位根検定

以下、これらについて簡単に説明しよう。

#### 3.4.1 季節性を含む時系列の単位根検定

月次や四半期の時系列が単位根を含むかどうかの検定は、周期性を考慮したモデルを考える必要がある。そのためのモデルとしては、 $m$  を周期として、

$$(37) \quad y_t = \sum_{j=1}^m (\alpha_j + \beta_j t) D_{jt} + \eta_t, \quad \eta_t = \rho_m \eta_{t-m} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d}(0, \sigma^2)$$

が基本的なものである。ここで、 $D_{jt}$  は  $m$  個のダミー変数であり、 $t-j$  が  $m$  の倍数ならば 1、さもなければ 0 をとる変数である。すなわち、確定的なトレンドは季節ごとに独自の切片と傾きをもつことが想定されている。もちろん、モデルの変形として、 $\alpha_j$  や  $\beta_j$  が 0 に制約されているモデルや、確定的トレンドが季節間で不変なモデルを考えることもできる。

このモデルに対して、検定問題

$$H_0 : \rho_m = 1 \quad \text{vs.} \quad H_1 : \rho_m < 1$$

を考えよう。この場合も、非季節的な  $m = 1$  の場合と同様の単位根検定統計量を考えることができる。ここでは、次の 3 つの検定統計量  $R_{m1}$ ,  $R_{m2}$ ,  $R_{m3}$  を考えよう。

$$R_{m1} = T(\hat{\rho}_m - 1), \quad \hat{\rho}_m = \frac{\sum_{t=m+1}^T \hat{\eta}_{t-m} \hat{\eta}_t}{\sum_{t=m+1}^T \hat{\eta}_{t-m}^2}$$
$$R_{m2} = \frac{1}{T} \frac{\sum_{t=1}^T \hat{\eta}_t^2}{\sum_{t=m+1}^T (\hat{\eta}_t - \hat{\eta}_{t-m})^2}, \quad R_{m3} = \frac{1}{T} \frac{\sum_{t=1}^T \tilde{\eta}_t^2}{\sum_{t=1}^T (\tilde{\eta}_t - \tilde{\eta}_{t-m})^2}$$

ここで、 $\hat{\eta}$  は OLS 残差、 $\tilde{\eta}$  は GLS 残差を表す。  $R_{m1}$  は DF 係数統計量、  $R_{m2}$  は DW 統計量、  $R_{m3}$  は LBIU 統計量である。

局所対立仮説  $\rho_m = 1 - (c/T)$  のもとで、これらの検定統計量の漸近分布に関して、次の結果を得る（詳しくは Tanaka (1996, Chap. 9) を参照されたい）。

$$(38) \quad R_{m1} \Rightarrow \frac{m \sum_{i=1}^m \int_0^1 Z_i(r) dY_i(r)}{\sum_{i=1}^m \int_0^1 Z_i^2(r) dr}$$

$$(39) \quad R_{m2} \Rightarrow \frac{1}{m^2} \sum_{i=1}^m \int_0^1 Z_i^2(r) dr$$

$$(40) \quad R_{m3} \Rightarrow \frac{1}{m^2} \sum_{i=1}^m \int_0^1 [Y_i(r) - r Y_i(1)]^2 dr$$

ここで、

$$Y_i(r) = e^{-cr} \int_0^r e^{cs} dW_i(s)$$

$$Z_i(r) = Y_i(r) - \int_0^1 Y_i(s) ds - 12(r - 1/2) \int_0^1 (s - 1/2) Y_i(s) ds$$

であり、 $W_i(r)$  ( $i = 1, \dots, m$ ) は互いに独立な  $m$  個の標準ブラウン運動である。したがって、 $Y_i(r)$  と  $Z_i(r)$  ( $i = 1, \dots, m$ ) は、それぞれ、互いに独立な  $m$  個の O-U 過程とトレンド調整された O-U 過程である。以上の結果に基づいて、漸近的な局所検出力を求めることができる。

Dickey-Hasza-Fuller (1984) には、シミュレーションに基づく係数検定統計量や  $t$  検定統計量などの帰無分布の分位点や検出力が報告されている。

### 3.4.2 MA 部分の単位根検定

単位根検定とは、通常は AR 部分に関する検定であるが、MA 部分の単位根検定についても考えることができる。その主たる必要性は、過剰階差が起きているかどうかを調べるものである。以下で見るように、それは結局のところ、AR 部分の単位根検定と解釈することも可能である。

MA 部分の単位根検定のための基本的なモデルは、MA(1) モデル

$$(41) \quad y_t = \varepsilon_t - \alpha \varepsilon_{t-1}, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

であり、 $y_t$  は観測値そのものでも階差変換後の値でもよい。検定問題は、

$$H_0 : \alpha = 1 \quad \text{vs.} \quad H_1 : \alpha < 1$$

である。

系列  $\{y_t\}$  が階差系列の場合、 $H_0$  のもとでは過剰階差となる。すなわち、原系列は AR 部分に単位根をもたない。他方、 $H_1$  のもとでは適切な階差変換であり、原系列は AR 部分に単位根をもっている。したがって、階差系列に関する MA 部分の単位根検定は、AR 部分に「単位根なし」という帰無仮説を、「単位根あり」という対立仮説に対して検定するものであり、通常とは逆向きの検定となる。

(41) の誤差項  $\{\varepsilon_t\}$  に正規性を仮定すると、観測値ベクトル  $y = (y_1, \dots, y_T)'$  に対する未知パラメータ  $\alpha$  と  $\sigma^2$  の対数尤度関数  $L(\alpha, \sigma^2)$  は、

$$(42) \quad L(\alpha, \sigma^2) = -\frac{T}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |\Omega(\alpha)| - \frac{1}{2\sigma^2} y' \Omega^{-1}(\alpha) y$$

で与えられる。ここで、

$$\Omega(\alpha) = \begin{pmatrix} 1 + \alpha^2 & -\alpha & & & 0 \\ -\alpha & 1 + \alpha^2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & -\alpha \\ 0 & & & -\alpha & 1 + \alpha^2 \end{pmatrix}$$

である。

検定統計量として自然なものは、 $\alpha$  の MLE (最尤推定量)  $\hat{\alpha}$  に基づく係数検定であろう。しかし、 $\hat{\alpha}$  の帰無分布、すなわち、反転不可能な MA(1) モデルにおける  $\hat{\alpha}$  の分布はわかっていない。ただし、AR の場合の係数推定量と異なり、 $T(\hat{\alpha} - 1)$  の極限帰無分布は、負の部分で連続な密度と原点を正の確率でとる分布の混合となることがわかっている (Tanaka-Satchell (1989), Davis-Dunsmuir (1996))。

実際、 $\alpha$  に集約された対数尤度を  $l(\alpha)$  とすると、

$$(43) \quad l(\alpha) = L(\alpha, \hat{\sigma}^2) = -\frac{T}{2} \log(2\pi y' \Omega^{-1}(\alpha) y / T) - \frac{1}{2} \log |\Omega(\alpha)| - \frac{T}{2}$$

となる。ただし、 $\hat{\sigma}^2 = y' \Omega^{-1}(\alpha) y / T$  は、 $\alpha$  が与えられた場合の MLE である。したがって、

$$\frac{dl(\alpha)}{d\alpha} = \frac{T}{2} \frac{y' \Omega^{-1} \Omega \Omega^{-1} y}{y' \Omega^{-1} y} - \frac{1}{2} \text{tr}(\Omega^{-1} \Omega) = 0, \quad \Omega = \Omega(1)$$

となることがわかる．このことから，尤度関数が  $\alpha = 1$  を極大値とする確率は，漸近的に，

$$\begin{aligned}\lim_{T \rightarrow \infty} P(\hat{\alpha} = 1 | \alpha = 1) &= \lim_{T \rightarrow \infty} P\left(\left.\frac{d^2 l(\alpha)}{d\alpha^2}\right|_{\alpha=1} < 0\right) \\ &= P\left(\sum_{n=1}^{\infty} \frac{\varepsilon_n^2}{n^2 \pi^2} < \frac{1}{6}\right) = 0.6574\end{aligned}$$

となる．すなわち， $T(\hat{\alpha} - 1)$  の極限帰無分布は，原点を確率 0.6574 でとるような分布に従うことになる．なお，負の値に対しては連続な密度をもつが，極限確率変数の明示的な表現は得られていない．

以下では，対数尤度関数の  $\alpha$  に関する偏導関数に基づく検定を考える．上述のように，1 階の偏導関数は  $H_0$  のもとで恒等的に 0 となるので，2 階の偏導関数

$$\begin{aligned}\left.\frac{\partial^2 L(\alpha, \sigma^2)}{\partial \alpha^2}\right|_{\alpha=1, \sigma^2=\hat{\sigma}^2} &= -\frac{1}{2} \text{tr}(-I_T + 2\Omega^{-1}) - T \frac{y'(\Omega^{-1} - \Omega^{-2})y}{y'\Omega^{-1}y} \\ &= -\frac{T(T+5)}{6} + T \frac{y'\Omega^{-2}y}{y'\Omega^{-1}y}\end{aligned}$$

を考える．そして，統計量として，

$$(44) \quad S_T = \frac{1}{T^2} \left( \left.\frac{\partial^2 L(\alpha, \sigma^2)}{\partial \alpha^2}\right|_{\alpha=1, \sigma^2=\hat{\sigma}^2} + \frac{T(T+5)}{6} \right) = \frac{1}{T} \frac{y'\Omega^{-2}y}{y'\Omega^{-1}y}$$

を定義して， $S_T$  が大きいときに  $H_0$  を棄却する検定を考える．この検定は LBIU である．統計量  $S_T$  は，局所対立仮説  $\alpha = 1 - (c/T)$  のもとで，次の分布に収束する (Tanaka (1990))．

$$(45) \quad S_T \Rightarrow \sum_{n=1}^{\infty} \left( \frac{1}{n^2 \pi^2} + \frac{c^2}{n^4 \pi^4} \right) Z_n^2, \quad \{Z_n\} \sim \text{i.i.d.N}(0, 1)$$

また，この場合の検定の検出力の包絡線は次のように求めることができる．まず，単純対立仮説  $\alpha_1 = 1 - (\theta/T)$  に対する最強力不変検定は，

$$(46) \quad M_T(\theta) = T \frac{y'\Omega^{-1}y - y'\Omega^{-1}(\alpha_1)y}{y'\Omega^{-1}y}$$

が大きいときに  $H_0$  を棄却するものである．そして，真の値が  $\alpha = 1 - (c/T)$  のもとで， $M_T(\theta)$  の分布収束に関して次の結果を得る．

$$(47) \quad M_T(\theta) \Rightarrow M(c, \theta) = \theta^2 \sum_{n=1}^{\infty} \frac{n^2 \pi^2 + c^2}{n^2 \pi^2 (n^2 \pi^2 + \theta^2)} Z_n^2 - \theta$$

これより、与えられた有意水準  $\gamma$  のもとで、 $M(0, c)$  の上側  $100\gamma\%$  点  $x_\gamma(c)$  が求められる。そして、 $P(M(c, c) > x_\gamma(c))$  を計算すれば、 $\alpha = 1 - (c/T)$  において漸近的に達成可能な検出力の上限が得られることになる。

3.2 節の **図 3-2** には、AR の場合とともに、MA 部分の単位根検定に関して、LBIU 検定の検出力曲線および包絡線が描かれている。MA の場合の検定の方が、帰無仮説からの乖離に対して敏感であることがわかる。ただし、LBIU 検定は、AR の場合と同様に、帰無仮説から離れると振る舞いが悪くなることを見てとれる。この点を改良した検定を考えることもできる。さらに、(41) の MA(1) モデルを拡張して、誤差項  $\{\varepsilon_t\}$  が定常過程に従う場合の単位根検定を考えることもできる (Saikkonen-Luukkonen (1993))。また、定数項や 1 次のトレンドを含むモデル、および季節性を考慮したモデルで議論することも可能であるが、ここでは省略する (Tanaka (1996, Chap. 10) 参照)。

### 3.4.3 定常性の検定

次のモデルを考えよう。

$$(48) \quad y_t = \alpha + \beta t + \gamma_t + \varepsilon_t, \quad \gamma_t = \gamma_{t-1} + \xi_t, \quad \gamma_0 = 0$$

ここで、 $\{\varepsilon_t\}$  と  $\{\xi_t\}$  は、それぞれが  $\text{i.i.d.}(0, \sigma_\varepsilon^2)$ ,  $\text{i.i.d.}(0, \sigma_\xi^2)$  の誤差項であり、互いに独立である。また、 $\sigma_\varepsilon^2$  は正、 $\sigma_\xi^2$  は非負値である。 $\{\gamma_t\}$  は、 $\sigma_\xi^2$  が正ならばランダム・ウォーク、0 ならば恒等的に 0 である。(48) のモデルは、**状態空間モデル** の特殊な場合であり、第 1 方程式は**観測方程式**、第 2 方程式は**状態方程式**と呼ばれる。

このモデルに対して、検定問題

$$H_0 : \lambda = \frac{\sigma_\xi^2}{\sigma_\varepsilon^2} = 0 \quad \text{vs.} \quad H_1 : \lambda > 0$$

を考える。 $H_0$  のもとでは  $\gamma_t \equiv 0$  となるので、 $\{y_t\}$  は単位根をもたない。しかし、 $H_1$  のもとでは単位根をもつことになる。したがって、この検定は「単位根なし」の帰無仮説を「単位根あり」の対立仮説に対して検定するものであり、通常単位根検定とは逆向きの検定である。その意味で、この検定は、前節で説明した MA 部分の単位根検定と同様の意味合いをもっている。

この場合の自然な検定は、 $\lambda$  の MLE  $\hat{\lambda}$  に基づくものであろう。しかし、 $H_0$  のもとでの  $\hat{\lambda}$  の分布は、反転不可能な MA(1) モデルの係数の MLE と同様

に、未知である。ただし、 $T^2 \hat{\lambda}$  の極限帰無分布は、正の部分で連続な密度と原点を正の確率でとる分布との混合となることがわかっている。

以下では、対数尤度の偏導関数に基づく検定方式を考える。そのために、まず、パラメータ  $\lambda, \sigma_\varepsilon^2, \alpha, \beta$  の尤度関数を求める。(48) から、 $y_t = \alpha + \beta t + \varepsilon_t + \xi_1 + \dots + \xi_t$  となるので、誤差項  $\{\varepsilon_t\}$  と  $\{\xi_t\}$  に正規性を仮定すると、観測値ベクトル  $y = (y_1, \dots, y_T)'$  は次の分布をもつことがわかる。

$$(49) \quad y = \alpha e + \beta d + C \xi + \varepsilon \sim N(\alpha e + \beta d, \sigma_\varepsilon^2 (I_T + \lambda C C')),$$

ここで、 $e = (1, \dots, 1)' : T \times 1$ ,  $d = (1, 2, \dots, T)'$ ,  $\xi = (\xi_1, \dots, \xi_T)'$ ,  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_T)'$ . そして、 $I_T$  は  $T \times T$  の単位行列である。また、 $C$  は下三角行列で、その  $(i, j)$  要素は  $i \geq j$  のとき 1, その他の場合は 0 である。したがって、対数尤度は

$$(50) \quad L(\lambda, \phi) = -\frac{T}{2} \log(2\pi\sigma_\varepsilon^2) - \frac{1}{2} \log |I_T + \lambda C C'| \\ - \frac{1}{2\sigma_\varepsilon^2} (y - \alpha e - \beta d)' (I_T + \lambda C C')^{-1} (y - \alpha e - \beta d)$$

で与えられる。ただし、 $\phi = (\sigma_\varepsilon^2, \alpha, \beta)'$  は攪乱母数である。

これより、

$$\left. \frac{\partial L(\lambda, \phi)}{\partial \lambda} \right|_{H_0} = -\frac{T(T+1)}{4} + \frac{T}{2} \frac{(y - \hat{\alpha} e - \hat{\beta} d)' C C' (y - \hat{\alpha} e - \hat{\beta} d)}{(y - \hat{\alpha} e - \hat{\beta} d)' (y - \hat{\alpha} e - \hat{\beta} d)}$$

が得られる。ただし、 $\hat{\alpha}$  と  $\hat{\beta}$  は、モデル (48) における OLS 推定量である。

そして、検定統計量として、

$$(51) \quad S_T = \frac{2}{T^2} \left( \left. \frac{\partial L(\lambda, \phi)}{\partial \lambda} \right|_{H_0} + \frac{T(T+1)}{4} \right) \\ = \frac{1}{T} \frac{(y - \hat{\alpha} e - \hat{\beta} d)' C C' (y - \hat{\alpha} e - \hat{\beta} d)}{(y - \hat{\alpha} e - \hat{\beta} d)' (y - \hat{\alpha} e - \hat{\beta} d)}$$

を考える。 $S_T$  が大きいときに帰無仮説を棄却する検定は LBI (Locally Best Invariant: 局所最良不変) となる。さらに、局所対立仮説  $\lambda = c^2/T^2$  のもとで、 $S_T$  の極限分布に関して次のことが成り立つ (Nabeya-Tanaka (1988), Tanaka (1996, Chap. 10) 参照)。

$$(52) \quad S_T \Rightarrow \int_0^1 \int_0^1 [K(r, s) + c^2 L(r, s)] dW(r) dW(s)$$

ここで、 $\{W(r)\}$  は標準ブラウン運動である。また、

$$K(r, s) = \min(r, s) - r s - 3r s (1 - r)(1 - s),$$

$$L(r, s) = \int_0^1 K(r, u) K(u, s) du$$

(52) の極限分布は、次の形の MA(1) モデル

$$(53) \quad y_t = \mu + \varepsilon_t - \alpha \varepsilon_{t-1}, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma_\varepsilon^2)$$

に対して得られる LBIU 単位根検定統計量の  $\alpha = 1 - (c/T)$  のもとでの極限分布と同一となることが示される。この同等性の直観的な理由は、状態空間モデル (48) が

$$(54) \quad \Delta y_t = \beta + \xi_t + \Delta \varepsilon_t$$

と書き換えられ、 $\{\Delta y_t\}$  が (53) と同様に、0 でない平均をもつ MA(1) モデルとみなすことができることによる。いずれにしろ、(52) の極限分布を評価することにより、LBI 検定の局所検出力が求められる。

状態空間モデル (48) は、確定的トレンドの有無により、さらにいくつかの検定が考えられる (Tanaka (1996, Chap. 10))。また、誤差項  $\{\varepsilon_t\}$  を定常過程に拡張することにより、定常性を帰無仮説とする検定が提案されている (Kwiatkowski-Phillips-Schmidt-Shin (1992))。

#### 3.4.4 フラクショナルな和分次数の検定

単位根検定は、通常は原系列が I(1) か I(0) であるかを判断するために行われる。すなわち、和分の次数  $d$  が 1 か 0 かを調べるものである。

これに対して、 $d$  を実数に拡張した場合の I( $d$ ) 過程を考えることもできる。そのために、観測値の系列  $\{y_t\}$  に対して、フラクショナル差分  $(1 - L)^d$  を次のように定義する。

$$(55) \quad (1 - L)^d y_t = \sum_{j=0}^{\infty} \binom{d}{j} (-L)^j y_t = \sum_{j=0}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)\Gamma(j + 1)} y_{t-j}$$

ここで、 $\Gamma(x)$  はガンマ関数である。なお、最右辺の無限展開は、実際には  $y_t = 0$  ( $t \leq 0$ ) であるから、有限項展開となる。

フラクショナル差分  $(1 - L)^d$  を施した系列が定常な ARMA( $p, q$ ) 過程であるとき、原系列  $\{y_t\}$  は、ARFIMA( $p, d, q$ ) 過程であるという。ARFIMA( $p, d, q$ )

過程は、 $d$  が  $1/2$  未満であれば定常、 $-1/2$  より大きければ反転可能であることが知られている (Hosking (1981)). 他方、 $d$  が  $1/2$  以上ならば非定常となる. そして、 $T \rightarrow \infty$  のときの  $y_T$  の確率的な大きさについては次のことが成り立つ.

$$y_T = \begin{cases} O_p(1) & (d < 1/2) \\ O_p(\sqrt{\log T}) & (d = 1/2) \\ O_p(T^{d-1/2}) & (d > 1/2) \end{cases}$$

ARFIMA( $p, d, q$ ) モデルの最も簡単な例として、ARFIMA( $0, d, 0$ ) を考えよう.

$$(56) \quad (1-L)^d y_t = \varepsilon_t \iff y_t = \sum_{j=0}^{t-1} \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)} \varepsilon_{t-j}$$

ここで、 $\{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$  であり、 $y_t$  に関して解いた右側の MA 表現は、 $d < 1/2$  ならば無限和による表現も可能である. なお、 $d < 1/2$  は定常性の条件であることに注意されたい. ただし、以下では  $d \geq 1/2$  の非定常の場合も扱うので、このように有限和を使うことにする.

以上の設定のもとで、検定問題として、

$$H_0 : d = d_0 \quad \text{vs.} \quad H_1 : d > d_0$$

を考える. ここで、対立仮説は左片側あるいは両側にとることも可能である.

検定方式としては、誤差項に正規性を仮定した上で、対数尤度の導関数に基づくものを考えよう. このとき、検定統計量として、

$$(57) \quad S_T = \frac{1}{\sqrt{T}} \left. \frac{\partial L(d, \sigma^2)}{\partial d} \right|_{H_0} / \sqrt{\frac{\pi^2}{6}} = \sqrt{T} \sum_{j=1}^{T-1} \frac{1}{j} \hat{\rho}_j / \sqrt{\frac{\pi^2}{6}}$$

を得る (Tanaka (1999)). ただし、

$$\hat{\rho}_j = \frac{\sum_{t=j+1}^T \hat{\varepsilon}_{t-j} \hat{\varepsilon}_t}{\sum_{t=1}^T \hat{\varepsilon}_t^2}, \quad \hat{\varepsilon}_t = (1-L)^{d_0} y_t$$

である.  $\hat{\rho}_j$  は、残差に基づくラグ  $j$  の自己相関である.

右側対立仮説に対して、 $S_T$  が大きいときに  $H_0$  を棄却する検定は、漸近的に LBI である. そして、局所対立仮説  $d = d_0 + (\delta/\sqrt{T})$  ( $\delta > 0$ ) のもとで、統計量の極限分布は次のようになる.

$$(58) \quad S_T \Rightarrow N\left(\sqrt{\frac{\pi^2}{6}} \delta, 1\right)$$

したがって、有意水準  $\alpha$  のもとでの  $N(0,1)$  の上側  $100\alpha\%$  点を  $z_\alpha$  とするとき、局所検出力は、漸近的に

$$P(S_T > z_\alpha) \rightarrow P\left(Z > z_\alpha - \sqrt{\frac{\pi^2}{6}} \delta\right)$$

となる。ここで、 $Z \sim N(0,1)$  である。

ここでは、ARFIMA(0,d,0) における  $d$  の値に関する検定を説明したが、ARFIMA( $p, d, q$ ) の場合に拡張することもできる。詳しくは、Robinson (1994), Tanaka (1999) を参照されたい。

### 3.4.5 構造変化を含む時系列の単位根検定

単位根検定に影響を与える要因として、**構造変化** (structural change) の問題がある。今、構造変化が起きた時点を  $T_B$  とするとき、きわめて単純な例としてレベル・シフトを表す次の場合を考えることができる。

$$(59) \quad f(t) = \begin{cases} \alpha_0 & (t \leq T_B) \\ \alpha_0 + \alpha_1 & (t > T_B) \end{cases}$$

ここでは、以下の説明を簡単にするためにレベル・シフトだけを想定しているが、傾きの変化も考慮したトレンド・シフトの場合を考えることも、もちろん可能である。

ダミー変数

$$(60) \quad D_t(T_B) = \begin{cases} 0 & (t \leq T_B) \\ 1 & (t > T_B) \end{cases}$$

を定義することにより、レベル・シフトの構造変化は、次のようにコンパクトに表現できる。

$$(61) \quad f(t) = \alpha_0 + \alpha_1 D_t(T_B)$$

このようなレベル・シフトを含む時系列に対する単位根検定のためのモデルは、

$$(62) \quad y_t = \alpha_0 + \alpha_1 D_t(T_B) + \eta_t, \quad \eta_t = \rho \eta_{t-1} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \sigma^2)$$

となり、これは、次のように書き換えることができる。

$$(63) \quad \Delta y_t = a + b D_{t-1}(T_B) + \alpha_1 I_t(T_B + 1) + (\rho - 1) y_{t-1} + \varepsilon_t$$

ここで、 $a = \alpha_0(1 - \rho)$ ,  $b = \alpha_1(1 - \rho)$  であり、 $H_0$  のもとでは、ともに 0 となるパラメータである。また、 $I_t(T_B + 1)$  は、

$$(64) \quad I_t(T_B + 1) = \begin{cases} 0 & (t \neq T_B + 1) \\ 1 & (t = T_B + 1) \end{cases}$$

で定義される一点ダミーであるので、漸近的には無視できる項である。

検定方式として、 $\rho$  の OLS 推定量  $\hat{\rho}$  に基づく検定を考えよう。 $\hat{\rho}$  の漸近分布を導出するために、次の仮定をおく。

**仮定：**  $T_B$  の値と標本サイズ  $T$  の比は、漸近的に一定値  $\lambda$  ( $0 < \lambda < 1$ ) に収束する。すなわち、

$$\lim_{T \rightarrow \infty} \left( \frac{T_B}{T} \right) = \lambda \quad (0 < \lambda < 1)$$

であり、しかも、 $\lambda$  は既知である。

以上の設定のもとで、DF 係数統計量  $T(\hat{\rho} - 1)$  は次の極限帰無分布をもつ。ただし、以下において、 $\Sigma$  は  $1 \leq t \leq T$  に関する和、 $\Sigma^*$  は  $T_B + 1 \leq t \leq T$  に関する和を表す。

$$(65) \quad T(\hat{\rho} - 1) = T \times \frac{\begin{vmatrix} \Sigma 1 & \Sigma^* 1 & \Sigma \Delta y_t \\ \Sigma^* 1 & \Sigma^* 1 & \Sigma^* \Delta y_t \\ \Sigma y_{t-1} & \Sigma^* y_{t-1} & \Sigma y_{t-1} \Delta y_t \end{vmatrix}}{\begin{vmatrix} \Sigma 1 & \Sigma^* 1 & \Sigma y_{t-1} \\ \Sigma^* 1 & \Sigma^* 1 & \Sigma^* y_{t-1} \\ \Sigma y_{t-1} & \Sigma^* y_{t-1} & \Sigma y_{t-1}^2 \end{vmatrix}}$$

$$\Rightarrow \frac{\begin{vmatrix} 1 & 1 - \lambda & W(1) \\ 1 - \lambda & 1 - \lambda & W(1) - W(\lambda) \\ \int_0^1 W(r) dr & \int_\lambda^1 W(r) dr & \int_0^1 W(r) dW(r) \end{vmatrix}}{\begin{vmatrix} 1 & 1 - \lambda & \int_0^1 W(r) dr \\ 1 - \lambda & 1 - \lambda & \int_\lambda^1 W(r) dr \\ \int_0^1 W(r) dr & \int_\lambda^1 W(r) dr & \int_0^1 W^2(r) dr \end{vmatrix}}$$

この結果は、レベル・シフトだけを想定した場合のものである。より一般的な場合については、Perron (1989) を参照されたい。また、この結果は、構造変化が起きる時点が既知であることを前提にしている点で制約的である。未知

とした場合については、Zivot-Andrews (1992), Vogelsang-Perron (1998) などで議論されている。

#### 4. 共和分過程

この章では、2つの I(1) 系列間の回帰について、2通りの場合を考える。1つは見せかけの回帰であり、もう1つは真の回帰、すなわち、共和分の場合である。そして、見せかけの回帰か、真の回帰かを判断する統計的方法について説明する。

##### 4.1 見せかけの相関と回帰

非定常な I(1) 系列間で相関や回帰関係を考える場合には、通常の統計的判断が適用できない場合がある。また、もっともらしい結果が得られたとしても、注意が必要となる。これが、I(1) の世界における**見せかけの相関** (spurious correlation), あるいは**見せかけの回帰** (spurious regression) と総称される現象である。

このことを説明するために、2次元のランダム・ウォーク系列

$$(66) \quad \Delta z_t = \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

を考えよう。 $\Sigma$  が対角行列であるから、2つのランダム・ウォーク  $\{x_t\}$  と  $\{y_t\}$  は互いに無相関である。

このとき、まず、 $\{x_t\}$  と  $\{y_t\}$  の相関係数

$$r_{xy} = \frac{\sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^T (x_t - \bar{x})^2 \sum_{t=1}^T (y_t - \bar{y})^2}}$$

を考えよう。2つの系列は無相関であるから、通常は、標本相関係数は0に確率収束すると考えられる。しかし、今の場合には、I(1) 系列の特殊性から、次の結果が成り立つ。

$$r_{xy} \Rightarrow \frac{\int_0^1 \tilde{W}_1(r) \tilde{W}_2(r) dr}{\sqrt{\int_0^1 \tilde{W}_1^2(r) dr \int_0^1 \tilde{W}_2^2(r) dr}}$$

ここで、

$$\tilde{W}_i(r) = W_i(r) - \int_0^1 W_i(s) ds, \quad (i = 1, 2)$$

であり、 $W_1(r)$  と  $W_2(r)$  は互いに独立な標準ブラウン運動である。この結果より、互いに無相関な  $I(1)$  系列の相関係数は 0 には収束せずに、退化しない分布をもつことがわかる。これは見せかけの相関である。

見せかけの相関を一步進めると、見せかけの回帰が得られる。 $\{y_t\}$  を  $\{x_t\}$  に回帰した式

$$(67) \quad y_t = \hat{\alpha} + \hat{\beta}x_t + \hat{u}_t, \quad (t = 1, \dots, T)$$

を考えよう。ただし、 $\hat{\alpha}$  と  $\hat{\beta}$  は OLS 推定量である。

(67) 式において、残差系列  $\{\hat{u}_t\}$  が定常的ならば、 $\{x_t\}$  と  $\{y_t\}$  は共和分関係にあり、回帰は意味がある。しかし、2 つの  $I(1)$  系列は互いに無相関なので、共和分関係はありえない。したがって、残差系列  $\{\hat{u}_t\}$  にも  $I(1)$  性が残っている。

互いに無相関な 2 つの変数が定常的ならば、回帰係数  $\hat{\beta}$  は 0 に確率収束する。しかし、今の場合の説明変数は  $I(1)$  系列の  $\{x_t\}$  である。このとき、次の分布収束が成り立つ。

$$(68) \quad \hat{\beta} \Rightarrow \frac{\sigma_y \int_0^1 \tilde{W}_1(r) \tilde{W}_2(r) dr}{\sigma_x \int_0^1 \tilde{W}_1^2(r) dr}$$

この結果から、相関係数と同様に、 $\hat{\beta}$  は 0 に確率収束せずに、退化しない分布をもつことがわかる。 $\hat{\beta}$  の極限分布の平均まわりの  $k$  次のモーメントを  $\mu_k$  とすれば、

$$\mu_1 = 0, \quad \mu_2 = \frac{0.3965 \sigma_y^2}{\sigma_x^2}, \quad \mu_3 = 0, \quad \mu_4 = 4.0838 \mu_2^2$$

を得る (Tanaka (1993))。分布は原点に関して対称であるが、尖度は  $(\mu_4/\mu_2^2) - 3 = 1.0838$  となるので、正規分布よりも尖った分布となる。

図 4-1 には、 $\hat{\beta}/\sqrt{\mu_2}$  の極限分布 (平均 0, 分散 1) が破線で図示されている。実線は  $N(0, 1)$  の密度であり、上述のように、 $N(0, 1)$  よりも尖った形状をしていることがわかる。なお、点線の曲線は次節で説明する共和分係数の推定量の分布である。

図 4-1

(67) の回帰式に関連する統計量は、 $\hat{\beta}$  以外にも、 $\hat{\alpha}$ ,  $t$  統計量, 決定係数, DW 統計量などがあり, 通常回帰とは異なる結果をもたらす. この事実を, Granger-Newbold (1974) はシミュレーションにより発見し, Phillips (1986) は上記のような方法で理論的に解明した. その主要な結果は次の通りである.

- (a)  $\hat{\alpha} = O_p(\sqrt{T})$ ,  $\hat{\beta} = O_p(1)$
- (b) 係数の  $t$  統計量  $= O_p(\sqrt{T})$
- (c) 残差の 2 乗和  $\sum_{t=1}^T \hat{u}_t^2 = O_p(T^2)$
- (d) 決定係数  $R^2 = O_p(1)$
- (e) Durbin-Watson 統計量  $DW = O_p(1/T)$

(66) のモデルでは, 2 つのランダム・ウォーク系列が互いに無相関であることを仮定したが, 相関がある場合でも, 一般には見せかけの回帰となる. 回帰が意味をもつのは共和分関係が成り立つ場合であり, ランダム・ウォーク系列間では, 誤差項の共分散行列  $\Sigma$  が正則でない場合である. それは, 結局のところ, 2 つの系列のパスが同一直線上にある場合に限る. 一般の  $I(1)$  系列では, 共和分の条件はもっと緩やかになるが, この点については次節で議論する.

見せかけの回帰は, 純粋な  $I(1)$  系列に線形トレンドをあてはめる場合にも起きるし, 三角関数のような周期関数をあてはめる場合でも起きる. 詳しくは, Durlauf-Phillips (1988), Phillips (1998), Tanaka (2001) を参照されたい.

これらの回帰は,  $\{y_t\}$  が  $I(1)$  系列である限り無意味なものである. しかし,  $t$  統計量や決定係数が有意性を示唆して, もっともらしい回帰をもたらす. 特に, 説明変数が三角関数であっても有意な回帰を与えることは驚くべき事実である. ただし, DW 統計量のオーダーが  $1/T$  であることから, 残差に非常に強い正の系列相関があることを示唆し, 回帰が満足の行くものではないというシグナルを発している. このことは, 回帰の残差系列が単位根を含む  $I(1)$  系列であることを意味しており, あてはめた回帰式が見せかけであることの証左にほかならない.

## 4.2 共和分係数の推定

本節では, 2 つの  $I(1)$  系列間の回帰が意味のある場合, すなわち, 共和分関係がある場合を考える. そのために, 2.5 節の議論を思い起こそう.

今、2つの I(1) 系列  $\{x_t\}$  と  $\{y_t\}$  が次のような 2次元の I(1) 過程に従うものとする。

$$(69) \quad \Delta z_t = \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} = C(L) \varepsilon_t = [C(1) + \Delta \tilde{C}(L)] \varepsilon_t$$

ここで、 $\{\varepsilon_t\} \sim \text{i.i.d.}(0, I_2)$  である。また、 $\{C_j\}$  は、

$$(70) \quad \sum_{j=0}^{\infty} j |C_j| < \infty, \quad C(1) = \sum_{j=0}^{\infty} C_j = \begin{pmatrix} A'_1 \\ A'_2 \end{pmatrix} \neq \begin{pmatrix} 0' \\ 0' \end{pmatrix}$$

をみたすものとする。

以上の設定のもとで、 $\{x_t\}$  と  $\{y_t\}$  が共和分関係となるのは、 $C(1)$  のランクが 1 に落ちるときである。このとき、2.5 節で説明した通り、 $\beta = A'_2 A_1 / A'_1 A_1$  に対して、(69) は、

$$(71) \quad y_t = \beta x_t + g'(L) \varepsilon_t, \quad \Delta x_t = C'_1(L) \varepsilon_t$$

と変形される。ここで、 $g'(L) = \tilde{C}'_2(L) - \beta \tilde{C}'_1(L)$  であり、 $\tilde{C}'_i(L)$  は  $\tilde{C}(L)$  の第  $i$  行である。また、 $C'_1(L)$  は  $C(L)$  の第 1 行である。

モデル (71) は、見せかけの回帰とは異なり、意味のある共和分回帰であり、回帰係数  $\beta$  の推測は重要となる。まず、 $\beta$  の OLS 推定量  $\hat{\beta}_{OLS}$  は、

$$(72) \quad \hat{\beta}_{OLS} = \left( \sum_{t=1}^T x_t^2 \right)^{-1} \sum_{t=1}^T x_t y_t = \beta + \left( \sum_{t=1}^T x_t^2 \right)^{-1} \sum_{t=1}^T x_t g'(L) \varepsilon_t$$

である。

もし、(71) の共和分回帰が非常に単純な構造をしている場合、例えば、

$$(73) \quad y_t = \beta x_t + \varepsilon_{2t}, \quad \Delta x_t = \varepsilon_{1t}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \text{i.i.d.}(0, I_2)$$

の場合には、 $T$  で規準化することにより、次の分布収束が得られる。

$$(74) \quad T(\hat{\beta}_{OLS} - \beta) = \left( \frac{1}{T^2} \sum_{t=1}^T x_t^2 \right)^{-1} \frac{1}{T} \sum_{t=1}^T x_t \varepsilon_{2t} \Rightarrow \frac{\int_0^1 W_1(r) dW_2(r)}{\int_0^1 W_1^2(r) dr}$$

ここで、 $W(r) = (W_1(r), W_2(r))'$  は 2次元の標準ブラウン運動である。この結果から、一致性のオーダーとして、 $\hat{\beta} - \beta = O_p(1/T)$  が得られる。ここでは、モデル (73) において、説明変数  $x_t$  と誤差項  $\varepsilon_{2t}$  が互いに独立であること

を仮定したが、この結果は独立性によるものではない。下に示すように、相関があっても、LSE は同じオーダーの一致性をもつことがわかる。

(74) の極限分布は原点に関して対称であり、分散は  $\sigma_1^2 = 5.5629$  となることが示される (Tanaka (1993))。前節の図 4-1 には、見せかけの回帰係数推定量の極限分布の他に、この場合の共分散係数推定量  $T(\hat{\beta}_{OLS} - \beta)/\sigma_1$  の極限分布が点線で図示されている。N(0,1) や見せかけの回帰係数推定量よりも尖度が大きいことがわかる。

しかし、(71) のような一般の場合、すなわち、誤差項が従属的で、しかも、説明変数と相関をもつ場合には、推定量の漸近分布は複雑になる。具体的に、 $\hat{\beta}$  の漸近分布を導出するために、次の 2 次元の I(1) 系列を考える。

$$(75) \quad \Delta \tilde{z}_t = \begin{pmatrix} \Delta x_t \\ \Delta w_t \end{pmatrix} = \begin{pmatrix} C_1'(L) \\ g'(L) \end{pmatrix} \varepsilon_t$$

さらに、長期の共分散行列

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} = \begin{pmatrix} A_1' A_1 & A_1' g(1) \\ g'(1) A_1 & g'(1) g(1) \end{pmatrix}$$

を定義する。そして、 $\Omega$  の分解を

$$\Omega = D D', \quad D = \begin{pmatrix} \omega_{11}^{1/2} & 0 \\ \omega_{21} \omega_{11}^{-1/2} & \omega_{22.1}^{1/2} \end{pmatrix}, \quad \omega_{22.1} = \omega_{22} - \omega_{21} \omega_{11}^{-1} \omega_{12}$$

によって定義する。

このとき、次の分布収束が成り立つ。

$$\begin{aligned} \frac{1}{T^2} \sum_{t=1}^T \tilde{z}_t \tilde{z}_t' &\Rightarrow D \int_0^1 W(r) W'(r) dr D' \\ \frac{1}{T} \sum_{t=1}^T \tilde{z}_t \Delta \tilde{z}_t' &\Rightarrow D \int_0^1 W(r) dW'(r) D' + \Lambda \end{aligned}$$

ただし、 $\{W(r)\}$  は 2 次元の標準ブラウン運動であり、

$$\Lambda = \sum_{k=0}^{\infty} E(\Delta \tilde{z}_t \Delta \tilde{z}_{t+k}') = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}$$

である。この事実と CMT により、次の結果を得る。

$$(76) \quad T(\hat{\beta}_{OLS} - \beta) \Rightarrow \frac{Q_1 + Q_2 + \lambda_{12}}{R}$$

ここで,

$$\begin{aligned} Q_1 &= \omega_{12} \int_0^1 W_1(r) dW_1(r) \\ Q_2 &= \sqrt{\omega_{11} \omega_{22.1}} \int_0^1 W_1(r) dW_2(r) \\ R &= \omega_{11} \int_0^1 W_1^2(r) dr \end{aligned}$$

(76) の極限確率変数は、3つの部分  $Q_1/R$ ,  $Q_2/R$ ,  $\lambda_{12}/R$  から構成されている。この中で、 $Q_1/R$  は1変量の単位根分布に相当する部分、 $\lambda_{12}/R$  は分布シフトに相当する部分であり、いずれも分布のゆがみや非対称性をもたらす要因となる。

これに対して、 $\Omega$  と  $\Lambda$  の一致推定量  $\hat{\Omega}$  と  $\hat{\Lambda}$  が得られるならば、次の推定量

$$\hat{\beta}_{FM} = \left( \sum_{t=1}^T x_t^2 \right)^{-1} \left[ \sum_{t=1}^T x_t (y_t - \Delta x_t \hat{\omega}_{12} / \hat{\omega}_{11}) - T (\hat{\lambda}_{12} - \hat{\lambda}_{11} \hat{\omega}_{12} / \hat{\omega}_{11}) \right]$$

を定義することができる。FM は fully modified の略である。このとき、 $\hat{\beta}_{FM}$  は、次の分布収束に従うことが示される (Phillips-Hansen (1990))。

$$T(\hat{\beta}_{FM} - \beta) \Rightarrow Q_2/R$$

極限分布は原点に関して対称であり、最も単純なモデル (73) における OLS 推定量と同様の分布であることが了解されよう。

### 4.3 回帰の残差に基づく共和分検定

共和分に関する統計的問題は、2つの I(1) 系列が本当に共和分の関係にあるかどうかを検定することである。ここでは、Engle-Granger (1987) により提案された方法を説明する。

まず、一方の I(1) 系列  $\{y_t\}$  を他方の I(1) 系列  $\{x_t\}$  に回帰して、残差  $\hat{\eta}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$  を求める。残差系列は、2つの系列が共和分していなければ単位根をもち、共和分関係にあれば単位根をもたないと考えられる。

そこで、通常の単位根検定の場合と同様にして、 $\Delta \hat{\eta}_t$  を  $\hat{\eta}_{t-1}$  と  $\Delta \hat{\eta}_t$  のラグに回帰して、回帰関係

$$(77) \quad \Delta \hat{\eta}_t = \delta \hat{\eta}_{t-1} + \sum_{j=1}^p \hat{\phi}_j \Delta \hat{\eta}_{t-j} + \hat{v}_t$$

を求める。このとき、「共和分関係なし」の帰無仮説  $H_0: \delta = 0$  を、「共和分あり」の対立仮説  $H_1: \delta < 0$  に対して検定する統計量として、 $\hat{\delta}$  の  $t$  値  $t_{\hat{\delta}}$  を使うことができる。ただし、 $\hat{\delta}$  は残差から計算されているので、通常の単位根分布とは異なる分布をもつことに注意されたい。この場合の極限帰無分布の分位点は、Engle-Granger (1987), Davidson-MacKinnon (1993) などに報告されている。

図 2-1 の TOPIX と為替レートの対数値系列では、 $p = 6$  の場合に、 $\hat{\delta}$  の  $t$  値として、TOPIX を説明変数とする回帰の残差では  $-2.01$ 、為替レートを説明変数とする回帰の残差では  $-2.92$  が得られた。これらは、Davidson-MacKinnon (1993, p. 722) に報告されている 10% 有意点  $-3.04$  を越えない値なので、2 つの系列が共和分関係にあるとはいえない、と判断できる。

回帰の残差に基づく共和分検定としては、この他に、ノンパラメトリック修正による方法 (Phillips-Ouliaris (1990)) がある。これらの方法の欠点は、検定結果が、いずれを説明変数あるいは被説明変数にするかに依存する点にある。この点を回避した方法は次章で説明する。

## 5. 多変量時系列と共和分

前章では 2 つの  $I(1)$  時系列間に限定して考えたが、本章では、一般の多変量  $I(1)$  時系列を対象として共和分関係を考察する。以下、主として Johansen (1995) の議論に沿って説明する。

### 5.1 共和分のシステム推定

$q$  個の成分からなるベクトル値時系列  $\{y_t\}$  が、 $\text{VAR}(p)$  モデル (次数  $p$  のベクトル自己回帰モデル)

$$(78) \quad y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \Sigma)$$

に従うものとする。ここで、各々の成分は  $I(1)$  か  $I(0)$  であるとする。また、誤差項の共分散行列  $\Sigma$  は  $q$  次の正値定符号とする。

このモデルは、行列値のラグ多項式  $B(L)$  を使って、次のように表すことができる。

$$(79) \quad B(L) y_t = \varepsilon_t, \quad B(L) = I_q - B_1 L - \cdots - B_p L^p$$

ここで、特性方程式  $|B(x)| = 0$  の根、すなわち、 $B(x)$  の行列式の零根は、すべて絶対値が 1 以上であり、絶対値が 1 となる根は 1 のみ、すなわち、単位根のみであると仮定する。このとき、個々の系列は高々  $I(1)$  であることから、

$$\text{特性方程式の単位根の個数} = q - \text{rank}(B(1))$$

が成り立つ。

以上の設定のもとで、(78) の  $\text{VAR}(p)$  を次のように変形してみよう。

$$(80) \quad \Delta y_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} = C(L) \varepsilon_t$$

このとき、(79) と (80) より、

$$B(L) \Delta y_t = \Delta \varepsilon_t = B(L) C(L) \varepsilon_t$$

となるから、 $B(L) C(L) = \Delta I_q = (1 - L) I_q$  となり、特に、

$$B(1) C(1) = 0$$

を得る。

上の最後の関係と前節の議論から、 $B(1)$  の各行が共和分ベクトルの可能性をもつことがわかる。このとき、 $B(1)$  のランクを**共和分ランク**という。それは、 $q$  変量時系列  $\{y_t\}$  に存在する 1 次独立な共和分ベクトルの最大数であり、その値は 0 以上、 $q$  以下である。

共和分ランクが 0、すなわち、 $B(1)$  が零行列の場合は共和分関係は存在しない。それは、 $C(1)$  が正則であることと同等である。なぜなら、このとき、 $B(L) = (1 - L) \tilde{B}(L)$  であり、 $C(1) = \tilde{B}^{-1}(1)$  となるからである。ここで、 $\tilde{B}(1)$  は、時系列が高々  $I(1)$  であることから正則である。他方、共和分ランクが正のときは共和分関係が存在する。ただし、ランクが  $q$  の場合は  $\{y_t\}$  そのものが定常過程となる。

Johansen (1995) は、共和分ランクが  $r$  のときの  $\text{VAR}(p)$  モデルを推定するために、まず、(79) の  $B(L)$  を次のように表現した。

$$B(L) = B(1)L + B(L) - B(1)L = B(1)L + \Delta \Gamma(L)$$

ここで,

$$\Gamma(L) = I_q - \Gamma_1 L^1 - \cdots - \Gamma_{p-1} L^{p-1}, \quad \Gamma_j = - \sum_{i=j+1}^p B_i$$

である. この表現を使って, (78) は次のように変形できる.

$$(81) \quad \Delta y_t = \gamma \alpha' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

ここで,  $\gamma \alpha' = -B(1)$  であり,  $\gamma$  と  $\alpha$  は  $q \times r$  の行列で, ランクはともに  $r$  である.

(81) は **誤差修正モデル** (ECM: Error Correction Model) と呼ばれる. 右辺第 1 項の  $\{\alpha' y_{t-1}\}$  は I(1) 系列の線形結合であるが, それ以外の項が定常であるから,  $r$  が正である限り定常であり, 原系列の長期的な均衡関係からの乖離を表している. このように, 共和分関係が存在する場合には, 階差系列に加えてレベルの変数も混在することになり, ARIMA モデルでは表現不可能な形をしている. 行列  $\alpha$  は **共和分行列** と呼ばれる. 他方,  $\alpha'$  の前にかかる行列  $\gamma$  は **調整行列**, あるいは **負荷行列** と呼ばれる.

(81) の誤差修正モデルに含まれるパラメータ全体をベクトル  $\theta$  で表そう. 誤差項  $\{\varepsilon_t\}$  に正規性を仮定すると,  $\theta$  の MLE は,  $\{\Delta y_t\}$  に対する対数尤度

$$(82) \quad L(\theta) = -\frac{qT}{2} \log(2\pi) - \frac{T}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^T (z_{0t} - \gamma \alpha' z_{1t} - \Gamma z_{2t})' \Sigma^{-1} (z_{0t} - \gamma \alpha' z_{1t} - \Gamma z_{2t})$$

を最大にする値となる. ここで,

$$z_{0t} = \Delta y_t, \quad z_{1t} = y_{t-1}, \quad z_{2t} = (\Delta y'_{t-1}, \dots, \Delta y'_{t-p+1})', \quad \Gamma = (\Gamma_1, \dots, \Gamma_{p-1})$$

である.

共和分ランクが 0 の場合,  $\alpha = \gamma = 0$  であり,  $L(\theta)$  の最大値は

$$(83) \quad L_0 = -\frac{qT}{2} \log(2\pi) - \frac{T}{2} \log |\hat{\Sigma}| - \frac{qT}{2}$$

である. ただし,

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T (z_{0t} - \hat{\Gamma} z_{2t}) (z_{0t} - \hat{\Gamma} z_{2t})', \quad \hat{\Gamma} = \sum_{t=1}^T z_{0t} z'_{2t} \left( \sum_{t=1}^T z_{2t} z'_{2t} \right)^{-1}$$

他方，共和分ランク  $r$  が正という条件のもとで  $L(\theta)$  を最大化するために，次の統計量を定義する．

$$S_{ab} = \frac{1}{T} \sum_{t=1}^T \zeta_{at} \zeta'_{bt}$$

$$\zeta_{at} = z_{at} - \sum_{t=1}^T z_{at} z'_{2t} \left( \sum_{t=1}^T z_{2t} z'_{2t} \right)^{-1} z_{2t}$$

このとき，まず， $\alpha$  を与えられたものとする， $L(\theta)$  の最大化は，次の量

$$(84) \quad \left| \hat{\Sigma}(r) \right| = \left| S_{00} - S_{01} \alpha (\alpha' S_{11} \alpha)^{-1} \alpha' S_{10} \right|$$

$$= |S_{00}| \frac{|\alpha' (S_{11} - S_{10} S_{00}^{-1} S_{01}) \alpha|}{|\alpha' S_{11} \alpha|}$$

を最小化することに帰着する（詳細は Johansen (1995), Hatanaka (1996) を参照されたい）．ここで，最後の等号は，次の関係式から得られる．

$$\left| \begin{pmatrix} S_{00} & S_{01} \alpha \\ \alpha' S_{10} & \alpha' S_{11} \alpha \end{pmatrix} \right| = |S_{00}| \left| \alpha' S_{11} \alpha - \alpha' S_{10} S_{00}^{-1} S_{01} \alpha \right|$$

$$= |\alpha' S_{11} \alpha| \left| S_{00} - S_{01} \alpha (\alpha' S_{11} \alpha)^{-1} \alpha' S_{10} \right|$$

(84) の最小値は，

$$|S_{00}| \prod_{j=1}^r (1 - \hat{\lambda}_j)$$

となることが示される．ただし， $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_r$  は，固有値問題

$$S_{10} S_{00}^{-1} S_{01} \hat{V} = S_{11} \hat{V} \hat{\Lambda}, \quad \hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_r)$$

の解である．また， $\hat{V} = (\hat{v}_1, \dots, \hat{v}_r)$  は対応する固有ベクトルからなる行列である．最小値を与える  $\alpha$  は，

$$\hat{\alpha} = (\hat{v}_1, \dots, \hat{v}_r), \quad \hat{\alpha}' S_{11} \hat{\alpha} = I_r$$

とすることができる．後者の条件は規準化則である．

以上より， $\text{rank}(\alpha) = r$  のもとで，対数尤度関数 (82) の最大値は，

$$L(\hat{\theta}) = -\frac{qT}{2} \log(2\pi) - \frac{T}{2} \log |S_{00}| - \frac{T}{2} \sum_{j=1}^r \log(1 - \hat{\lambda}_j) - \frac{qT}{2}$$

で与えられることになる．

ここで考えた誤差修正モデルは、係数行列がフル・ランクでない場合を扱うものである。このようなモデルは、一般に**ランク落ちモデル**と呼ばれている。ここで説明した Johansen の方法は、回帰モデルの場合に展開された Anderson (1951) の議論に準拠している。

## 5.2 共和分ランクの検定

前節では、共和分ランクを既知として推定問題を考えた。しかし、一般には共和分ランクは未知であり、その値を決定することは重要な統計的問題である。

共和分ランクに関する検定のためのモデルとして、前節のように、VAR( $p$ )を変換した誤差修正モデル

$$(85) \quad \Delta y_t = \gamma \alpha' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

を考えよう。時系列  $\{y_t\}$  は  $q$  変量で、各成分は高々  $I(1)$  に従っているものとする。

このとき、検定問題

$$H_0 : \text{rank}(\alpha) = r < q \quad \text{vs.} \quad H_1 : \text{rank}(\alpha) = q$$

を考える。これは、係数行列  $\gamma \alpha'$  のランク制約に関する検定である。帰無はランク落ちの制約仮説であり、対立仮説はフル・ランクという意味で無制約である。 $q$  次元ベクトル時系列は、帰無仮説のもとでは  $r$  個の 1 次独立な共和分関係をもち、対立仮説では定常となる。特に、 $q = 1$  の場合は 1 変量の単位根検定に帰着することが了解されよう。したがって、この検定は単位根検定の拡張と考えられる。

検定統計量としては、誤差項  $\{\varepsilon_t\}$  に正規性を仮定して、尤度比統計量を考えることができる。 $L(\theta)$  を対数尤度とすれば、尤度比検定は、前節の議論を使って、

$$J_T = -2(L(\hat{\theta}_0) - L(\hat{\theta}_1)) = -T \sum_{j=r+1}^q \log(1 - \hat{\lambda}_j)$$

が大きいために  $H_0$  を棄却するものとなる。ただし、 $\hat{\theta}_0$  は  $H_0$  のもとでの MLE、 $\hat{\theta}_1$  は  $H_1$  のもとでの MLE である。ここで、 $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_r$  は、 $H_0$  と  $H_1$  で共通の値となる。

尤度比統計量  $J_T$  の極限帰無分布について、Johansen (1988) は次のことを示した。

$$(86) \quad J_T \Rightarrow \text{tr} \left( \int_0^1 dW(r) W'(r) \left( \int_0^1 W(r) W'(r) dr \right)^{-1} \int_0^1 W(r) dW'(r) \right)$$

ここで、 $\{W(r)\}$  は  $q-r$  次元の標準ブラウン運動である。統計量がトレースで表されるので、この検定は**トレース検定**と呼ばれる。シミュレーションから得られた極限分布の分位点が Johansen (1995) の Table 15.1 に与えられている。

Johansen (1995) は、上記の検定問題の他に、

$$H_0 : \text{rank}(\alpha) = r < q \quad \text{vs.} \quad H_1 : \text{rank}(\alpha) = r + 1$$

を考えた。この場合の尤度比検定は、 $-T(1 - \hat{\lambda}_{r+1})$  が大きいときに  $H_0$  を棄却するものとなり、帰無分布は漸近的に (86) の極限として現れる行列の固有値の最大値に収束する。この意味で、この場合の検定は**最大固有値検定**と呼ばれ、Osterwald-Lenum (1992) に有意点が数表化されている。

### 5.3 さまざまな拡張

(78) で扱った VAR( $p$ ) モデルは、定数項や1次のトレンドを含むモデル、あるいは季節性を考慮したモデルなどに拡張できる。そして、それぞれの場合に対して、さまざまな検定が提案されている。

まず、最も簡単な拡張として、定数項を追加したモデル

$$(87) \quad y_t = \mu + B_1 y_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t, \quad \{\varepsilon_t\} \sim \text{i.i.d.}(0, \Sigma)$$

を考えよう。このモデルは、5.1 節の議論を使って、

$$(88) \quad \begin{aligned} \Delta y_t &= \sum_{j=0}^{\infty} C_j (\mu + \varepsilon_{t-j}) = C(1) \mu + C(L) \varepsilon_t \\ &= \mu + \gamma \alpha' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t \end{aligned}$$

と表現することができる。

このような定数項を含む誤差修正モデルに対しても、 $\mu = 0$  の場合と同様に推定や検定の問題を考えることができる。例えば、前節の検定問題に対しては、前と同様の尤度比検定が導出される。ただし、(81) のモデルと異なり、(87) に

おいては、定数ベクトル  $\mu$  の性質により、帰無分布は場合分けして考える必要がある。

このことを説明するために、共和分行列  $\alpha$  と調整行列  $\gamma$  に対して、 $\alpha'_\perp \alpha = 0$ ,  $\gamma'_\perp \gamma = 0$  となるランク  $q-r$  の  $q \times (q-r)$  行列  $\alpha_\perp$  と  $\gamma_\perp$  を定義する。このとき、Johansen (1991) は、(88) 式に現れる長期分散に関連した行列  $C(1)$  が、

$$(89) \quad C(1) = \alpha_\perp \left( \gamma'_\perp \tilde{C}(1) \alpha_\perp \right)^{-1} \gamma'_\perp$$

と表されることを示した。ただし、 $\tilde{C}(1)$  は、B-N 分解  $C(L) = C(1) + (1-L)\tilde{C}(L)$  から得られる行列である。

このことから、 $\gamma'_\perp \mu = 0$  ならば、 $C(1)\mu = 0$  となり、原系列  $\{y_t\}$  はトレンド項を含まないことになる。この点が、統計量の分布を場合分けする理由となる。

検定統計量の分布を記述するために、 $(q-r) \times (q-r)$  の確率行列

$$(90) \quad R = \int_0^1 dW(r) F'(r) \left( \int_0^1 F(r) F'(r) dr \right)^{-1} \int_0^1 F(r) dW'(r)$$

を定義しよう。ここで、 $\{W(r)\}$  は  $q-r$  次元標準ブラウン運動、 $\{F(r)\}$  は、以下で定義される  $q-r$  あるいは  $q-r+1$  次元確率過程である。このとき、前節と同様のトレース検定統計量の極限帰無分布は  $\text{tr}(R)$  となる。

ここで、 $\gamma'_\perp \mu = 0$  ならば、 $F(r)$  は  $q-r+1$  次元で、

$$F_j(r) = W_j(r) \quad (j = 1, \dots, q-r)$$

$$F_{q-r+1}(r) = 1$$

となる。この場合の分布表は、Johansen (1995) の Table 15.2 に与えられている。他方、 $\gamma'_\perp \mu \neq 0$  ならば、 $F(r)$  は  $q-r$  次元で、

$$F_j(r) = W_j(r) - \int_0^1 W_j(s) ds \quad (j = 1, \dots, q-r-1)$$

$$F_{q-r}(r) = r - \frac{1}{2}$$

となる。この場合の分布表は、Johansen (1995) の Table 15.3 に与えられている。

Johansen (1995) は、VAR(p) モデルが定数項だけでなく 1 次のトレンドも含む場合についても考察している。すなわち、

$$(91) \quad \Delta y_t = \mu_0 + \mu_1 t + \gamma \alpha' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

の形の ECM に対して、トレース検定統計量の極限帰無分布を求めている。それは、前と同様、(90) 式の  $R$  を使って、 $\text{tr}(R)$  となる。ただし、 $\gamma'_\perp \mu_1 = 0$  ならば、 $F(r)$  は  $q - r + 1$  次元で、

$$F_j(r) = W_j(r) - \int_0^1 W_j(s) ds \quad (j = 1, \dots, q - r)$$

$$F_{q-r+1}(r) = r - \frac{1}{2}$$

となる。この場合の分布表は、Johansen (1995) の Table 15.4 に与えられている。他方、 $\gamma'_\perp \mu_1 \neq 0$  ならば、 $F(r)$  は  $q - r$  次元で、

$$F_j(r) = W_j(r) - a_j - b_j r \quad (j = 1, \dots, q - r - 1)$$

$$F_{q-r}(r) = r^2 - a - b r$$

となる。ここで、 $a_j$  と  $b_j$  は、正規方程式

$$\int_0^1 W_j(r) dr = a_j + b_j \int_0^1 r dr$$

$$\int_0^1 r W_j(r) dr = a_j \int_0^1 r dr + b_j \int_0^1 r^2 dr$$

の解であり、確率変数である。他方、 $a$  と  $b$  は、方程式

$$\int_0^1 r^2 dr = a + b \int_0^1 r dr$$

$$\int_0^1 r^3 dr = a \int_0^1 r dr + b \int_0^1 r^2 dr$$

の解、すなわち、 $a = -1/6$ 、 $b = 1$  である。この場合の分布表は、Johansen (1995) の Table 15.5 に与えられている。

本節の冒頭で述べたように、共和分ランクの検定は、以上の他にも、多くのモデルに対して考案されている。例えば、Johansen-Schaumburg (1999) は、季節性を含むモデルを考察し、Jeganathan (1999) は、誤差項がフラクショナルな  $I(d)$  過程の場合の共和分検定を議論している。これらの詳細については、Hubrich-Lütkepohl-Saikkonen (2001) を参照されたい。

[参考文献]

- Anderson, T. W. (1951) "Estimating linear restrictions on regression coefficients for multivariate normal distributions," *Annals of Mathematical Statistics*, Vol. 22, pp. 327-351.
- Billingsley, P. (1968) *Convergence of Probability Measures*, New York: Wiley.
- Box, G. E. P. and Jenkins, G. M. (1976) *Time Series Analysis: Forecasting and Control*, Second Edition, San Francisco: Holden-Day.
- Chan, N. H. and Wei, C. Z. (1988) "Limiting distributions of least squares estimates of unstable autoregressive processes," *Annals of Statistics*, Vol. 16, pp. 367-401.
- Csörgő, K. and Horváth, L. (1993) *Weighted Approximations in Probability and Statistics*, New York: Wiley.
- Davidson, R. and MacKinnon, J. G. (1993) *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.
- Davis, R. A. and Dunsmuir, W. T. M. (1996) "Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle," *Econometric Theory*, Vol. 12, 1-29.
- Dickey, D. A. and Fuller, W. A. (1979) "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, Vol. 74, 427-431.
- Dickey, D. A., Hasza, D. P., and Fuller, W. A. (1984) "Testing for unit roots in seasonal time series," *Journal of the American Statistical Association*, Vol. 79, 355-367.
- Durlauf, S. N. and Phillips, P. C. B. (1988) "Trends versus random walks in time series analysis," *Econometrica*, Vol. 56, 1333-1354.

- Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996) "Efficient tests for an autoregressive unit root test," *Econometrica*, Vol. 64, 813-836.
- Engle, R. F. and Granger, C. W. J. (1987) "Co-integration and error correction: Representation, estimation, and testing," *Econometrica*, Vol. 55, pp. 251-276.
- Fuller, W. A. (1996) *Introduction to Statistical Time Series*, Second Edition, New York: Wiley.
- Granger, C. W. J. and Newbold, P. (1974) "Spurious regressions in econometrics," *Journal of Econometrics*, Vol. 2, 111-120.
- Hatanaka, M. (1996) *Time-Series-Based Econometrics*, Oxford: Oxford University Press.
- Hosking, J. R. M. (1981) "Fractional differencing," *Biometrika*, Vol. 68, 165-176.
- Hubrich, K., Lütkepohl, H., and Saikkonen, P. (2001) "A review of systems cointegration tests," *Econometric Reviews*, Vol. 20, 247-318.
- Jeganathan, P. (1999) "On asymptotic inference in cointegrated time series with fractionally integrated errors," *Econometric Theory*, Vol. 15, 583-621.
- Johansen, S. (1988) "Statistical analysis of cointegrating vectors," *Journal of Economic Dynamics and Control*, Vol. 12, 231-254.
- Johansen, S. (1991) "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica*, Vol. 59, 1551-1580.
- Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

- Johansen, S. and Schaumburg, E. (1999) "Likelihood analysis of seasonal cointegration," *Journal of Econometrics*, Vol. 88, 301-339.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) "Testing the null hypothesis of stationarity against the alternative of a unit root," *Journal of Econometrics*, Vol. 54, 159-178.
- Nabeya, S. (1999) "Asymptotic moments of some unit root test statistics in the null case," *Econometric Theory*, Vol. 15, 139-149.
- Nabeya, S. and Tanaka, K. (1988) "Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative," *Annals of Statistics*, Vol. 16, 218-235.
- Nabeya, S. and Tanaka, K. (1990a) "A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors," *Econometrica*, Vol. 58, 145-163.
- Nabeya, S. and Tanaka, K. (1990b) "Limiting powers of unit-root tests in time-series regression," *Journal of Econometrics*, Vol. 46, 247-271.
- Osterwald-Lenum, M. (1992) "A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics," *Oxford Bulletin of Economics and Statistics*, Vol. 54, 461-471.
- Perron, P. (1989) "The great crash, the oil price shock, and the unit root hypothesis," *Econometrica*, Vol. 57, 1361-1401.
- Phillips, P. C. B. (1986) "Understanding spurious regressions in econometrics," *Journal of Econometrics*, Vol. 33, 311-340.
- Phillips, P. C. B. (1987) "Time series regression with a unit root," *Econometrica*, Vol. 55, 277-301.
- Phillips, P. C. B. (1998) "New tools for understanding spurious regressions," *Econometrica*, Vol. 66, 1299-1325.

- Phillips, P. C. B. and Hansen, B. E. (1990) "Statistical inference in instrumental variables regression with I(1) processes," *Review of Economic Studies*, Vol. 57, 99-125.
- Phillips, P. C. B. and Ouliaris, S. (1990) "Asymptotic properties of residual based tests for cointegration," *Econometrica*, Vol. 58, 165-193.
- Phillips, P. C. B. and Solo, V. (1992) "Asymptotics for linear processes," *Annals of Statistics*, Vol. 20, 971-1001.
- Robinson, P. M. (1994) "Efficient tests of nonstationary hypotheses," *Journal of the American Statistical Association*, Vol. 89, 1420-1437.
- Tanaka, K. (1990) "Testing for a moving average unit root," *Econometric Theory*, Vol. 6, 433-444.
- Tanaka, K. (1993) "An alternative approach to the asymptotic theory of spurious regression, cointegration, and near cointegration," *Econometric Theory*, Vol. 9, 36-61.
- Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*, New York: Wiley.
- Tanaka, K. (1999) "The nonstationary fractional unit root," *Econometric Theory*, Vol. 15, 549-582.
- Tanaka, K. (2001) "K-asymptotics associated with deterministic trends in the integrated and near-integrated processes," *Japanese Economic Review*, Vol. 52, 35-63.
- Tanaka, K. and Satchell, S. E. (1989) "Asymptotic properties of the maximum-likelihood and nonlinear least-squares estimators for noninvertible moving average models," *Econometric Theory*, Vol. 5, 333-353.
- Vogelsang, T. J. and Perron, P. (1998) "Additional tests for a unit root allowing for a break in the trend function at an unknown time," *International Economic Review*, Vol. 39, 1073-1100.

Zivot, E. and Andrews, D. W. (1989) "Further evidence on the great crash, the oil price shock, and the unit root hypothesis," *Journal of Business and Economic Statistics*, Vol. 10, 251-270.

表 3-1  $T(\hat{\rho} - 1)$  の極限帰無分布の特性値

モデル	1% 点	5% 点	10% 点	平均	標準偏差	歪度	尖度
A	-13.695	-8.039	-5.714	-1.781	3.180	-2.270	8.369
B	-20.626	-14.094	-11.251	-5.379	4.511	-1.551	4.015
C	-29.359	-21.711	-18.245	-10.246	6.033	-1.204	2.373

表 3-2  $t_{\hat{\rho}-1}$  の極限帰無分布の特性値

モデル	1% 点	5% 点	10% 点	平均	標準偏差	歪度	尖度
A	-2.58	-1.95	-1.62	-0.423	0.981	0.250	0.083
B	-3.42	-2.86	-2.57	-1.533	0.840	0.218	0.334
C	-3.96	-3.41	-3.13	-2.181	0.750	0.057	0.371

図2-1 TOPIXと為替レートの対数値系列

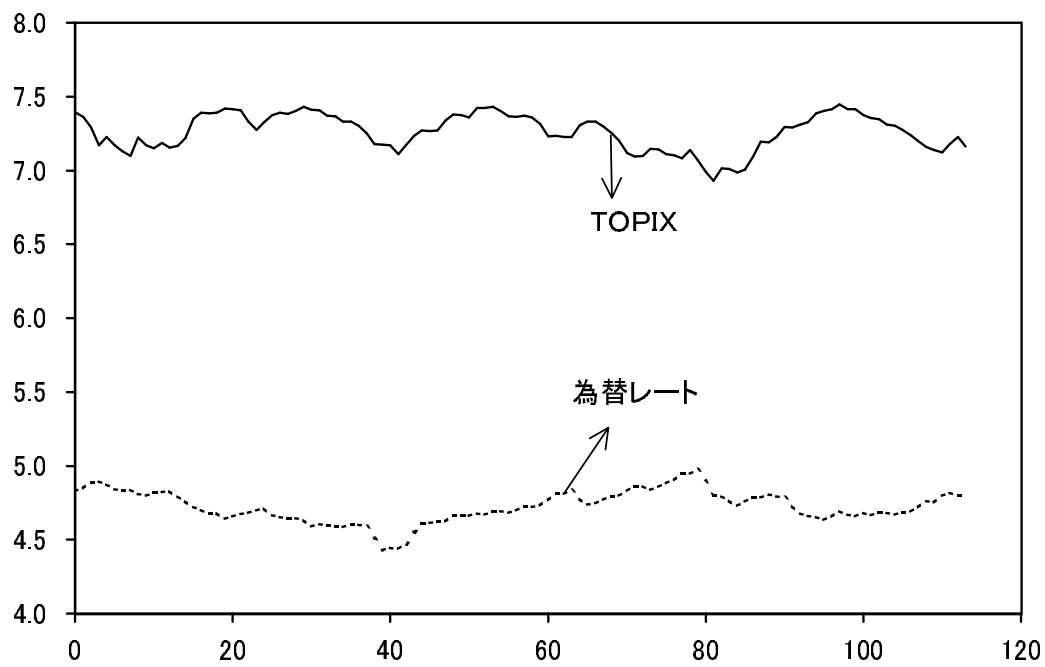


図 2-2 自己相関係数

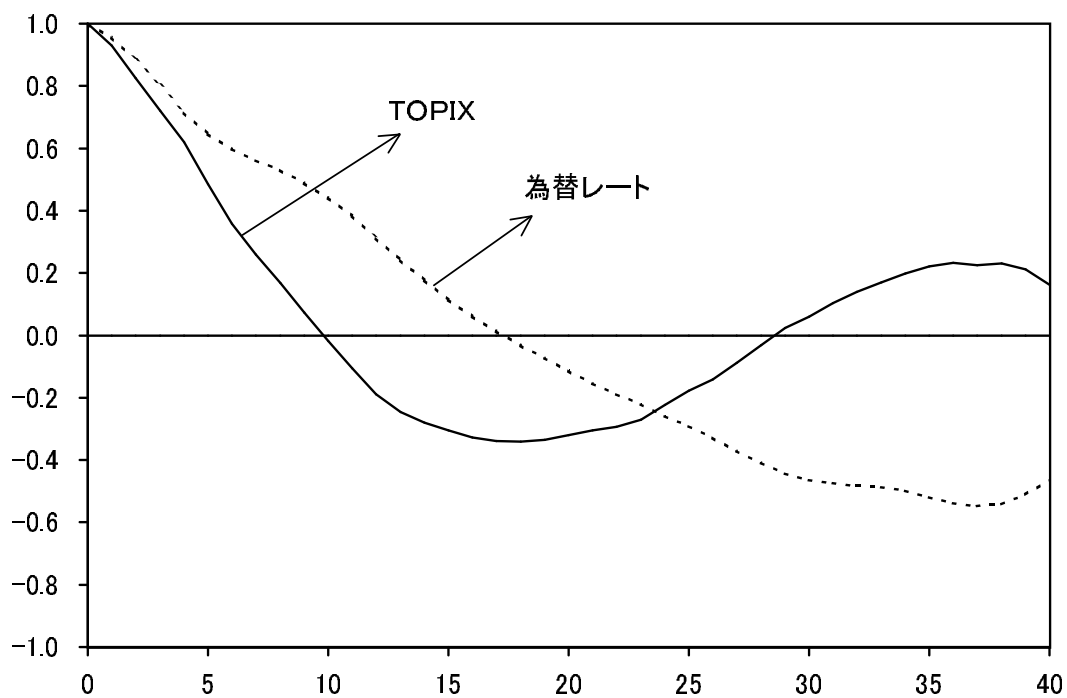


図2-3 TOPIXと為替レート対数値の階差系列

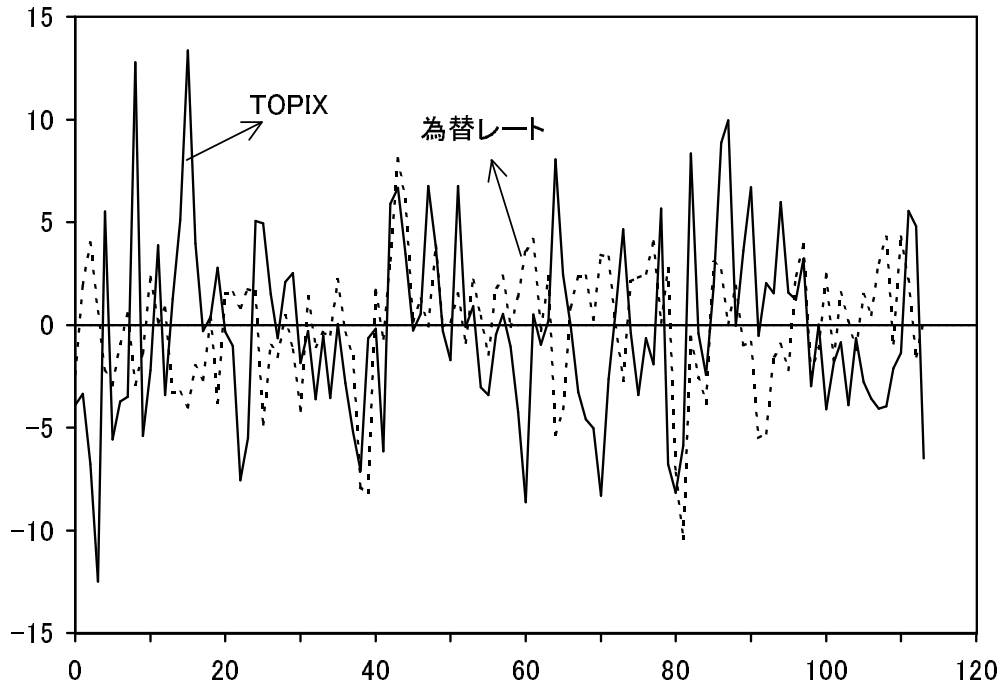


図2-4 部分和過程

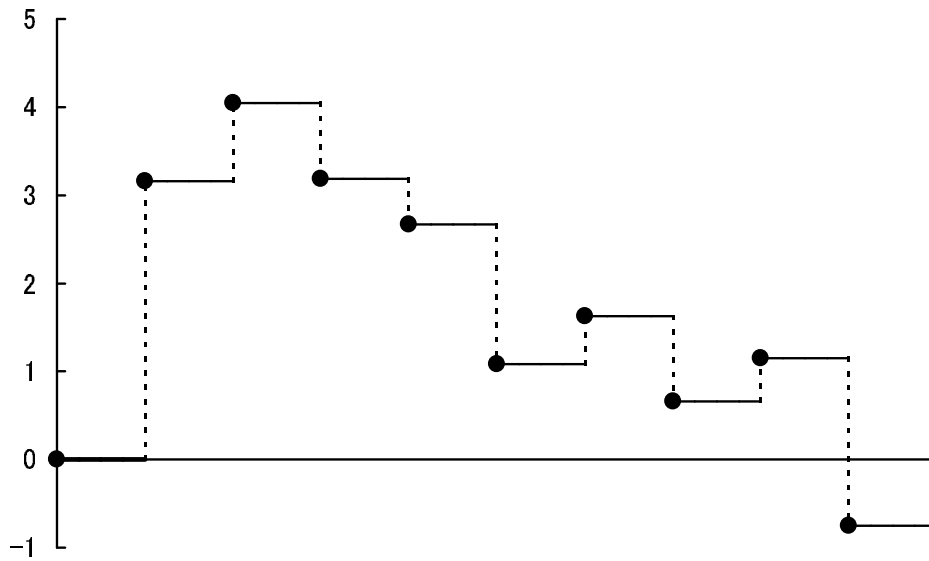


図2-5 ブラウン運動

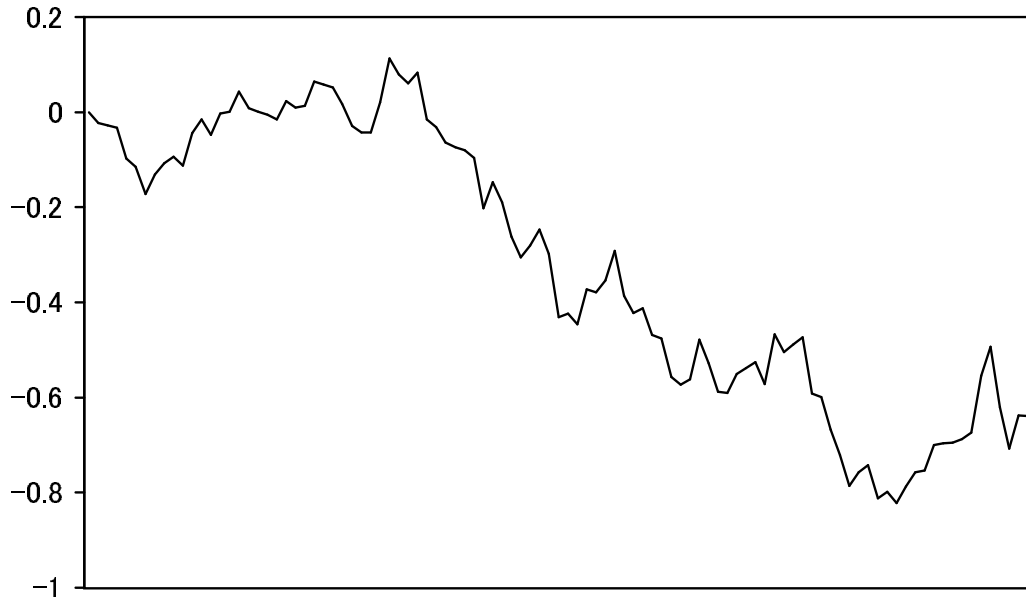


図3-1  $T(\hat{\rho} - 1)$ の極限帰無分布

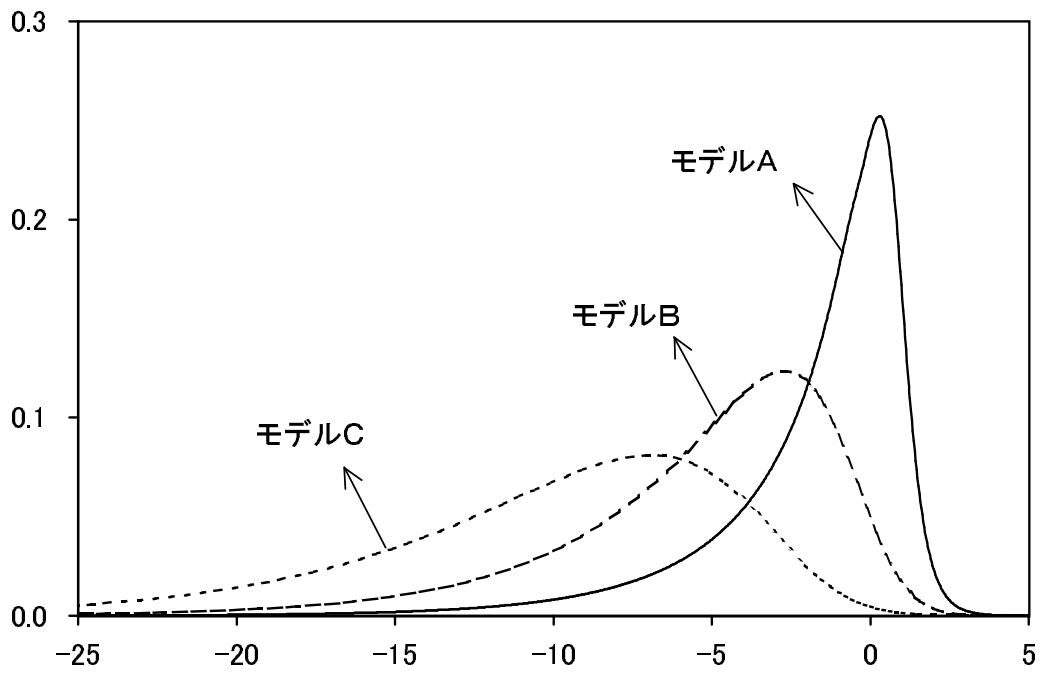


図3-2 単位根検定の局所検出力

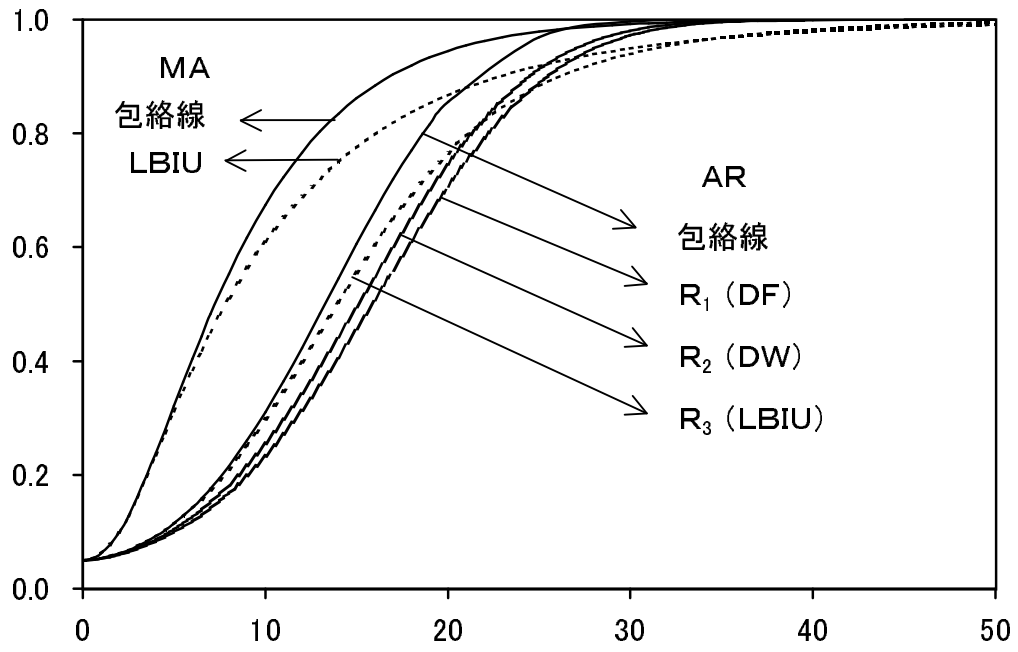


図4-1 共和分関係の有無による係数推定量の分布

