広島経済大学経済研究論集
第41巻第４号　2019年３月
http://dx.doi.org/10.18996/keizai2019410404

広島経済大学経済学会
2018年度　第４回研究集会〔2018年12月６日（木）〕報告要旨

経済における因果性の探索

前　川　功　一*

1.　は　し　が　き

本稿は，2018年12月６日（木）明徳館 4F で開催された広島経済大学経済学会2018年度第４回研究集会において報告した「経済における因果性の探索―カクテルパーティー風に―」の補完的要約である。筆者は広島経済大学創立五十周年記念論文集（2017年７月発行）に「非ガウス型構造 VAR モデルによる因果序列の探索」と題する論文［4］を著わした。しかしこの論文のタイトルは，計量経済学の専門家でない方々には，全く理解不能であったらしく，研究集会の幹事役の野北先生から研究集会で，非専門家に分かりやすくこの論文を解説して欲しいという依頼を受けた。アインシュタインは「６歳の子供に説明できなければ，理解したとは言えない（*If you can't explain it to a six year old, you don't understand it yourself.*）」という至言を残した。６歳は極端であるが，本学に在籍する先生方に説明できないようでは，私自身が本当には理解できていないのかもしれない。そこでどこまで平易に説明できるかにチャレンジするつもりで研究集会での報告を引き受けることにした。そして経済学を専門としない先生方に理解して頂くために，専門用語や数式による説明を可能な限り控え，イラストや例え話を多用して分かりやすく説明することを心がけた。しかし例え話による説明には限界もあり，講演終了後，数人の先生から例え話と経済現象との関係が分かりにくいというご指摘も頂いた。そこで本誌の紙面を借りて，このご指摘にお答えするために多少数式を交えながら補足説明させて頂こうと思う。

2.　経済時系列モデルによる経済因果性の探索

経済システムに外部からショックが与えられると，そのショックは次々に経済活動に波及的な効果を及ぼす。すなわちある経済現象 A が原因になり次の現象 B が起こり，B が次の現象 C の原因となるという一連の因果序列が存在すると考えられる。例として，日本銀行の非伝統的金融緩和政策の効果が次々に波及するような場合が挙げられる。しかし現実には正しい因果序列は判然としないので，多くの論者がそれぞれの立場から様々な意見を述べ合い，侃々諤々の議論が論壇を賑やかすことになる。議論を戦わせることは，より良い結論を得るために必要なことではあるが，詳細なデータ分析に基づいた議論は多くない。どちらかというと論者の主観的主張で終わっている場合が少なくない。筆者は近年，計量経済学の立場からこの問題に対して，「データに語らせる」あるいは「虚心坦懐にデータに耳を傾ける」ことに重きを置く多変量経済時系列モデルによるアプローチを試みつつある。筆者の論文［4］はその一端をまとめたものである。そこで用いた「データに語らせる」立場に立つ多変量経済時系列モデルと

＊ 広島経済大学大学院経済学研究科教授

は一言でいうと，現在の経済状況を示すデータは過去のデータに依存していて，遠い過去の影響は次第に薄れていくと想定したモデルである。式を使って表すと

$$y_t = \Gamma_1 y_{t-1} + \cdots + \Gamma_s y_{t-s} + \cdots + \Gamma_p y_{t-p} + \varepsilon_t \quad (1)$$

となる。ここで y_t は時点 t における k 個の経済変数を含むベクトル

$$y_t = \left(y_{1t}, \cdots, y_{kt}\right)',$$

また Γ_s は y_{t-s} が y_t に及ぼす効果を示す $k \times k$ の係数行列である。また ε_t は時点 t における確率的撹乱項を表す。外生的な経済ショックも ε_t の一種として扱われる。この ε_t に対してはある仮定が置かれるがそれについては後述する。このようなモデルを p 次の多変量自己回帰モデル（Vector Autoreregessive model）といい，略して VAR（p）とあらわされる。(1) 式で y_t は，現時点 t から p 期前までの $y_{t-1}, \cdots, y_{t-p}$ に依存する形になっている。しかしこのモデルには，現時点 t における $y_t = \left(y_{1t}, \cdots, y_{kt}\right)$ に含まれる変数 $y_{1t}, \cdots, y_{kt}$ の相互間の同期間内依存関係（または依存構造）が含まれていない。それを含むように拡張したモデルを**構造 VAR** モデルという。式で示せば，

$$\begin{aligned} y_t = \Gamma_0 y_t + \Gamma_1 y_{t-1} + \cdots + \Gamma_s y_{t-s} + \cdots \\ + \Gamma_p y_{t-p} + \varepsilon_t \end{aligned} \quad (2)$$

と表される。右辺に $\Gamma_0 y_t$ という項が含まれる点が (1) 式と異なる。ここで係数行列 $\Gamma_0, \cdots, \Gamma_p$ の値は未知である。また確率的撹乱項 ε_t は観測できない。経済変数 y_t の時系列データ $y_t, y_{t-1}, \cdots, y_{t-p}$ が得られたなら統計的手法を使って係数行列 $\Gamma_0, \cdots, \Gamma_p$ を推定することが出来る。(2) 式において現時点の変数ベクトル y_t を左辺に集めて

$$y_t = B^{-1}\Gamma_1 y_{t-1} + \cdots + B^{-1}\Gamma_p y_{t-p} + B^{-1}\varepsilon_t \quad (3)$$

という VAR（p）の形に変形することが出来る。ここで $B = I - \Gamma_0$ である。さらに $A_1 = B^{-1}\Gamma_1 \cdots$, $A_p = B^{-1}\Gamma_p$ と置けば (3) 式は

$$y_t = A_1 y_{t-1} + \cdots, + A_p y_{t-p} + u_t \quad (4)$$

となる。(4) 式は構造 VAR モデルの誘導形と呼ばれる。(3) 式に現れる撹乱項 u_t と (2) 式の撹乱項 ε_t との間には，

$$u_t = B^{-1}\varepsilon_t \quad (5)$$

という関係式が成立つ。u_t は ε_t を含んでいるので観測できない。さてここで (4) 式を最小２乗法で推定し A_i の推定値を $\hat{A}_i$ と置き

$$\hat{u}_t = y_t - \hat{A}_1 y_{t-1} - \cdots - \hat{A}_p y_{t-p}$$

を計算する。$\hat{u}_t$ は誘導形残差と呼ばれ，観測できない u_t の推定値として用いることが出来る。以上で (4) 式の VAR（p）モデルが推定できたのであるが，本来の目的は (4) 式の背後にある (3) 式を推定することである。(4) 式の推定結果から逆算すれば (3) 式の推定値を計算できそうに見えるが，一般的にはこの逆算によって (3) 式を一意に推定できない。このような場合，(3) 式は識別不能という。識別性については後述するので，ここでは単に ε_t が正規分布に従うという伝統的な仮定の下では (4) 式の推定結果から (3) 式の推定結果が得られない場合があると述べるに留めておく。ところが機械学習の分野で使われている独立成分分析（Independent Component Analysis，以下 ICA と略記）という手法を使うと，ε_t に正規性を仮定しなければ識別が可能になることが知られている。

3. 独立成分分析（ICA）とは

ICA は機械学習やニューラルネットワークの分野で開発された手法であるが，構造 VAR モデルと数学的構造に共通点があるので，最近

ICAの経済時系列分析への応用論文が少しずづ増えつつある。そこでまず初めにICAの概略を拙稿［4］に沿って紹介する。ICAの解説については参考文献［1］を参照されたい。

いま n 個の独立に非正規分布に従う n 個の確率変数 $s_1, s_2, \cdots, s_n$ を考える。ICAのイメージを掴みやすくするために，これらはオーケストラの n 種類の楽器から発せられた音声信号としよう。これらの音声信号 $s_1, s_2, \cdots, s_n$ が空中でミックスされホールに響き渡る。このミックスされた響きが n 個の座席で音波 $x_1, x_2, \cdots, x_n$ として観測される状況を考える（下の図1を参照)。この状況は，カクテルパーティ会場のざわめきの中から話し相手の声を聞き分ける状況にも例えられる。この現象はカクテルパーティー効果と呼ばれているが，研究集会での講演の副題を“カクテルパーティー風に”とした理由はここにある。ここで各座席で観測された音声信号のベクトルを $x = (x_1, x_2, \cdots, x_n)'$，音源（各楽器）の信号を $s = (s_1, s_2, \cdots, s_n)'$ で示せば，x と s の関係は式を用いて

$$x = As \tag{7}$$

と表すことが出来る。ここに A は，$n \times n$ のフルランク行列で混合行列と呼ばれる。独立成分分析の問題とは，各座席で観測された音波 x

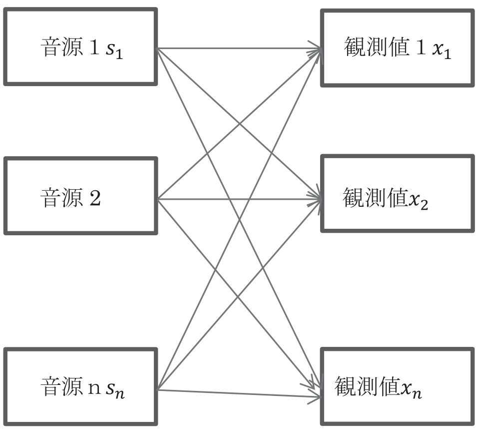

図1　ICAのイメージ図

をいかに s と A に分離・抽出するかという問題である。混合行列 A が既知であれば，観測値 x が得られたとき

$$s = A^{-1}x \tag{6}$$

の関係から s が復元される。ICAとは，音源信号 $s_1, s_2, \cdots, s_n$ は独立な非正規分布に従うという仮定の下で，観測された信号 x から混合行列 A と音源信号 s を推定する手法である。

次に，非正規性と独立性の仮定の役割を説明する。まず識別性の問題を考える。今二つのModel AとModel Bがあり，観測値データ $x_1, x_2, \cdots, x_n$ が手元にあるとする。そしてこの観測値データがモデルAから生成されたのか，モデルBから生成されたのかを識別したいとしよう。それが識別できない場合，この二つのモデルは観測値上同等なモデルであると言われる。一般に何らかの制約を置かなければ識別不能という事態が生じる。経済分析では，しばしば経済理論又は経済制度上の先験的知識を使って，係数にゼロ制約を課すことによって識別性を回復するという手段が取られる。先験情報を有効に活用することは大切ではあるが，それに頼り過ぎると弊害もある。それを避けたいというのが「データに語らせる」立場である。ICAは，そのような先験的制約を課す代わりに非正規性と独立性の仮定を置くことによって識別可能性を獲得しようとする手法である。Moneta et al（2013）は簡単な例を使って，なぜこの二つの仮定を置くことによってモデルの識別が可能になるかについて幾何学的に説明している。ただし非正規性以外にも若干の仮定が必要であるが，この点に関しては拙稿［4］の補論を参照されたい。

4.　ICAによる非ガウス型SVARモデルの推定と因果序列

この節ではまず初めに，ICAの基本モデルと

構造 VAR モデルの数学的構造の類似性を示し，ICA がなぜ構造 VAR の推定に応用できるのかを示す。前節で示したように構造 VAR モデル（2）式の撹乱項 ε_t と誘導形（4）式の撹乱項 u_t の間には

$$u_t = B^{-1}\varepsilon_t \tag{5}$$

という関係が成立している。この式は ICA における（7）式：$x = As$ と数学的には同じ構造を持っている。したがって $u_t = B^{-1}\varepsilon_t$ を，ICA の文脈の中で解釈すると，次のようにいうことが出来る。独立な成分 $\varepsilon_{1t},\varepsilon_{2t},\cdots,\varepsilon_{nt}$ を要素とするベクトル ε_t が係数行列 B^{-1} によって混合された結果として u_t が構成される。実際に u_t は観測不能であるから，u_t の代わりに誘導形残差 $\hat{u}_t$ を B^{-1} と ε_t に分解する。その際 B^{-1} が下側三角行列の形に分解できれば，因果序列が存在することになる。その理由を簡単な３変数の構造 VAR モデルを使って説明する。B^{-1} が下記のような下側３角行列であれば，$u_t = B^{-1}\varepsilon_t$ を要素に遡って書けば

$$u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix} = B^{-1}\varepsilon_t = \begin{bmatrix} 1 & 0 & 0 \\ \beta^{21} & 1 & 0 \\ \beta^{31} & \beta^{32} & 1 \end{bmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon_{1t} \\ \beta^{21}\varepsilon_{1t} + \varepsilon_{2t} \\ \beta^{31}\varepsilon_{1t} + \beta^{32}\varepsilon_{2t} + \varepsilon_{3t} \end{pmatrix}$$

となる。この式は次のように解釈することが出来る。いま経済システム（2）の外部からショック ε_{1t} が与えられると，そのショックは u_{1t} を経て y_{1t} に伝搬する。さらにそれは $u_{2t} = \beta^{21}\varepsilon_{1t} + \varepsilon_{2t}$ の関係式を通して u_{2t} へ，したがって y_{2t} へ伝搬し，さらに $u_{3t} = \beta^{31}\varepsilon_{1t} + \beta^{32}\varepsilon_{2t} + \varepsilon_{3t}$ の関係式を通して u_{3t} へ，したがって y_{3t} へ伝搬する。このことは最初に与えられた外部ショック ε_{1t} は

$$y_{1t} \to y_{2t} \to y_{3t}$$

の順序を経て次々に伝搬していく因果序列が存在することを意味する。このような序列をデータから計算可能であるためには，撹乱項ベクトル ε_t の構成要素は次の二つの仮定：

仮定１：$\varepsilon_{1t},\cdots,\varepsilon_{kt}$ は非正規布に従う

仮定２：$\varepsilon_{1t},\cdots,\varepsilon_{kt}$ は統計的に独立

仮定３：Γ_0 は下三角行列（従って B も下三角行列）

を満たしていなければならないことが証明される（例えば参考文献［1］を見よ）。

実際のデータを使って因果序列を求めるためのアルゴリズムのパッケージとして LiNGAM（Linear Non Gaussian Acyclic Model）がある。そこで使われている計算手順の概略を以下に示しておく。

（1）観測値データ $y_{1t},\cdots,y_{kt},t=1,\ldots.,T$ を収集する

（2）データから構造 VAR モデルの誘導形（4）を推定し，誘導形残差 $\hat{u}_t$ を求める，

（3）その誘導形残差 $\hat{u}_t$ を ICA アルゴリズムに入力し，独立成分 $\hat{\varepsilon}_t$ と係数行列 $\hat{B}^{-1}$ に分解し，Γ_0 を $\hat{\Gamma}_0 = I - \hat{B}$ によって推定する。

（4）仮定３を満たすために，$\hat{B}$ を可能な限り下側三角行列に近づくように変形する（首尾よくそのように変形できた場合は因果序列が検出できたことになり，そのように変形できなかったときは因果序列が検出できなかったと判断される）。

LiNGAM には与えられた誘導形残差 $\hat{u}_t$ に対して，係数行列 B^{-1} が最も下側三角行列に近い行列を求める一種の最適化のアルゴリズムが組み込まれている。ここで仮定３について注意が必要である。この仮定は数学の定理の証明で用いられるような仮定というよりは，最適化（下側三角行列に最も近い行列を求める）過程

における到達目標的な意味で使われている。LiNGAM は，因果序列があると仮定して，それにどこまで迫れるかを見ようとしている。したがって計算してみたらうまく下側三角行列が得られることもあるし，得られない場合もある。そして首尾よく三角化に成功したときは LiNGAM の計算終了後に，メッセージ“Done! Causal B nicely triangular. No problems to report here”が出力され，最適化に失敗した場合は“WARNING: Causal B only somewhat triangular!”が出力される。最適化が成功した後，その結果をクロスチェックしたい場合は仮説３（因果序列は存在する）を帰無仮説，対立仮説（因果序列は存在しない）とする（尤度比検定等の）統計的仮説検定を行うことも可能である。

なおここでいう因果序列とは同期間内（時点 t）での因果序列を意味するのであるが，t 期間内の因果序列は，誘導形（2）式の係数 $B^{-1}\Gamma_j$ を通して過去の y_{t-j} にも影響を及ぼす。特に B^{-1} が下側三角行列で，y_{t-i} の係数 Γ_{1-i} もまた下側三角行列であれば $B^{-1}\Gamma_{t-i}$ も下側三角行列となるので，同期内の因果序列構造は時間を超えて保存される。

5. む す び

研究集会では，非専門家の方々に分かりやすく説明するために，専門用語や数式表現を避け，イラストと例え話を多用して説明した。そして独立成分分析をオーケストラの奏でる響きから個別の楽器の音声を分離するという例え話を使って説明した。その例え話とそれに続く独立成分分析の経済分析への応用との関連性が分かりにくいという指摘があったので，本稿はその部分を補足的に説明した。補足のポイントは，独立成分分析と経済の因果序列の分析の間には（（5）式と（6）式に見られるような）共通の数学的構造があるという点である。この点を例え話風に説明するとかえって分かりにくいので，数式を交えながら補足説明した次第である。

最後に，拙稿［4］で例示として挙げた日本銀行の非伝統的金融緩和政策の因果序列に触れて本稿を締めくくりたい。この例題は，日本のマクロ経済変数の2012年１月から2014年12月の期間の月次データを用いて，日銀の量的緩和政策（マネタリーベース増加政策）がどのような因果序列を経てマクロ経済変数に波及して行ったかを実証的に検証する試みである。検証の結果，マネタリーベースの増加 ⇒ 予想インフレ率の上昇 ⇒ 消費者物価指数の上昇 ⇒ 総合消費の増加 ⇒ 有効求人倍率の増加，という金融変数から実体経済への一連の因果序列が観察された。これは単なる例示にしか過ぎないものの，拙稿［4］で取り上げた手法の有効性に希望を抱かせる結果といえよう。しかしそこでの分析にはいくつかの未解決の問題とさらに検討を要する問題が残されているので，今後それらについて分析を深めていきたい。

参 考 文 献

［1］Hyvärinen, A., Juha Karhunen, Erkki Oja（2001），Independent Component Analysis, John wiley & sons, Inc.（日本語訳：詳解独立成分分析，根本幾・川勝真喜訳　東京電機大学出版局，2005年）

［2］Moneta, Allesio, Doris Entner, Patrik O. Hoyer, Alex Coad（2013），Causal Inference by Independent Component Analysis: Theory and Applications. Oxford Bulletin of Economics and Statistics, 75, 5

［3］前川功一，Amirullah, S. H.（2016），ICA 分析による因果序列の検出―インドネシア・ルピアの為替レート分析―，広島経済大学研究双書　第44冊「東アジアの経済成長の持続可能性について」，広島経済大学　地域経済研究所

［4］前川功一（2017），非ガウス型構造 VAR モデルによる因果序列の探索―日本の量的金融緩和政策の分析を事例として―，広島経済大学創立50周年記念論文集（上巻）