

応用計量経済学(9)

横浜市立大学商学部

松浦 克己

大阪大学国際公共政策研究科

Colin McKenzie

第9章 時系列モデル

1 VARモデル

実証分析を行う目的の一つに予測がある。その有力な方法は第6章で取り上げた連立方程式モデルである。またARモデルやARMAモデルも予測に用いられる。1980年代以降はSimsによって提唱されたベクトル自己回帰モデル(vector autoregression, VAR)が連立方程式モデルと並んでよく用いられるようになった。特に金融資本市場の分析ではVARモデルは有力な分析予測手段となっている。

VARモデルでは、複数変数間の影響を取り上げる(ARMAモデルとはこの点で異なる)。そのとき内生変数と外生変数を予め区別しない(この意味で連立方程式モデルと異なる¹⁾)。いわばその取り上げられた複数の変数のラグでモデルを説明し、予測を行おうというものである。

(VARモデルの考え方)

VARモデルの推定・検定に当たっては、データが定常であることが前提である(単位根を持たない)。以下の議論ではこの前提は充たされているものとする²⁾。

最初に次のような構造型モデルを考える。

$$y_t = A_1 + B_{10}x_t + B_{11}x_{t-1} + C_{11}y_{t-1} + u_{1t} \quad (9.1)$$

$$x_t = A_2 + B_{21}x_{t-1} + C_{20}y_t + C_{21}y_{t-1} + u_{2t} \quad (9.2)$$

$$u_{1t} \sim \text{IID}(0, \sigma_y^2) \quad u_{2t} \sim \text{IID}(0, \sigma_x^2),$$

$$E(u_{1t}, u_{2t}) = \sigma_{xy}$$

y_t と x_t の誘導型は

$$y_t = a_1 + b_{11}x_{t-1} + c_{11}y_{t-1} + e_{1t} \quad (9.3)$$

$$x_t = a_2 + b_{21}x_{t-1} + c_{21}y_{t-1} + e_{2t} \quad (9.4)$$

と書くことができる。9.3)式にある y_t と9.4)式にある x_t を説明するために1期前の x_{t-1} と y_{t-1} のみを利用するので、9.3)式と9.4)式は1次ベクトル自己回帰モデル(VAR(1))と呼ばれている。VARの特徴の一つは説明変数のほとんどが先決内生変数ということである。このとき e_{1t}, e_{2t} は以下のようである。

$$e_{1t} = \frac{u_{1t} + B_{10}u_{2t}}{1 - B_{10}C_{20}} \quad e_{2t} = \frac{u_{2t} + C_{20}u_{1t}}{1 - B_{10}C_{20}}$$

u_{1t} と u_{2t} はIIDであるから、 $E(e_{1t}) = 0$ 、 $E(e_{2t}) = 0$ となる。

e_{1t} の分散は

$$V(e_{1t}) = (u_{1t} + B_{10}u_{2t})'(1 - B_{10}C_{20})^{-2} \\ = (\sigma_y^2 + B_{10}^2 \sigma_x^2 + 2B_{10} \sigma_{xy}) / (1 - B_{10}C_{20})^2$$

となり一定である。また自己共分散は

$$\text{Cov}(e_{1t}, e_{1t-k}) = \frac{E(u_{1t} + B_{10}u_{2t})(u_{1t-k} + B_{10}u_{2t-k})}{(1 - B_{10}C_{20})^2} = 0$$

であるから、系列相関はない。 e_{2t} についても全く

¹⁾ この違いを強調して伝統的な連立方程式モデルを計量経済モデル、VARモデルを時系列モデルということがある。

²⁾ VARモデルの推計に先だって、単位根の検定を必ず行う必要がある。

同様に証明することができる。e_{1t}とe_{2t}はIIDとなる。

e_{1t}とe_{2t}の共分散を考えると

$$\begin{aligned} \text{Cov}(e_{1t}, e_{2t}) &= \frac{E(u_{1t} + B_{10}u_{2t})(u_{2t} + C_{20}u_{1t})}{(1 - B_{10}C_{20})^2} \\ &= \frac{B_{10} \sigma_x^2 + C_{20} \sigma_y^2 + (1 + C_{20}B_{10}) \sigma_{xy}}{(1 - B_{10}C_{20})^2} \end{aligned} \quad (9.5)$$

となる。9.5)式の分母は一般には0とならないであろう。そうすると9.3)式と9.4)式は第7章で説明したSURモデルの特殊ケースとして考えたら良い。

しかし、

$$B_{10} = C_{20} = \sigma_{xy} = 0$$

という特殊ケースにおいては、9.5)式は0となる。すなわち当期のx_tは当期のy_tに影響せず、当期のy_tは当期のx_tに影響せず、u_{1t}とu_{2t}は相関しないというこのケースでは、方程式間の誤差項に相関がないので、9.3)式と9.4)式を各々OLSで推計すればBLUEとなる。また9.3)式と9.4)式を比較すると、方程式間で説明変数が全て共通である。これは第7章で見たSURをOLSで推計できる特殊ケースとなっていることが分かる(第7章参照)。

2変量のVARモデルを次のように書くことができる。ラグの長さを仮に2期だとする。この場合VAR(2)と呼ぶ。

$$y_t = a_1 + b_{11}x_{t-1} + b_{12}x_{t-2} + c_{11}y_{t-1} + c_{12}y_{t-2} + e_{1t} \quad (9.6)$$

$$x_t = a_2 + b_{21}x_{t-1} + b_{22}x_{t-2} + c_{21}y_{t-1} + c_{22}y_{t-2} + e_{2t} \quad (9.7)$$

$$e_{1t} \sim \text{IID}(0, \sigma_y^2) \quad e_{2t} \sim \text{IID}(0, \sigma_x^2)$$

さらに内生変数がm個、ラグの長さがpのVARモデル(VAR(p))を

$$\begin{aligned} y_{1t} &= a_1 + b_{11}y_{1t-1} + \dots + b_{1p}y_{1t-p} + c_{11}y_{2t-1} + \dots \\ &+ c_{1p}y_{2t-p} + \dots + n_{11}y_{mt-1} + \dots + n_{1p}y_{mt-p} + e_{1t} \end{aligned}$$

$$\begin{aligned} y_{2t} &= a_2 + b_{21}y_{1t-1} + \dots + b_{2p}y_{1t-p} + c_{21}y_{2t-1} + \dots \\ &+ c_{2p}y_{2t-p} + \dots + n_{21}y_{mt-1} + \dots + n_{2p}y_{mt-p} + e_{2t} \end{aligned} \quad (9.8)$$

$$\begin{aligned} y_{mt} &= a_m + b_{m1}y_{1t-1} + \dots + b_{mp}y_{1t-p} + c_{m1}y_{2t-1} + \dots \\ &+ c_{mp}y_{2t-p} + \dots + n_{m1}y_{mt-1} + \dots + n_{mp}y_{mt-p} + e_{mt} \end{aligned}$$

と表すことができる。

VARモデルを連立方程式モデルと比べた利点は、ラグ付き変数のみを利用するので予測が容易だということである。9.6)式をOLSで推計し、a₁, b₁₁, b₁₂, c₁₁, c₁₂の推定量を $\hat{a}_1, \hat{b}_{11}, \hat{b}_{12}, \hat{c}_{11}, \hat{c}_{12}$ とし、t+1期を予測したいとすると、

$$\hat{y}_{t+1} = \hat{a}_1 + \hat{b}_{11}x_t + \hat{b}_{12}x_{t-1} + \hat{c}_{11}y_t + \hat{c}_{12}y_{t-1} \quad (9.9)$$

のようにt期とt-1期の情報を用いればよい。

これに対し通常の連立方程式モデルは、当期における変数間の相互作用が考慮される。たとえば9.1)式と9.2)式の係数の推定量を $\hat{\cdot}$ で表記すると、y_{t+1}の予測は9.1)式より

$$\hat{y}_{t+1} = \hat{A}_1 + \hat{B}_{10}x_{t+1} + \hat{B}_{11}x_t + \hat{C}_{11}y_t \quad (9.10a)$$

のようにt+1期のx_{t+1}の情報が必要となる。t期にt+1期のxに関する情報は得られないので、これを予測する必要がある。例えば9.2)式より

$$\hat{x}_{t+1} = \hat{A}_2 + \hat{B}_{21}x_t + \hat{C}_{20}y_{t+1} + \hat{C}_{21}y_t \quad (9.10b)$$

を予測する。しかしその予測自体がy_{t+1}の予測に依存する。問題はかなり複雑である。VARモデルはこの点で簡便となっている。

(説明変数とラグの長さの選択)

VARモデルは内生変数と外生変数を区別しない(言い換えれば各左辺の変数のラグは、全ての式の右辺に現れる)。方程式間で説明変数は共通であるから、ラグの長さを共通にするケースが多い。3変数で4期のラグを取ると、定数項を含め39個のパラメータを推計する必要がある。9.9)式のように変数をm個、ラグの長さをpとするとm+m²pのパラメータを推定する必要がある。変数の数やラグの長さを負や増やすと自由度が急速

に低下する。サンプル期間が短い(サンプル数が少ない)場合、これは深刻な問題となる。他方でpの長さが短いと過小定式化の誤りが生じることがある。

mを所与とした場合の、ラグの長さの問題を考えてみよう。

ラグの長さに関わらず同一サンプル期間について、VARモデルを推計する。

VARモデルで情報量基準のAICとSBCは

$$AIC = -2L/T + 2n/T$$

$$SBC = -2L/T + n(\log T)/T$$

で計算される。ここでnは推定されるパラメータ数、Tはサンプル数。

$$L = -mT/2(1 + \log 2) - T/2 \log | |$$

は推計されるVARモデルの誤差項の分散共分散行列の値である。

このAICまたはSBCの値を最小にするラグの長さを選べばよい³⁾。もちろん仮説検定の考え方に従ってラグを決めることができる。

(Eviewsによる推計例)

金融政策と経済成長の関係はマクロ経済の大きなテーマである。その一環として通貨需要関数の安定性が問われてきた。ここでは簡単な通貨需要関数をVARモデルで推計してみよう。通貨需要としてはM2+CD(未季節調整)、所得(経済活動)の代理変数として鉄工業生産指数(IIPと表記)、金利としてコールレート(有担保、CALL1と表記)をとりあげる。なお、M2+CDは消費者物価で実質化した対数値で、この変数をM2CDと表記する。

M2CD、IIP、コールレートの単位根検定を第8章で解説した要領で行った。その結果トレンド

と定数項を含み、ラグの長さを4期として、1回の階差をとるとADF統計量は

$$M2CD(-9.114), IIP(-14.486)$$

$$CALL1(-5.199)$$

であり、いずれもI(1)変数であることが分かった。

なお分析期間は1975年1月～1998年12月である(サンプル数を確保するためにデータ自体は1970年1月から読み込んである)。

$$M2CD_t = a_1 + b_{11} M2CD_{t-1} + \dots + b_{1p}$$

$$M2CD_{t-p} + c_{11} IIP_{t-1} + \dots + c_{1p}$$

$$IIP_{t-p} + d_{11} CALL1_{t-1} + \dots + d_{1p}$$

$$CALL1_{t-p} + e_{1t}$$

$$IIP_t = a_2 + b_{21} M2CD_{t-1} + \dots + b_{2p}$$

$$M2CD_{t-p} + c_{21} IIP_{t-1} + \dots + c_{2p}$$

$$IIP_{t-p} + d_{21} CALL1_{t-1} + \dots + d_{2p}$$

$$CALL1_{t-p} + e_{2t}$$

$$CALL1_t = a_3 + b_{31} M2CD_{t-1} + \dots + b_{3p}$$

$$M2CD_{t-p} + c_{31} IIP_{t-1} + \dots + c_{3p}$$

$$IIP_{t-p} + d_{31} CALL1_{t-1} + \dots + d_{3p}$$

$$CALL1_{t-p} + e_{3t}$$

というVAR(p)モデルを推定しよう⁴⁾。

Quick/Estimate VAR

を選択する。次のような画面がでてくるであろう(図1参照)。

VAR specificationにUnrestricted VARとVector Error Correction

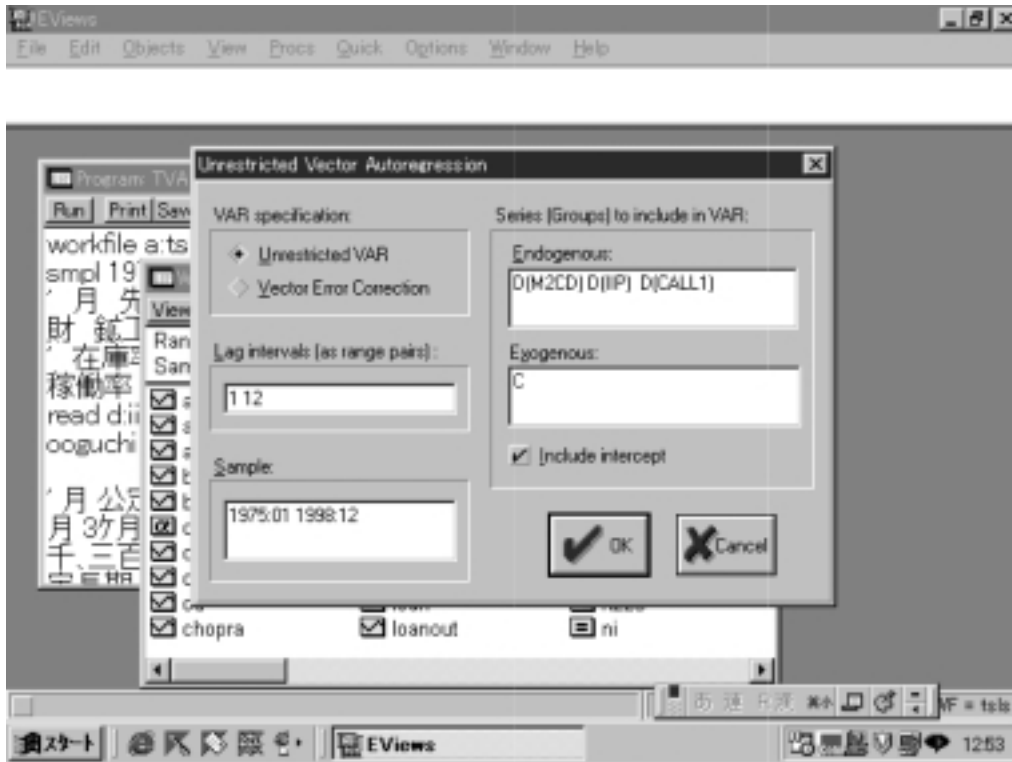
の二通りの選択がある。ここではUnrestricted VARを選択する(Vector Error CorrectionはError Correctionモデルである)。

Endogenousに内生変数を入れる。ここでは単位根検定の結果を受けて、

³⁾ VARモデルの場合にも、Error Correctionを考えることはできる。VAR Error Correction Model, VECMといわれることがある。第8章で説明したように共和分検定の統計量はm=6までしか開発されていない。そのためにVARモデルでも変数の数は6個以下とされるのが通常である。ただしOsterwaldとLenucciはVARモデルのケースについて最大11個までの臨界値を与えている。

⁴⁾ M2CD、IIP、CALL1の三つの変数の間には共和分関係が存在しないと仮定している。

図9.1 VARモデルの入力例



$D(M2CD)$ $D(IIP)$ $D(CALL1)$
と階差をとって入れている。

Exogenousに外生変数を入れる。ここでは既定値のc (定数項) を入れている⁵⁾。

Lag intervals (as range pairs)
でラグの長さを聞いてくる。最長18期のラグから推計を開始して、12期がAIC基準を最小にしたので、ここではラグの長さを12とし、1 12と入力している。

Sampleで分析期間を聞いてくる。ここでは1975年1月から1998年12月としている。

3変数でラグが12期であるから、推計すべきパラメータは $3 + 3^2 * 12 = 111$ 個である。容易に想像されるように、各変数については多重共線関係が頻発している。このようなケースでは個別の係数の有意水準などは通常問題とされない。係数が

111個もあるので、個別の係数の経済的意味が問われることもない(実際問題無理である)。

VARモデルの経済的な解釈は、因果性の検定やインパルス応答関数、分散分析などで行われる。

結果は表9.1に示すとおりである(紙幅の関係で大幅に省略してある)。

表の最下欄にあるAkaike Information CriteriaがVARモデルのラグの長さの選択に使われたものである。Determinant Residual CovarianceがVARモデルの誤差項の分散共分散の行列式である。

その上の欄に報告されているR squared, Adj. R squaredは各方程式の決定係数などである。これも報告されないことがある。

⁵⁾ VARモデルで使われる定数項以外の外生変数にタイムトレンドと季節ダミーがある。

表9.1 Vector Autoregression Estimates

Sample : 1975 : 01 1998 : 12			
Included observations : 288			
Standard errors & t statistics in parentheses			
	Δ(M2CD)	Δ(IIP)	Δ(CALL1)
Δ(M2CD(-1))	0.006631 (0.05312) (0.12484)	-13.14594 (15.0009) (-0.87634)	0.418919 (2.33449) (0.17945)
Δ(M2CD(-2))	-0.087157 (0.05275) (-1.65219)	8.738316 (14.8980) (0.58654)	-0.754545 (2.31847) (-0.32545)
Δ(M2CD(-12))	0.512651 (0.05079) (10.0939)	25.00701 (14.3431) (1.74348)	-0.532661 (2.23213) (-0.23863)
Δ(IIP(-1))	-0.000128 (0.00018) (-0.70107)	-0.407664 (0.05151) (-7.91441)	0.010197 (0.00802) (1.27206)
Δ(IIP(-12))	-0.000174 (0.00018) (-0.94507)	0.564735 (0.05210) (10.8402)	0.008157 (0.00811) (1.00610)
Δ(CALL1(-1))	0.001520 (0.00142) (1.07359)	0.595691 (0.39989) (1.48963)	0.062393 (0.06223) (1.00257)
Δ(CALL1(-12))	-0.001773 (0.00143) (-1.27094)	0.347068 (0.40349) (0.86016)	-0.160096 (0.06279) (-2.54959)
C	0.001402 (0.00062) (2.26224)	-0.297864 (0.17504) (-1.70172)	-0.060048 (0.02724) (-2.20441)
R squared	0.808423	0.943793	0.192816
Adj. R squared	0.780946	0.935732	0.077045
Sum sq. resid	0.008880	708.2142	17.15196
S.E. equation	0.005948	1.679753	0.261409
F statistic	29.42167	117.0741	1.665491
Log likelihood	1087.066	-538.2235	-2.452130
Akaike AIC	-7.292127	3.994608	0.273973
Schwarz SC	-6.821538	4.465196	0.744562
Mean dependent	0.003708	0.151042	-0.029514
S.D. dependent	0.012708	6.625944	0.272101
Determinant Residual Covariance	4.51E 06		
Log Likelihood	546.4911		
Akaike Information Criteria	-3.024244		
Schwarz Criteria	-1.612478		

2 因果性の検定

(Grangerの意味での因果関係)

時系列モデルである変数(y)が他の変数(x)に影響を及ぼす、あるいは逆に影響しないという検定は、他の条件を一定としてyの過去の値がxの変動について説明力を持つか、あるいは全く説明力を持たないかで行なわれる。この考え方をランジェーの意味での因果性(Granger causality)あるいは因果関係という⁶⁾。

具体的には9.6)式で、帰無仮説を $b_{11} = b_{12} = 0$ として、その帰無仮説が棄却できなければxはGrangerの意味でyと因果関係が無いという。逆に $b_{11} = b_{12} = 0$ の帰無仮説が棄却されれば、xはGrangerの意味でyと因果関係があるという。同様に9.7)式で $c_{21} = c_{22} = 0$ の帰無仮説が棄却できなければ、yはGrangerの意味でxと因果関係は無い。 $c_{21} = c_{22} = 0$ の帰無仮説が棄却されればyはGrangerの意味でxと因果関係がある。

この検定は通常のF検定で行うことができる。制約のない9.6)式を推計し、その残差平方和をRSS₀とする。 $b_{11} = b_{12} = 0$ の制約をつけた

$$y_t = a_1 + c_{11}y_{t-1} + c_{12}y_{t-2} + e_{1t} \quad (9.6)'$$

の残差平方和をRSS_rとする。nを標本数とし、F検定は

$$F = \frac{(RSS_r - RSS_0) / 2}{RSS_0 / (n - 5)}$$

で行われる(このケースでは制約の数は2個、パラメータは5個である)。

この検定をGranger testという(9.7)式につ

6) TodaとPhillipsが示したように、変数が単位根をもちかつ変数の間に共積分関係が存在すると、因果関係の有無を検定することは複雑となる。ここで単位根と共積分関係が存在しないと仮定する。

7) 三つの変数(x, y, z)の場合には、因果関係を調べる方法が二つある。例えば、yはGrangerの意味でxと因果関係を有するか否かについて、本文のように

$$x_t = a_1 + b_{11}x_{t-1} + \dots + b_{1p}x_{t-p} + c_{11}y_{t-1} + \dots + c_{1p}y_{t-p} + e_{1t} \quad (s1)$$

において $c_{11} = \dots = c_{1p} = 0$ を検定にすれば良い。しかしz_tを考慮したモデル

$$x_t = a_1 + b_{11}x_{t-1} + \dots + b_{1p}x_{t-p} + c_{11}y_{t-1} + \dots + c_{1p}y_{t-p} + d_{11}z_{t-1} + \dots + d_{1p}z_{t-p} + e_{1t} \quad (s2)$$

において $c_{11} = \dots = c_{1p} = 0$ を検定することもできる。この二つの検定結果は異なる場合があることに注意が必要である。EVIEWWSはs1)に関してGrangerテストを実施する。

いても全く同様に行うことができる)。3変数以上の場合も同様の手順を繰り返せばよい)。

なおGrangerの因果関係(因果性)は、いわばyのラグ(xのラグ)がx(y)を説明するために役に立つことを意味している。我々が日常用語で使う因果関係とは異なる。その混同を避けるために必ず「Grangerの意味で」という形容詞がつけられる。

(EviewsでのGrangerテスト)

これをEviewsで見てみよう。表9.1の結果が表示されているとする。

Procs/Make Endogenous Group
を選択する。次に

View/Granger Causality
を選択する。Lags to includeと聞いてくるので、十分長いラグを入れる。ここでは12期を選択した。結果は表9.2に掲げるとおりである。

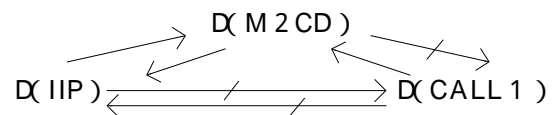
三つのペアについてGrangerテストの結果が表示されている。

まずD(IIP)はD(M2CD)にGrangerの意味での因果性は無いという帰無仮説は5%水準で棄却されている(F値は2.21、p値は0.01)。またD(M2CD)はD(IIP)にGrangerの意味での因果性は無いという帰無仮説も1%水準で棄却されている(F値は2.68、p値は0.002)。

D(CALL1)はD(M2CD)にGrangerの意味での因果性は無いという帰無仮説は5%水準で棄却されている(F値は1.95、p値は0.03)。逆にD(M2CD)がD(CALL1)とGrangerの意味での因果性は無いという帰無仮説は棄却されない(F値は1.30、p値は0.22)。

D(CALL1)はD(IIP)とGrangerの意味での因果性は無いという帰無仮説は棄却されない(F値は0.74、p値は0.71)。またD(IIP)がD(CALL1)とGrangerの意味での因果性は無いという帰無仮説も棄却されない(F値は1.03、p値は0.42)。

このGrangerの意味での因果関係をまとめて図示すると次のようである。



このGrangerテストの結果はVARモデルの分析では、通例報告される。

Grangerテストと類似するものにSimsテストがある。これは先行ラグ(lead lag)の影響を見るものである。リードラグを加えた次のモデルを考えF検定するものである。xがyに対しGrangerの意味で因果関係があれば、xのリードはyに影響しているであろう。例えば以下のものである。

$$y_t = c_2x_{t+2} + c_1x_{t+1} + c_0x_t + c_{-1}x_{t-1} + c_{-2}x_{t-2} + e_{3t}$$

表9.2

Pairwise Granger Causality Tests			
Date : 11/19/99 Time : 11 : 11			
Lags : 12			
Null Hypothesis:	Obs	F Statistic	Probability
D(IIP) does not Granger Cause D(M2CD)	288	2.21368	0.01152
D(M2CD) does not Granger Cause D(IIP)		2.67909	0.00204
D(CALL1) does not Granger Cause D(M2CD)	288	1.94858	0.02927
D(M2CD) does not Granger Cause D(CALL1)		1.30006	0.21820
D(CALL1) does not Granger Cause D(IIP)	288	0.74424	0.70733
D(IIP) does not Granger Cause D(CALL1)		1.02994	0.42147

9.11)

$$H_0: c_1 = c_2 = 0$$

帰無仮説が棄却されなければ、xはGrangerの意味でyと因果性はない。

(ブロック外生性)

VARモデルは、データからモデルを語らせるという側面があり、どの変数がモデルに含まれるべきかは必ずしも先験的に明らかではない。通貨、生産、金利という3変量のモデルでさえ推計すべきパラメータは111個であった。一変数加えるだけでパラメータは大きく増える。逆に必要な変数を落とせば、その推計結果に一致性もない。このようにある変数がモデルに含まれるべきか、除かれるべきかは推定に大きな影響を及ぼす。

たとえば3変量(x, y, z)のVARモデルでzがxに対してもyに対してもGrangerの意味での因果性を持たなければ、zをモデルから除外することが考えられる。逆にGrangerの意味での因果性を持っていればモデルに加えるべきこととなる。この検定をブロックの外生性(block exogeneity)をみる因果関係テスト(block causality test, block exogeneity test)という。たとえば

$$y_t = a_{11} + b_{11}y_{t-1} + c_{11}x_{t-1} + d_{11}z_{t-1} + e_{1t} \quad 9.12a)$$

$$x_t = a_{12} + b_{12}y_{t-1} + c_{12}x_{t-1} + d_{12}z_{t-1} + e_{12} \quad 9.12b)$$

$$z_t = a_{13} + b_{13}y_{t-1} + c_{13}x_{t-1} + d_{13}z_{t-1} + e_{13} \quad 9.12c)$$

に対して、 $d_{11} = d_{12} = 0$ の制約を加えた

$$y_t = a_{21} + b_{21}y_{t-1} + c_{21}x_{t-1} + e_{2t} \quad 9.13a)$$

$$x_t = a_{22} + b_{22}y_{t-1} + c_{22}x_{t-1} + e_{2t} \quad 9.13b)$$

を考える。制約のないVARモデル(9.12a)式、9.12b)式と9.12c)の三式)の誤差項の分散共分散の行列を Σ 、制約のあるモデル(9.13a)式、9.13b)式と9.12c)の三式)のそれを Σ^d とする。制約が有効であれば Σ と Σ^d に余り差はないはずである。制約が有効でなければ両者は乖離するであろう。この考え方を利用して、ブロックの外生性テストは次の尤度比検定で行われる。

$$T(\text{Log}|\Sigma^d| - \text{Log}|\Sigma|) \quad 9.14)$$

Tは用いられたサンプル数。9.14)式が自由度2の χ^2 分布に従う(除かれたのは9.12a)式、9.12b)式における z_{t-1} である。ラグがpであれば制約の数は2pであるから自由度は2pとなる)。

3 インパルス反応関数

VARモデルで各変数間の影響を分析する方法の一つに、ある式の誤差項に与えられた衝撃(イノベーション、innovation)がその変数や他の変数にどのように伝搬しているかを見る方法がある。

$$y_t = a_1 + b_{21}y_{t-1} + c_{21}x_{t-1} + e_{1t} \quad V(e_{1t}) = \sigma_y^2 \quad 9.13a)$$

$$x_t = a_2 + b_{22}y_{t-1} + c_{22}x_{t-1} + e_{2t} \quad V(e_{2t}) = \sigma_x^2 \quad 9.13b)$$

を再び考える。 $\text{Cov}(e_{1t}, e_{2t}) = 0$ と仮定する。 y_t, x_t の初期値 y_0, x_0 の値を0とする。イノベーションを与えない($e_{1t} = e_{2t} = 0 \quad t$)場合の y_t, x_t の値を y_t^0, x_t^0 と表記する。第一期の値は $y_1^0 = a_1, x_1^0 = a_2$ となり、次期には $y_2^0 = a_1 + b_{21}a_1 + c_{21}a_2, x_2^0 = a_2 + b_{22}a_1 + c_{22}a_2$ となる。その上で $e_{11} = \sigma_y, e_{21} = 0$ のイノベーションを与える。(または $e_{11} = 0, e_{21} = \sigma_x$ のイノベーションを与える)。なおイノベーションの大きさは通常1標準偏差とされている。イノベーションを与える場合の y_t, x_t の値を y_t^a, x_t^a と表記する。そうすると第1期の値は $y_1^a = a_1 + \sigma_y, x_1^a = a_2$ (または $x_1^a = a_2 + \sigma_x, y_1^a = a_1$)となる。次期には $y_2^a = a_1 + b_{21}(a_1 + \sigma_y) + c_{21}a_2, x_2^a = a_2 + b_{22}(a_1 + \sigma_y) + c_{22}a_2$ (または $x_2^a = a_2 + b_{22}a_1 + c_{22}(a_2 + \sigma_x), y_2^a = a_1 + b_{21}a_1 + c_{21}(a_2 + \sigma_x)$)となる。イノベーションを与えないケースと与えるケースの差を計算することで、イノベーション効果を見ることが出来る。その差を $y_t^d = y_t^a - y_t^0, x_t^d = x_t^a - x_t^0$ と表記すると、上記のケースでは $y_1^d = \sigma_y, x_1^d = 0$ (又は $x_1^d = \sigma_x, y_1^d = 0$)、 $y_2^d = b_{21}\sigma_y, x_2^d = c_{21}\sigma_y$ (又は $x_2^d = c_{22}\sigma_x, y_2^d = c_{21}\sigma_x$)となる。9.13a)式より

$$y_t^d = e_{1t}^d$$

$$y_t^d = b_{21}y_{t-1}^d + c_{21}x_{t-1}^d \quad t = 2, 3 \dots$$

9.13b) 式より

$$x_t^d = e_{2t}^d$$

$$x_t^d = b_{22}y_{t-1}^d + c_{22}x_{t-1}^d \quad t = 2, 3 \dots$$

となる。これをインパルス反応関数 (impulse response functions) という。

9.13a) 式と9.13b) 式の誤差項が相関を持つ場合がある (この場合でも方程式間で説明変数が共通であるから、SUREをOLSで推定して良いケースであることに変わりはない)。すなわち $Cov(e_{1t}, e_{2t}) = k \neq 0$ というケースである。

このとき直交イノベーション (orthogonal innovations) を考える。

k / σ_y^2 を9.13a) 式の両辺に乘じ、9.13b) 式に代入し整理すると、

$$x_t = \left(a_2 - \frac{k}{\sigma_y^2} a_1 \right) + \frac{k}{\sigma_y^2} y_t + \left(b_{22} - b_2 \frac{k}{\sigma_y^2} \right) y_{t-1}$$

$$+ \left(c_{22} - \frac{k}{\sigma_y^2} c_{21} \right) x_{t-1} + u_t \quad 9.14)$$

ここで $u_t = e_{2t} - (k / \sigma_y^2) e_{1t}$ である。

u_t と e_{2t} の共分散を考えると

$$Cov(e_{1t}, u_t) = E(e_{1t}e_{2t}) - (k / \sigma_y^2) E(e_{1t}^2)$$

$$= k - k = 0 \quad 9.15)$$

となるので、 u_t と e_{1t} は相関しないことが分かる。イノベーション分析を行う際9.13a)式と9.14)式を利用する。イノベーション (u_t と e_{1t}) が直交しているため、9.13a) 式と変換された9.14) 式にイノベーションを与えて、その影響を見ることができる。

ただしその影響の程度はVARモデルの変数の並べ方に依存し、変数の並べ方によって結果が異なってくる。上記のケース (9.13a) 式と9.13b) 式) でいえば9.13a) 式と9.14) 式の代わりに、9.13b) 式と k / σ_y^2 を9.13b) 式の両辺に乘じ、9.13a) 式に代入したものを利用することもあり得る。

図9.2 インパルス反応関数の入力例

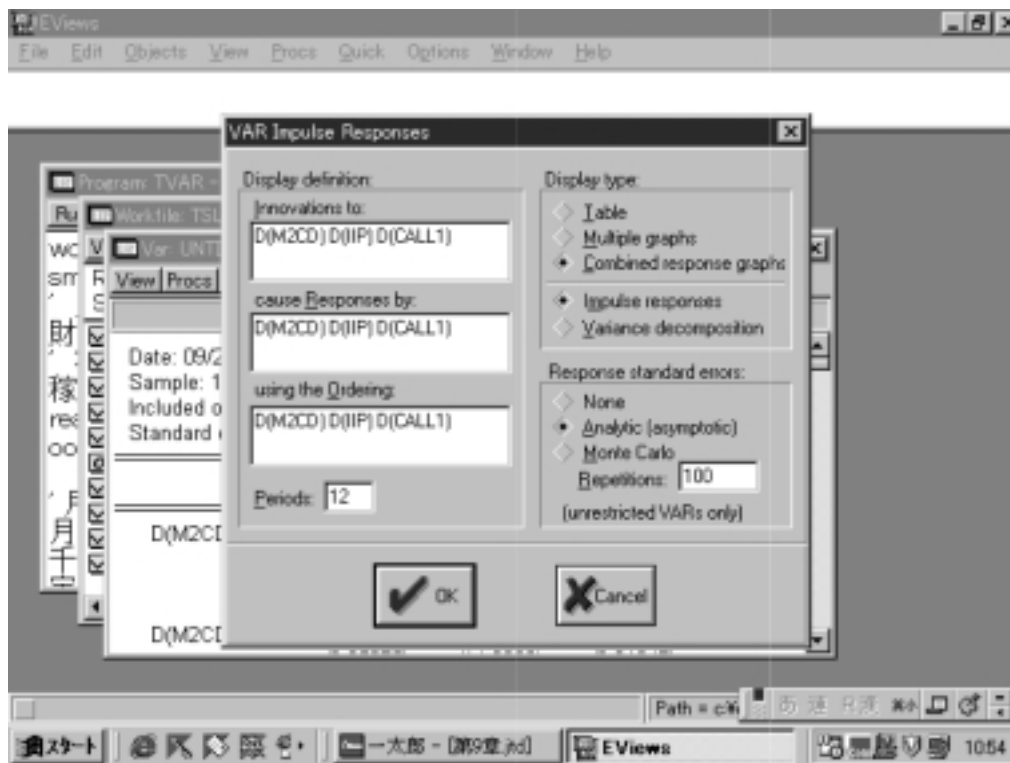


図9.3 インパルス反応分析

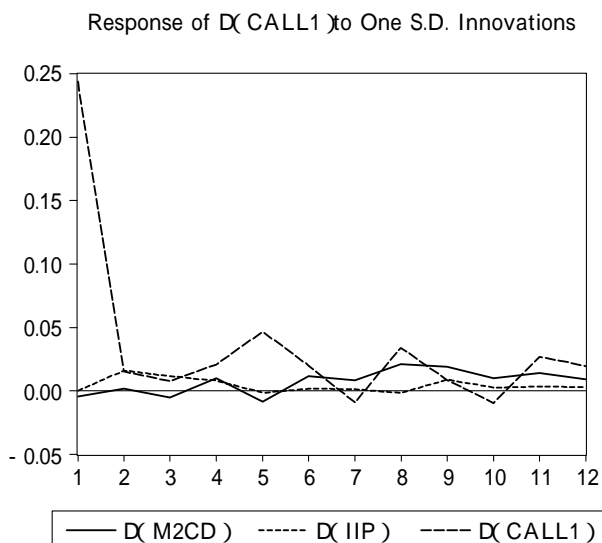
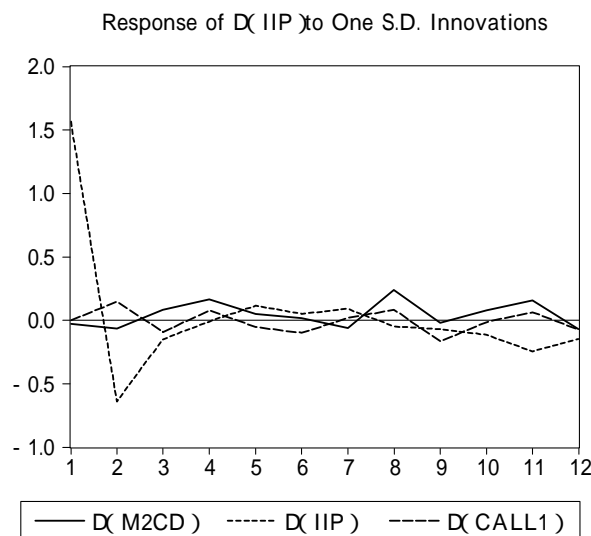
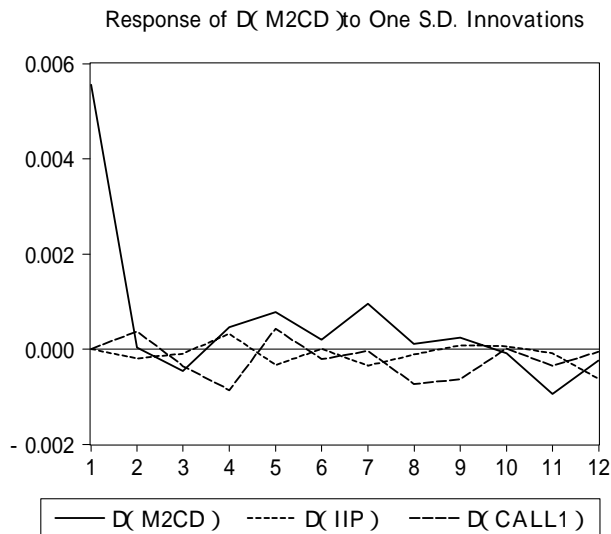


表9.3 Impulse Response to One S.D. Innovations

Response of Δ (M2CD):			
Period	Δ (M2CD)	Δ (IIP)	Δ (CALL1)
1	0.005553 (0.00023)	0.000000 (0.00000)	0.000000 (0.00000)
2	3.38E-05 (0.00025)	-0.000200 (0.00027)	0.000371 (0.00032)
12	-0.000242 (0.00023)	-0.000629 (0.00018)	-5.65E-05 (0.00033)
Response of Δ (IIP):			
Period	Δ (M2CD)	Δ (IIP)	Δ (CALL1)
1	-0.029781 (0.09240)	1.567862 (0.06533)	0.000000 (0.00000)
2	-0.063523 (0.08690)	-0.639071 (0.08041)	0.145348 (0.09129)
12	-0.070267 (0.06859)	-0.146730 (0.04768)	-0.077365 (0.09929)
Response of Δ (CALL1):			
Period	Δ (M2CD)	Δ (IIP)	Δ (CALL1)
1	-0.004479 (0.01438)	0.000149 (0.01438)	0.243999 (0.01017)
2	0.001743 (0.01218)	0.015997 (0.01179)	0.015224 (0.01419)
12	0.009371 (0.01036)	0.002900 (0.00882)	0.019646 (0.01397)
Ordering: Δ (M2CD) Δ (IIP) Δ (CALL1)			

変数がm個であれば、その組合せはm!となる(3変数であれば6組、4変数であれば24組ある)。そこで重要と考えられる順に変数を並べ、影響の程度をみることが多い。ただしこの方法はデータをしてモデルを語らしめるといふ時系列モデルの趣旨から外れるところがある⁸⁾。しかし直感的(視覚的)に影響の程度が把握しやすいことから、インパルス反応関数もVARモデルでは必ずと言って良いほど報告される。

(Eviewsでのインパルス反応関数)

Eviewsで反応関数を見てみよう。表9.1の結果が表示されているとする。

View/Impulse Var Decomposition

を選択する。VAR Impulse Responsesの画面がでてくる(図9.2参照)。

Innovation toの欄に影響を与える変数名を入れる。

cause Responses byの欄にイノベーションを与えられる変数名を入れる。この二つの欄の順番は推計や予測には関係がない。印刷(画面表示)の順番がこれに従うだけである。

using the Orderingの欄にショックを与える順番を入れる。この順番は上述したように、インパルスの反応に実際の影響を与えるので注意を要する。単に通りの順番の組合せではなく(変数がm個の場合m!通りあるが)、少なくとも複数の順番の組合せを実施して、大きな差がないかどうか、仮に大きな差があればその大きな差のあるケースも併せて報告することが望ましい。

Impulse responsesとVariance decompositionとあるうち前者を選択する(後者は分散分解である)。

Display typeにインパルス反応の値がほしい場合はTableを、グラフがほしい場合はMultiple graphsまたはCombined response graphsを選択する。

Periodsの欄にラグの長さを入れる。ここでは12期を選択している。

結果は図9.3と表9.3(抜粋)に掲げるとおりである(なおこのケースではショックを与える順番によって結果はほとんど変わることがなかった)。

D(M2CD)のインパルス反応を見ると、自己に対しては1期に大きいが、2期にほぼ0となり、5、7期にやや大きくなっている。D(IIP)に対しては4期でやや効果が見られる。D(CALL1)に対しては、1期にはほとんど影響がないが2期にはややプラスに影響している。

表9.3の最下欄には変数の順番が(Ordering)が示されている。順番によって結果が異なるので、論文ではこの順番も報告する必要がある。

4 予測誤差の分散分解

インパルス反応関数では、ある変数に与えられたショックがどのように影響しているかを見た。その変動にどの変数が、どれだけ寄与しているかを知ることができれば、変数間の相互関係をより詳しく見ることもできるであろう。その目的で提案されたのが予測誤差の分散分解(forecast error variance decomposition)である。9.13a)、9.13b)式より次のようなMAプロセスを考える⁹⁾。

$$x_t = u + \sum_{i=1}^n (\alpha_i e_{1t-i} + \beta_i e_{2t-i}) \quad (9.16)^{10)}$$

α_i は係数である。9.13a)式と9.13b)式より $V(e_{1t}) = \sigma_y^2$ と $V(e_{2t}) = \sigma_x^2$ であることが分かる。

n期先の値は $x_{t+n} = u + \sum_{i=1}^n (\alpha_i e_{1t+n-i} + \beta_i e_{2t+n-i})$ である。t期において形成される x_{t+n} の期待値を $E(x_{t+n})$ と書くとする、 $E(x_{t+1}) = u + \sum_{i=1}^n (\alpha_i e_{1t+n-i} + \beta_i e_{2t+n-i})$ となるのでn期先の予測誤差は

$$x_{t+n} - E(x_{t+n}) = \sum_{i=0}^{n-1} (\alpha_i e_{1t+n-i} + \beta_i e_{2t+n-i}) \quad (9.17)$$

である。9.17)式を書き直すと

8) この順番の並べ方で結果が大きく変わることがあるということは、VARモデルに対する基本的な批判の根拠となっている。またそのために政策的な意味が取りにくくなる。それにもかかわらずVARモデルは、Grangerの意味での因果関係で変数間の影響の伝搬経路(先後関係)が把握できること、視覚的に影響の程度を捉えられるところから、経済分析では重要な地位を占めるに至っている。

9) 9.16)式を得るために9.13b)式にある y_{t-1} の代わりに9.13a)式より $y_{t-1} = a_1 + b_{21}y_{t-2} + c_{21}x_{t-2} + e_{1t-1}$ を、 x_{t-1} の代わりに9.13b)式より $x_{t-1} = a_2 + b_{22}y_{t-2} + c_{22}x_{t-2} + e_{2t-1}$ を代入すると

$$x_t = a_2 + b_{22}(a_1 + b_{21}y_{t-2} + c_{21}x_{t-2} + e_{1t-1}) + c_{22}(a_2 + b_{22}y_{t-2} + c_{22}x_{t-2} + e_{2t-1}) + e_{2t} \quad (s3)$$

が得られる。この代入を繰り返す。s3)式より9.16)式の係数とs3)式の係数の関係は $\alpha_0 = 0$ 、 $\alpha_1 = b_{22}$ 、 $\alpha_2 = 1$ 、 $\alpha_3 = c_{22}$ となる。

10) インパルス反応を計算するためにも、このモデルは便利である。

$$\begin{aligned}
 X_{t+n} - E(X_{t+n}) = & \quad 1_0 e_{1t+n} + 1_1 e_{1t+n-1} + \dots + 1_{n-1} e_{1t+1} \\
 & + 2_0 e_{2t+n} + 2_1 e_{2t+n-1} + \dots + 2_{n-1} e_{2t+1}
 \end{aligned}
 \tag{9.18}$$

となる。Cov(e_{1t} , e_{2t}) = 0を仮定し、n期先のxの予測誤差の分散を x_n^2 と表すと

$$\begin{aligned}
 x_n^2 = & \quad x^2 (1_0^2 + 1_1^2 + \dots + 1_{n-1}^2) \\
 & + y^2 (2_0^2 + 2_1^2 + \dots + 2_{n-1}^2)
 \end{aligned}
 \tag{9.19a}$$

となる。 1_{ij}^2 は非負である。この予測誤差を e_{2t} と e_{22} のショックに分解し

$$RVC_{x_n} = \frac{x^2 (1_0^2 + 1_1^2 + \dots + 1_{n-1}^2)}{x_n^2}
 \tag{9.19b}$$

$$RVC_{y_n} = \frac{y^2 (2_0^2 + 2_1^2 + \dots + 2_{n-1}^2)}{x_n^2}
 \tag{9.19c}$$

を得ることができる。9.19b)式はx自身のショックの割合を示し、9.19c)式はyのショックの割合を示している。この割合を相対的分散寄与率 (Relative variance contribution, RVC) ということがある。

短期には、自己のショックでRVCの大半は説明されるが、ラグが長くなるに従って他の変数のショックが高くなることが多い。そこで変数間の変動の影響を把握するためにはラグの長さが十分長いことが必要である¹¹⁾。

(Eviewsでの予測例)

Eviewsで予測誤差の分散分解を行うと手順は、インパルス反応関数と同じである。VAR推定結果が表示されているとする。

Impulse Responses と Variance Decomposition あるうち後者を選択する。

ここではラグの長さを36期としている。

結果は図9.4、表9.4に示すとおりである。

D(IIP)についてみると、第2期では99.1%が自己のショックによるが、36期では89.3%となり、M2CDが8.6%、CALL1が2.1%となっている。

(VECM)

なおVARモデルの延長に誤差修正項を加えた

図9.4 Variance Decomposition of D(M2CD)

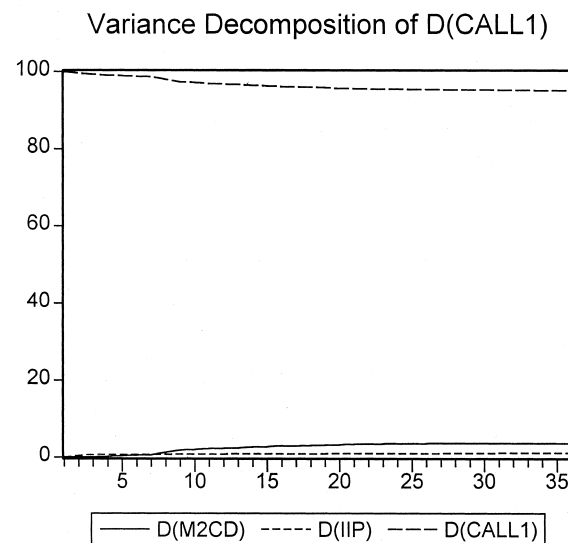
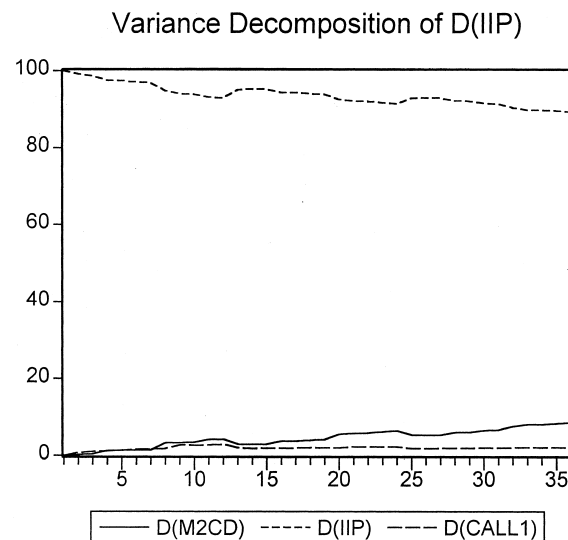
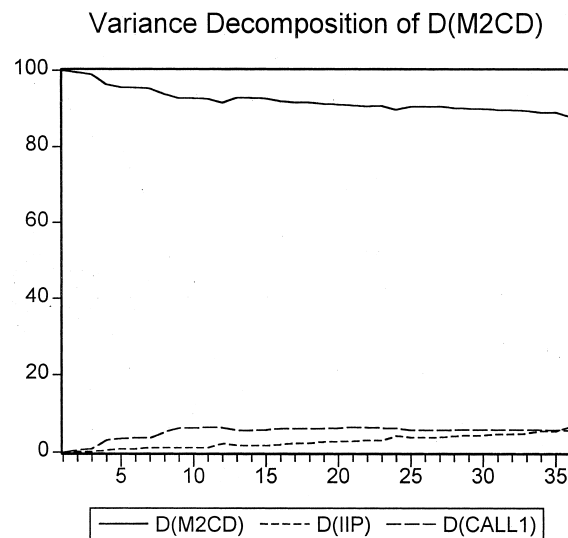


表9.4 Variance Decomposition

Variance Decomposition of D (M2CD):				
Period	S.E.	D(M2CD)	D(IIP)	D(CALL1)
1	0.005553	100.0000	0.000000	0.000000
2	0.005569	99.42702	0.129313	0.443665
36	0.007904	87.87271	6.528113	5.599178
Variance Decomposition of D (IIP):				
Period	S.E.	D(M2CD)	D(IIP)	D(CALL1)
1	1.568145	0.036066	99.96393	0.000000
2	1.700779	0.170156	99.09951	0.730332
36	2.913066	8.598201	89.32135	2.080449
Variance Decomposition of D (CALL1):				
Period	S.E.	D(M2CD)	D(IIP)	D(CALL1)
1	0.244040	0.033680	3.74E-05	99.96628
2	0.245043	0.038465	0.426194	99.53534
36	0.267667	3.673532	1.124942	95.20153
Ordering: D (M2CD) D (IIP) D (CALL1)				

Vector Error Correction モデル (VECM) と同時点間の誤差項の相関を考慮する構造VARモデル (Structural VAR) がある。行列の知識が不可欠なので本シリーズの程度を超える¹²⁾。そこでVECMについては第8章のECMとの関係で、直感的な意味だけを説明しておきたい。

モデルが3変数以上の場合、変数の数をm個とすると、共和分関係はm - 1個まで存在しうる。Engle Grangerの共和分検定では、共和分の数に1個と限定した。しかし共和分の数には未知であり、0個 (共和分は存在しないケース) から最大m - 1個までありうる。JohansenとJuseliusはこの共和分の個数 (共和分階数、rank of cointegration) の検定と共和分ベクトルの未知パラメータ (誤差修正項の係数) を同時に求める最尤法を開発した。ECMの場合と同じくVARモデルについても、変

数間に共和分が存在するならば、その変数間の相互依存関係を考慮することが望ましい。たとえば、本章の例でいえば、M2CD, IIP, CALL1はいずれも単位根を有するI(1)変数である。この3変数について

$$u_t = M2CD - a_0 - a_1 IIP - a_2 CALL1 \quad u_t \sim K(0) \quad (9.20)$$

となる関係が存在すれば、(9.20)式の関係性をVARモデルに取り込むことが望ましい (無論(9.20)式を充たすような共和分関係がなければ、今まで述べてきたVARモデルで足りる)。均衡からの乖離を是正する長期均衡関係を考慮することができるからである。JohansenとJuseliusは共和分の個数を検定する尤度比検定を開発した (検定は最大固有値検定量とトレース検定量とがある)。仮に共和分が2個存在するとされたならば

¹¹⁾ 上記の説明ではCov(e_{1t}, e_{2t}) = 0と仮定したが、Cov(e_{1t}, e_{2t}) ≠ 0の場合には予測誤差の分散分解もやはり変数の並べ方によって代わってくる。その点には留意が必要である。

¹²⁾ VECMに興味のある読者は森棟 [1999] の第11章、Johnston and DiNardo [1997] のCh 9 - 3を参照されたい。構造VARについてはEnders [1995] のCh 5 - 10 ~ 12を参照されたい。

$$\begin{aligned}
M2CD_t &= a_1 + b_{11} M2CD_{t-1} + \dots + b_{1p} \\
&\quad M2CD_{t-p} + c_{11} IIP_{t-1} + \dots + c_{1p} \\
IIP_{t-p} + d_{11} CALL1_{t-1} + \dots + d_{1p} \\
&\quad CALL1_{t-p} + f_{11} \hat{u}_{1t-1} + f_{12} \hat{u}_{2t-1} + e_{1t} \\
IIP_t &= a_2 + b_{21} M2CD_{t-1} + \dots + b_{2p} \\
&\quad M2CD_{t-p} + c_{21} IIP_{t-1} + \dots + c_{2p} \\
IIP_{t-p} + d_{21} CALL1_{t-1} + \dots + d_{2p} \\
&\quad CALL1_{t-p} + f_{21} \hat{u}_{1t-1} + f_{22} \hat{u}_{2t-1} + e_{2t} \\
CALL1_t &= a_3 + b_{31} M2CD_{t-1} + \dots + b_{3p} \\
&\quad M2CD_{t-p} + c_{31} IIP_{t-1} + \dots + c_{3p} \\
IIP_{t-p} + d_{31} CALL1_{t-1} + \dots + d_{3p} \\
&\quad CALL1_{t-p} + f_{31} \hat{u}_{1t-1} + f_{32} \hat{u}_{2t-1} + e_{3t}
\end{aligned}$$

を推計するものである¹³⁾。

ここで f_{11} , f_{12} 等は誤差修正項 (\hat{u}_{1t-1} と \hat{u}_{2t-1})
にかかると、送付ベクトル (loading vector)

ということは前章で見たとおりである。

VECMにおけるインパルス反応関数や予測誤差の分散分解は、上述のVARモデルと同様に行えばよい。

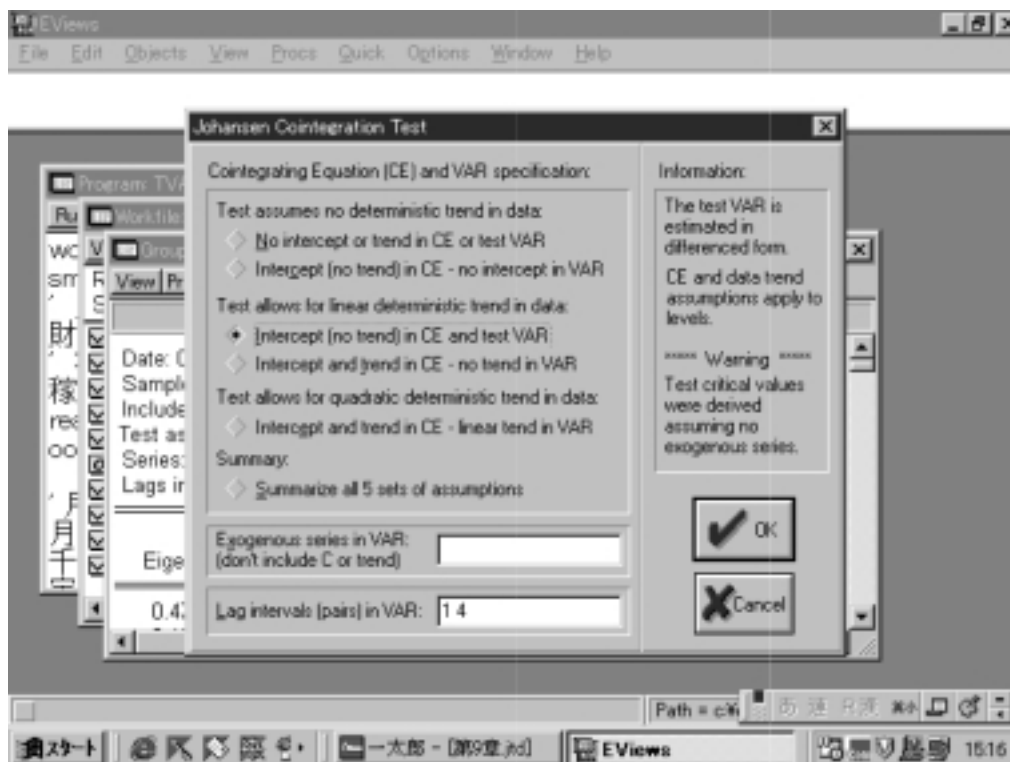
5 ARIMA

第5章でAR、MA、ARMAを取り上げた。第8章でデータの定常性、単位根を取り上げた。ここではこれらを含めたBox、Jenkinsによって提唱されたARIMA (Autoregressive Integrated Moving Average) について簡単に説明する。ARIMAモデルは短期の予測によく用いられるものである。

ある変数が $I(d)$, $AR(p)$, $MA(q)$ で表されたとする (p, d, q は非負の整数である)。すな

¹³⁾ 図1のVector Error Correctionがこれに当たる。共和分個数の検定には、`coint(c, 4) M2CD IIP CALL 1`とすることで行うことができる。cointが共和分検定のコマンドである。a, b, c, d, e, sのオプションがある。これはデータシリーズにトレンドや定数項を含むか否かで選択する (cはデータに線形トレンドを含み、定数項は共和分を含むが、共和分にベクトルには線形トレンドを含まないケースである)。共和分検定のための画面とVECMの入力画面を下に掲げておく。興味のある読者は試されると良い。

注図1 共和分検定の画面



わちd次で和分され、p次の自己回帰とq次の移動平均過程を持つとする。

$$\begin{aligned} \Delta^d y_t = & a + b_1 \Delta y_{t-1} + b_2 \Delta y_{t-2} + \dots + b_p \Delta y_{t-p} \\ & + e_t + c_1 e_{t-1} + c_2 e_{t-2} + \dots + c_q e_{t-q} \end{aligned} \quad (9.21)$$

これを次数 (p, d, q) のARIMAモデルとい、ARIMA(p, d, q) と表記する。

d = 0 であれば、ARIMA(p, 0, q) である(すなわちARMA(p, q))、q = 0 であればARIMA(p, d, 0)、p, d, qが各1であればARIMA(1, 1, 1) である。

BoxとJenkinsはARIMAモデルの分析を、同定

(identification) 推定(estimation) 診断(diagnostic checking) 予測(forecasting) の手順で行っている。

(同定)

まず、正しいp, d, qの次数の選択を行う。データは定常であることを要するから単位根の検定を行い、階差定常あるいはトレンド定常となるようなdを選択する。これは第8章で説明したプロセスである。

定常化されたデータについて標本自己相関(auto correlation、標本コレログラム、correlo-

14) コレログラムは標本共分散を標本分散で基準化したものであった。第5章で説明したように

$$d_j = \frac{(y_t - \bar{y})(y_{t-j} - \bar{y})}{(y_t - \bar{y})^2}$$

系列相関がないという帰無仮説の下で、 $t = d_k / \sqrt{1/n}$ が標準正規分布に従う。

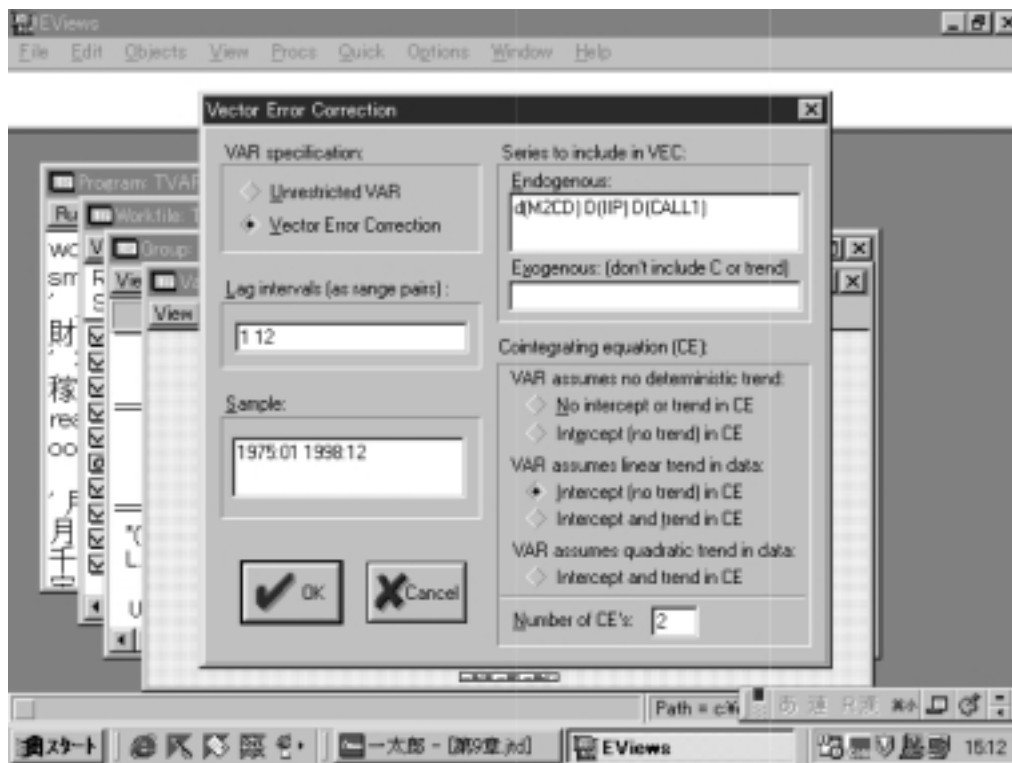
標本偏自己相関 (d_{kk})

$$\begin{aligned} d_{kk} &= d_j \quad k = 1 \\ &= \frac{d_k - \sum_{j=1}^{k-1} d_{k-j} d_j}{1 - \sum_{j=1}^{k-1} d_j^2} \quad k > 1 \end{aligned}$$

ただし、 $d_{kj} = d_{k-1, j-1} = d_{k-1, k-j}$ 、 $j = 1, 2, \dots, k-1$ である。

系列相関がないという帰無仮説の下で $t = d_{kk} / \sqrt{1/n}$ が標準正規分布に従う。

注図2 VECMの入力例



gram、acf)と標本偏自己相関(partial autocorrelation, pacf)によりAR, MAの次数を判断する¹⁴⁾。これは第5章で述べたARMAモデルのプロセスである。

AR(p)であれば、pacfはラグp次までゼロではないが、p+1次以降は0となる。acfは0ではないが、ラグが長くなると共に小さくなる。

MA(q)であれば、acfはq次までは0ではないが、q+1次以降は0となる。pacfは0ではないがラグが長くなると小さくなる。

ARMA(p, q)であればacfはq次までは0ではない、pacfもp次までは0ではない。いずれもラグが長くなると共に小さくなる。

しかし、第5章で説明したようにAR, MA, ARMAの選択は実際には必ずしも容易でない場合がある。そのような場合は、acf, pacfによらずAIC, SBCの情報量基準により、次数を選択することが考えられる。

Box = Jenkinsは次数の選択に当たり「けちの原理」(principle of parsimony)を提唱し、できるだけ推定すべきパラメータを少なくするためにp, qの次数を低く選ぶ方が好ましいとしている。そのため実際にはp, qの次数として1, 2が選ばれることが多い¹⁵⁾。

また分散不均一の問題を回避するために、対数を取ったりBox-Cox変換が行われることがある。 y_t について分散不均一が疑われる場合、新たに $Ly_t = \text{Log}(y_t)$ と変数を定義するか、Box-Cox変換

$$y_t^* = (y_t - 1) / c \quad y_t > 0 \quad 0 < c < 1 \quad (9.22)$$

を行うものである。ただこの(9.22)式で変換した場合、 y_t^* の経済的意味をとることは容易ではない。むしろ後述するARCHモデルによる方が望ましい。

(推定)

モデルが定式化されれば、ARIMAモデルの推計を行う。ARMAモデルの推計法は第5章で示した

$$y_t = a + by_{t-1} + e_t, \quad e_t = u_t + cu_{t-1}, \quad V(u_t) = \sigma^2, \quad t = 1, \dots, n$$

において

$$\text{Log} L = -(n/2) \text{Log}(\sigma^2) - (1/2\sigma^2) \sum_{t=1}^n u_t^2 + 1/2 \text{Log} |w| \quad (9.23)$$

$$\text{Log} |w| = \text{Log}[(1 - c^{2n+2}) / (1 - c^2)]$$

で求められる¹⁶⁾。

(診断)

推定結果について、残差が系列相関を有するか否か、分散不均一を有するか否かの検定を行う。残差がホワイトノイズであれば、同定は正しく推計は成功したということになる(得られたパラメータが経済的に意味を有するということは当然の前提である)。

系列相関についてはBreush-GodfreyのLMテストやBox-PierceのQテスト、Ljung-Boxテストで行う。自己ラグを含むARIMAや複雑なARMAプロセスではLMテストによる。分散不均一についてはGoldfeld-QuandtテストやBreush-Paganテスト、Whiteテストを行う(第5, 6章参照)。

もう一つの重要な診断は構造変化の有無である。これはChowテストの応用である(第3章参照)。

¹⁵⁾ 「けちの原理」については疑問無しとしない。系列相関の問題を誤れば、その推定量は有効性を持たないだけでなく、ラグ付き内生変数を含む場合には一貫性さえないことを第5章で示した。データが定常であるということは、

$$\begin{aligned} E(y_t) &= \mu \quad \text{for all } t \\ V(y_t) &= \sigma^2 \quad \text{for all } t \\ \text{Cov}(y_t, y_{t+k}) &= \rho_k \quad \text{for all } t \text{ and } k, k \geq 0 \end{aligned}$$

の条件を満たすことである。これらが1次や2次の次数で充たされるとは限らない。むしろ十分に長い次数を取って検定を行う方が妥当である。結果として低い次数が選ばれることがあるだけである。

¹⁶⁾ MA部分について反転可能である $|c| < 1$ が必要である。

9 21) 式のサンプルを2分割し、各サンプルを利用して9 21) 式を推定し、その残差平方和を RSS_{u1} , RSS_{u2} とする。全てのデータを利用した9 21) 式の推計の残差平方和を RSS_r とすると、

$$F = \frac{(RSS_r - RSS_{u1} - RSS_{u2}) / (1 + p + q)}{(RSS_{u1} + RSS_{u2}) / (T - 2(1 + p + q))} \quad 9 22)$$

が自由度 $(1 + p + q, T - 2(1 + p + q))$ のF分に従う (1 は定数項である)。

(予測)

ARIMAモデルの主要な目的の一つは短期の予測である。モデルの同定も正しく、診断も問題がなかったとしよう。その時は予測のプロセスに進む。

t 期において利用可能な情報を用いて形成される h 期先の $(t+h)$ 期の予測

$$y_{t+h/t} = E(y_{t+h}) \quad 9 23)$$

を考える。予測はできるだけ誤差が少ないことが望ましいであろうから、

$$E(y_{t+h/t} - y_{t+h})^2 \quad 9 24)$$

を最小にする (minimum mean square error forecast, MSE) ことを考える。

まず AR(1) モデルを考える。以下の推計結果が得られたとする。

$$y_t = a + by_{t-1} + e_t \quad 9 25)$$

e_t は IID $(0, \sigma_e^2)$ とする。1 期先の y は

$$y_{t+1} = a + by_t + e_{t+1} \quad 9 25)$$

である。 e_{t+1} は t + 1 期の誤差項である。将来の誤差項であるからその値は未知である。誤差項は互いに相関しないのであれば e_{t+1} を予測するために t 期までの情報は役に立たないので、将来の誤差項についてはその期待値 $E(e_{t+h}) = 0$ で置き換える。

9 25) より

$$y_{t+1/t} = a + by_t \quad 9 25)$$

となる。2 期先の予測は 1 期先の予測値と $E_t(e_{t+2}) = 0$ を用い

$$y_{t+2/t} = a + by_{t+1/t}$$

となる。以下繰り返すことにより

$$y_{t+h/t} = a + by_{t+h-1/t} \quad 9 26)$$

を得ることができる。

予測誤差は 1 期先で

$$w_{t+1} = y_{t+1} - y_{t+1/t} = a + by_t + e_{t+1} - (a + by_t) = e_{t+1}$$

となり、その分散は

$$V(w_{t+1}) = V(e_{t+1}) = \sigma_e^2$$

となる。2 期先については

$$\begin{aligned} w_{t+2} &= y_{t+2} - y_{t+2/t} = a + by_{t+1} + e_{t+2} - (a + by_{t+1/t}) \\ &= e_{t+2} + b(y_{t+1} - y_{t+1/t}) = e_{t+2} + bw_{t+1} = e_{t+2} + be_{t+1} \end{aligned}$$

$$V(w_{t+2}) = V(e_{t+2} + be_{t+1}) = \sigma_e^2 + b^2 \sigma_e^2 = \sigma_e^2 (1 + b^2)$$

以下繰り返すことにより

$$V(w_{t+h}) = \sigma_e^2 (1 + b^2 + b^4 + \dots + b^{2(h-1)}) \quad 9 27)$$

を得ることができる。

次に MA(1) を考える。

$$y_t = c + u_t + du_{t-1} \quad 9 28)$$

u_t は IID $(0, \sigma_u^2)$ とする。1 期先の y は次のようである。

$$y_{t+1} = c + u_{t+1} + du_t$$

y_{t+1} の MSE 予測と予測誤差は、

$$y_{t+1/t} = E(y_{t+1}) = c + du_t$$

$$w_{t+1} = y_{t+1} - y_{t+1/t} = u_{t+1}$$

である。その分散は次のとおりである。

$$V(w_{t+1}) = V(u_{t+1}) = \sigma_u^2$$

2 期先は

$$y_{t+2} = c + u_{t+2} + du_{t+1}$$

なので、MSE 予測は

$$y_{t+2/t} = E(y_{t+2}) = c$$

となる。

$$\begin{aligned} V(y_{t+2} - y_{t+2/t}) &= V(u_{t+2} + du_{t+1}) = \sigma_u^2 + d^2 \sigma_u^2 \\ &= (1 + d^2) \sigma_u^2 \end{aligned}$$

である。代入を繰り返すことにより h 期先は

$$y_{t+h/t} = E(y_{t+h}) = c \quad 9 29)$$

$$V(y_{t+h} - y_{t+h/t}) = (1 + d^2) \sigma_u^2 \quad 9 30)$$

である。

ARMA(1,1) についてみると

$y_t = a + by_{t-1} + u_t + du_{t-1}$ 9 31)
 u_t はIID(0 , σ_u^2) とする。1 期先の y_t 、MSE予測、
 予測誤差、予測誤差の分散は次の通りである。

$$y_{t+1} = a + by_t + u_{t+1} + du_t$$

$$y_{t+1/t} = E(a + by_t + u_{t+1} + du_t) = a + by_t + du_t$$

$$w_{t+1} = y_{t+1} - y_{t+1/t} = u_{t+1}$$

$$V(w_{t+1}) = \sigma_u^2$$

である。2 期先は

$$y_{t+2} = a + by_{t+1} + u_{t+2} + du_{t+1}$$

$$y_{t+2/t} = E(a + by_{t+1} + u_{t+2} + du_{t+1}) = a + bE(y_{t+1}) \\ = a + by_{t+1/t}$$

$$w_{t+2} = y_{t+2} - y_{t+2/t} = u_{t+2} + du_{t+1} + bw_{t+1}$$

$$V(w_{t+2}) = \sigma_u^2((b+d)^2 + 1)$$

となる。h期先も同様に求めることができる。

$$y_{t+h/t} = a + by_{t+h/1/t} \quad 9 32)$$

予測誤差の分散はh として

$$V(w_{t+h}) = \sigma_u^2 \left(\frac{1 + 2bd + d^2}{1 - b^2} \right) \quad 9 33)$$

で求められる。一般にARMA(p , q) のh期先の
 予測は

$$y_{t+h/t} = a + b_1 y_{t+h-1/t} + b_2 y_{t+h-2/t} + \dots + b_h y_{t/t} + \dots \\ + d_p y_{t+h-p} + d_h u_t + d_{h+1} u_{t-1} + \dots + d_q u_{t+h-q} \quad 9 34)$$

で与えられる。9 34) 式から明らかなように
 ARMAモデルでは、その予測は予測期間が長い
 と標本の平均に近づくことになる。このために長
 期の予想には使われずもっぱら短期の予想に用い
 られる。

ARIMAモデルの予測を考えよう。具体的に
 ARIMA(1 , 1 , 1) を仮定する。

$$y_t = a + b y_{t-1} + e_t + ce_{t-1} \quad 9 35)$$

1 期先の y_t は

$$y_{t+1} = a + b y_t + e_{t+1} + ce_t$$

なので、予測は

$$y_{t+1/t} = a + b y_t + ce_t$$

である。2 期先の y_t は

$$y_{t+2} = a + b y_{t+1} + e_{t+2} + ce_{t+1}$$

なので、予測は

$$y_{t+2/t} = a + b y_{t+1/t} \\ = a + b(a + b y_t + ce_t) \\ = (1 + b) a + b^2 y_t + bce_t$$

となる。

卸売物価指数 (WPI) についてARIMAモデル
 を推定してみる(期間は1975年1月~1998年6月) 。
 まず単位根の検定を行ったところ、定数項とトレ
 ンドが共に無い形でラグが1期の1回の階差をと
 る(DWPIと表記) とADF統計量は - 6 .705であっ
 た。階差をとることによりデータは定常となった。
 そのacfとpacfをみる。

View/Correlogramを選択しみてみる (ラグの
 長さは36期) 。結果は図9 .5 にあるとおりである。
 acfとpacfの形からAR(3) を選ぶ¹⁷⁾。ARIMA (3 ,
 1 , 0) を推定する。

equation ariwpi.LS D(WPI) c
 D(WP(- 1)) D(WP(- 2)) D(WP(- 3))
 結果は表9 .5 に示すとおりである¹⁸⁾¹⁹⁾。

¹⁷⁾ AIC基準やSBC基準で、複数の組合せについて検討することが望ましいが、ここでは省略する。

¹⁸⁾ 同じ推定結果を得るために次のプログラムも使用可能である。

equation ariwpi.LS D(WPS) c AR(1) AR(2) AR(3)

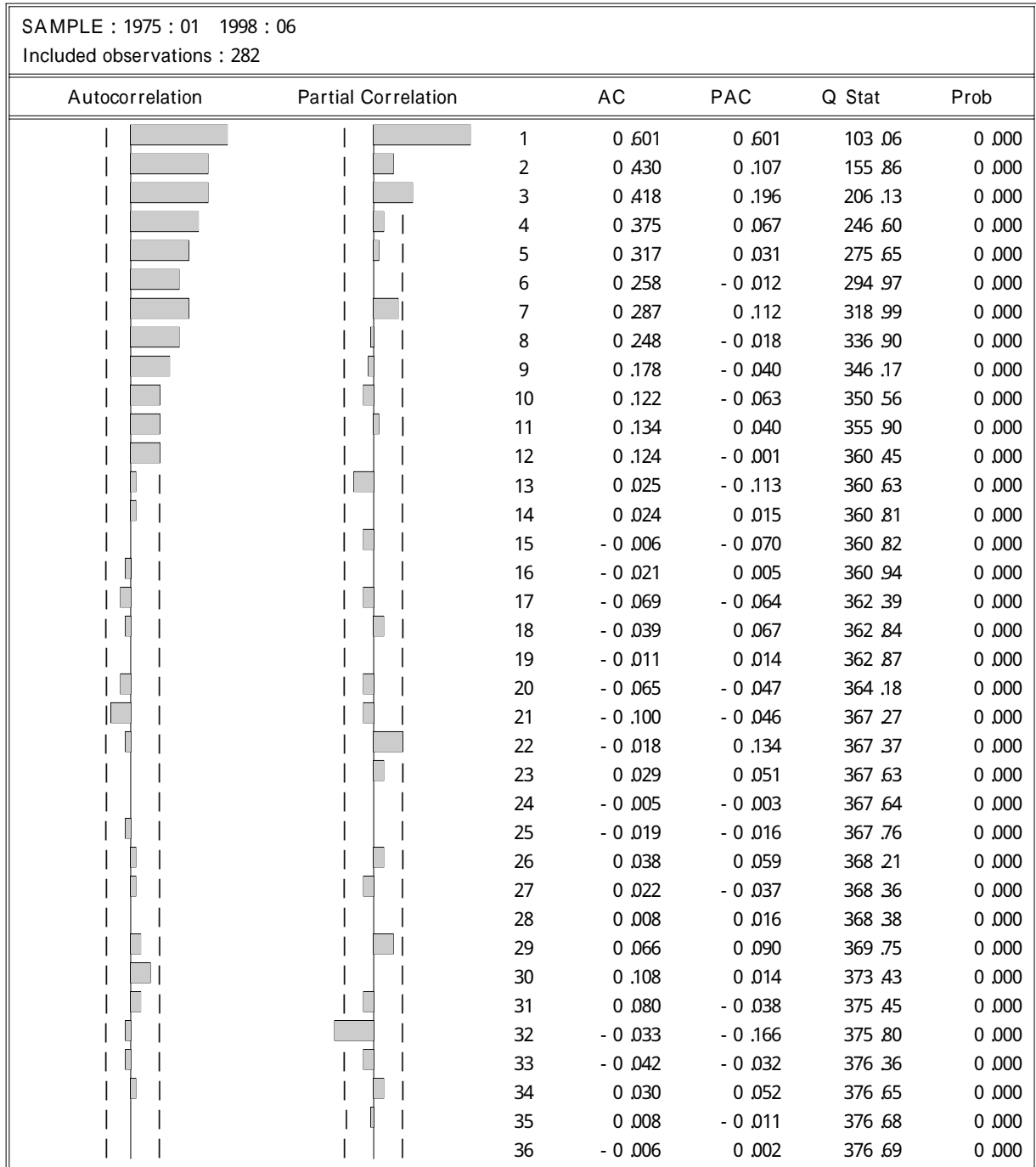
しかしこのようにARIMA(3 , 1 , 0) を推定するとEviewsではWhiteの分散不均一テストは計算不可能になる。ARIMA
 (3 , 1 , 0) の方がARIMA(3 , 1 , 1) より妥当ということを確認するために下記のプログラムでARIMA(3 , 1 , 1) モ
 デルを推定すると

equation ariwpi2.LS D(WPI) c D(WP(- 1)) D(WP(- 2)) D(WP(- 3)) MA(1)

MAの係数のp値は0 .060となる。有意水準5 %で検定するとMAの係数はゼロと判断される。

¹⁹⁾ D(WPI) は、series dwpi = wpi - wp(- 1) のようにも定義できる。ARIMAモデルなどでは予測の関係からD(WPI) とする
 方が便宜である。というのはEviewsではARIMAの予測を前者の場合WPIとして認識するからである。後者は階差をとった形で
 予測することになる。

図9.5 Correlogram of D(WPI)



残差の検定を行う。ラグのあるモデルなので Breusch GodfreyのLMテストを行う（第5章参照）。

View/Residual Tests/Serial Correlation LM Test を選択する。ラグの長さは2期とした。統計量は

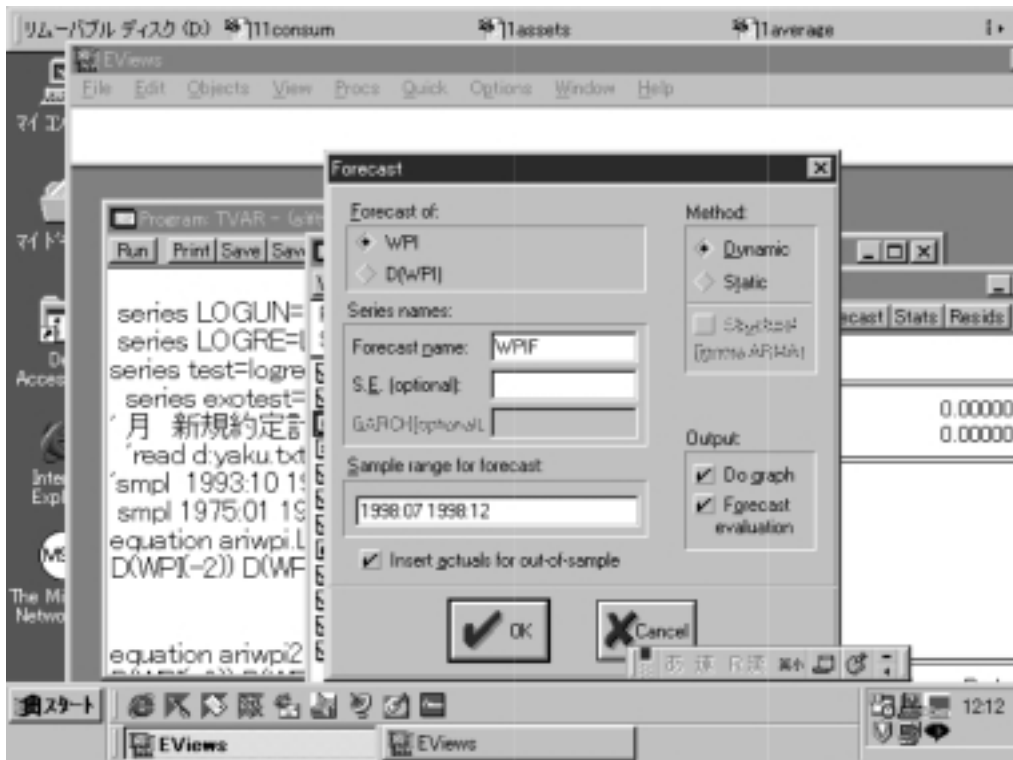
1.233 (p値=0.540) であった。系列相関の問題はないようである。次にWhiteの分散不均一テスト(クロス項有り)を行う。

View/Residual Tests/White Heteroscedasticity Test を選択する。統計量は79.216 (p値=0.000) で

表9.5 ARIMAモデルの推計例

Dependent Variable : D (WPI)				
Method : Least Squares				
Sample : 1975 : 01 1998 : 06				
Included observations : 282				
Variable	Coefficient	Std. Error	t Statistic	Prob.
C	0.009809	0.032299	0.303712	0.7616
D(WPI(-1))	0.518756	0.058877	8.810890	0.0000
D(WPI(-2))	-0.000531	0.066587	-0.007972	0.9936
D(WPI(-3))	0.194381	0.058859	3.302504	0.0011
R squared	0.392659	Mean dependent var		0.033333
Adjusted R squared	0.386105	S.D. dependent var		0.691089
S.E. of regression	0.541478	Akaike info criterion		1.625055
Sum squared resid	81.50921	Schwarz criterion		1.676713
Log likelihood	-225.1327	F statistic		59.91099
Durbin Watson stat	2.020983	Prob(F statistic)		0.000000

図9.6 Forecastの入力例

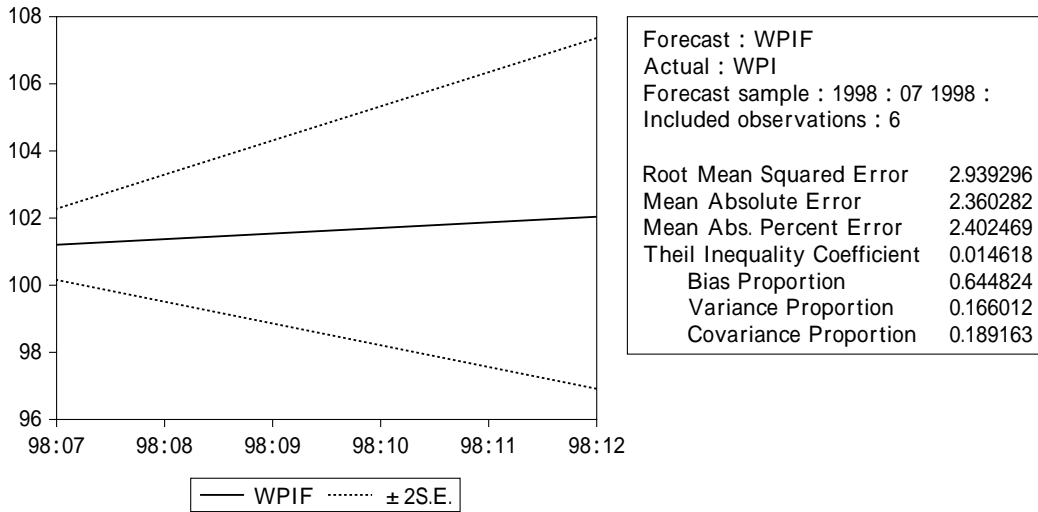


あった。分散は不均一である（とりあえずこの問題は無視する）。

次に予測に移ろう。ツールバーのForecastを選択する（図9.6参照）。

予測期間を聞いてくる（ここでは1998年7月から12月としている）。予測をWPIについて行うかD(WPI)について行うかの選択を聞いてくる。ここではWPIを選択する。OUTPUTとしてはグ

表9.6 予測の結果



ラフと数値評価の双方を選択する。

結果は表9.6に掲げるとおりである。

点線で±2標準偏差の信頼区間が表示されている。数値評価の内容は次のように計算されたものである(Theil Inequality Coefficientについては省略する)。

Root Mean Squared Error (平均平方誤差)

$$[(1/h) \sum_{s=1}^h (y_{T+s/T} - y_{T+s})^2]^{1/2} \quad 9.36a)$$

Mean Absolute Error (平均絶対誤差)

$$(1/h) \sum_{s=1}^h |y_{T+s/T} - y_{T+s}| \quad 9.36b)$$

Mean Absolute Percentage Error (平均絶対%誤差)

$$(1/h) \sum_{s=1}^h |(y_{T+s/T} - y_{T+s})/y_{T+s}| \quad 9.36c)$$

予測誤差の平均平方和 (Mean Squared Forecast Error) は次のように分解することができる。

$$\sum_{s=1}^h (y_{T+s/T} - y_{T+s})^2/h = (\bar{y}_{T/T} - \bar{y})^2 + (\bar{y}_{T/T} - \bar{y})^2 + 2(1 - r_{y/Ty}) \bar{y}_{T/T} \bar{y} \quad 9.37)$$

\bar{y} と $\bar{y}_{T/T}$ はそれぞれ予測期間中の y_{T+s} と $y_{T+s/T}$ の平均値、 \bar{y} と $\bar{y}_{T/T}$ はそれぞれ予測期間中の y_{T+s} と $y_{T+s/T}$ の標準偏差、 $r_{y/Ty}$ は y_{T+s} と $y_{T+s/T}$ の相関係数とする。9.37)式を用いて予測誤差の要因を求める。

予測の平均と実際の値の平均との乖離 (Bias Proportion)

$$\frac{(\bar{y}_T - \bar{y})^2}{\sum_{s=1}^h (y_{T+s/T} - y_{T+s})^2/h} \quad 9.38a)$$

予測の変動と実際の変動の差 (Variance Proportion)

$$\frac{(\bar{y}_{T/T} - \bar{y})^2}{\sum_{s=1}^h (y_{T+s/T} - y_{T+s})^2/h} \quad 9.38b)$$

残余の予測誤差 (Covariance Proportion)

$$\frac{2(1 - r_{y/Ty}) \bar{y}_{T/T} \bar{y}}{\sum_{s=1}^h (y_{T+s/T} - y_{T+s})^2/h} \quad 9.38c)$$

この例では比較的予測が当たっているように見える(分散不均一の問題はあるが)。

6 ARCH

EngleやBollerslevによって提唱された自己回帰条件付き分散不均一モデル(Autoregressive Conditional Heteroscedasticity, ARCH)について簡単に触れたい。これは回帰モデルの誤差項(u_t)の分散(σ_t^2)が過去の誤差の大きさに依存するという可能性を考えるものである。

$$\sigma_t^2 = a_0 + b_1 u_{t-1}^2 + b_2 u_{t-p}^2 + \dots + b_p u_{t-p}^2 \quad b_i \geq 0 \quad 9.39)$$

もちろん $b_1 = b_2 = \dots = b_p = 0$ であれば分散均一である。9.39)式はARCH(p)過程と呼ばれてい

る。9.39) 式の分散は前の誤差によってもたらされる。ARCH(p) 過程は次のようにも書かれる。

$$u_t = e(a_0 + b_1 u_{t-1}^2 + b_2 u_{t-2}^2 + \dots + b_p u_{t-p}^2)^{1/2} \quad (9.40)$$

e_t は u_t と独立している IID 過程である。9.40) 式の特殊のケースとして ARCH(1) 過程が得られる。ARCH(1) であれば $u_t = e(a_0 + b_1 u_{t-1}^2)^{1/2}$ である。 u_t は平均 0 である²⁰⁾。

ARCH の検定は以下の手順による。

$y_t = a + b x_t + u_t$ を OLS で推計し、残差 \hat{u}_t を求める。

$$\hat{u}_t^2 = a_0 + b_1 \hat{u}_{t-1}^2 + b_2 \hat{u}_{t-2}^2 + \dots + b_p \hat{u}_{t-p}^2 + w_t \quad (9.41)$$

を OLS で推計し、9.41) 式の決定係数を R^2 とし、標本数を n とする。帰無仮説

$$H_0 : b_1 = b_2 = \dots = b_p = 0 \quad (\text{ARCH 効果なし})$$

を検定する。帰無仮説が正しく標本数が多い場合は LM 検定統計量 nR^2 が自由度 p の χ^2 統計量に従う。しかし 9.41) 式より p の実際の特定は難しい。

ARCH モデルの拡張したモデルとして次の GARCH(p, q) (Generalized ARCH) がある。

$$\hat{u}_t^2 = a_0 + b_1 \hat{u}_{t-1}^2 + b_2 \hat{u}_{t-2}^2 + \dots + b_p \hat{u}_{t-p}^2 + c_1 \hat{u}_{t-1}^2 + c_2 \hat{u}_{t-2}^2 + \dots + c_q \hat{u}_{t-q}^2 \quad (9.42)$$

$q = 0$ (すなわち $c_1 = c_2 = \dots = c_q = 0$) であれば、ARCH(p) モデルになる。GARCH(p, q) モデルにおいて条件付き分散 (\hat{u}_t^2) は p 個のラグ付き誤差項の自乗 (\hat{u}_{t-s}^2) と q 個のラグ付き条件付き分散 (\hat{u}_{t-s}^2) に依存する。

GARCH(1, 1) であれば 9.42) 式は

$$\hat{u}_t^2 = a_0 + b_1 \hat{u}_{t-1}^2 + c_1 \hat{u}_{t-1}^2$$

になる。上式を書き直すと

$$\hat{u}_t^2 = a_0 / (1 - c_1) + b_1 (u_{t-1}^2 + c_1 u_{t-2}^2 + \dots + c_1^{k-1} u_{t-k}^2 + \dots)$$

を得ることができる。すなわち GARCH(1, 1) モデルを ARCH() として表現することができる。ただし係数は三つしかない。ARCH と GARCH モデルを最尤法で推計する²¹⁾。

WPI の ARIMA モデルの残差のテストでは分散不均一であった。GARCH の存在が疑われる。表 9.5 が表示されているとする。

View/Residual Test/ARCH LM Test を選択する。ARCH(1) の有無を検定するために次数を 1 とする。結果は表 9.7 に示すとおりである。LM 統計量は 7.17 (p 値は 0.007) であるから、ARCH 効果がないという帰無仮説は棄却される。次の GARCH(1, 1) モデル

$$WPI_t = a_0 + a_1 WPI_{t-1} + a_2 WPI_{t-2} + a_3$$

$$WPI_{t-3} + u_t$$

$$\hat{u}_t^2 = b_0 + b_1 \hat{u}_{t-1}^2 + c_1 \hat{u}_{t-1}^2$$

を推計するコマンドは次の通りである。

equation archwpi.ARCH(1, 1) D(WPI) D(WPI(-1)) D(WPI(-2)) D(WPI(-3))

ARCH と指定することで GARCH(p, q) モデルの推計を行う。() の中で p と q の次数を指定している。結果は表 9.8 に掲げるとおりである。

下欄に Variance equation として $\hat{u}_t^2 = b_0 + b_1 \hat{u}_{t-1}^2 + c_1 \hat{u}_{t-1}^2$ の推計結果が報告されている。表 9.8 より b_1 と c_1 の推定値はそれぞれ 0.153 (p 値 = 0.000) と 0.760 (p 値 = 0.000) となるので ARCH 効果 (\hat{u}_{t-1}^2) と GARCH 効果 (\hat{u}_{t-1}^2) が必要ということが分かる。

²⁰⁾ 証明は省略するが、 $a_0 > 0$ 、 $|a_1| < 1$ 、 $E(u_{t-1} u_t) = 0$ である。

²¹⁾ ARCH, GARCH の詳細については Davidson and MacKinnon [1993] Ch16-4, Hendry [1995] Ch15-8 を参照されたい。

表9.7 ARCHテスト

ARCH Test :				
F statistic	7.301941	Probability	0.007310	
Obs*R squared	7.166719	Probability	0.007427	
Test Equation: Dependent Variable:RESID ^ 2 Method : Least Squares Sample (adjusted) : 1975 : 02 1998 : 06 Included observations : 281 after adjusting endpoints				
Variable	Coefficient	Std. Error	t Statistic	Prob.
C	0.243653	0.046710	5.216276	0.0000
RESID ^ 2 (- 1)	0.159677	0.059091	2.702210	0.0073
R squared	0.025504	Mean dependent var	0.289858	
Adjusted R squared	0.022012	S.D. dependent var	0.736811	
S.E. of regression	0.728656	Akaike info criterion	2.211863	
Sum squared resid	148.1323	Schwarz criterion	2.237759	
Log likelihood	-308.7668	F statistic	7.301941	
Durbin Watson stat	1.981737	Prob (F statistic)	0.007310	

表9.8 GARCHの推計

Dependent Variable : :D (WPI) Method : ML ARCH Sample : 1975 : 01 1998 : 06 Included observations : 282 Convergence achieved after 25 iterations				
	Coefficient	Std. Error	z Statistic	Prob.
C	-0.012915	0.032314	-0.399657	0.6894
D(WPI(- 1))	0.453950	0.071676	6.333377	0.0000
D(WPI(- 2))	0.003156	0.090201	0.034988	0.9721
D(WPI(- 3))	0.146074	0.069528	2.100924	0.0356
Variance Equation				
C	0.027481	0.014903	1.844055	0.0652
ARCH(1)	0.153168	0.043783	3.498355	0.0005
GARCH(1)	0.759610	0.072169	10.52545	0.0000
R squared	0.382472	Mean dependent var	0.033333	
Adjusted R squared	0.368999	S.D. dependent var	0.691089	
S.E. of regression	0.548971	Akaike info criterion	1.532647	
Sum squared resid	82.87638	Schwarz criterion	1.623049	
Log likelihood	-209.1032	F statistic	28.38732	
Durbin Watson stat	1.861617	Prob (F statistic)	0.000000	

参考文献

VARモデルについては

山本拓 [1988] 『経済の時系列分析』の8～10章

VECMについては

森棟 [1999] 公夫 『計量経済学』東洋経済新報社の11章

Harris, R. [1995] *Using Cointegration Analysis in Econometric Modelling*のCh 5 , 6 が分かりやすい。

ARIMAモデル、GARCHについては

Enders, W. [1995] *Applied Econometric Time Series*のCh 2 , 3

Davidson, R. and J. MacKinnon [1993] *Estimation and Inference in Econometrics*のCh16 4

Hendry, D. [1995] *Dynamic Econometrics*, Oxford University PressのCh15 9、Ch16参照。