

第2章の解答

[1] (a) ε_t はホワイトノイズであるから、 w_t の期待値は

$$E[w_t] = \frac{1}{4}(E[\varepsilon_t] + E[\varepsilon_{t-1}] + E[\varepsilon_{t-2}] + E[\varepsilon_{t-3}]) = 0$$

となる。また、 $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$ を所与としたとき、 w_t の条件付き期待値は

$$\begin{aligned} E_{t-2}w_t &= E[w_t | \varepsilon_{t-3} = \varepsilon_{t-2} = 1] \\ &= \frac{1}{4}(E_{t-2}\varepsilon_t + E_{t-2}\varepsilon_{t-1} + 1 + 1) = \frac{1}{2} \end{aligned}$$

(b) 分散の性質から、 $\text{var}(w_t) = E[w_t^2] - E[w_t]^2$ となる。したがって、 w_t の分散は

$$\begin{aligned} \text{var}(w_t) &= \frac{1}{16}E[(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})^2] - 0 \\ &= \frac{1}{16}(E[\varepsilon_t^2] + E[\varepsilon_{t-1}^2] + E[\varepsilon_{t-2}^2] + E[\varepsilon_{t-3}^2]) = \frac{4\sigma^2}{16} = \frac{\sigma^2}{4} \end{aligned}$$

となる ($E[w_t] = 0$ 、 $i \neq j$ のとき $E[\varepsilon_i \varepsilon_j] = 0$ に注意)。

同様に、 $\varepsilon_{t-3} = \varepsilon_{t-2} = 1$ を所与としたとき、 w_t の条件付き分散は

$$\text{var}(w_t | \varepsilon_{t-3} = \varepsilon_{t-2} = 1) = E_{t-2}w_t^2 - (E_{t-2}w_t)^2$$

となる。これは ε_t がホワイトノイズであることに注意すると

$$\begin{aligned} &\frac{1}{16}E_{t-2}(\varepsilon_t^2 + \varepsilon_{t-1}^2 + \varepsilon_{t-2}^2 + \varepsilon_{t-3}^2 + 2\varepsilon_t\varepsilon_{t-1} + 2\varepsilon_t\varepsilon_{t-2} + 2\varepsilon_t\varepsilon_{t-3} + 2\varepsilon_{t-1}\varepsilon_{t-2} + 2\varepsilon_{t-1}\varepsilon_{t-3} \\ &\quad + 2\varepsilon_{t-2}\varepsilon_{t-3}) - \left(\frac{1}{2}\right)^2 \end{aligned}$$

と表せる ((a)から $E_{t-2}w_t = 1/2$ に注意)。したがって、

$$\begin{aligned} &\frac{1}{16}E_{t-2}(\varepsilon_t^2 + \varepsilon_{t-1}^2 + 1 + 1 + 2\varepsilon_t\varepsilon_{t-1} + 2\varepsilon_t + 2\varepsilon_t + 2\varepsilon_{t-1} + 2\varepsilon_{t-1} + 2) - \frac{1}{4} \\ &= \frac{1}{16}(\sigma^2 + \sigma^2 + 1 + 1 + 2) - \frac{1}{4} = \frac{\sigma^2}{8} \end{aligned}$$

(c) ① 共分散の性質から

$$\text{cov}(w_t, w_{t-1}) = E[w_t w_{t-1}] - E[w_t]E[w_{t-1}]$$

となる。ここで $E[w_t] = E[w_{t-1}] = 0$ であるから、

$$\begin{aligned} \text{cov}(w_t, w_{t-1}) &= \frac{1}{16}E[(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4})] \\ &= \frac{1}{16}(E[\varepsilon_{t-1}^2] + E[\varepsilon_{t-2}^2] + E[\varepsilon_{t-3}^2]) = \frac{3\sigma^2}{16} \end{aligned}$$

となる。式展開では、交差項の期待値はすべて0 ($i \neq j$ のとき $E[\varepsilon_i \varepsilon_j] = 0$) となることを用

いた。

② 分散の性質から

$$\text{cov}(w_t, w_{t-2}) = E[w_t w_{t-2}] - E[w_t]E[w_{t-2}]$$

となる。ここで $E[w_t] = E[w_{t-2}] = 0$ であるから、

$$\begin{aligned}\text{cov}(w_t, w_{t-2}) &= \frac{1}{16} E[(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4} + \varepsilon_{t-5})] \\ &= \frac{1}{16} (E[\varepsilon_{t-2}^2] + E[\varepsilon_{t-3}^2]) = \frac{2\sigma^2}{16} = \frac{\sigma^2}{8}\end{aligned}$$

となる。式展開では、交差項の期待値は 0 となることを用いた。

③ 分散の性質から

$$\text{cov}(w_t, w_{t-5}) = E[w_t w_{t-5}] - E[w_t]E[w_{t-5}]$$

となる。ここで $E[w_t] = E[w_{t-5}] = 0$ であるから、

$$\text{cov}(w_t, w_{t-5}) = \frac{1}{16} E[(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3})(\varepsilon_{t-5} + \varepsilon_{t-6} + \varepsilon_{t-7} + \varepsilon_{t-8})] = 0$$

となる。ここで交差項しか存在しないため、期待値はすべて 0 となっている。

[2] 単純化のため、 $a_0 = 0$ として証明しよう。このとき、(2.20)式から解は

$$y_t = \sum_{i=0}^{\infty} c_i \varepsilon_{t-i}$$

となる。これを差分方程式(2.19)の y_{t-i} に代入すると

$$\begin{aligned}c_0 \varepsilon_t + c_1 \varepsilon_{t-1} + c_2 \varepsilon_{t-2} + c_3 \varepsilon_{t-3} + \dots &= a_1 [c_0 \varepsilon_{t-1} + c_1 \varepsilon_{t-2} + c_2 \varepsilon_{t-3} + \dots] \\ &\quad + a_2 [c_0 \varepsilon_{t-2} + c_1 \varepsilon_{t-3} + c_2 \varepsilon_{t-4} + \dots] \\ &\quad + a_3 [c_0 \varepsilon_{t-3} + c_1 \varepsilon_{t-4} + c_2 \varepsilon_{t-5} + \dots] + \dots \\ &\quad + a_p [c_0 \varepsilon_{t-p} + c_1 \varepsilon_{t-p-1} + c_2 \varepsilon_{t-p-2} + \dots] + \varepsilon_t\end{aligned}$$

となる。上式の右辺の全項を左辺に移し、類似項同士をまとめると、

$$\begin{aligned}(c_0 - 1)\varepsilon_t + (c_1 - a_1 c_0)\varepsilon_{t-1} + (c_2 - a_1 c_1 - a_2 c_0)\varepsilon_{t-2} + (c_3 - a_1 c_2 - a_2 c_1 - a_3 c_0)\varepsilon_{t-3} \\ + \dots + (c_p - a_1 c_{p-1} - a_2 c_{p-2} - \dots - a_p c_0)\varepsilon_{t-p} \\ + (c_{p+1} - a_1 c_p - a_2 c_{p-1} - \dots - a_p c_1)\varepsilon_{t-p+1} + \dots = 0\end{aligned}$$

となる。

ここで $\{\varepsilon_t\}$ のとりうる全ての値に対し、上式が等式で成立するには、各項の係数が全て 0 であればよい。ここで ε_{t-p} の係数が 0 なら

$$c_p - a_1 c_{p-1} - a_2 c_{p-2} - \dots - a_p c_0 = 0$$

となり、これを c_p について解くと、

$$c_p = a_1 c_{p-1} + a_2 c_{p-2} + \dots + a_p c_0$$

である。同様に、 ε_{t-p+1} の係数が0なら

$$c_{p+1} - a_1 c_p - a_2 c_{p-1} - \dots - a_p c_1 = 0$$

となり、これを c_{p+1} について解くと、

$$c_{p+1} = a_1 c_p + a_2 c_{p-1} + \dots + a_p c_1$$

となる。ここで係数 c_i は、 p 次の差分方程式に従っている。

[3] (a) ① $t-2$ 期までの情報には y_{t-2} が含まれており、 $E_{t-2}y_{t-2} = y_{t-2}$ となる。このため、 $t-2$ 期までの情報を所与としたとき、 y_t の期待値は

$$E_{t-2}y_t = E_{t-2}(a_0 + a_2 y_{t-2} + \varepsilon_t) = a_0 + a_2 y_{t-2}$$

② $t-1$ 期までの情報には y_{t-2} が含まれており、 $E_{t-1}y_{t-2} = y_{t-2}$ となる。このため、 $t-1$ 期までの情報を所与としたとき、 y_t の期待値は

$$E_{t-1}y_t = E_{t-1}(a_0 + a_2 y_{t-2} + \varepsilon_t) = a_0 + a_2 y_{t-2}$$

③ t 期までの情報には y_t が含まれており、 $E_t y_t = y_t$ となる。このため、 t 期までの情報を所与としたとき、 y_{t+2} の期待値は

$$E_t y_{t+2} = E_t(a_0 + a_2 y_t + \varepsilon_{t+2}) = a_0 + a_2 y_t$$

④ y_t の特殊解は

$$y_t = \frac{a_0}{1 - a_2} + \varepsilon_t + a_2 \varepsilon_{t-2} + a_2^2 \varepsilon_{t-4} + a_2^3 \varepsilon_{t-6} + a_2^4 \varepsilon_{t-8} + \dots$$

となる（これが正しいことを、繰り返し代入法を用いて自分で導出しよう）。このため、

$$E[y_t] = \frac{a_0}{1 - a_2}$$

となる。また、共分散は

$$\begin{aligned} cov(y_t, y_{t-1}) &= E[(y_t - E[y_t])(y_{t-1} - E[y_{t-1}])] \\ &= E[(\varepsilon_t + a_2 \varepsilon_{t-2} + a_2^2 \varepsilon_{t-4} + a_2^3 \varepsilon_{t-6} + \dots)(\varepsilon_{t-1} + a_2 \varepsilon_{t-3} + a_2^2 \varepsilon_{t-5} + a_2^3 \varepsilon_{t-7} + \dots)] = 0 \end{aligned}$$

となる。ここで共分散は全て交差項として表せるため、その期待値は0となる。

⑤ ④と同様に、

$$\begin{aligned} cov(y_t, y_{t-2}) &= E[(y_t - E[y_t])(y_{t-2} - E[y_{t-2}])] \\ &= E[(\varepsilon_t + a_2 \varepsilon_{t-2} + a_2^2 \varepsilon_{t-4} + a_2^3 \varepsilon_{t-6} + \dots)(\varepsilon_{t-2} + a_2 \varepsilon_{t-4} + a_2^2 \varepsilon_{t-6} + a_2^3 \varepsilon_{t-8} + \dots)] \\ &= E[a_2 \varepsilon_{t-2}^2 + a_2^3 \varepsilon_{t-4}^2 + a_2^5 \varepsilon_{t-6}^2 + a_2^7 \varepsilon_{t-8}^2 + \dots] \\ &= a_2 E[\varepsilon_{t-2}^2 + a_2^2 \varepsilon_{t-4}^2 + a_2^4 \varepsilon_{t-6}^2 + a_2^6 \varepsilon_{t-8}^2 + \dots] \\ &= a_2 \sigma^2 (1 + a_2^2 + a_2^4 + a_2^6 + \dots) = \frac{a_2 \sigma^2}{1 - a_2^2} \end{aligned}$$

コラム：インパルス応答関数

1章5節の例8では、GDPの例を用いて、インパルス応答関数とは何かを説明した。繰り返しになるが、「インパルス応答関数とは、1単位のショックが生じたとき、それが*i*期先の*y*にどのような影響を与えるか」を記述したものである。ここでは、AR(1)過程

$$y_t = a_1 y_{t-1} + \varepsilon_t$$

のインパルス応答関数を求めてみよう。

初期条件を $y_0 = 0$ とし、1期に1単位のショックが生じたとする ($\varepsilon_1 = 1$)。初期のショックの影響を見たいため、それ以外のショックはすべて0とする ($\varepsilon_2 = \varepsilon_3 = \dots = 0$)。まず、 $y_0 = 0, \varepsilon_1 = 1$ から、 $y_1 = a_1 \times 0 + 1 = 1$ となる。次に、 $y_1 = 1, \varepsilon_2 = 0$ から、 $y_2 = a_1 \times 1 + 0 = a_1$ となる。同様に、 $y_2 = a_1, \varepsilon_3 = 0$ から、 $y_3 = a_1 \times a_1 + 0 = a_1^2$ と計算できる。一般的には、

$$y_i = a_1^i$$

となる。したがって、 $|a_1| < 1$ であれば、*i*が大きくなるにつれて、ショックの影響は小さくなっていく。

インパルス応答関数は別方法でも求められる。まず、AR(1)をMA(∞)と表現する。

$$y_t = \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + a_1^3 \varepsilon_{t-3} + \dots + a_1^i \varepsilon_{t-i} + \dots$$

これは $t = 1$ とすると、 $y_1 = \varepsilon_1 + a_1 \varepsilon_0 + a_1^2 \varepsilon_{-1} + a_1^3 \varepsilon_{-2} + \dots$ を意味する。ショック ε_1 の y_1 への同時点の効果は、 y_1 を ε_1 で偏微分したものであり

$$\frac{\partial y_1}{\partial \varepsilon_1} = 1$$

となる。これは ε_1 が1単位増価すると、 y_1 は1単位上昇することを意味する。次に、 $t = 2$ とすると $y_2 = \varepsilon_2 + a_1 \varepsilon_1 + a_1^2 \varepsilon_0 + a_1^3 \varepsilon_{-1} + \dots$ となり、 y_2 を ε_1 で偏微分すると、

$$\frac{\partial y_2}{\partial \varepsilon_1} = a_1$$

となる。こうした計算を繰り返すと、

$$\frac{\partial y_i}{\partial \varepsilon_1} = a_1^i$$

が得られる。この関数は、次の関係から簡単に導出できる。

$$\frac{\partial y_i}{\partial \varepsilon_1} = \frac{\partial y_t}{\partial \varepsilon_{t-i}}$$

つまり、インパルス応答関数は、MA(∞)で表した

$$y_t = \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + a_1^3 \varepsilon_{t-3} + \dots + a_1^i \varepsilon_{t-i} + \dots$$

を様々な ε_{t-i} で偏微分したものである。これは単に ε_{t-i} の係数 a_1^i に過ぎないのである。

(b) 既に述べたとおり、 y_t の特殊解は

$$y_t = \frac{a_0}{1-a_2} + \varepsilon_t + a_2\varepsilon_{t-2} + a_2^2\varepsilon_{t-4} + a_2^3\varepsilon_{t-6} + a_2^4\varepsilon_{t-8} + \dots$$

となる。したがって、インパルス応答関数は、

$$\begin{aligned} \frac{\partial y_t}{\partial \varepsilon_t} &= 1, & \frac{\partial y_{t+1}}{\partial \varepsilon_t} &= \frac{\partial y_t}{\partial \varepsilon_{t-1}} = 0, & \frac{\partial y_{t+2}}{\partial \varepsilon_t} &= \frac{\partial y_t}{\partial \varepsilon_{t-2}} = a_2, \\ \frac{\partial y_{t+3}}{\partial \varepsilon_t} &= \frac{\partial y_t}{\partial \varepsilon_{t-3}} = 0, & \frac{\partial y_{t+4}}{\partial \varepsilon_t} &= \frac{\partial y_t}{\partial \varepsilon_{t-4}} = a_2^2, & \dots \end{aligned}$$

となる。詳しくは、コラム：インパルス応答関数を参照されたい。

[4] (a) この差分方程式の同次部分は

$$y_t = 0.75y_{t-1} - 0.125y_{t-2}$$

となる。右辺の全項を左辺に移動させ、試行解 $y_t = A\alpha^t$ を代入すると

$$A\alpha^t - 0.75A\alpha^{t-1} + 0.125A\alpha^{t-2} = 0$$

となる。ここで両辺を $A\alpha^{t-2}$ で割ると、

$$\alpha^2 - 0.75\alpha + 0.125 = 0$$

となる。この式を満たす特性根 α は0.5、0.25となる。同次解は

$$y_t^h = A_1(0.5)^t + A_2(0.25)^t$$

と表せる。ただし、 A_1 、 A_2 は任意の定数である。

次に、未定係数法によって特殊解を求めよう。試行解は $y_t = b_0 + \sum c_i \varepsilon_{t-i}$ とする。ここで差分方程式

$$y_t = a_0 + 0.75y_{t-1} - 0.125y_{t-2} + \varepsilon_t$$

に、試行解を代入すると

$$\begin{aligned} b_0 + c_0\varepsilon_t + c_1\varepsilon_{t-1} + c_2\varepsilon_{t-2} + \dots &= a_0 + 0.75(b_0 + c_0\varepsilon_{t-1} + c_1\varepsilon_{t-2} + c_2\varepsilon_{t-3} + \dots) \\ &\quad - 0.125(b_0 + c_0\varepsilon_{t-2} + c_1\varepsilon_{t-3} + c_2\varepsilon_{t-4} + \dots) + \varepsilon_t \end{aligned}$$

となる。試行解が特殊解であるためには、上式は等式で成立しなければならない。上式の右辺の全項を左辺に移し、類似項同士をまとめると、

$$\begin{aligned} (b_0 - a_0 - 0.75b_0 + 0.125b_0) + (c_0 - 1)\varepsilon_t + (c_1 - 0.75c_0)\varepsilon_{t-1} \\ + (c_2 - 0.75c_1 + 0.125c_0)\varepsilon_{t-2} + (c_3 - 0.75c_2 + 0.125c_1)\varepsilon_{t-3} + \dots = 0 \end{aligned}$$

を得る。定数 b_0 は、 $b_0 - a_0 - 0.75b_0 + 0.125b_0 = 0$ を解くことで

$$b_0 = \frac{a_0}{1-0.75+0.125} = \frac{a_0}{0.375}$$

となる。同様に、 $c_0 - 1 = 0$ から $c_0 = 1$ 、 $c_1 - 0.75c_0 = 0$ から $c_1 = 0.75$ となる。また、 $i > 1$ の範囲では

$$c_i = 0.75c_{i-1} - 0.125c_{i-2}$$

となる（例えば、 $c_2 = 0.4375$ 、 $c_3 = 0.2344$ 、 $c_4 = 0.1211$ となる）。

(b) インパルス応答関数は、特殊解の係数によって与えられる（コラム参照）。

$$\partial y_t / \partial \varepsilon_t = 1, \quad \partial y_{t+1} / \partial \varepsilon_t = \partial y_t / \partial \varepsilon_{t-1} = 0.75, \quad \partial y_{t+2} / \partial \varepsilon_t = \partial y_t / \partial \varepsilon_{t-2} = 0.4375$$

特性根は両方とも正かつ1より小さいため、インパルス応答関数は単調に減衰していく。

(c) 一般解は同次解と特殊解の和である。したがって、一般解は

$$y_t = A_1(0.5)^t + A_2(0.25)^t + \frac{a_0}{0.375} + \varepsilon_t + 0.75\varepsilon_{t-1} + 0.4344\varepsilon_{t-2} + \dots$$

となる。このため、0期と1期において、

$$y_0 = A_1 + A_2 + \frac{a_0}{0.375} + \varepsilon_0 + 0.75\varepsilon_{-1} + 0.4344\varepsilon_{-2} + \dots$$

$$y_1 = A_1(0.5) + A_2(0.25) + \frac{a_0}{0.375} + \varepsilon_1 + 0.75\varepsilon_0 + 0.4344\varepsilon_{-1} + \dots$$

となる。ここで、同次部分が0となるように（つまり、 $A_1 = A_2 = 0$ ）、初期条件 y_0 、 y_1 を選択すればよい。つまり、初期条件を

$$y_0 = \frac{a_0}{0.375} + \varepsilon_0 + 0.75\varepsilon_{-1} + 0.4344\varepsilon_{-2} + \dots$$

$$y_1 = \frac{a_0}{0.375} + \varepsilon_1 + 0.75\varepsilon_0 + 0.4344\varepsilon_{-1} + \dots$$

とすれば、 $A_1 = A_2 = 0$ となる。したがって、一般解は

$$y_t = \frac{a_0}{0.375} + \varepsilon_t + 0.75\varepsilon_{t-1} + 0.4344\varepsilon_{t-2} + \dots$$

(d) ユール＝ウォーカー方程式を用いて、コレログラムを求められる。第2章5節のAR(2)過程の自己相関関数の議論を思い出してほしい。定数項 a_0 は、コレログラムに影響を与えないため、 $a_0 = 0$ とする。ここで、 $a_1 = 0.75$ 、 $a_2 = -0.125$ である。まず、定義から $\rho_0 = 1$ となる。また、

$$\rho_1 = \frac{a_1}{1 - a_2} = \frac{0.75}{1 + 0.125} = 0.67$$

となる。そして、(2.28)式 $\rho_s = 0.75\rho_{s-1} - 0.125\rho_{s-2}$ を用いて、 $\rho_2 = 0.375$ 、 $\rho_3 = 0.1979$ 、 $\rho_4 = 0.1055$ 、 $\rho_5 = 0.0514$ 、...を得る。

[5] ここでは(b)を最初に証明し、(a)はその特殊ケースとして証明する。

(b) 証明の仕方は2通りある。まず、未決定係数法を用いて証明する。ここで y_t は

$$y_t = a_0^* + a_2^*t + e_t$$

$$e_t = a_1 e_{t-1} + \varepsilon_t$$

で表せる仮定し、その係数を特定しよう。まず、 y_t の式の両辺に $1 - a_1L$ を掛けると

$$(1 - a_1L)y_t = (1 - a_1L)a_0^* + (1 - a_1L)a_2^*t + (1 - a_1L)e_t$$

となる。ここで $(1 - a_1L)e_t = \varepsilon_t$ から

$$y_t - a_1 y_{t-1} = (1 - a_1)a_0^* + a_2^*[t - a_1(t - 1)] + \varepsilon_t$$

$$= [(1 - a_1)a_0^* + a_1 a_2^*] + a_2^*(1 - a_1)t + \varepsilon_t$$

となる。ここで $a_0 = (1 - a_1)a_0^* + a_1 a_2^*$ 、 $a_2 = a_2^*(1 - a_1)$ という関係が成立すると仮定すると、上式は

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \varepsilon_t$$

と表現できる。 $a_2 = a_2^*(1 - a_1)$ を a_2^* について解くと、

$$a_2^* = \frac{a_2}{1 - a_1}$$

となる。この式を $a_0 = (1 - a_1)a_0^* + a_1 a_2^*$ に代入して a_0^* について解くと

$$a_0^* = \frac{a_0}{1 - a_1} - \frac{a_1 a_2}{(1 - a_1)^2}$$

が得られる。したがって、 $y_t = a_0 + a_1 y_{t-1} + a_2 t + \varepsilon_t$ は、

$$y_t = a_0^* + a_2^* t + e_t$$

$$e_t = a_1 e_{t-1} + \varepsilon_t$$

と同じ確率過程である。 e_t は定常であるから、 y_t はトレンド $a_0^* + a_2^* t$ の周りで定常となる。

ラグオペレータを使っても証明できる。AR(1)過程は

$$(1 - a_1 L)y_t = a_0 + a_2 t + \varepsilon_t$$

であり、さらに両辺に $(1 - a_1 L)^{-1}$ を掛けると、

$$y_t = \frac{a_0}{1 - a_1} + \frac{a_2}{1 - a_1 L} t + \frac{\varepsilon_t}{1 - a_1 L}$$

となる。ここで $(1 - a_1 L)e_t = \varepsilon_t$ から、 $e_t = \varepsilon_t / (1 - a_1 L)$ となる。また、

$$\begin{aligned} \frac{a_2}{1 - a_1 L} t &= a_2(1 + a_1 L + a_1^2 L^2 + a_1^3 L^3 + \dots)t \\ &= a_2(t + a_1(t - 1) + a_1^2(t - 2) + a_1^3(t - 3) + \dots) \\ &= a_2(1 + a_1 + a_1^2 + a_1^3 + \dots)t - a_1 a_2(1 + 2a_1 + 3a_1^2 + \dots) \\ &= \frac{a_2}{1 - a_1} t - \frac{a_1 a_2}{(1 - a_1)^2} \end{aligned}$$

となる¹。したがって、

$$y_t = \left(\frac{a_0}{1 - a_1} - \frac{a_1 a_2}{(1 - a_1)^2} \right) + \frac{a_2}{1 - a_1} t + e_t$$

(a) これは(b)の特殊ケースである。 $a_2 = 0$ とすると、 $a_0^* = a_0 / (1 - a_1)$ 、 $a_2^* = 0$ となる。

(c) ここで $y_t = 1 + 0.95y_{t-1} + 0.01t + \varepsilon_t$ としよう。このとき、 $a_0 = 1$ 、 $a_1 = 0.95$ 、 $a_2 = 0.01$

¹ $X = 1 + 2a_1 + 3a_1^2 + \dots$ としよう。ここで、両辺に a_1 を掛けると、 $a_1 X = a_1 + 2a_1^2 + 3a_1^3 + \dots$ となる。したがって、 $X - a_1 X = 1 + a_1 + a_1^2 + a_1^3 + \dots$ となる。また、 $1/(1 - a_1) = 1 + a_1 + a_1^2 + a_1^3 + \dots$ から、 $X = 1/(1 - a_1)^2$ となる。

から、

$$a_0^* = \frac{1}{1-0.95} - \frac{0.95 \times 0.01}{(1-0.95)^2} = 16.2, \quad a_2^* = \frac{0.01}{1-0.95} = 0.2$$

となる。したがって、これは

$$y_t = 16.2 + 0.02t + e_t, \quad e_t = 0.95e_{t-1} + \varepsilon_t$$

として表せる。

[6] 情報量規準を単調変換（数値の順序を保持する変換）しても、選ばれる次数に変化はない。たとえば、AIC に定数を足したり引いたり、正の定数を掛けたり、割ったりしても、選ばれる次数は変わらない（負の定数を掛けると、大小関係が逆転するので、単調変換ではないことに注意してほしい）。これは情報量規準の表現方法が多数存在する理由となる。

以下では、(b)を証明してから、(a)を証明していく。

(b) まず、AIC' と AIC が同じラグ次数を選択することを示す。AIC' は

$$\ln(SSR/T) + 2n/T = \ln(SSR) + 2n/T - \ln(T)$$

である。ここで T を掛けると

$$T\ln(SSR/T) + 2n - T\ln(T)$$

となる。ここで $T\ln(T)$ は正の定数であるため、情報量規準の式から $T\ln(T)$ を除いても次数の選択に影響を与えない。これはまさに AIC に他ならない。

$$T\ln(SSR/T) + 2n$$

したがって、AIC' と AIC は同じ次数を選択することがわかった。

次に、BIC' と BIC が同じラグ次数を選択することを示す。 $\ln(SSR/T) + n\ln(T)/T$ に T を掛けると、 $T\ln(SSR) + n\ln(T) - T\ln(T)$ となる。ここで $T\ln(T)$ は定数であるため除いても次数の選択に影響を与えない。したがって、BIC' と BIC は同じ次数を選択する。

(a) ここで $AIC^* = -2\ln(L)/T + 2n/T$ となる（ただし、 $\ln(L)$ は対数尤度の最大値）。第2章の補足 2.1 から、対数尤度は

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^T \varepsilon_t^2$$

となる。たとえば、AR(p)過程

$$y_t = a_0 + a_1y_{t-1} + a_2y_{t-2} + \cdots + a_p y_{t-p} + \varepsilon_t$$

であれば、

$$\varepsilon_t = y_t - a_0 - a_1y_{t-1} - a_2y_{t-2} - \cdots - a_p y_{t-p}$$

となる。最尤法では、この ε_t を、対数尤度の式に代入し、対数尤度を最大にするパラメータの組み合わせを選ぶ。したがって、対数尤度の最大値は、パラメータに最尤推定量を代入した

$$\ln L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2}\sum_{t=1}^T \varepsilon_t^2$$

となる。ここで、 $\hat{\varepsilon}_t$ は残差であり、 $\hat{\sigma}^2 = \sum \hat{\varepsilon}_t^2 / T$ となる。したがって、右辺第3項は

$$\frac{1}{2\hat{\sigma}^2} \sum_{t=1}^T \hat{\varepsilon}_t^2 = \frac{1}{2(\sum_{t=1}^T \hat{\varepsilon}_t^2 / T)} \sum_{t=1}^T \hat{\varepsilon}_t^2 = \frac{T}{2}$$

となり、定数となる。また、 $-\frac{T}{2} \ln(2\pi)$ も定数である。対数尤度の最大値から、定数を除いて考えると、第2項だけが残る。

$$-\frac{T}{2} \ln \hat{\sigma}^2 = -\frac{T}{2} \ln \left(\frac{\sum_{t=1}^T \hat{\varepsilon}_t^2}{T} \right)$$

ゆえに、 $AIC^* = -2\ln(L)/T + 2n/T$ は、

$$AIC^* = \ln \left(\frac{\sum_{t=1}^T \hat{\varepsilon}_t^2}{T} \right) + \frac{2n}{T} = \ln \left(\sum_{t=1}^T \hat{\varepsilon}_t^2 \right) - \ln(T) + \frac{2n}{T}$$

と書き換えられる。ここで右辺第2項が定数であるため、

$$\ln \left(\sum_{t=1}^T \hat{\varepsilon}_t^2 \right) + \frac{2n}{T}$$

を考えても同じ次数が選択される。ここで定数 T を掛けると、

$$T \ln(SSR) + 2n$$

となる（ただし、残差平方和 $\sum_{t=1}^T \hat{\varepsilon}_t^2$ を SSR と表記している。同様にして、 BIC^* も証明できる。

[7] ここでは、 $AIC' = \ln(SSR/T) + 2n/T$ 、 $BIC' = \ln(SSR/T) + n \ln(T)/T$ を用いて、ラグ次数の選択を行う。

(a) T が非常に大きいと、 BIC' により選択される次数は1となることを示す。このためには、 T が非常に大きいと、① $P\{p^{BIC}=0\}=0$ 、② $P\{p^{BIC}=2\}=0$ が成立することを示せばよい（確率の和は1から、 $P\{p^{BIC}=1\}=1$ が成立する）。

① $P\{p^{BIC}=0\}=0$

$AR(p)$ における BIC' の値を $BIC'(p)$ 、 SSR の値を $SSR(p)$ と表記する。このとき、 BIC' が0次のラグを選ぶためには、 $BIC'(0) < BIC'(1)$ が成立しなければならない。つまり、

$$[\ln(SSR(0)/T) + (1+0)\ln(T)/T] - [\ln(SSR(1)/T) + (1+1)\ln(T)/T] < 0$$

となる。ここで $BIC'(p) = \ln(SSR(p)/T) + (1+p)\ln(T)/T$ に注意してほしい。上式を書き換えると、

$$\ln(SSR(0)/T) - \ln(SSR(1)/T) - \ln(T)/T < 0$$

となる。ここで、 $SSR(0)$ とは、 y_t の偏差平方和である($SSR(0) = \sum (y_t - \bar{y})^2$)。このため、 T が大きくなると、 $SSR(0)/T$ は y_t の分散 σ_y^2 に収束していく。また、

$$SSR(1) = \sum_{t=1}^T (y_t - \hat{a}_0 - \hat{a}_1 y_{t-1})^2$$

であるため、 T が大きくなると(真のモデルは $AR(1)$ である)、 $SSR(1)/T$ は ε_t の分散 σ^2 に収束していく。最後に、 T が大きくなると $\ln(T)/T$ は0に収束する。以上から、大標本では

$$\ln(\sigma_y^2) - \ln(\sigma^2) < 0$$

となる。しかし、 $\sigma_y^2 > \sigma^2$ であることから、この式は成立しない。つまり、大標本では $P\{p^{BIC}=0\} = P\{\ln(\sigma_y^2) - \ln(\sigma^2)\} = 0$ となる。

② $P\{p^{BIC}=2\}=0$

BIC' が 2 次のラグを選ぶためには、 $BIC'(2) < BIC'(1)$ が成立しなければならない。

$$[\ln(SSR(2)/T) + (1+2)\ln(T)/T] - [\ln(SSR(1)/T) + (1+1)\ln(T)/T] < 0$$

上式を整理すると、

$$-\ln\left(\frac{SSR(1)}{SSR(2)}\right) + \frac{\ln(T)}{T} < 0$$

となる。さらに両辺に T を掛けると、

$$-T \ln\left(\frac{SSR(1)}{SSR(2)}\right) + \ln(T) < 0$$

となり、これを書き換えると

$$-T \ln\left(1 + \frac{SSR(1) - SSR(2)}{SSR(2)}\right) + \ln(T) < 0$$

となり、さらにこれは

$$-T \ln\left(1 + \frac{1}{T-2} \frac{SSR(1) - SSR(2)}{SSR(2)/(T-2)}\right) + \ln(T) < 0$$

と書ける。ここで帰無仮説 $a_2 = 0$ (y_{t-2} の係数が 0) とした F 統計量は、

$$F = \frac{SSR(1) - SSR(2)}{SSR(2)/(T-2)}$$

であるから、先の式は

$$-T \ln\left(1 + \frac{F}{T-2}\right) + \ln(T) < 0$$

となる。ここで T が十分に大きいと、 $F/(T-2)$ は非常に小さい値となるため、

$$-T \ln\left(1 + \frac{F}{T-2}\right) + \ln(T) = -T \frac{F}{T-2} + \ln(T) = -F + \ln(T) < 0$$

となる (T が非常に大きければ、 $\ln(1 + F/(T-2)) = F/(T-2)$ 、また $T/(T-2) = 1$ で近似できる)²。ここで T が大きくなると、 $\ln(T)$ も大きくなっていくため、上式は成立しない。

したがって、 $P\{p^{BIC}=2\} = P\{-F + \ln(T)\} = 0$ となる。

(b) ここで① $P\{p^{aic}=0\}=0$ 、② $P\{p^{aic}=2\}=0.16$ を証明し、確率の和は 1 であることから、③ $P\{p^{aic}=1\}=0.84$ を説明する。この結果から、AIC に一致性がないこと、長めの次数を選ぶ傾向があることが確認できる。

① $P\{p^{aic}=0\}=0$

² ε が小さな値であれば、 $\ln(1+\varepsilon) \doteq \varepsilon$ となる。証明は、『入門実践する統計学』の章末の数学付録を参照されたい。

これは(a)の① $P\{p^{BIC}=0\}=0$ において、 $\ln(T)$ を2に置き換えたうえで、同様の式展開をすれば証明できる。

② $P\{p^{aic}=2\}=0.16$

AIC'が2次のラグを選ぶためには、 $AIC'(2) < AIC'(1)$ が成立していなければならない。

$$[\ln(SSR(2)/T) + (1+2)2/T] - [\ln(SSR(1)/T) + (1+1)2/T] < 0$$

ここで、(a)の②と同様の式展開をすると、

$$-F + 2 < 0$$

を得る。つまり、 $P\{p^{aic}=2\} = P\{2 < F\}$ となる。 F は、帰無仮説 $a_2 = 0$ とした F 統計量であり、これは T が大きいとき、自由度1の χ^2 分布に従う³。自由度1の χ^2 確率変数が2より大きくなる確率は、エクセルで「=CHIDIST(2,1)」と代入すれば求められる。これは0.157であり、四捨五入すると0.16と確認できる。

③ $P\{p^{aic}=1\}=0.84$

$P\{p^{aic}=0\}=0$ 、 $P\{p^{aic}=2\}=0.16$ であり、確率の和は1から $P\{p^{aic}=1\}=0.84$ となる。つまり、大標本では、AIC'が短めの次数を選ぶ確率は0であり、また真の次数を選ぶ確率は84%と高い。しかし、AIC'では、真の次数1より高めの次数を選ぶ確率は14%もある。

[8] ARMA過程の利点は、高次のAR過程であっても、低次のARMA過程によって表せる点にある。たとえば、ARMA(1,1)過程 $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1}$ を考える。このとき、 $y_{t-1} = a_0 + a_1 y_{t-2} + \varepsilon_{t-1} + \beta_1 \varepsilon_{t-2}$ から、 $\varepsilon_{t-1} = y_{t-1} - a_1 y_{t-2} - \beta_1 \varepsilon_{t-2}$ となる。したがって、

$$\begin{aligned} y_t &= a_0 + a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \\ &= a_0 + a_1 y_{t-1} + \varepsilon_t + \beta_1 [y_{t-1} - a_1 y_{t-2} - \beta_1 \varepsilon_{t-2}] \\ &= a_0 + (a_1 + \beta_1) y_{t-1} - a_1 \beta_1 y_{t-2} - \beta_1^2 \varepsilon_{t-2} + \varepsilon_t \end{aligned}$$

となる。さらに、 $\varepsilon_{t-2} = y_{t-2} - a_1 y_{t-3} - \beta_1 \varepsilon_{t-3}$ を代入すると、

$$\begin{aligned} y_t &= a_0 + (a_1 + \beta_1) y_{t-1} - a_1 \beta_1 y_{t-2} - \beta_1^2 \varepsilon_{t-2} + \varepsilon_t \\ &= a_0 + (a_1 + \beta_1) y_{t-1} - a_1 \beta_1 y_{t-2} - \beta_1^2 [y_{t-2} - a_1 y_{t-3} - \beta_1 \varepsilon_{t-3}] + \varepsilon_t \\ &= a_0 + (a_1 + \beta_1) y_{t-1} - \beta_1 (a_1 + \beta_1) y_{t-2} + a_1 \beta_1^2 y_{t-3} + \beta_1^3 \varepsilon_{t-3} + \varepsilon_t \end{aligned}$$

となる。これを繰り返していくと、AR(∞)過程

$$y_t = a_0 + (a_1 + \beta_1) y_{t-1} - \beta_1 (a_1 + \beta_1) y_{t-2} + a_1 \beta_1^2 y_{t-3} + \dots + \varepsilon_t$$

となる（式展開では $|\beta_1| < 1$ とした）。これはARMA(1,1)過程であっても、パラメータ a_1 、

³ F 分布の復習をしよう（『入門実践する統計学』の9章を参照されたい）。 $W_1 \sim \chi^2(n_1)$ 、 $W_2 \sim \chi^2(n_2)$ としたとき、 $[W_1/n_1] / [W_2/n_2]$ は自由度 (n_1, n_2) の F 分布に従う。ここで $n_2 \rightarrow \infty$ なら W_2/n_2 は1に収束する。よって、自由度 $(1, T)$ の F 分布で T が無大になるなら、これは自由度1の χ^2 分布となる。

β_1 の値を変えれば、複雑な高次の AR 過程を近似できる可能性を意味している。

[9] ここで $E[d_t] = 0$ であるから、 $E[d_t d_{t-i}] = \gamma_i$ となる。また $i > q$ の範囲では $\gamma_i = 0$ と仮定する。 $E[d_t] = 0$ から、平均の期待値も 0 となる。

$$E[\bar{d}] = 0$$

このため、平均の分散は

$$\text{var}(\bar{d}) = E[\bar{d}^2] = \frac{1}{H^2} E[(d_1 + d_2 + d_3 + \dots + d_H)^2]$$

となる。上式を展開すると（少し計算は面倒であるが）、

$$\begin{aligned} & \frac{1}{H^2} E \begin{bmatrix} d_1^2 + d_1 d_2 + d_1 d_3 + d_1 d_4 + \dots + d_1 d_H + \\ d_2 d_1 + d_2^2 + d_2 d_3 + d_2 d_4 + \dots + d_2 d_H + \\ d_3 d_1 + d_3 d_2 + d_3^2 + d_3 d_4 + \dots + d_3 d_H + \\ \dots \\ d_H d_1 + d_H d_2 + d_H d_3 + d_H d_4 + \dots + d_H^2 \end{bmatrix} \\ &= \frac{1}{H^2} \begin{bmatrix} \text{var}(d_1) + \text{cov}(d_1, d_2) + \text{cov}(d_1, d_3) + \text{cov}(d_1, d_4) + \dots + \text{cov}(d_1, d_H) + \\ \text{cov}(d_1, d_2) + \text{var}(d_2) + \text{cov}(d_2, d_3) + \text{cov}(d_2, d_4) + \dots + \text{cov}(d_2, d_H) + \\ \text{cov}(d_1, d_3) + \text{cov}(d_2, d_3) + \text{var}(d_3) + \text{cov}(d_2, d_4) + \dots + \text{cov}(d_3, d_H) + \\ \dots \\ \text{cov}(d_1, d_H) + \text{cov}(d_2, d_H) + \text{cov}(d_3, d_H) + \text{cov}(d_4, d_H) + \dots + \text{var}(d_H) \end{bmatrix} \\ &= \frac{1}{H^2} \begin{bmatrix} \gamma_0 & +\gamma_1 & +\gamma_2 & +\gamma_3 & +\dots & +\gamma_{H-1} & + \\ \gamma_1 & +\gamma_0 & +\gamma_1 & +\gamma_2 & +\dots & +\gamma_{H-2} & + \\ \gamma_2 & +\gamma_1 & +\gamma_0 & +\gamma_1 & +\dots & +\gamma_{H-3} & + \\ \dots & & & & & & \\ \gamma_{H-1} & +\gamma_{H-2} & +\gamma_{H-3} & +\gamma_{H-4} & +\dots & +\gamma_0 & \end{bmatrix} \end{aligned}$$

となる。ここで γ_0 は対角線上に H 個並んでいることが分かる。また、 γ_1 は対角線から 1 つ外れたところに 2 列並んでいるため $2(H-1)$ 個ある。同様に、 γ_2 は対角線から 2 つ外れたところに 2 列並んでいるため $2(H-2)$ 個ある。このようにまとめると、上式は簡潔に

$$\begin{aligned} & \frac{1}{H^2} (H\gamma_0 + 2(H-1)\gamma_1 + 2(H-2)\gamma_2 + 2(H-3)\gamma_3 + \dots + 2\gamma_{H-1}) \\ &= \frac{1}{H} \left\{ \gamma_0 + 2 \left(1 - \frac{1}{H}\right) \gamma_1 + 2 \left(1 - \frac{2}{H}\right) \gamma_2 + 2 \left(1 - \frac{3}{H}\right) \gamma_3 + \dots + 2 \left(1 - \frac{H-1}{H}\right) \gamma_{H-1} \right\} \end{aligned}$$

と表現できる。ここで、 $i > q$ の範囲で $\gamma_i = 0$ であるため、上式は

$$\frac{1}{H} \left\{ \gamma_0 + 2 \left(1 - \frac{1}{H}\right) \gamma_1 + 2 \left(1 - \frac{2}{H}\right) \gamma_2 + \dots + 2 \left(1 - \frac{q}{H}\right) \gamma_q \right\} = \frac{1}{H} \left\{ \gamma_0 + 2 \sum_{i=1}^q \left(1 - \frac{i}{H}\right) \gamma_i \right\}$$

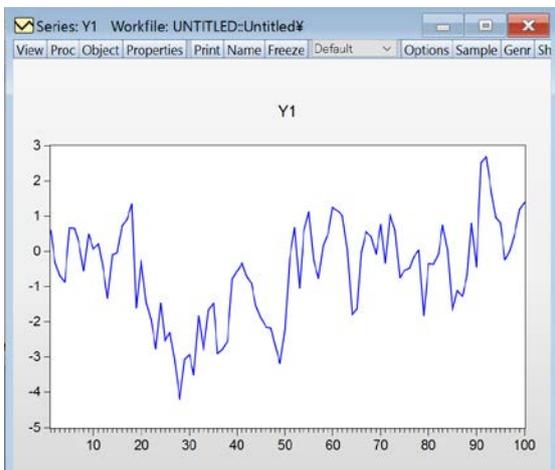
と展開できる。ここで H が十分に大きければ i/H はほぼ 0 となるため、上式は

$$\frac{1}{H} \left\{ \gamma_0 + 2 \sum_{i=1}^q \gamma_i \right\}$$

<実証分析>

ここではEViewsを用いて実証分析を行う。他の統計パッケージを用いても、ほぼ同じ結果が得られる。

[10] (a) y_1 を図にすると以下となる。EViewsでは、系列 y_1 をダブルクリックして選択し、series ウィンドウを表示させた後、メニューバーから「View」→「Graph」を選択すると図を表示できる。平均は時間を通じて安定しており、ショックの分散も安定しているように見える。以上から、 y_1 は定常であると思われる（定常性を確認するための検定については、4章で詳しく学習する）。



(b) 表 2.2 を再現するためには、コマンドとして、AR(1)なら

```
ls y1 y1(-1)
```

と入力し、さらに ε_{t-12} を追加したいなら

```
ls y1 y1(-1) ma(12)
```

と入力すればよい。そうすると結果は以下のようになる。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y1(-1)	0.790470	0.062441	12.65955	0.0000

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y1(-1)	0.793523	0.063756	12.44627	0.0000
MA(12)	-0.022643	0.103884	-0.217969	0.8279
SIGMASQ	0.859270	0.124012	6.928936	0.0000

Statistic	Value	Mean dependent var	S.D. dependent var
R-squared	0.553814	-0.582649	1.394797
Adjusted R-squared	0.544518		
S.E. of regression	0.941340		2.746873
Sum squared resid	85.06769		2.825513
Log likelihood	-132.9702		2.778691
Durbin-Watson stat	2.155823		

Inverted MA Roots	.73	.63+.36i	.63-.36i	.36+.63i	.36-.63i

AR(1)の結果は完全に一致しているが、 ε_{t-12} を追加したケースは ε_{t-12} の係数が-0.0325ではなく-0.022643となっている。違いは非常に小さいので気にしないでもよい。

また、EViews の回帰分析では、係数の推定値のほかに残差平方和（SSR）、情報量基準（AIC、BIC）などのモデルの妥当性を表す指標も表示される。SSR は「Sum Squared resid」、AIC は「Akaike info criterion」、BIC は「Schwarz criterion」の値を見ればよい。SSR はほぼ同じ値となっているが、AIC と BIC は表 2.2 の結果と異なる。これは教科書の定義と EViews 上の定義が異なるためである。

修正 Q 統計量は Equation ウィンドウ上で「View」→「Residual Diagnostics」→「Correlogram Q-Statistics」を選択すると確認できる。Q 統計量は以下の結果となり（Q 統計量は「Q-Stat」、p 値は「Prob」）、表 2.2 とほぼ同じ結果が得られる。

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*		
		1	-0.097	-0.097	0.9645	0.326			1	-0.100	-0.100	1.0198	
		2	0.004	-0.005	0.9665	0.617			2	0.000	-0.010	1.0198	0.313
		3	0.006	0.006	0.9700	0.809			3	0.008	0.007	1.0265	0.599
		4	-0.045	-0.044	1.1787	0.882			4	-0.043	-0.042	1.2214	0.748
		5	-0.046	-0.055	1.4055	0.924			5	-0.049	-0.058	1.4725	0.831
		6	0.091	0.082	2.2988	0.890			6	0.091	0.081	2.3558	0.798
		7	-0.086	-0.071	3.1072	0.875			7	-0.087	-0.072	3.1869	0.785
		8	0.166	0.154	6.1387	0.632			8	0.165	0.153	6.1737	0.520
		9	0.066	0.093	6.6287	0.676			9	0.068	0.094	6.6840	0.571
		10	-0.030	-0.009	6.7297	0.751			10	-0.032	-0.009	6.7969	0.658
		11	0.180	0.189	10.432	0.492			11	0.180	0.189	10.462	0.401
		12	-0.034	-0.002	10.566	0.566			12	-0.020	0.013	10.506	0.486
		13	-0.162	-0.141	13.615	0.401			13	-0.164	-0.138	13.630	0.325
		14	-0.061	-0.121	14.052	0.446			14	-0.061	-0.122	14.064	0.369
		15	0.119	0.134	15.743	0.399			15	0.116	0.128	15.675	0.334
		16	-0.001	0.023	15.743	0.471			16	-0.001	0.025	15.675	0.404
		17	0.051	-0.024	16.063	0.519			17	0.051	-0.021	15.987	0.454
		18	-0.055	-0.046	16.435	0.562			18	-0.054	-0.049	16.349	0.499
		19	0.023	-0.019	16.503	0.623			19	0.022	-0.020	16.410	0.564
		20	0.002	-0.012	16.504	0.685			20	0.003	-0.018	16.412	0.630
		21	0.093	0.138	17.623	0.673			21	0.092	0.135	17.503	0.620
		22	-0.002	0.059	17.624	0.728			22	-0.000	0.062	17.503	0.680
		23	0.048	0.021	17.933	0.761			23	0.048	0.018	17.805	0.717
		24	-0.152	-0.127	21.024	0.637			24	-0.153	-0.127	20.923	0.586
		25	0.046	0.048	21.306	0.675			25	0.046	0.052	21.209	0.626

(c) AR(2)で推定するためには、コマンドとして

ls y1 y1(-1 to -2)

と入力すればよい。そうすると以下の結果となる。

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y1(-1)	0.710157	0.101249	7.013958	0.0000
Y1(-2)	0.105101	0.101497	1.035516	0.3030

R-squared	0.561503	Mean dependent var	-0.585600
Adjusted R-squared	0.556936	S.D. dependent var	1.401658
S.E. of regression	0.932987	Akaike info criterion	2.719346
Sum squared resid	83.56458	Schwarz criterion	2.772100
Log likelihood	-131.2479	Hannan-Quinn criter.	2.740684
Durbin-Watson stat	2.011668		

また、Q 統計量を計算するため、View→Residual Diagnostics→Correlogram Q-Statistics とすると、以下の画面が表示される。

Date: 10/26/17 Time: 16:00
Sample: 1 100
Included observations: 98
Q-statistic probabilities adjusted for 2 dynamic regressors

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.019	-0.019	0.0382	0.845
		2	-0.045	-0.046	0.2485	0.883
		3	-0.015	-0.017	0.2714	0.965
		4	-0.073	-0.076	0.8205	0.936
		5	-0.065	-0.070	1.2659	0.938
		6	0.055	0.045	1.5912	0.953
		7	-0.070	-0.078	2.1247	0.953
		8	0.156	0.153	4.7910	0.780
		9	0.074	0.096	5.3944	0.799
		10	-0.018	0.001	5.4292	0.861
		11	0.169	0.188	8.6600	0.653
		12	-0.032	-0.017	8.7776	0.722
		13	-0.187	-0.142	12.817	0.462
		14	-0.066	-0.095	13.318	0.502
		15	0.123	0.147	15.106	0.444
		16	0.025	0.023	15.182	0.511
		17	0.023	-0.041	15.245	0.578
		18	-0.046	-0.045	15.502	0.627
		19	0.008	-0.025	15.509	0.690
		20	0.009	-0.006	15.519	0.746
		21	0.090	0.138	16.546	0.738
		22	0.023	0.051	16.615	0.784
		23	0.024	0.001	16.690	0.824
		24	-0.151	-0.133	19.700	0.714
		25	0.032	0.063	19.834	0.755
		26	0.044	-0.000	20.099	0.787

Q-Stat の列をみると、 $Q(8)=4.79$ 、 $Q(16)=15.11$ 、 $Q(24)=19.70$ となる（多少のずれはあるが、教科書とほぼ同じ値である）。また、対応する p 値(Prob の列)をみると、全て 0.1 を上回っており、有意水準 10% でも系列相関がないという仮説を棄却できない。

(d) ARMA(1,1)であれば $ls\ y1\ y1(-1)\ ma(1)$ と入力すればよい。以下が推定結果である。

Equation: UNTITLED Workfile: UNTITLED:Untitled*

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

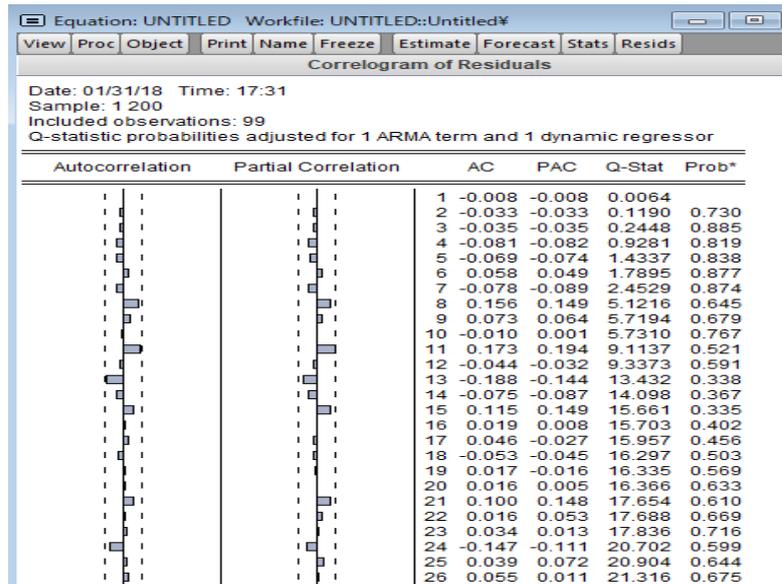
Dependent Variable: Y1
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/31/18 Time: 17:26
Sample: 2 100
Included observations: 99
Convergence achieved after 4 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y1(-1)	0.843735	0.067068	12.58036	0.0000
MA(1)	-0.141707	0.133656	-1.060242	0.2917
SIGMASQ	0.849649	0.120041	7.077966	0.0000

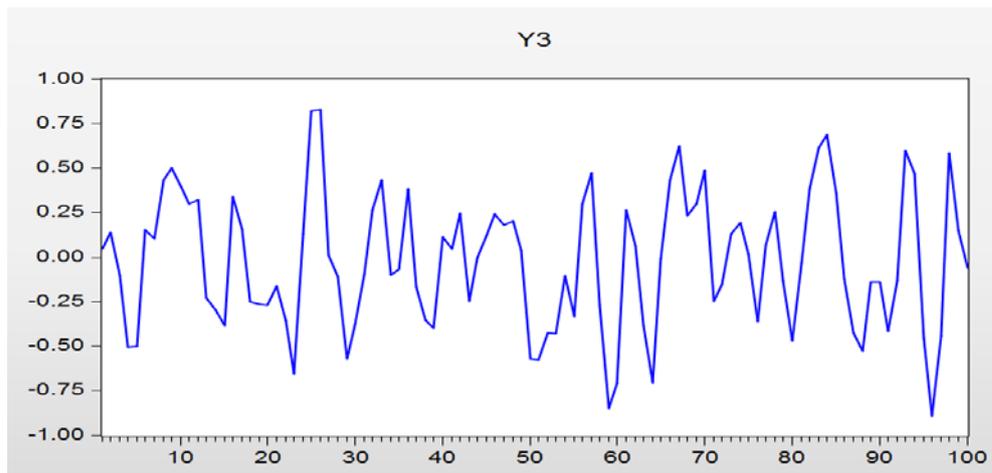
R-squared	0.558809	Mean dependent var	-0.582649
Adjusted R-squared	0.549617	S.D. dependent var	1.394797
S.E. of regression	0.936056	Akaike info criterion	2.735757
Sum squared resid	84.11529	Schwarz criterion	2.814397
Log likelihood	-132.4199	Hannan-Quinn criter.	2.767574
Durbin-Watson stat	1.982613		

Inverted MA Roots	.14
-------------------	-----

また、Q 統計量を計算するため、Equation Window において、View→Residual Diagnostics→Correlogram Q-Statistics を選択すると、以下の画面が表示される。残差の ACF、PACF も同じ画面から確認できる（ACF は「AC」、PACF は「PAC」として表示される）。ACF、PACF とも 2 標準誤差（点線の内側）に収まっている。Q-Stat の列をみると、 $Q(8)=5.12$ 、 $Q(16)=15.70$ 、 $Q(24)=20.70$ となる。また、対応する p 値(Prob の列)をみると、全て 0.1 を上回っており、有意水準 10% でも系列相関がないという帰無仮説を棄却できない。



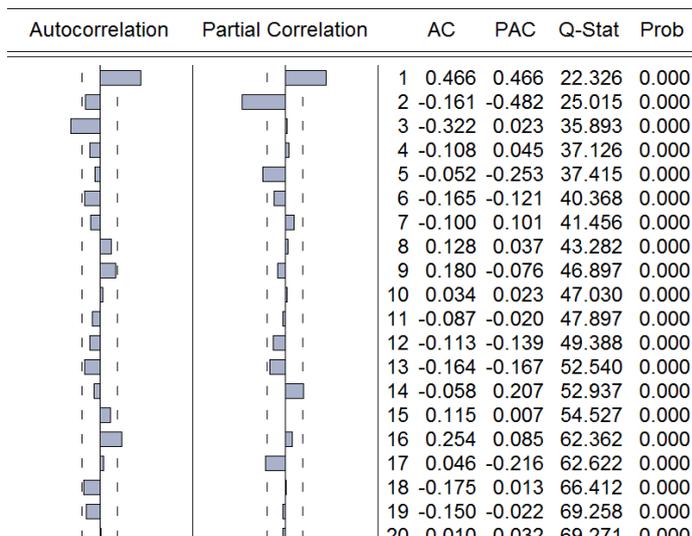
[11] (a) y_3 を図示すると以下ようになる。平均は安定しており、ショックの分散も一定のように見える。



また、Series Window で、View→Correlogram とすると、以下が得られる。これをみると、ACF は単調減少ではないため、単なる AR(1)ではなさそうである⁴。PACF をみると、2 次まで尖りがあるが、それ以降はほぼ 0 であり、AR(2)の可能性を疑わせる⁵。

⁴ AR(2)の理論上の ACF は 0 に減衰し（振動してもよい）、PACF は次数 2 までは尖りがあり、それ以上は 0 となる。実際の値を見ると、ACF は単調減少ではないため、単なる AR(1)ではなさそうである。

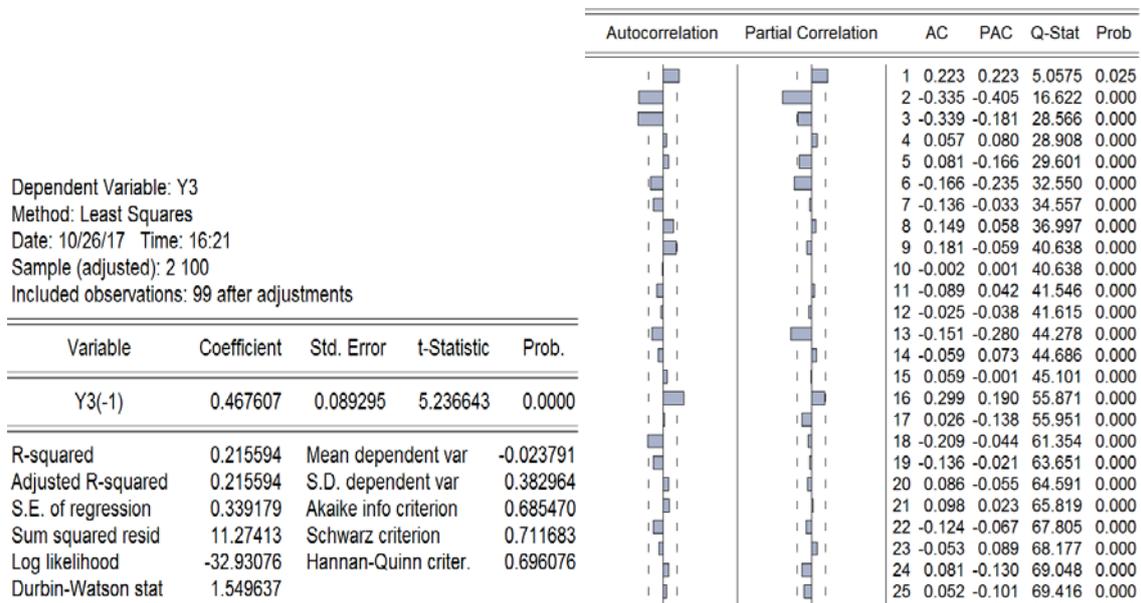
⁵ 注意深くみると、ACF はラグ 16 で、PACF はラグ 14 と 17 で大きくなっている。データは AR(2)から発生しているため、こうした結果は偶然の産物に過ぎない。しかし、そのことを知らない分析者は AR(2)ではなく、こうした特徴を考慮するモデルを推定しようとするかもしれない。



(b)AR(1)モデルとして推定するため、

ls y3 y3(-1)

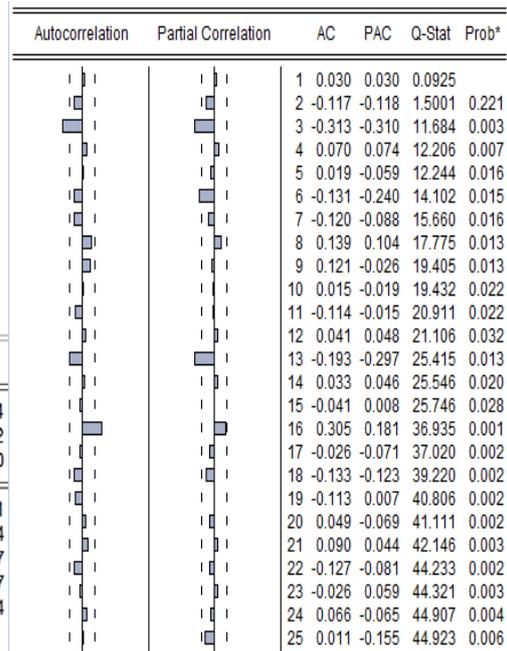
と入力すると左下の画面となる。右下の画面では、残差の AC、PAC、Q-stat を示している。Q 統計量は大きな値をとっており（対応する p 値は 0.05 を下回る）、系列相関がある可能性を示唆している。したがって、AR(1)は不適切なモデルといえる。



(c) ARMA(1,1)で推定すると以下となる。多少、係数の値は異なるがほぼ同じ結果がえられる。また、残差の Q 統計量を調べると大きな値をとっており、系列相関が残るため、不適切なモデルといえる。

Dependent variable: Y3
 Method: ARMA Maximum Likelihood (BFGS)
 Date: 10/26/17 Time: 16:25
 Sample: 2 100
 Included observations: 99
 Convergence achieved after 6 iterations
 Coefficient covariance computed using outer product of gradients

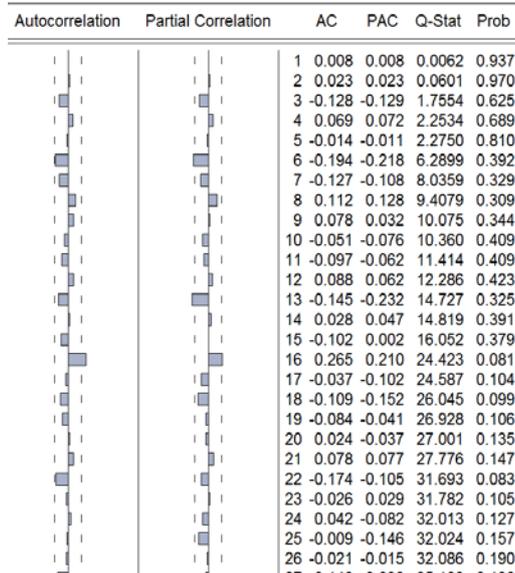
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y3(-1)	0.190075	0.152415	1.247090	0.2154
MA(1)	0.501373	0.149853	3.345756	0.0012
SIGMASQ	0.097438	0.019257	5.059915	0.0000
R-squared	0.328849	Mean dependent var	-0.023791	
Adjusted R-squared	0.314866	S.D. dependent var	0.382964	
S.E. of regression	0.316990	Akaike info criterion	0.572867	
Sum squared resid	9.646343	Schwarz criterion	0.651507	
Log likelihood	-25.35690	Hannan-Quinn criter.	0.604684	
Durbin-Watson stat	1.934114			



(d) 最後に、AR(2)を推定すると左下の画面となる。推定値は、真のデータ生成過程 $y_t = 0.7y_{t-1} - 0.49y_{t-2} + \varepsilon_t$ とほぼ同じである。また、右下の画面では、残差の AC、PAC、Q-stat を示している。Q 統計量は小さく有意になっていない。したがって、系列相関が残っているとはいえ、適切なモデルであるといえる。

Dependent Variable: Y3
 Method: Least Squares
 Date: 10/26/17 Time: 16:28
 Sample (adjusted): 3 100
 Included observations: 98 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y3(-1)	0.692389	0.089516	7.734825	0.0000
Y3(-2)	-0.480873	0.089577	-5.368291	0.0000
R-squared	0.396296	Mean dependent var	-0.025483	
Adjusted R-squared	0.390007	S.D. dependent var	0.384561	
S.E. of regression	0.300350	Akaike info criterion	0.452460	
Sum squared resid	8.660171	Schwarz criterion	0.505215	
Log likelihood	-20.17056	Hannan-Quinn criter.	0.473798	
Durbin-Watson stat	1.969341			

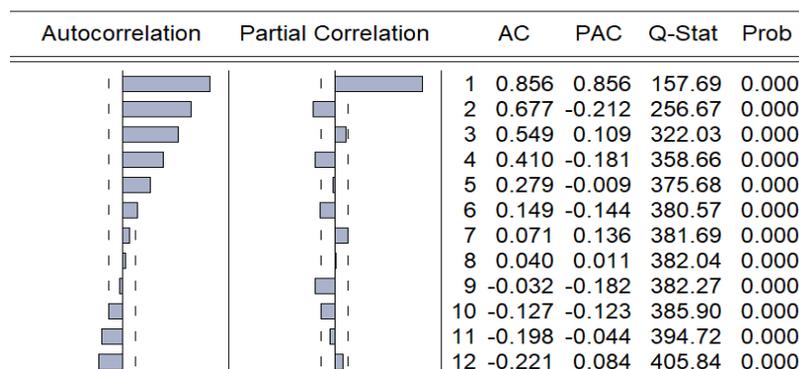


[12] QUARTERLY.XLS を開いて EViews にデータを取り込もう。まず、

genr s=r5-tbill

として、スプレッドの系列 s を作る。

まず、コレログラムを調べてみよう。Workfile Window から s を選択し、View→Correlogram とし、Lags to include を 12 (図 2.8 は 12 まで考慮しているため) として OK をおす。



この結果は、教科書の図 2.8 とほぼ同じであることを確認してもらいたい。教科書の図では、0 次の ACF と PACF が表記されているが、これらは常に 1 であるため、EViews では掲載されていない。

この s を用いて表 2.4 を再現するためには、以下のコマンドを入力すればよい（ただし、教科書と同様、標本期間は全て 1961Q4～2012Q4 とする）。ただし、特定の標本期間で推定するため、メニューバーの「Quick」→「Estimate Equation」を選択して Equation Estimation ウィンドウで推計式を入力する。「Equation specification」にコマンドを入力し、「Sample」で期間（この場合は 1961q4 2012q4）を入力すればよい⁶。

AR(7)なら

s c s(-1 to -7)

とする（定数項の値を一致させたいなら ls s c ar(1 to 7)とすればよい）。同様にして、AR(2)、AR(6)も推定できる。p=1,2,7 なら、

s c s(-1 to -2) s(-7)

とする。ARMA(1,1)なら

s c s(-1) ma(1)

とすればよい。同様に、ARMA(2,1)も推定できる。ARMA(2,(1,7))なら (p=2、q=1,7)、

s c s(-1 to -2) ma(1) ma(7)

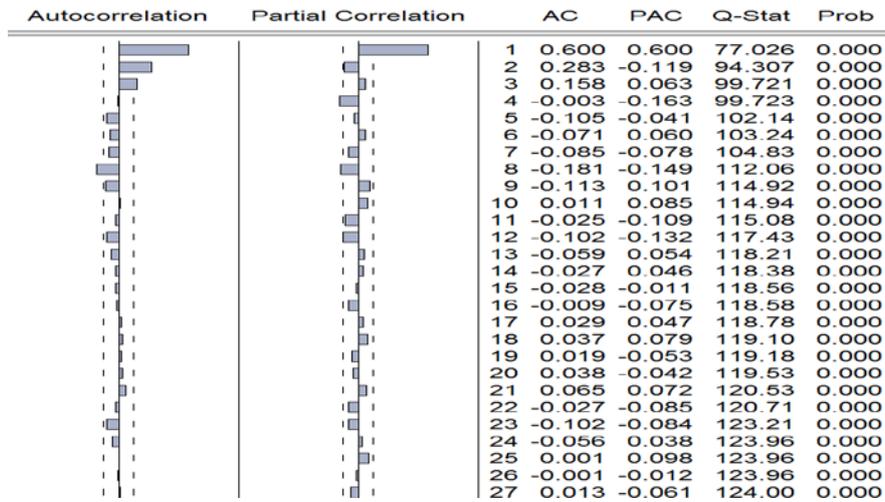
となる。値は多少異なるが、表 2.4 を簡単に再現できる。

[13] (a)

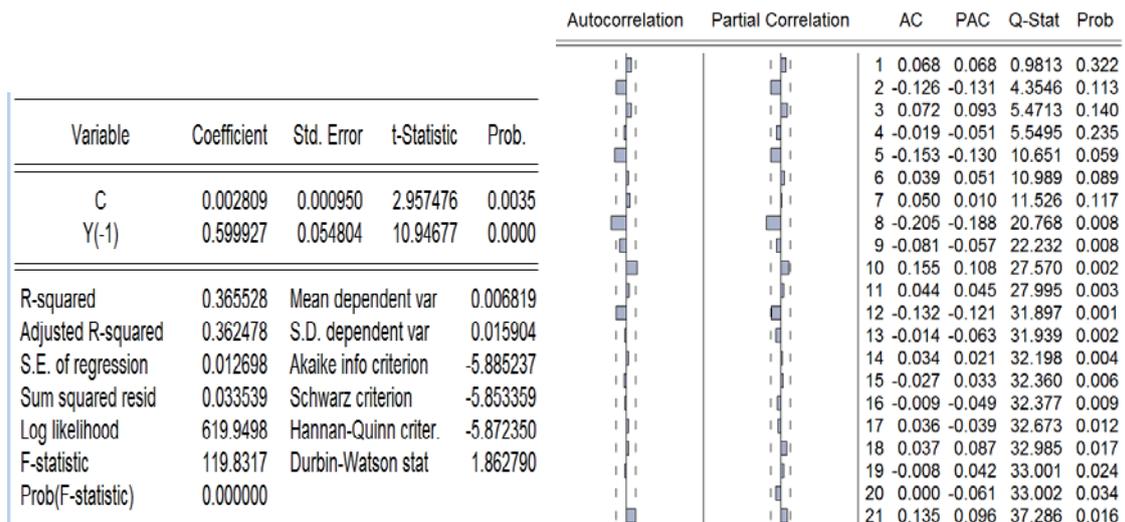
genr y=log(indprod)- log(indprod(-1))

として y を作る。そして、ACF と PACF は以下となる。ACF は急激に減少しており、PACF は 1 次で尖りがあり、他はほぼ 0 である（8 次でだけ大きな値をとっている）。以上から、AR(1)過程の可能性が疑われるが、懸念としては 8 次で ACF と PACF が 0 から乖離していることである。

⁶ なお、Equation Estimation ウィンドウでは推計方法を選択するため、「ls」など推計方法を表すコマンドは入力しない。



(b)低次の ACF が高いことから AR(1)として推定すると、左下の画面となる。また、残差の ACF と PACF をみると、ともに 8 次で大きな値をとっている。



(c) ここで y_{t-8} を説明変数に加えるため

ls y c y(-1) y(-8)

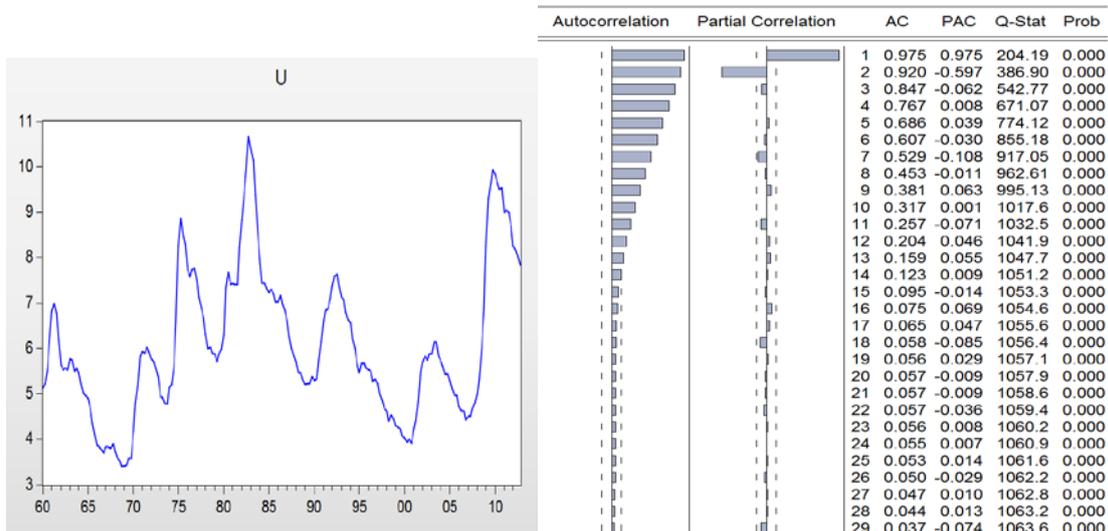
と入力すると左下の画面が得られる。 y_{t-8} の係数は有意である。また、AIC、BIC、SSRは全て、AR(1)よりも小さくなっているため、モデルのフィットは改善しているといえる。また、残差の系列相関を調べると、8次の系列相関は小さくなっている。懸念としては、10次の系列相関が大きくなっていること、また、鉱工業生産指数の分析において、なぜ8次のラグが有意で2~7次のラグが有意ではないのか分からない。

Variable	Coefficient	Std. Error	t-Statistic	Prob.	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
Dependent Variable: Y											
Method: Least Squares											
Date: 10/26/17 Time: 17:03											
Sample (adjusted): 1962Q2 2012Q4											
Included observations: 203 after adjustments											
C	0.003618	0.001015	3.565184	0.0005			1	0.084	0.084	1.4483	0.229
Y(-1)	0.588308	0.055857	10.53244	0.0000			2	-0.097	-0.105	3.3870	0.184
Y(-8)	-0.131225	0.053174	-2.467817	0.0144			3	0.083	0.102	4.8176	0.186
							4	-0.043	-0.073	5.2027	0.267
							5	-0.174	-0.147	11.559	0.041
							6	0.016	0.030	11.615	0.071
							7	0.039	0.011	11.943	0.102
							8	-0.115	-0.095	14.751	0.064
							9	0.029	0.038	14.934	0.093
							10	0.200	0.155	23.570	0.009
							11	0.077	0.079	24.855	0.010
							12	-0.132	-0.134	28.631	0.004
							13	-0.026	-0.049	28.784	0.007
							14	0.016	0.026	28.843	0.011
							15	-0.083	-0.012	30.368	0.011
							16	-0.031	-0.030	30.586	0.015
							17	0.023	-0.030	30.703	0.022
							18	0.042	0.087	31.103	0.028
							19	-0.013	-0.007	31.142	0.039
							20	0.007	-0.070	31.152	0.053
							21	0.162	0.142	37.134	0.016
							22	-0.002	0.027	37.135	0.023
							23	-0.104	-0.059	39.629	0.017
							24	-0.047	-0.099	40.135	0.021
R-squared	0.380153	Mean dependent var	0.006742								
Adjusted R-squared	0.373954	S.D. dependent var	0.015480								
S.E. of regression	0.012248	Akaike info criterion	-5.952226								
Sum squared resid	0.030003	Schwarz criterion	-5.903262								
Log likelihood	607.1509	Hannan-Quinn criter.	-5.932417								
F-statistic	61.33009	Durbin-Watson stat	1.830974								
Prob(F-statistic)	0.000000										

(d)失業率を u と定義する。

genr u=unemp

そして、失業率を図にするとあてもなく変動し、平均が一定でないようにみえるため、系列 u は非定常であるかもしれない。また、ACF をみると非常にゆっくりと低下している（定常であるに於ては収束速度が遅すぎるようにみえる）。以上の結果から、定常かどうかははっきりとはいえないが、非定常の可能性を疑わせる。



(e) AR(2)モデルを推定し、その残差の ACF、PACF を求めた。推定の際には、

ls u c s(-1 to -2)

ではなく、

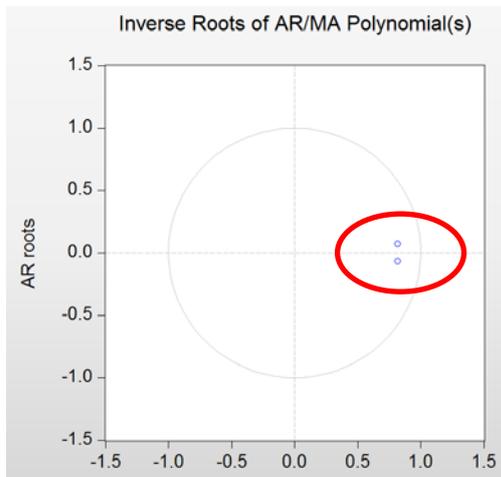
ls u c ar(1 to 2)

としておこう（これは ARMA structure を用いるために必要である）。

Equation: UNTITLED Workfile: UNTITLED::Untitled¥					Date: 10/26/17 Time: 17:39
View Proc Object Print Name Freeze Estimate Forecast Stats Resids					Sample: 1960Q1 2012Q4
Dependent Variable: U					Included observations: 212
Method: ARMA Maximum Likelihood (BFGS)					Q-statistic probabilities adjusted for 2 ARMA terms
Date: 10/26/17 Time: 17:28					
Sample: 1960Q1 2012Q4					
Included observations: 212					
Convergence achieved after 4 iterations					
Coefficient covariance computed using outer product of gradients					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	6.084676	0.567983	10.71278	0.0000	
AR(1)	1.640403	0.037382	43.88180	0.0000	
AR(2)	-0.677992	0.038594	-17.56745	0.0000	
SIGMASQ	0.060439	0.004660	12.97102	0.0000	
R-squared	0.976700	Mean dependent var	6.081981		
Adjusted R-squared	0.976364	S.D. dependent var	1.614394		
S.E. of regression	0.248197	Akaike info criterion	0.090007		
Sum squared resid	12.81314	Schwarz criterion	0.153339		
Log likelihood	-5.540748	Hannan-Quinn criter.	0.115604		
F-statistic	2906.359	Durbin-Watson stat	1.967838		
Prob(F-statistic)	0.000000				
Inverted AR Roots	.82-.07i	.82+.07i			

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.015	0.015	0.0488		
2	-0.010	-0.010	0.0695		
3	0.075	0.076	1.3064	0.253	
4	-0.100	-0.103	3.4943	0.174	
5	-0.097	-0.093	5.5480	0.136	
6	0.117	0.116	8.5819	0.072	
7	0.136	0.152	12.703	0.026	
8	-0.155	-0.165	18.039	0.006	
9	0.022	-0.013	18.147	0.011	
10	0.101	0.113	20.434	0.009	
11	-0.037	0.036	20.747	0.014	
12	-0.046	-0.090	21.230	0.020	
13	0.021	-0.056	21.333	0.030	
14	-0.018	0.041	21.411	0.045	
15	0.005	0.087	21.416	0.065	
16	-0.042	-0.132	21.820	0.082	
17	0.071	0.027	22.983	0.085	
18	-0.035	0.028	23.274	0.107	
19	-0.023	0.020	23.395	0.137	
20	0.024	-0.043	23.528	0.171	
21	0.047	0.040	24.055	0.194	
22	-0.011	0.032	24.085	0.239	
23	-0.009	0.004	24.107	0.288	
24	0.014	-0.057	24.157	0.339	
25	0.011	0.028	24.186	0.384	

ここで ACF、PACF は小さな値をとっている。また、Q 統計量も全般的に小さな値をとり、対応する p 値も 5% を上回る（ただし、次数 7~14 で 5% を下回る）。特性根(Inverted AR Roots)は、 $0.82 \pm 0.07i$ となっている。Equation Window から View→ARMA structure→Roots を選択すると特性根を図示することもできる。



この特性根は、第 1 章 5 節のケース 3 に該当している。この図をみると、特性根が単位円内にあるが、単位円のかかなり近くに根が存在しているのが分かる。

- (f) 失業率の階差を AR(1)モデルで推定すると、係数は 0.976 とかなり 1 に近い値をとる。また、残差の Q 統計量を求めると、大きな値をとっていることが確認できる。これは系列相関の可能性を示唆し、AR(1)は不適切なモデルである。また、SSR、AIC、BIC の値と比べても、AR(1)では、どれも AR(2)より大きい値をとっているため、当てはまりの悪いモデルといえる。

Equation: UNTITLED Workfile: UNTITLED::Untitled¥

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: U
 Method: Least Squares
 Date: 10/26/17 Time: 17:36
 Sample (adjusted): 1960Q1 2012Q3
 Included observations: 211 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.132394	0.090533	1.462376	0.1451
U(1)	0.976146	0.014378	67.89122	0.0000
R-squared	0.956623	Mean dependent var	6.073697	
Adjusted R-squared	0.956415	S.D. dependent var	1.613710	
S.E. of regression	0.336893	Akaike info criterion	0.671330	
Sum squared resid	23.72084	Schwarz criterion	0.703102	
Log likelihood	-68.82536	Hannan-Quinn criter.	0.684173	
F-statistic	4609.217	Durbin-Watson stat	0.669315	
Prob(F-statistic)	0.000000			