

ADF-GLS 検定とその用例

坂野 慎 哉

I はじめに

時系列データを用いて回帰分析を行うとき、分析に先立って、利用されるデータの系列に単位根が含まれているかどうか、検定を行ってチェックすることが多い。これは、単位根を含む時系列データを説明変数、被説明変数として回帰分析を行った場合、たとえそれらの変数が独立であったとしても回帰係数の推定量は0に収束しないという、「見せかけ回帰」という現象が起りうるからである。もちろん、独立な変数同士を回帰分析するのであれば、係数推定値が0に近い値をとらないことは望ましくない現象である。

上記のような単位根の検定の手法として代表的なものに、Fuller [1976] および Dickey and Fuller [1979] によって与えられた、Dickey-Fuller 検定（以下 DF 検定と略す）や、その拡張である、Augmented Dickey-Fuller 検定（以下 ADF 検定と略す）がある。ある時系列データ y_1, \dots, y_T が、次のような2つの式で示されるデータ生成過程（data generating process：以下 DGP と略す）から発生していると考えられているとする。

$$y_t = d_t + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (t=1, \dots, T)$$
 (1)
ここで、 $\{d_t\}$ は確定的（すなわち確率変数でない）要素であり、たとえば定数項や、「トレンド変数」と呼ばれる、時間とともに1ずつ増大していく変数とその係数からなる項がそれにあたる。 $\{v_t\}$ は、平均0で定常な観測されない過程とする。DF 検定や ADF 検定は、(1)における

α が1に等しいという帰無仮説を、 α が1より小さいという対立仮説に対して検定する検定手法である。帰無仮説が棄却されると、検定の対象となる系列は単位根を含まないと判断される。帰無仮説が棄却されなければ、その系列が単位根を含まないと判断することはできない。

さて、DF 検定や ADF 検定、とくに ADF 検定は広く使われている単位根検定であるが、その検出力、すなわち対立仮説が正しいという条件のもとで帰無仮説を棄却する確率が、低いことが知られている。もちろん、対立仮説が正しいければ帰無仮説は誤りであり、帰無仮説が誤っているときに帰無仮説を棄却できないのでは、その検定結果は誤判断となるから、検定にとって検出力が低いことは望ましくない性質である。

ADF 検定（もしくは DF 検定）の上記の欠点を改善した検定手法はいくつか考案されているが、Elliott, Rothenberg, and Stock [1996]（以下、ERS と略す）による ADF-GLS 検定（ERS 自身は「DF-GLS 検定」と記している）はその1つである。

小稿は、森棟 [1999] の第9章までの知識のみを前提として、ADF-GLS 検定の用法について説明する。以下、小稿の構成を述べる。第II章では、小稿の理解に必要な最低限の ADF 検定の概要について復習する。第III章では、ADF-GLS 検定の手法について具体的に説明する。第IV章では、ADF-GLS 検定や ADF 検定を行う上で必要となる、ラグ次数の選択の方法について述べる。第V章では、ADF-GLS 検定を用いて我が国のマクロ経済データを分析した

例を示し、同じデータ系列を ADF 検定を用いて分析した結果と比較する。

II ADF 検定の概要

本章では、ADF-GLS 検定の説明に必要な ADF 検定の用法の知識について、復習しておく。

ある時系列データ y_1, \dots, y_T が、小稿第 I 章の (1) においてよりも確定的要素 $\{d_t\}$ をより具体的にした、次のような 2 つの式で示される DGP から発生していると考えられているとする。

$$y_t = \beta_0 + \beta_1 t + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (2)$$

ここで、 β_0, β_1 は未知パラメータであり、 t はトレンド変数である。 $\beta_1 t$ は「線形トレンド項」と呼ばれる。 $\{v_t\}$ については、独立で同一の分布をする平均 0、分散一定の観測されない過程であると仮定する。これは、(1) におけるよりも強い仮定である。第 I 章で述べたように、ADF 検定や DF 検定は、帰無仮説 $H_0: \alpha = 1$ を、対立仮説 $H_1: \alpha < 1$ に対して検定する。

いま (2) について、その右側の式の両辺から $\{u_{t-1}\}$ を引き、次のように書き換える。

$$y_t = \beta_0 + \beta_1 t + u_t, \quad \Delta u_t = a_0 u_{t-1} + v_t \quad (3)$$

ここで、 $\Delta u_t \equiv u_t - u_{t-1}$ 、 $a_0 \equiv \alpha - 1$ である。よって、 $\alpha = 1$ ならば $a_0 = 0$ であり、帰無仮説 $H_0: a_0 = 0$ を、対立仮説 $H_1: a_0 < 0$ に対して検定する検定が、(3) における単位根検定となる。したがって、仮に (3) における $\{u_t\}$ が観測可能な系列だとすると、(3) の右側の式を、右辺の $\{v_t\}$ が誤差項である、定数項のない単純回帰モデルと見ることができ、 t 統計量を検定統計量とする a_0 の有意性検定（以下これを「 t 検定」と呼ぶことにする）を行うことにより、単位根検定を行うことができる。その場合、検定は左片側検定になる。

実際には、(3) の左側の式の β_0, β_1 が未知パ

ラメータであることにより、 $\{u_t\}$ は観測可能な系列ではない。よって、 $\{u_t\}$ に代わる観測可能な系列が必要になる。そのために、(3) の左側の式を、 $\{u_t\}$ を誤差項とする単純回帰モデルとみなし、通常の最小二乗法（Ordinary Least Squares Method: 以下 OLS と略す）によって、 β_0, β_1 の OLS 推定量 $\hat{\beta}_0, \hat{\beta}_1$ をそれぞれ求め、それらから OLS 残差の系列 $\{\hat{u}_t\}$ を求める。すなわち、

$$\hat{u}_t \equiv y_t - \hat{\beta}_0 - \hat{\beta}_1 t \quad (4)$$

である。この $\{\hat{u}_t\}$ は OLS 残差であるから観測可能である。そして、(3) の右側の式の $\{u_t\}$ を、(4) で定義される $\{\hat{u}_t\}$ で置き換えた式

$$\Delta \hat{u}_t = a_0 \hat{u}_{t-1} + \text{error} \quad (5)$$

を考える。ただし、(5) 右辺の「error」は適当な誤差項を表すものとする。(3) の左側の式において誤差項とみなされた $\{u_t\}$ に、たとえ系列相関があつたりその分散が均一でなかったとしても、 $\hat{\beta}_0$ や $\hat{\beta}_1$ はそれぞれ β_0 と β_1 の不偏推定量であり一致推定量であるので、 $\{\hat{u}_t\}$ は $\{u_t\}$ のよい推定値とみなすことができ、単位根検定は、(5) における a_0 の t 検定によって行える。この t 検定が DF 検定である。ただし、このときの a_0 の t 統計量（DF 検定を提案した Dickey and Fuller [1979] はこの t 統計量を $\hat{\tau}_\tau$ と記したので、小稿でも以下この記法を用いる）は、帰無仮説 $H_0: a_0 = 0$ のもとで通常の t 分布に従わず、ある特殊な分布をする。そのためこの DF 検定においては、与えられた有意水準に対応した検定の臨界値を求めるために、特別な分布表が必要となる。それは Fuller [1976] によって与えられたが、その後 Fuller [1996] によって若干修正されている。森棟 [1999] 319 ページには、自由度 20 の t 分布の密度関数のグラフと、 $\hat{\tau}_\tau$ の漸近的な密度関数のグラフを対比した図が与えられている。

以上で説明した検定の方法は、時系列データ y_1, \dots, y_T の DGP が (2)、あるいは同じことだが

(3)のように、定数項と線形トレンド項の両方を含んでいると想定されている場合のものである。しかし、 y_1, \dots, y_T の DGP に線形トレンド項がないと想定できる場合、すなわち、

$$y_t = \beta_0 + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (6)$$

もしくは(3)と同様に $\Delta u_t \equiv u_t - u_{t-1}$, $a_0 \equiv \alpha - 1$ として

$$y_t = \beta_0 + u_t, \quad \Delta u_t = a_0 u_{t-1} + v_t \quad (7)$$

と想定されている場合にも、検定の手続きは同様である。具体的には、(6)もしくは(7)の右側の式を、定数項しかない回帰モデルとみなして OLS 推定し、 β_0 の OLS 推定量 $\hat{\beta}_0$ (これは実際には y_t の標本平均と等しくなる) を用いて

$$\hat{u}_t \equiv y_t - \hat{\beta}_0 \quad (8)$$

から得られる残差の系列 $\{\hat{u}_t\}$ を用いて(5)と同じ式を考え、その式における a_0 の t 検定を行う。Dickey and Fuller [1979] はこの場合の t 統計量を $\hat{\tau}_\mu$ と記したので、小稿でも以下この記法を用いるが、この $\hat{\tau}_\mu$ の $H_0: a_0 = 0$ のもとでの分布は、通常の t 分布とも、上述の $\hat{\tau}_\tau$ の分布とも異なる。そのため、検定の臨界値を求めるためには、やはりこの場合のための特別な分布表が必要となる。それも Fuller [1996] が与えている。この検定も DF 検定と呼ばれる。

さらに、 y_1, \dots, y_T の DGP に線形トレンド項だけでなく定数項もないと想定できる場合、すなわち $y_t = \alpha y_{t-1} + v_t$ とできる場合、もしくは $\Delta y_t \equiv y_t - y_{t-1}$, $a_0 \equiv \alpha - 1$ として

$$\Delta y_t = a_0 y_{t-1} + v_t \quad (9)$$

と想定されている場合には、(9)を直接 OLS 推定し、 a_0 の t 検定を行う。Dickey and Fuller [1979] はこの t 統計量を $\hat{\tau}$ と記したので、小稿でも以下この記法を用いるが、この $\hat{\tau}$ の $H_0: a_0 = 0$ のもとでの分布は、通常の t 分布とも、上述の $\hat{\tau}_\tau$ の分布とも異なる。この場合の検定の臨界値を求めるための特別な分布表も、やはり Fuller [1996] が与えている。この場合も DF 検定と呼ばれる。

これまでは、 y_1, \dots, y_T の DGP における $\{v_t\}$ が、独立で同一の分布をする平均 0、分散一定の観測されない過程であると仮定していた。次に、この仮定をゆるめ、 $\{v_t\}$ が次数 p の定常な AR 過程であると想定できる場合を考える。すなわち、 $\{\eta_t\}$ がホワイト・ノイズ (平均 0、分散一定、自己共分散が全て 0 の確率変数の系列) であるとするとき、

$$v_t = \phi_1 v_{t-1} + \phi_2 v_{t-2} + \dots + \phi_p v_{t-p} + \eta_t \quad (10)$$

となっていると想定できるとする。(10)は、ある実数 k について $L^k v_t \equiv v_{t-k}$ (ただし $L^1 \equiv L$) という意味を持つラグ演算子 L を含む多項式 (ラグ多項式) を用い、

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) v_t = \eta_t \quad (11)$$

とも書ける。ただし、 $\{v_t\}$ が定常であるためには、 $\phi_1, \phi_2, \dots, \phi_p$ を係数に持つ多項式

$$1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0 \quad (12)$$

は、全ての根が絶対値で 1 より大きくなくてはならないことが知られている。

y_1, \dots, y_T の DGP において、 $\{v_t\}$ 以外はこれまでと同様に

$$y_t = \beta_0 + \beta_1 t + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (13)$$

となっているとするとき、(13)の右側の式はラグ演算子を用いて $(1 - \alpha L)u_t = v_t$ と書けるから、(11)は

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - \alpha L)u_t = \eta_t \quad (14)$$

のようにも書ける。そして(14)は、次のように書き直せることが知られている。

$$\begin{aligned} \Delta u_t = & a_0 u_{t-1} + a_1 \Delta u_{t-1} + a_2 \Delta u_{t-2} \\ & + \dots + a_p \Delta u_{t-p} + \eta_t \end{aligned} \quad (15)$$

ただし、 $a_0 = (\alpha - 1)(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ であり、 a_1, a_2, \dots, a_p は対応して適当に定義されるパラメータである。

上述のように、 $\{v_t\}$ が定常と仮定されている場合、多項式(12)は全ての根が絶対値で 1 より大きいことから、(12)の左辺の z に 1 を代入した式は 0 とはなりえないはずである。すなわち、 $1 - \phi_1 - \phi_2 - \dots - \phi_p \neq 0$ である。したがっ

て、(15)の a_0 が0となるのは $\alpha=1$ のとき以外あり得ず、よって、 $\{v_t\}$ のAR次数 p が既知であつてかつ $\{u_t\}$ が可観測であるならば、この場合もこれまでと同様、 a_0 についてのt検定が、DGPが単位根を含むかどうかの検定にもなっていることがわかる。

実際には、やはりこれまでと同様、(15)における観測不能な $\{u_t\}$ を(13)の左側の式をOLS推定して得られる残差 $\{\hat{u}_t\}$ に代えた、

$$\Delta\hat{u}_t = a_0\hat{u}_{t-1} + a_1\Delta\hat{u}_{t-1} + a_2\Delta\hat{u}_{t-2} + \dots + a_p\Delta\hat{u}_{t-p} + \text{error} \quad (16)$$

における右辺第1項の係数 a_0 についてt検定を行うのだが、このt検定がADF検定である。このときの a_0 のt統計量は、帰無仮説 $H_0: a_0=0$ のもとでの極限分布が $\hat{\tau}$ と同じになるので、この検定において臨界値を求める際には $\hat{\tau}$ の分布表が使える。

DF検定には、 y_1, \dots, y_T のDGPに線形トレンド項が含まれない場合の検定手法もあった。ADF検定においても同様に、DGPに線形トレンド項が含まれない場合の検定手法があり、その手続きは、(5)の代わりに(16)を用いる点が異なるほかは、線形トレンド項が含まれない場合のDF検定と同様である。このときの a_0 のt統計量は、帰無仮説 $H_0: a_0=0$ のもとでの極限分布が $\hat{\tau}_\mu$ と同じになるので、この検定において臨界値を求める際には $\hat{\tau}_\mu$ の分布表が使える。

さらにDF検定には、 y_1, \dots, y_T のDGPに線形トレンド項も定数項も含まれない場合の検定手法もあった。ADF検定においても同様に、DGPに線形トレンド項も定数項も含まれない場合の検定手法があり、それは、

$$\Delta y_t = a_0 y_{t-1} + a_1 \Delta y_{t-1} + a_2 \Delta y_{t-2} + \dots + a_p \Delta y_{t-p} + \text{error} \quad (17)$$

を直接OLS推定し、 a_0 のt検定を行うことにより行われる。(9)の代わりに(17)を用いる点が異なるほかは、DF検定と同様である。この

ときの a_0 のt統計量は、帰無仮説 $H_0: a_0=0$ のもとでの極限分布が $\hat{\tau}$ と同じになるので、この検定において臨界値を求める際には $\hat{\tau}$ の分布表が使える。

ここまでのADF検定の説明では、 $\{v_t\}$ のAR次数 p が既知と仮定されていた。 p は、ADF検定を行う際に推定される回帰式(16)や(17)における、右辺の階差をとったラグ項の最高ラグ次数でもあるので、これがわからなければADF検定はできない。そしてこの p は、実際には未知であるので、何らかの方法で選択しなくてはならない。この p の選択方法については、第IV章で述べる。

III ADF-GLS 検定の方法

小稿第II章と同様、ある時系列データ y_1, \dots, y_T が、次のような2つの式で示されるDGPから発生していると考えられているとする。

$$y_t = \beta_0 + \beta_1 t + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (18)$$

ただし、 $\{v_t\}$ は、次数 p の定常なAR過程であるとする。これまでと同様、帰無仮説 $H_0: \alpha=1$ を、対立仮説 $H_1: \alpha < 1$ に対して検定する検定を考えている。帰無仮説が棄却された場合、 y_1, \dots, y_T は単位根を含まない過程であると判断される。

前章で説明したように、ADF検定は2段階で行われる。DGPとして(18)を想定している場合なら、第1段階で(18)の左側の式のOLS残差を求め、第2段階ではそのOLS残差を用いて(16)を推定し、検定統計量を求める。ADF-GLS検定も同様で、その手続きは2段階である。

(18)の左側の式は、もし $\{u_t\}$ を誤差項と考えるならば、 t を説明変数とし、 β_0, β_1 を係数とする単純回帰モデルとみなせる。このとき、 β_0, β_1 を係数として持ちながら、誤差項が $\{u_t\}$ では

なく $\{v_t\}$ になるようにモデルを書き換えてみる。その書き換えは、次のように行える。(18)の左側の式の両辺に α をかけ、時間を 1 期前にずらすと、

$$\alpha y_{t-1} = \alpha \beta_0 + \alpha \beta_1(t-1) + \alpha u_{t-1} \quad (19)$$

となる。そしてこの (19) を (18) の左側の式から辺引いた式は、(18) の右側の式より、

$$y_t - \alpha y_{t-1} = \beta_0(1-\alpha) + \beta_1\{t-\alpha(t-1)\} + v_t \quad (20)$$

と書ける。 α は検定の対象でありもちろん未知であるが、いま既知であると仮定すると、(20) 左辺の $y_t - \alpha y_{t-1}$ 、右辺の $1-\alpha$ と $t-\alpha(t-1)$ はいずれも求めることができ、これらをそれぞれ 1 つの変数とみなすことにすれば、 $\{v_t\}$ が適当な性質を持つ誤差項と考えることができる場合、(20) を OLS 推定して β_0 、 β_1 の推定値を求めることができる。

ところで、仮に $\{u_t\}$ が (18) の左側の式の誤差項とみなすことができ、しかもそれが回帰モデルの誤差項の古典的仮定を満たしているとする、この式を OLS 推定して得られる β_0 、 β_1 の推定量の分散は、 β_0 、 β_1 の線形不偏推定量の分散の中で最も小さいものとなる。一方、(18) 右側の式の $\{v_t\}$ が独立な平均 0、分散一定の正規確率変数の系列であるとき、誤差項 $\{u_t\}$ は AR(1) となるが、その場合には、(18) の左側の式を OLS 推定して得られる β_0 、 β_1 の推定量の分散はそのような性質を持つ保証がない。しかし、 α が既知の場合に、上述のように (18) の左側の式を (20) のように変換 (Cochrane-Orcutt 変換と呼ばれる) してから OLS 推定して求めた β_0 、 β_1 の推定量の分散は、線形不偏推定量の分散の中で最も小さいものになる。この推定法は、一般化最小二乗法 (Generalized Least Squares Method: GLS と略される) と総称される推定手法の一種である。

ここでの y_1, \dots, y_T の DGP においては、 $\{v_t\}$ は定常な AR(p) と想定されているので、(20) は上述の場合とは前提が異なるが、ADF-GLS 検

定ではまず、形式的に GLS と同じ手続きを行って β_0 、 β_1 を推定する。すなわち、(20) における未知の α を $\bar{\alpha} \equiv 1 + \bar{c}/T$ で置き換え (\bar{c} はある定数)、さらに (20) の $\{v_t\}$ を古典的な誤差項に代えた次の式を、OLS で推定するのである。

$$y_t - \bar{\alpha} y_{t-1} = \beta_0(1-\bar{\alpha}) + \beta_1\{t-\bar{\alpha}(t-1)\} + error \quad (21)$$

(21) の「error」は誤差項を示す。OLS 推定にあたっては、 $y_t - \bar{\alpha} y_{t-1}$ 、 $1-\bar{\alpha}$ 、 $t-\bar{\alpha}(t-1)$ は、それぞれ 1 つの変数とみなされる。ただし、1 期目のデータは、それぞれ y_1 、1、 t (すなわち 1) とする。ERS は DGP が (18) の場合、 \bar{c} の値を -13.5 にすることを奨めている。

上記の方法で (21) を推定して得られる β_0 、 β_1 の推定値を、それぞれ $\tilde{\beta}_0$ 、 $\tilde{\beta}_1$ とおく。これらの推定値を用いて、

$$y_t^d \equiv y_t - \tilde{\beta}_0 - \tilde{\beta}_1 t \quad (22)$$

なる式で系列 $\{y_t^d\}$ を求める。(22) の右辺は、通常の残差を求める式

$$y_t - \bar{\alpha} y_{t-1} - \tilde{\beta}_0(1-\bar{\alpha}) - \tilde{\beta}_1\{t-\bar{\alpha}(t-1)\} \quad (23)$$

とは異なる。ここまでが ADF-GLS 検定の第 1 段階といえる部分である。

ADF 検定では第 2 段階において、第 1 段階で求めた OLS 残差からなる式 (16) を OLS 推定したが、同様に ADF-GLS 検定では、第 1 段階で求めた $\{y_t^d\}$ からなる次の式を OLS 推定する。

$$\Delta y_t^d = a_0 y_{t-1}^d + a_1 \Delta y_{t-1}^d + a_2 \Delta y_{t-2}^d + \dots + a_p \Delta y_{t-p}^d + error \quad (24)$$

ここで、 p は $\{v_t\}$ の AR 次数である。 p は通常未知であるから、何らかの方法で選択する必要があるが、その方法については第 IV 章で述べる。(24) の a_0 が 0 かどうかを検定する t 検定が、ADF-GLS 検定である。ただし、このときの t 検定統計量は、ADF 検定のときと同様に通常の t 分布には従わないし、さらには ADF 検定の検定統計量 $\hat{\tau}_c$ の分布とも異なる。そのため、ADF-GLS 検定の臨界値を求めるためには独自

の分布表が必要になるが、それはERSの「TABLE1」の「C」の部分にある。 T が無限大の場合について一部引用しておく、左から1%点 = -3.48, 2.5%点 = -3.15, 5%点 = -2.80, 10%点 = -2.57である。

次に、 y_1, \dots, y_T のDGPに線形トレンド項が含まれていない場合、すなわちDGPが

$$y_t = \beta_0 + u_t, \quad u_t = \alpha u_{t-1} + v_t \quad (25)$$

となっている場合の、ADF-GLS検定の方法を紹介する。ここで、(25)の $\{v_t\}$ は、やはり次数 p の定常なAR過程であるとする。

DGPに線形トレンド項が含まれる場合、第1段階では、DGP(18)に対応して(21)をたててOLS推定するというのであった。いまの場合の第1段階では、同様にしてDGP(25)から対応する式

$$y_t - \bar{\alpha} y_{t-1} = \beta_0 (1 - \bar{\alpha}) + \text{error} \quad (26)$$

をたてて、 $y_t - \bar{\alpha} y_{t-1}$ 、 $1 - \bar{\alpha}$ をそれぞれ1つの変数とみなしてOLS推定する。ここで、やはり $\bar{\alpha} \equiv 1 + \bar{c}/T$ であるが、ERSはこの場合には \bar{c} を-7とすることを奨めている。(26)は、(21)からDGPにおける線形トレンド項を除いたものに対応している。(26)をOLS推定して得られる β_0 の推定値を $\hat{\beta}_0$ とおき、DGPに線形トレンド項が含まれる場合における(22)に対応する式、

$$y_t^d \equiv y_t - \hat{\beta}_0 \quad (27)$$

で系列 $\{y_t^d\}$ を求める。ここまでが第1段階である。

第2段階においては、線形トレンド項が含まれる場合における(24)と同じ式を、(27)から求めた $\{y_t^d\}$ を用いてたて、OLS推定する。このときの(24)の a_0 が0かどうかを検定するt検定が、DGPに線形トレンド項が含まれない場合のADF-GLS検定である。このときのt検定統計量は、ADF検定のときと同様に通常のt分布には従わないが、線形トレンド項が含まれる場合のADF-GLS検定とは異なり、その極限

分布は、DGPに線形トレンド項も定数項も含まれない場合のADF検定の検定統計量 $\hat{\tau}$ の極限分布と同じになる。そのため、検定の臨界値を求めるにあたっては、Fuller [1976]が作成した $\hat{\tau}$ の表が使える。一部引用しておく(もちろん T は無限大)、左から1%点 = -2.58, 2.5%点 = -2.23, 5%点 = -1.95, 10%点 = -1.62である。

黒住 [2008]では、ADF-GLS検定の漸近的な検出力曲線とADF検定の漸近的な検出力曲線とを比較したグラフが提示されている。それを見ると、定数項と線形トレンド項の両方がDGPに含まれる場合でも、定数項のみがDGPに含まれる場合でも、ADF-GLS検定の漸近的な検出力のほうがADF検定のそれよりも高いことがはっきりと見て取れる。

なおERSは、 y_1, \dots, y_T のDGPに定数項も線形トレンドも含まれていない場合、ADF検定の検出力は十分高いことを明らかにした。ゆえにERSは、その場合の代替的検定手法は提案していない。

IV ラグ次数 p の選択

第II章や第III章で説明したように、ADF検定における(16)や(17)、ADF-GLS検定における(24)は、最高次数が p のラグ項を含んでいる。 p は $\{v_t\}$ のAR次数であるが、 $\{v_t\}$ は観測不能であるから p も未知であり、何らかの方法で選択する必要がある。

p を選択する手法の1つに、 p の「仮の値」 p^* をとりあえず決めて当該回帰式、たとえば(16)をOLS推定し、ラグ次数が p^* の係数に通常のt検定を行う、という方法がある。係数が有意でなければ p^* を1減らして(16)を再推定し、再びラグ次数が p^* の係数に通常のt検定を行う。この手続きを、ラグ次数が p^* の係数が有意になるまで繰り返し、有意になった時点での

p^* を p とする。

上記の選択手法ではラグ次数を減らしていくが、Hall [1994] は、もし上記の手法と逆にラグ次数を増やしていくというやり方をすると、ADF 検定の検定統計量の漸近分布は DF 検定の検定統計量の漸近分布に収束しないことを指摘している。

ここで、この選択手法を用いるときの、個々のラグ項の係数の有意性検定の有意水準（これを c とおく）の決め方について、森棟 [1999] の説明を多少補足しながら引用させていただく。有意水準は、帰無仮説が正しいという条件のもとで帰無仮説を棄却する確率であるから、帰無仮説が正しいという条件のもとで（すなわち個々のラグ項係数が有意ではないという条件のもとで）、当該ラグ項係数を有意でないとする正しい判断を下す確率は $1-c$ となる。検定を繰り返し行う上記の方法では p^* は順に減らしていくが、減らしていくことにより個々の検定は独立になることが知られている（森棟 [1999] 312 ページ）。そのため、全てのラグ項係数の有意性検定（たとえば (16) なら、 a_1, a_2, \dots, a_{p^*} 全ての有意性検定）を考えたとき、全てのラグ項係数が有意でないとする正しい判断をする確率は、独立な事象の積事象の確率の性質から $(1-c)^{p^*}$ となり、それゆえ、全ての係数の検定の有意水準を b （たとえば 0.05）とおくと、 $b=1-(1-c)^{p^*}$ となる。ところで、 $(1-c)^{p^*}$ を c の関数とみなしてマクローリン展開し、2 次以上の項を無視すると、

$$(1-c)^{p^*} \cong (1-0)^{p^*} - p^*(1-0)^{p^*-1}c = 1 - p^*c \quad (28)$$

となるから、結局個々のラグ項の係数の有意性検定の有意水準は、ほぼ $c=b/p^*$ とできる。

p を選択する手法として、赤池情報量基準（Akaike Information Criterion：以下 AIC と略す）や、ベイズ情報量基準（Bayesian In-

formation Criterion：以下 BIC と略す）を用いる方法もある。当該回帰式（たとえば (16)）を、ラグ次数をさまざまに変えて推定してその BIC を計算し、BIC が最小になるときの最高次のラグ次数を p として選択する。

TSP のバージョン 5.0 では、BIC は次の式で計算されている。

$$BIC = \frac{1}{2} \left\{ T(1 + \log 2\pi) + T \log \left(\frac{SSR}{T} \right) + (p+1) \log T \right\} \quad (29)$$

ここで、対数は自然対数であり、SSR は当該回帰式を OLS 推定したときの残差平方和である。一方、たとえば Hayashi [2000] では、BIC の計算式について、(29) とは異なる次の式を紹介している。

$$BIC = \log \left(\frac{SSR}{T} \right) + (p+1) \frac{\log T}{T} \quad (30)$$

もちろん、(29) と (30) では求まる BIC の値は異なる。BIC は複数個計算されたときの値の差のみに意味があり、値そのものには意味がないため、同一回帰式の BIC でもソフトウェアによって出力される値が異なることがあるし、同一のソフトウェアでもバージョンによって BIC の値が異なっていることすらある。たとえば、TSP でもバージョン 4.3A では、BIC は (30) で計算している。しかしそういう場合でも、たとえば (29) で計算される BIC と (30) で計算される BIC の場合、後者を BIC^* とおくと、

$$BIC = \frac{T}{2} \{ (1 + \log 2\pi) + BIC^* \} \quad (31)$$

となり、(31) 右辺の $1 + \log 2\pi$ は定数なので、両者は比例関係にあることになり、どちらの計算式によっても各回帰式の BIC の間の大小関係が異なるということはないし、それゆえ回帰式の選択結果も用いられる BIC の計算式によって変わるということはない。要するに、異なる計算式で計算された BIC 同士を比較しなければ問題は生じない。AIC についても同様である。

V 分析例

本章では、ADF-GLS 検定を用いた分析例を示す。

例として取り上げる時系列データは、日本の1974年1月期から2007年12月期までの鉱工業生産指数（月次、季節調整済み、 $T=408$ 、2000年=100）である。これは経済産業省の『経済産業統計』にて毎月報告されているが、小生は日経NEEDS-FAMEよりダウンロードした。データの期間を1974年1月期以降としたのは、それ以前のデータを含めると、系列が1973年末の第1次石油ショックによる経済構造の変化の影響を受けているかもしれず、構造変化を考慮しない通常のADF検定やADF-GLS検定で分析すると、検定結果に構造変化の影響が波及する恐れがあったためである。Perron [1989] は、DGPの定数項や線形トレンド項にシフトがある系列に対し、それらのシフトを考慮しない通常のDF検定を行うと、帰無仮説が棄却されづらくなることを指摘している。分析に先立ち、系列に分散安定化のため自然対数をとっておく。

比較のため、まずはADF検定を行う。

ADF検定を実際に行うにあたり、まずデータのレベル（＝階差をとらない状態）と、1階の階差とを、グラフにプロットしてみる。「図1」と「図2」は、それぞれレベルと1階階差のグラフである。

図1を見ると、この時系列にはトレンドが含まれているように見える。しかし、線形トレンドによるものか、単位根によるものか、あるいはその両方によるものなのか、この図を見ただけでは判断がつかない。そこで図2が必要になる。

たとえばいま、ある時系列データ y_1, \dots, y_T のDGPが小稿第II章(2)のようになっており、かつ帰無仮説 $H_0: \alpha=1$ が正しいと仮定しよう。すると、(2)から次の式が導ける。

$$\Delta y_t = \beta_1 + v_t \quad (32)$$

(32)より、レベル系列では線形トレンドの係数であるパラメータが、1階階差系列では定数項になることがわかる。このことから、図2のような1階階差系列をプロットしたグラフを眺め、プロットが横軸以外の目に見えない水平線

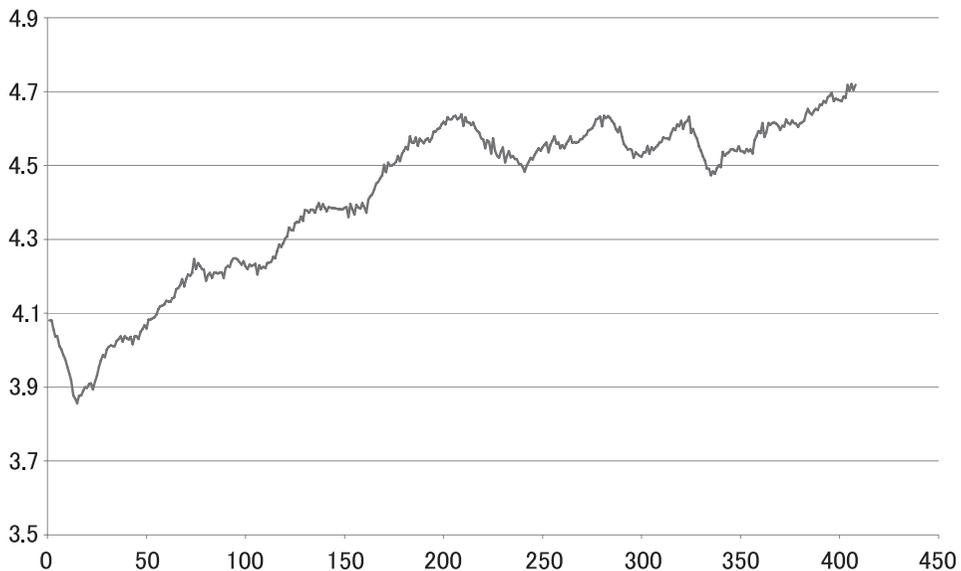


図1 鉱工業生産指数（自然対数値）

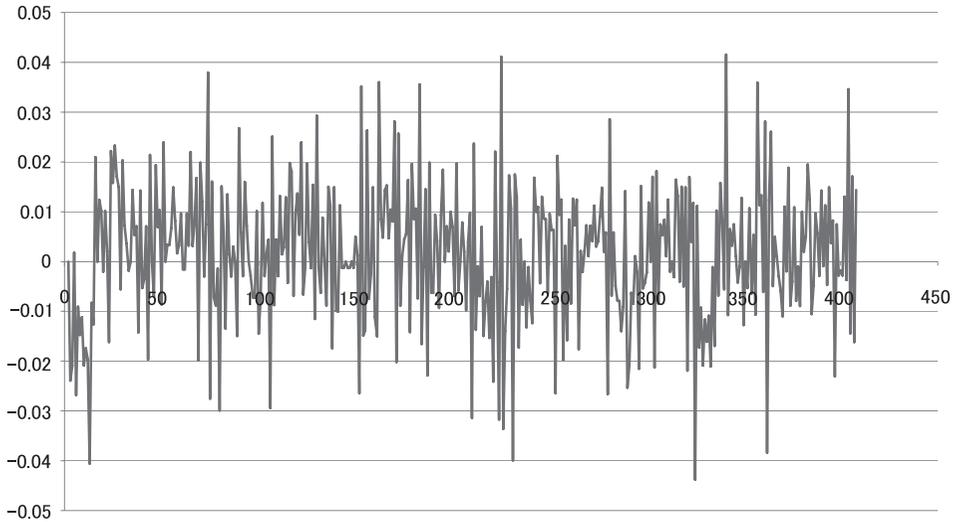


図2 鋁工業生産指数（自然対数値の1階階差）

表1 ADF検定：(16)式の推定

変数	$u(-1)$	$\Delta u(-1)$	$\Delta u(-2)$	$\Delta u(-3)$	$\Delta u(-4)$
係数	-0.004	-0.31	0.17	0.33	0.21
t値	-1.43	-6.3	3.4	6.9	4.3

の上下を行き来しているように見えている，すなわち定数項の存在が見取れるようであれば，レベルの系列には線形トレンドが含まれている可能性がある。実際に図2を見ると，系列の振れの中心は，ほぼ横軸と見てよいように思われる。よって，系列のDGPには線形トレンドは含まれておらず，定数項のみが含まれていると想定することにする。

検定の対象となる系列のDGPに，定数項のみが含まれると想定されることになったので，小稿第II章(6)の左側の式の β_0 をOLS推定し，OLS残差を求め，さらにそれを用いて小稿第II章(16)を推定する。(16)の推定にあたり，右辺のラグ項の p を選択する必要があるが，小稿ではその選択に，第IV章で取り上げたBICを用いる方法を使うことにする。推定に用いた統計ソフトはTSPである。

BICを比較しながら(16)を推定した結果，

$p=4$ とした。なお，(16)の最大ラグ次数の項の有意性検定を逐次行っていく p 選択の方法によっても，結果は同じになった。そして推定結果は表1のようになった。ただし表1においては， \hat{u}_{t-1} を「 $u(-1)$ 」， $\Delta\hat{u}_{t-1}$ を「 $\Delta u(-1)$ 」， $\Delta\hat{u}_{t-2}$ を「 $\Delta u(-2)$ 」，などと記している。小稿第II章で解説したように，変数 \hat{u}_{t-1} の係数のt統計量が，この場合のADF検定の検定統計量 \hat{t}_u であり，その値は-1.43であった。有意水準を5%とすると臨界値はFuller [1996]の表から-2.86と求まるので，帰無仮説は棄却されず，対数をとった鋁工業生産指数の系列が単位根を含んでいるという仮説を棄却できない。同表によれば，有意水準を10%としても臨界値は-2.57であり，やはり帰無仮説は棄却されない。

しかし，小稿第I章で述べたように，ADF検定は検出力が低いことが知られている。仮

表2 ADF-GLS検定：(24)式の推定

変数	$y(-1)$	$\Delta y(-1)$	$\Delta y(-2)$	$\Delta y(-3)$	$\Delta y(-4)$
係数	-0.00009	-0.32	0.2	0.3	0.2
t値	-1.8	-6.5	3.1	6.5	4

に、実は対立仮説が正しい、すなわち系列が単位根を含んでいなかったのだとしても、この検定がそれを「検出」できなかつただけかもしれないのである。このような背景から、ADF検定より検出力の高いADF-GLS検定を行う動機が生まれる。

そこで対数をとった鉱工業生産指数の系列をADF-GLS検定で分析してみる。上述の分析を援用し、DGPには定数項が含まれているが線形トレンド項は含まれていないと想定する。その場合には、小稿第三章で説明したように、 $\bar{c} = -7$ とした上で $y_t - \bar{\alpha}y_{t-1}$ 、 $1 - \bar{\alpha}$ をそれぞれ1つの変数として（ただし、1期目のデータは、それぞれ y_1 、1とする）(26)を推定し、その推定値を用いて(27)より系列 $\{y_t^d\}$ を求める。この $\{y_t^d\}$ を使い(24)を推定するとき、(24)の a_0 のt統計量がADF-GLS検定の検定統計量となる。

ADF検定の場合と同様、(24)の推定にあたっては、未知である p の選択も同時に行う必要がある。BICを比較しながら(24)を推定した結果、 $p=4$ とした。そして推定結果は表2のようになった。ただし表2においては、 y_{t-1}^d を「 $y(-1)$ 」、 Δy_{t-1}^d を「 $\Delta y(-1)$ 」、 Δy_{t-2}^d を「 $\Delta y(-2)$ 」、などと記している。変数 y_{t-1}^d の係数のt統計量が、この場合のADF検定の検定統計量であり、その値は-1.8であった。有意水準を5%とすると臨界値はERSの表から-1.95と求まるので、ADF検定の場合と同様に、帰無仮説は棄却されない。しかし、有意水準を10%にすると臨界値は-1.62となるので、その場合は帰無仮説が棄却される。上述の

ように、ADF検定においては有意水準を10%にしても帰無仮説は棄却されない。したがって、ADF検定より検出力の高いADF-GLS検定を行うと、有意水準によってはADF検定と異なる結果が得られるわけで、ADF検定の結果には疑問符がつくこととなった。

参考文献

- [1] Dickey, David A., and Wayne A. Fuller [1979] "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of American Statistical Association*, Vol. 74, pp. 427-431.
- [2] Elliott, Graham, Thomas J. Rothenberg and James H. Stock [1996] "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, Vol. 64, pp. 813-836.
- [3] Fuller, Wayne A. [1976] *Introduction to Statistical Time Series*, Wiley.
- [4] Fuller, Wayne A. [1996] *Introduction to Statistical Time Series* (2nd ed.), Wiley.
- [5] Hall, Alastair [1994] "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection," *Journal of Business and Economic Statistics*, Vol. 12, pp. 461-470.
- [6] Hamilton, James D. [1994] *Time Series Analysis*, Princeton University Press.
- [7] Hayashi, Fumio [2000] *Econometrics*, Princeton University Press.
- [8] 黒住英司 [2008] 「経済時系列分析と単位根検定：これまでの発展と今後の展望」『日本統計学会誌』第38巻、シリーズJ、第1号、39-57ページ。
- [9] 森棟公夫 [1999] 『計量経済学』東洋経済新報社。
- [10] Perron, Pierre [1989] "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, Vol. 57, pp. 1361-1401.