

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—

前 川 功 一¹⁾

1. はじめに

1980年代以前に、計量経済学の主流であった構造方程式モデルは、識別可能性を満たすために様々な先験的な係数制約を必要とした。しかしそれらの制約条件が恣意的であるとか、経済理論の根拠が薄弱であるなどと批判された結果、最近ではこのモデルはあまり使用されなくなった。それに代わるモデルとして、先験的な制約を可能な限り排除した時系列モデル、特に多変量自己回帰モデル（Vector Autoregressive (VAR) Model）が計量経済学の主流の座を占めるに至った。その後、現在に至るまで VAR モデルは計量経済学における標準的手法としての地位を維持している。VAR モデルの基本的立場は、経済理論あるいは実態経済に関する様々な経験的または主観的事前情報を極力排した単純な時系列モデルを用いて、虚心坦懐に「データが語る」ことに耳を傾けるという点にある。しかしながら、このようなアプローチに対しては、「理論無き計測」という批判を免れない。この批判に対処するため、VAR モデルに経済の構造を取り入れようとする方向が生まれた。それが構造 VAR モデルである（Structural VAR の頭文字をとって SVAR モデルと表記される）。SVAR モデルは、経済変数間の同期間内及び遅れのある相互依存関係を反映させようとするモデルであるが、かつての構造方程式と同様に識別性の問題を孕んでいる。そして SVAR モデルの識別性問題を解決するための様々な理論的枠組みが提案さ

れているが、最もよく知られている方法は、同期間内の構造を表す係数行列に再帰的と呼ばれる係数制約を置くことである。具体的にはこの係数行列が下側三角行列になると仮定することである。このような係数制約を課すことは、後述するように経済変数間の同期間内での因果序列 (Causal Order) を仮定することに等しい。しかしながら、このような強い仮定を置くことは、皮肉なことに構造方程式への批判として生まれた VAR モデルの「データに語らせる」という方向に逆行することになりかねない。

ところが最近、一部の研究者の間で、「データに語らせる」という時系列分析本来の立場を保ちつつ、独立成分分析 (Independent Component Analysis: ICA) という手法を SVAR モデルに応用することによって同期間内の変数間の因果序列を求める研究が行われている。後述するように ICA では、先験的な係数制約を置く代わりに、モデルの攪乱項に非正規性を仮定することによって、識別性の問題も回避しつつ自然な形で経済変数間の因果序列が探索される。経済時系列においては攪乱項に非ガウス性を仮定することは不自然ではない。また ICA と SVAR モデルの間には、共通する数学的構造が存在するので、ICA を計量経済学に導入することには一定の合理性がある。経済時系列へ ICA を応用する研究例として代表的な論文に Moneta, Entner, Hoyer and Coad (2013) がある²⁾。この論文は、アメリカの 4 半期マクロ経済時系列データを用いて、SVAR モデルの係数を、経済理論や経済制度に関する先験的情報を使わずに ICA を用いて変数間の因果序列の推定を試みたものである。

本稿の構成は以下の通りである。第 2 節で ICA の概略を紹介する。第 3 節で非ガウス型 SVAR モデルを導入した後、ICA を用いて SVAR を推定するための計算ソフト LiNGAM (3 節で詳述) を使って SVAR モデルの同時期内での変数間の因果序列を規定する係数行列 (下側三角行列) の推定の手順を示す。さらにモンテカルロ実験によって、LiNGAM の信頼性を検証する。最後に第 4 節で、LiNGAM を使って、日本の金融量的緩和政策の下における変数間の因果序列を実データに基づいて推定する。

2. ICA の概略

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り、参考文献に示されているように多数の理論的、応用的研究論文や解説的な文献が存在している。ICA は計量経済学の分野ではなじみが薄いのので、まず初めに本論と関係の深い部分の概略を前川・アミルラ（2016）（以下 M-A 論文と略記）に沿って説明し、ICA の詳細については参考文献に挙げた Hyvärinen, A., Juha Karhunen, Erkki Oja（2001）などに委ねる。

いま n 個の独立な分布に従う潜在変数 s_1, s_2, \dots, s_n があり、そのうち少なくとも $n-1$ 個は独立に非正規分布に従うものとする。これらの潜在変数の効果がミックスした結果として、 m 個 ($m \geq n$) の観測値 x_1, x_2, \dots, x_m が観測される状況を考える（下の図 1 を参照）。この状況は、式を用いて

$$x = As$$

と表すことが出来る。ここに $x = (x_1, x_2, \dots, x_m)'$, $s = (s_1, s_2, \dots, s_n)'$ である。また A は、フルランクを持つ $m \times n$ ($m \geq n$) の行列で混合行列と呼ばれる。独立成分分析の問題とは、与えられた観測値 x のみから s をいかに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

$$y = A^{-1}x \quad (1)$$

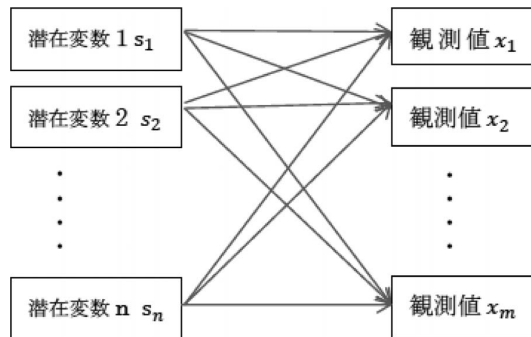


図 1 ICA のイメージ図

の関係から得られる y によって s は復元される。しかし A 行列は通常未知であるので、 y が限りなく s に近づくように、 A をデータから推定することが ICA の目的である。

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, \dots, x_m 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。伝統的な計量経済分析では、しばしば経済理論又は経済制度上の知識を使って、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。しかし ICA では、そのような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をもたらされることが示される。Moneta et al (2013) は、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簡単な例を用いて直感的に説明している。ただし非正規性以外にも若干の仮定が必要であるが、この点に関しては補論で解説する。また非正規性の仮定を用いる方法以外にも様々な識別性獲得の方法がある。Lütkepohl and Netsunajev (2017) は、それらの方法に関するサーベイを行っている。

3. 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 SVAR モデルの数学的構造の類似性を示し、ICA がなぜ SVAR の推定に応用できるのかを示す。

説明を簡略化するためと、後半の実証分析例で用いるモデルとの整合性を保つために、以下のような 1 期の遅れを持つ 5 変数から成る SVAR モデルを取り上げる。

$$y_t = By_t + \Gamma_1 y_{t-1} + \varepsilon_t \quad (1)$$

ここに

$$y'_t = (y_{1t}, y_{2t}, y_{3t}, y_{4t}, y_{5t}); \varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}),$$

である。さらに ε_t の要素は互いに統計的に独立な確率分布に従い、その分散共分散行列を Σ_ε とする。(1) 式は

$$\Gamma_0 y_t = \Gamma_1 y_{t-1} + \varepsilon_t,$$

と書き換えることができる。ここに $\Gamma_0 = I - B$ である。さらに書き換えると

$$y_t = \Gamma_0^{-1} \Gamma_1 y_{t-1} + \Gamma_0^{-1} \varepsilon_t = A y_{t-1} + u_t \quad (2)$$

が得られる。(2) 式は (1) 式の誘導形と呼ばれる。ここに $A = \Gamma_0^{-1} \Gamma_1$ である。また ε_t と u_t の関係は

$$u_t = \Gamma_0^{-1} \varepsilon_t, \quad (3)$$

となる。(3) 式は、次のように解釈することが出来る。独立な成分 ε_t が係数行列 Γ_0^{-1} によって混合された結果として u_t が観測される（一般的には u_t が観測されないで、その代りに以下では誘導形の最小 2 乗残差 \hat{u}_t を用いる）。独立成分分析 ICA とは、混合されたノイズ u_t が観測されたとき、ノイズ u_t を独立な成分 ε_t と混合行列 Γ_0^{-1} に分解する方法である。もし、 ε_t の正規性を仮定すると u_t の分布も正規分布になる。このような場合は後で述べるように識別性の問題が生じる。この問題を回避するために ICA では、以下のような仮定を置く。

仮定 1 : $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ のうち少なくとも 4 個は非正規分布に従う

仮定 2 : $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ は統計的に独立に分布する

仮定 3 : $y_{1t}, y_{2t}, y_{3t}, y_{4t}, y_{5t}$ は同期因果序列 (contemporaneous causal order)

または再帰的序列 (recursive order) を持つ。このとき Γ_0 は下側三角行列になる。

ここで (1) 式の係数行列 B が下側三角行列であるという仮定は、同期間内での因果序列を意味するだけでなく、誘導形 (2) 式の係数 $\Gamma_0^{-1} \Gamma_1$ を通して y_t と y_{t-1} の間の関係にも影響する。特に Γ_1 も下側三角行列であれば $\Gamma_0^{-1} \Gamma_1$ 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。 Γ_1 が下側三角行列でない場合、因果序列は時間を超えて保存はされないが、同期内の因果序列は異時点にも影響を与える。

非正規性の仮定より、歪度や尖度などの高次のモーメント情報が利用可能になり、それによって識別性の問題が回避される。

上で説明したICAを用いて誘導形残差 \hat{u}_t を $\hat{\varepsilon}_t$ と $\hat{\Gamma}_0^{-1}$ に分解して \hat{B} を求める手順をまとめれば次のようになる：

- (1) 観測値データ $y_{1t}, y_{2t}, y_{3t}, y_{4t}, y_{5t}, t = 1, \dots, T$ を収集する
- (2) データからSVARモデルの誘導形(2)を推定し、残差 \hat{u}_t を求める、
- (3) その残差 \hat{u}_t をICAアルゴリズムに投入し、独立成分(の推定値)
 $\hat{\varepsilon}_t$ と係数行列 $\hat{\Gamma}_0^{-1}$ に分解し、 $\Gamma_0 = I - B$ の関係から \hat{B} を求める。
- (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(うまくそのように変形できる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いときは因果序列は検出できないと判断される)。

以上の計算を自動的に行う計算パッケージにLiNGAM – Linear Non Gaussian Acyclic Model LiNGAMがある(<https://sites.google.com/site/sshimizu06/lingam>)。

LiNGAMはSVARモデルの係数行列 B が下側三角行列に可能な限り近くなるようなある種の最適化手法を用いて推定値行列 \hat{B} を出力する。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、計算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くまくだ側三角行列が得られたということもあるし、得られなかったという場合もありうる。そのためLiNGAMの計算終了後には、

“Done! Causal B nicely triangular. No problems to report here.”

または

“WARNING: Causal B only somewhat triangular!”

というメッセージが出力される。Moneta et al (2013) は, LiNGAM の計算過程の概略を以下のようにまとめている。

LiNGAM の計算過程のまとめ

1. 誘導形 (2) の係数行列を推定し, 残差 \hat{u} を計算する
2. ICA アルゴリズムを用いて \hat{u} を $\hat{u}=P\hat{\varepsilon}$ の形に分解し独立成分 $\hat{\varepsilon}$ を求める (得られた \hat{u} の非正規性をチェックする)
3. $\tilde{\Gamma}_0=P^{-1}$ と置き, $\tilde{\Gamma}_0$ の行を入れ替えて, $\Sigma_i 1/|\tilde{\Gamma}_{0_{ii}}|$ が最小になるような $\tilde{\Gamma}_0$ を見つける (linear matching problem)。ここに $\tilde{\Gamma}_{0_{ii}}$ は $\tilde{\Gamma}_0$ の (i, i) 対角要素である。
4. $\tilde{\Gamma}_0$ の各行をその行の対角要素で割り, 対角要素が1になるよう変形したものを $\hat{\Gamma}_0$ とする。
5. $\hat{B}=I-\hat{\Gamma}_0$ より B の推定値を得る
6. $Z\hat{B}Z'$ が可能な限り下側三角行列になうよう置換行列 Z を見つける。
7. 上側三角部分の要素の有意性が低い要素を0に置き換える等の処理を行う。
8. 上の処理の結果最終的に得られた \hat{B} を使って Γ を $\hat{\Gamma}=I-\hat{B}$ によって推定する。

しかし LiNGAM は多段階の計算過程を含むかなり複雑なアルゴムの集合体であるので, 非正規性の仮定の下で得られた推定値の有意性検定などの統計学的推論を進めることは困難である。事実 LiNGAM では最終的な推定結果に対する有意性検定は行われていない。この点を補うために前川 (2015) 及び前川・Amirullah (2016) はデータをから推定された係数推定値 $\hat{\Gamma}_1=(I-\hat{B})\hat{A}$ の有意性を調べるためにブートストラップ法を用いる検定を提案した。すなわち N 回リサンプルする毎に $\hat{\Gamma}_1$ を計算し, その結果得られた N 個の $\hat{\Gamma}_1$ の平均値 $\bar{\hat{\Gamma}}_1$ を Γ_1 の推定値とする。またブートストラップ法から得られる $\hat{\Gamma}_1$ の標準偏差 $SD_{\hat{\Gamma}}$ を使って t 値を

$$t = \frac{\bar{\Gamma}_1}{SD_{\bar{\Gamma}} / \sqrt{N}} \quad (7)$$

とする。これは厳密には t 分布に従わないであろうが、これを目安として用いる。そして大まかな判断基準として (rule of thumb) として、この t 値が 2.0 より大きければ有意、それより小さければ有意でないと見なし、有意でないと判定された推定値は 0 に置き換える。すなわち $t < 2.0$ ならば $\bar{\Gamma}_1 = 0$ と見なすことにする。係数行列 B の推定も同様に行うことができる。さらに次節に於いて、LiNGAM が示唆する因果序列を帰無仮説として最尤比検定を適用することによってクロスチェックを試みる。この手続きの有効性をモンテカルロ実験によって検証する。

4. モンテカルロ実験

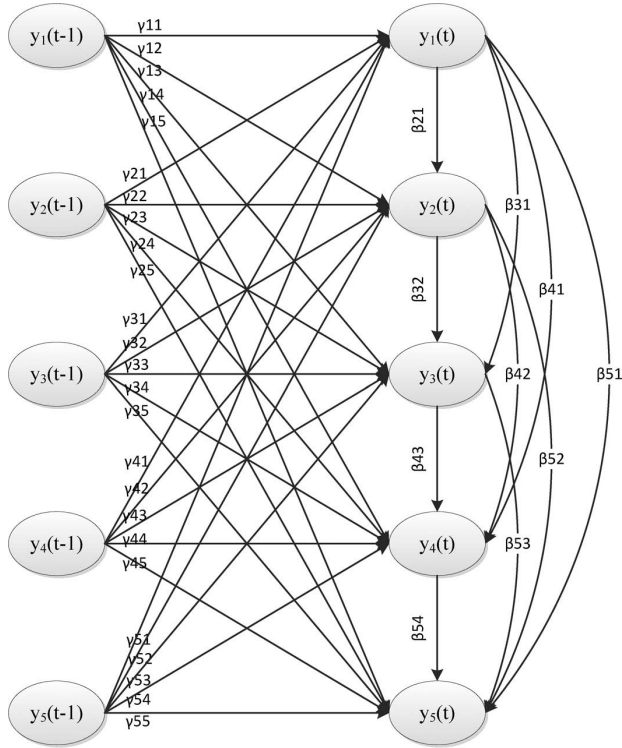
この節では説明を簡略化するために、因果序列が定まっている 5 変数から成るラグ 1 の非ガウス型 SVAR モデル

$$y_t = B y_t + \Gamma_1 y_{t-1} + \varepsilon_t \quad (8)$$

を考え、このモデルを使って LiNGAM のパフォーマンスをモンテカルロ実験によって評価する。実験に際して係数行列 B を以下のように下側三角行列に特定化する。また係数行列 Γ_1 も下側三角行列の場合とそうでない場合を扱う。すなわち係数行列 B と Γ_1 は

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & 0 & 0 \\ \beta_{51} & \beta_{52} & \beta_{53} & \beta_{54} & 0 \end{bmatrix} \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} & \gamma_{15} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} & \gamma_{25} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} & \gamma_{35} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} & \gamma_{45} \\ \gamma_{51} & \gamma_{52} & \gamma_{53} & \gamma_{54} & \gamma_{55} \end{bmatrix} \quad (9)$$

である。この時、変数間の同期間内での及び 1 期遅れを伴う因果関係は以下の図のように表される。



係数行列 B は同時点における因果序列を表わす。この場合の因果序列は $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \rightarrow y_{4t} \rightarrow y_{5t}$ である。 B が与えられると、定義により

$$\Gamma_0 = I - B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \beta_{21} & 1 & 0 & 0 & 0 \\ \beta_{21} & \beta_{21} & 1 & 0 & 0 \\ \beta_{21} & \beta_{21} & \beta_{21} & 1 & 0 \\ \beta_{21} & \beta_{21} & \beta_{21} & \beta_{21} & 1 \end{pmatrix}$$

となる。また誘導形の係数行列と攪乱項は

$$y_t = \Gamma_0^{-1} \Gamma_1 y_{t-1} + \Gamma_0^{-1} \varepsilon_t = A y_{t-1} + u_t \quad (2)$$

より $A = \Gamma_0^{-1} \Gamma_1$, $u_t = \Gamma_0^{-1} \varepsilon_t$ で与えられる。以下では、係数行列に関して次の二つのケースを考える。

Case A 同時点においても異時点においても因果序列がある場合。

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 & 0 \\ 0.2 & 0.3 & 0.2 & 0 & 0 \\ 0.1 & 0.2 & 0.2 & 0.1 & 0 \end{pmatrix} ; \Gamma_1 = \begin{pmatrix} 0.6 & 0 & 0 & 0 & 0 \\ 0.4 & 0.5 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.5 & 0 & 0 \\ 0.3 & 0.3 & 0.4 & 0.5 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.5 \end{pmatrix}$$

なお Case B として、同時点においては因果序列があるが、異時点間においてはそれがない場合として、以下のような想定も考察した。

Case B 同時点においては因果序列があるが、異時間においてはそれがない場合

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 & 0 \\ 0.2 & 0.3 & 0.2 & 0 & 0 \\ 0.1 & 0.2 & 0.2 & 0.1 & 0 \end{pmatrix} ; \Gamma_1 = \begin{pmatrix} 0.6 & 0.4 & 0.3 & 0.3 & 0.2 \\ 0.4 & 0.5 & 0.4 & 0.3 & 0.2 \\ 0.3 & 0.4 & 0.5 & 0.4 & 0.5 \\ 0.3 & 0.3 & 0.4 & 0.5 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.5 \end{pmatrix}$$

これらの二つのケースについて、以下の手順で人工的観測値 y_t を発生させる。

Γ_1 の想定の違いは以下の結果に大きな影響を与えなかったので、本稿では Case A のみの結果を報告する。

Step 1. 係数行列 B と Γ_1 の値を上のように特定化し、 $\Gamma_0 = I - B$ と $A = \Gamma_0^{-1}\Gamma_1$ を計算する。

Step 2. 非正規分布に従う攪乱項 $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ として自由度 5 の t 分布から独立な乱数を発生させる。得られた $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ を使って $u_t = \Gamma_0^{-1}\varepsilon_t$ を計算する。

Step 3. 上で得られた u_t を (2) に代入して y_t を発生させる。この y_t を“観測値”と見なす。

Step 4. 上で得られた“観測値” y_t を使って誘導形 (2) を最小 2 乗推定し、残差 \hat{u}_t を求める。

Step 5. 残差 \hat{u}_t を LiNGAM に与え係数行列の推定値 \hat{B} を求める

以下 Case A のモンテ・カルロ実験結果を示す。B と Γ_1 が上の Case A のように与えられると、 $\Gamma_0 = I - B$ と $A = \Gamma_0^{-1}\Gamma_1$ の関係から Γ_0^{-1} と A は

$$\Gamma_0^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.8 & 1 & 0 & 0 & 0 \\ 0.92 & 0.4 & 1 & 0 & 0 \\ 0.624 & 0.38 & 0.2 & 1 & 0 \\ 0.5064 & 0.318 & 0.2 & 0.1 & 1 \end{pmatrix}; A = \begin{pmatrix} 0.6 & 0 & 0 & 0 & 0 \\ 0.88 & 0.5 & 0 & 0 & 0 \\ 0.012 & 0.6 & 0.5 & 0 & 0 \\ 0.8864 & 0.57 & 0.5 & 0.5 & 0 \\ 0.727 & 0.477 & 0.35 & 0.25 & 0.5 \end{pmatrix}$$

となる。5 変数の場合、可能な因果序列の場合の数（以後、種類という）は $5! = 120$ 個ある。

Case A では、同時点内での真の因果序列は $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \rightarrow y_{4t} \rightarrow y_{5t}$ である。この真の序列を LiNGAM がどの程度正しく検出するかを、モンテ・カルロ実験で調べるために、以下のような実験を行った。Step 3 で得られた長さ $T=200$ の時系列「観測値」 y_t からブートストラップ法を用いて 1,000 回リサンプルし、Step 4～5 に従って、係数行列の推定値 \hat{B} を求め、因果序列を求める。その際、リサンプルされたサンプルは時系列構造を保存するために時間の順序に並び替える。リサンプルされた 1,000 個の時系列に対して LiNGAM を使って因果序列を求める。その結果、1,000 個の因果序列が得られる（これらの 1,000 個の中には重複するものが多数ある）。その結果を使って、次の二つの方法によって真の因果序列の推定を試みた。第一の方法は、1,000 回のブートストラップのうち、120 種類の順列がそれぞれ何回出現したかを調べる方法である。順列番号 90 から 120 までの検出された頻度は下の表 1 のようであった（順列 1～89 は頻度がほとんど 0 なのでここでは省略。）

この結果から、1,000 回中 $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \rightarrow y_{4t} \rightarrow y_{5t}$ という序列が 258 回検出され、検出頻度が最も高かったので、推定因果序列として $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \rightarrow y_{4t} \rightarrow y_{5t}$ を採用する。このように検出頻度が最も高い順列を用いる推定方法を「方法 1」または「的中率法」と呼ぶことにする。

表1 順列の検出頻度

番号	順列					乖離度 ^{注)}	検出頻度
90	2	4	1	3	5	10	2
91	2	1	5	3	4	8	10
92	2	1	5	4	3	10	8
93	2	1	3	5	4	4	24
94	2	1	3	4	5	2	57
95	2	1	4	5	3	8	16
96	2	1	4	3	5	4	14
97	1	5	3	4	2	18	1
98	1	5	3	2	4	14	5
99	1	5	4	3	2	20	3
100	1	5	4	2	3	18	3
101	1	5	2	3	4	12	7
102	1	5	2	4	3	14	2
103	1	3	5	4	2	14	1
104	1	3	5	2	4	10	5
105	1	3	4	5	2	12	1
106	1	3	4	2	5	6	15
107	1	3	2	5	4	4	53
108	1	3	2	4	5	2	99
109	1	4	5	3	2	18	2
110	1	4	5	2	3	16	3
111	1	4	3	5	2	14	1
112	1	4	3	2	5	8	9
113	1	4	2	5	3	10	10
114	1	4	2	3	5	6	27
115	1	2	5	3	4	6	73
116	1	2	5	4	3	8	40
117	1	2	3	5	4	2	113
118	1	2	3	4	5	0	258
119	1	2	4	5	3	6	48
120	1	2	4	3	5	2	69

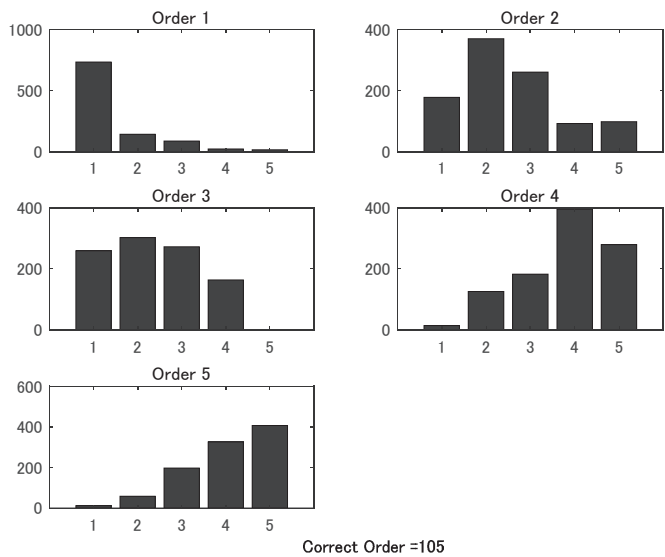
注) 乖離度の定義：真の因果序列（12345）と検出された

序列 (x_1, \dots, x_5) との偏差の2乗和 $\sum_{i=1}^5 (i - x_i)^2$ を両者の乖離度とした。

第2の方法は、下記の図3を使って説明する。この5つグラフの横軸は変数番号、縦軸は頻度をあらわす。また5つのヒストグラムに付されている Oder 1, Order 2, 等は順列の中の位置を示している。例えば上段左図

(Order 1) は、変数 1 は順列の 1 番目の場所 (Order 1) に 848 回、変数 2 は 139 回、変数 3 は 6 回、変数 4 は 4 回、変数 5 は 3 回出現したことを示している。図 3 のその他のパネルも同様に読む。この図のそれぞれのヒストグラムの頻度表はヒストグラムの下に表示されている。その結果は、変数 1 は順列の 1 番目の場所に、変数 2 は 2 番目の場所に、変数 3 は 3 番目の場所に、変数 4 は 4 番目の場所に、変数 5 は 5 番目の場所に現れる頻度が最も高いことが分かる。このことから、因果序列は (1 2 3 4 5)、すなわち $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t} \rightarrow y_{4t}$ である可能性が最も高いと判断することにする。このような方法を「方法 2」または「ヒストグラム法」と呼ぼう。上の実験結果では真の因果序列が正しく検出されたわけである。われわれは、このようなブートストラップ法による実験を何回も行ったが、「方法 1」または「的中率法」が真の序列を検出する頻度は、20~30%であったのに対して、「方法 2」または「ヒストグラム法」ではその頻度は 95~100%であった。このことから「方法 2」の方が優れているといつてよいであろう。なお実験では、Case A と Case B について、非正規分布に従う攪乱項 $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ として自由度 5 の t 分布から独立な乱数を発生させた。

ここでサンプルサイズについて触れておこう。実データによる経済時系列分析において現実的なサンプルサイズ (時系列の長さ) として $T=50$ と 200 の 2 通りを想定して実験を行った。また対照のため標本サイズが大きい場合として $T=1,000$ の場合も想定した。以下の図 2~4 にサンプルサイズごとのヒストグラム法の結果を示す。各グラフとも、横軸は変数番号を、縦軸は 1,000 回中の頻度である。下の表 2 から分かるように、「方法 1」の真の順列の的中率は T に関わらず必ずしも高くない。しかし「方法 2」のヒストグラムによる判定法ではいずれの場合も正しい順位を検出した。しかしながら $T=50$ の場合には、因果序列の探索結果が甘くなる傾向にある。 $T=50$ の場合の、方法 1 の結果は次の図 2 のようになった。因果序列の 3 位 (Order 3 のグラフ) からは変数 2 と変数 3 の頻度が接近しており序列が明確ではない。序列の 4 位と 5 位の位置には正しく変数 4 と変数

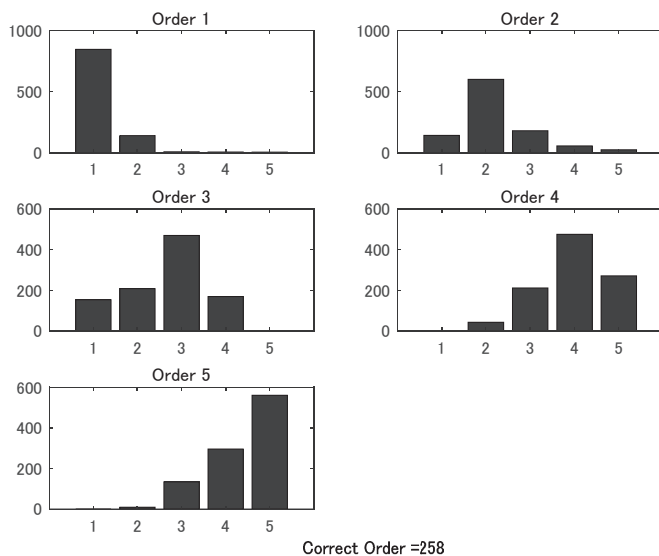


各グラフの度数分布

変数番号 →	1	2	3	4	5
Order 1	736	142	86	21	15
Order 2	178	371	261	92	98
Order 3	260	303	273	164	0
Order 4	14	126	183	397	280
Order 5	12	58	197	326	407

図 2 T=50の場合

5 が来ている。また真の因果序列が正しく検出された比率はわずか10%程度であった。この結果は LiNGAM によって正しい因果序列を検出するためには、T=50では不十分であることを示している。T=200の場合は比較的明瞭な結果が得られたが、T=1,000の大標本の場合は意外なことに順位 4 と 5 の区別が明確でなかった。事実シミュレーションでは順列 (1 2 3 4 5) が1,000回中200回選らばれた。この結果は本研究で行った多くのシミュレー

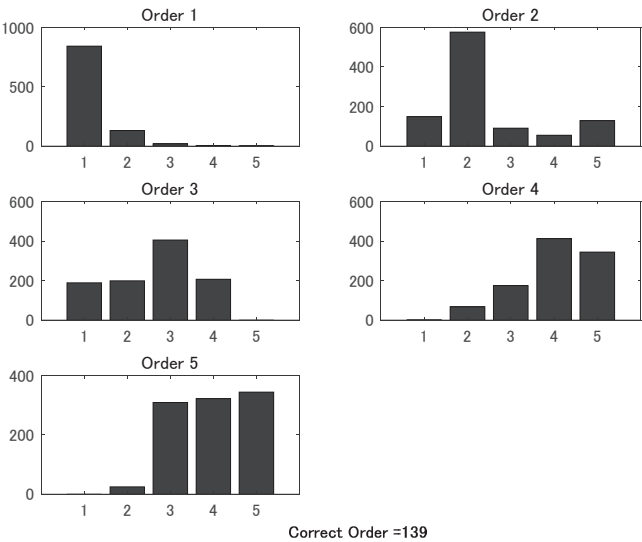


各グラフの度数分布

変数番号 →	1	2	3	4	5
Order 1	848	139	6	4	3
Order 2	141	602	179	55	23
Order 3	153	208	470	169	0
Order 4	0	42	211	476	271
Order 5	0	9	134	296	561

図3 T=200の場合

シミュレーション実験の中で、方法1（的中率法）と方法2（ヒストグラム法）の結果が一致しなかった唯一の例である。想定されたモデルがそういうモデルであったということなのかもしれない。この結果からT（標本サイズ）は少なくとも100以上は必要なのではないかと思われる。



各グラフの度数分布

変数番号	→	1	2	3	4	5
Order 1		842	132	20	4	2
Order 2		149	576	91	55	129
Order 3		189	199	405	207	0
Order 4		1	68	175	412	344
Order 5		0	25	309	322	344

図 4 T=1,000の場合

表 2 LiNGAM による真の因果序列検出率

T	50	200	1000
的中率	0.105	0.258	0.139

5. 最尤法による推定と検定

第 3 節の末尾で述べたように、LiNGAM によって得られた因果序列の推

測統計学的な有意性検定をおこなうことは困難であるので、本節では次のように2段階に分けて考えることにする。すなわち第一段階として LiN-GAM から得られた因果序列を因果序列の候補とみなし、第二段階としてその候補が示す因果序列（係数行列が下側三角行列）を帰無仮説とみなし、対立仮説は「因果序列は存在しない」とする。そしてモデル（1）から導かれる t 分布に基づく尤度関数を帰無仮説と対立仮説の下で評価し、尤度比を求め、尤度比検定をおこなう。その際、攪乱項に関する仮定はモデル（1）の仮定を変えない。すなわち攪乱項 $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}$ は独立にそれぞれ自由度5の t 分布に従うこと、及び時間的にも独立であると仮定する。

対数尤度関数の導出： 上のような仮定の下で得られる対数尤度関数の形を示しておく。いま ε_{it} の t 分布密度関数を $f_i(\varepsilon_{it}; \lambda_i, \sigma_i)$ とする。 λ_i, σ_i は自由度と標準偏差である。ここで $\varepsilon_t = \Gamma_0 u_t$ を代入すれば、対数尤度関数は次のように表される。

$$L_T(\theta) = T^{-1} \sum_{t=1}^T l_t(\theta)$$

ここに θ はモデル（1）の係数行列 B と誘導形に現れる係数行列 A に含まれるすべての未知パラメータと定数項 μ 、及び λ_i, σ_i を含むベクトルである。また $l_t(\theta)$ は次の式で表される。

$$l_t(\theta) = \sum_{i=1}^n \log f_i(\sigma_i^{-1} l'_i \Gamma_0 u_t; \lambda_i) - \log \det(B) - \sum_{i=1}^n \sigma_i$$

ここで l'_i は i 番目の要素が1、それ以外の要素は0となる $1 \times n$ の横ベクトルである。 l'_i を左から掛けることによって $\Gamma_0 u_t$ の i 番目の要素（すなわち ε_i ）が取り出される。この対数尤度関数を θ に関して最大にすることによって最尤推定量が求められる。しかしパラメータの数が多いので、Lanne, M., Meitz, M., Saikkonen, P. (2017) は2段階による最大化を推奨している。そのために、 $\theta = (\theta_1, \theta_2)$ のように2つの部分に分ける。ここに θ_2 は誘導形に含まれる係数行列 A の要素からなり、 θ_1 はそれ以外の未知パラメータを含む。まず第1段階として誘導形に最小2乗法を適用して残差 \hat{u}_t を求め

る（これによって係数行列の OLS 推定値 \hat{A} が求まり、 θ_2 が推定される）。この残差 \hat{u}_t を尤度関数の中の u_t に代入する。代入後の尤度関数を

$$\hat{L}_t(\theta_1) = T^{-1} \sum_{t=1}^T \hat{l}_t(\theta_1)$$

と置き、第 2 段階として $\hat{L}_t(\theta_1)$ を θ_1 に関して最大化することによって、 θ_1 の最尤推定量が得られ、すべての推定が終了する。以上の計算過程ではパラメータに何の制約も置かず、最尤推定が行われた。次に帰無仮説の制約の下で対数尤度を最大化する。そして制約なしと、制約付きの最大化された対数尤度の差を求めれば尤度比が得られる。

このようなフレームワークの下で LiNGAM が示した因果序列を帰無仮説とする尤度比検定を行うことによって、LiNGAM が検出した因果序列の妥当性が検証される。このような尤度比検定の有効性をモンテカルロ実験によって確かめてみる。本論で用いているモデル（1）は識別可能であるが、実験に際しては、この識別可能という事前情報を与えないで Lanne, Meitz, and Saikkonen（2017）が示した「識別性の枠組（Identification Scheme）」に沿って推定を進めた。この Identification Scheme については補論で紹介する。

次に、先に示した Case A のモデルを DGP として、時系列の長さ $T=200$ 、繰返し回数 100 回のモンテカルロ実験により MLE 推定と尤度比検定のパフォーマンスを調べた結果を下に示す。

係数行列 Γ_0^{-1} の下側三角部分の真の値（帰無仮説）を再掲すると

$$H_0: \Gamma_0^{-1} = \begin{pmatrix} 1.0000 & 0 & 0 & 0 & 0 \\ 0.8000 & 1.0000 & 0 & 0 & 0 \\ 0.9200 & 0.4000 & 1.0000 & 0 & 0 \\ 0.6240 & 0.3800 & 0.2000 & 1.0000 & 0 \\ 0.5064 & 0.3180 & 0.2200 & 0.1000 & 1.0000 \end{pmatrix}$$

である。対立仮説は

$$H_1: \Gamma_0^{-1} \text{ は下側三角行列ではない}$$

(「因果序列は存在しない」即ち上側三角部分に何の制約も置かない)

である。まず100回の繰り返し推定した結果、下側三角部分の真値とブートストラップ法による100回のリサンプリングで得られた100個の推定の平均値は、以下ようになった。

(1) 制約なしの場合

真 値 : 0.8000, 0.9200, 0.6240, 0.5064, 0.4000, 0.3800, 0.3180, 0.2000, 0.2200,
0.1000

平均値 : 0.7874, 0.9129, 0.6119, 0.4771, 0.3982, 0.3478, 0.2995, 0.2101, 0.2266,
0.1335

(2) 制約ありの場合

平均値 : 0.8013, 0.9241, 0.6241, 0.5022, 0.4063, 0.3722, 0.3178, 0.2002, 0.2164,
0.1161

(3) 制約なしの場合の上側三角部分の係数(真値はすべてゼロ)の推定結果の

平均値 : 0.0339, 0.0162, 0.0478, 0.0195, 0.0502, 0.0407, 0.0321, 0.0463, 0.0402,
0.0314

これらの結果を見ると最尤法による推定結果は概ね正確であることが視われる。ただし上記(3)は真値0に対する推定値が大きすぎるように見えるが、誤差は小数第2位の桁なのでさほど大きな誤差とは言えない(事実、このシミュレーションによれば真値=0の全ての係数に対してt値を計算したら、帰無仮説: 真値=0は全て棄却できなかった)。

以下に100回のリサンプリングから得られた100個の推定値の経験分布を図5, 6, 7に示しておく。全てのヒストグラムにおいて標本分布は真値の回りに分布していることが分る。

尤度比検定の結果は100回の繰り返し計算の結果、ちょうど95回帰無仮

説は棄却できなかった。したがって有意水準 5 % で LiNGAM が示唆した因果序列（帰無仮説）は採択されるという結果が得られた。検出力に関しても検証すべきであるが、ここで用いたモデルは小さいとはいえ、パラメータの数が多いので対立仮説の設定が難しいので今回はその検証は行な

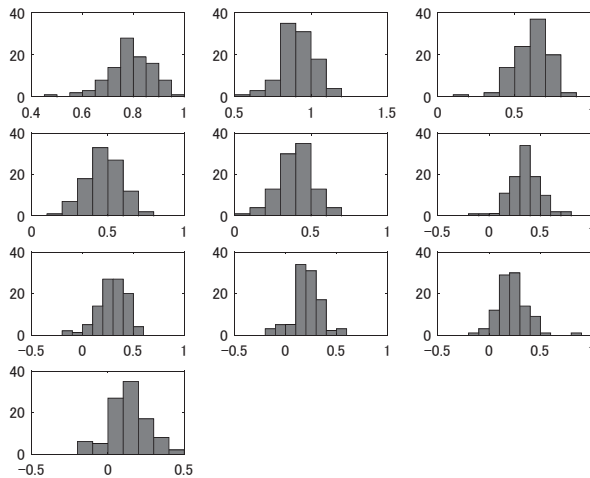


図 5 制約なしの場合における係数行列 B の下側三角部分の推定値の分布

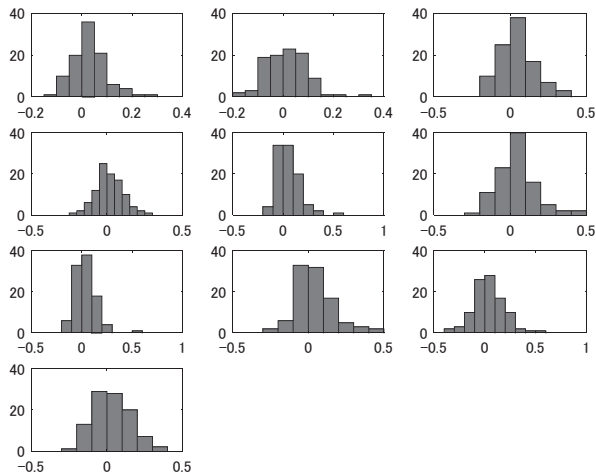


図 6 係数行列 B の上側三角部分の推定値の分布

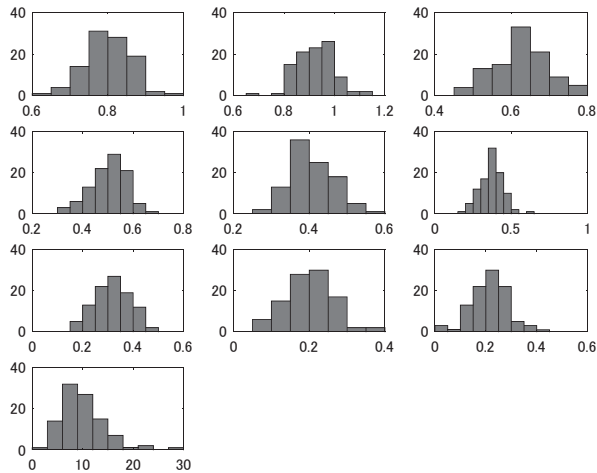


図7 制約ありの場合における係数行列 B の下側三角部分の推定値の分布

わなかった。検定法に関しては今後さらに、モンテカルロ実験を継続し検定の精度を高めたい。

6. 応用例 日本のマクロデータ 5 変数モデル

前川, 小村, 永田 (2015) (以下 MKN 論文) は, 金融の量的緩和政策 (マネタリーベースの量的緩和政策) の波及経路の分析を行うために, 次の 5 変数からなるラグ 1 の VAR モデルを用いた。推定に使用されたデータ期間は 2009 年 1 月～2014 年 12 月である。

変数番号 1 : マネタリーベース (対数値: $\ln mb$),

変数番号 2 : 予想インフレ率 ($arinf$), (物価連動国債価格から算出。補論 2 または前川, 小村, 永田 (2015) 参照)

変数番号 3 : 総合消費 (ci),

変数番号 4 : 有効求人倍率 (raj),

変数番号 5 : 消費者物価指数 (季調) (cpi_sa)

MKN 論文では VAR モデルを推定し, インパルス応答関数を参考にしながら経験則並びに経済学的観点を交えながら総合的に因果序列を考察すると

いう VAR 分析の標準的な推論をおこなった。

これに対して、本論ではデータ期間を延長し、SVAR モデルを用いて LiN-GAM によって、同時点内、及び異時点間の因果序列を推定した。データ期間は2012年1月から2014年12月間の月次データを用いた。またブートストラップにおけるリサンプル回数は10,000とした。以下に得られた結果の要点のみを示す。

まず誘導形（VAR モデル）の残差の非正規性を Jarque-Bella 検定を用いてチェックした。その結果5変数とも正規性は棄却されたので、LiN-GAM の前提は満たされた。

次に推定精度を調べるために、与えられたデータからブートストラップ法で10,000回リサンプルを行いBの推定を行った。ここではその結果得ら

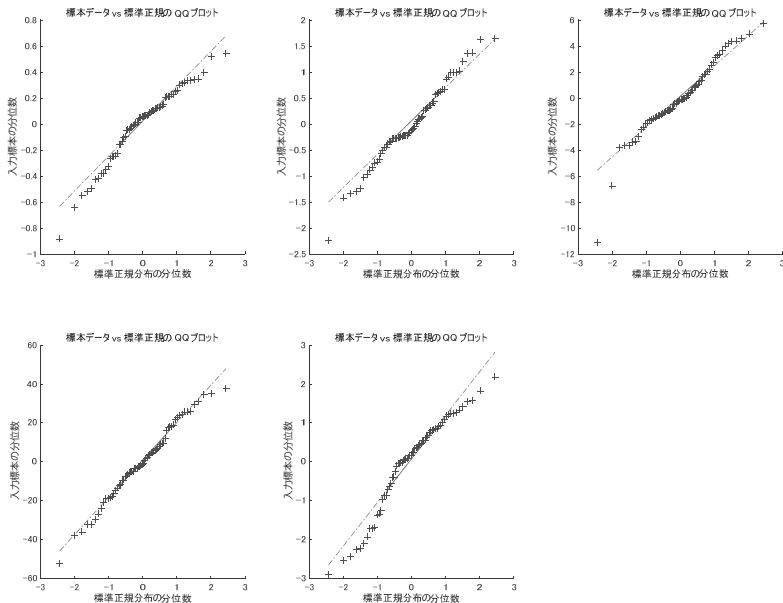


図8 Jarque-Bella 検定の結果

変数1: 6.12, 変数2: 0.287, 変数3: 25.15, 変数4: 0.627, 変数5: 5.35
変数2と4以外は、正規性が棄却される。

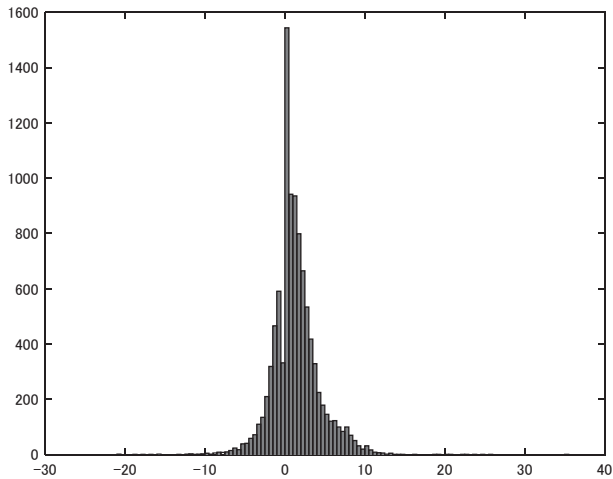


図9 係数 B (2,1) の推定値の経験分布
(10,000回のブートストラップ)
平均 = 1.2974, 標準偏差 = 2.9592

れた B (2,1) の推定値の標本分布のみを図9に示しておく。

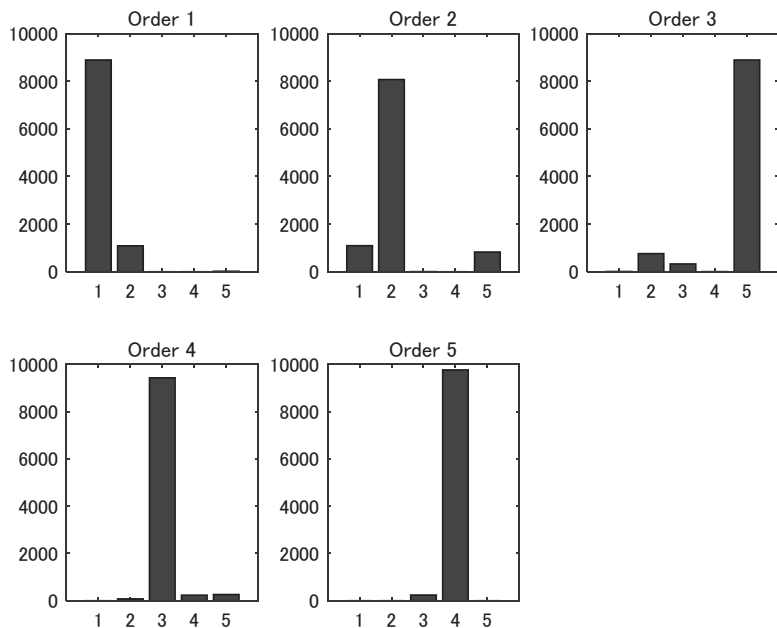
この結果から機械的に平均値の分布の t 値を計算し t 値 = 43.8425 を得た。したがって係数 B (2,1) は有意に 0 ではないと言える。最も厳しい基準（有意水準 1 %）で検定し、有意でなかった箇所は 0 に置き換えた結果は以下ようになった。

係数行列 B =

$$\begin{pmatrix} 0 & 0.0327 & 0 & 0 & 0 \\ 1.2974 & 0 & -0.0012 & 0 & -0.0285 \\ 13.9965 & -0.5075 & 0 & -0.0015 & -1.8652 \\ 79.1440 & 1.3214 & -0.3533 & 0 & -6.1145 \\ 6.9320 & -0.8138 & -0.0109 & 0 & 0 \end{pmatrix}$$

10,000回のリサンプルから得られた因果序列の頻度は、以下ようになった。

上の図の見方は、前節と同じである。Order k: 因果序列の順位，横軸：変数番号，縦軸：10,000回のリサンプリングにおける出現頻度。例えば左上の



横軸：1 = マネタリーベース，2 = 予想インフレ率，3 = 有効求人倍率，
4 = 総合消費，5 = 消費者物価指数

図10 MKN マクロ 5 変数 SVAR モデルの因果序列の探索結果

グラフは、10,000回のリサンプリングにおいて、変数1が因果序列の1番目に来た回数が約9,000回であったことを示している。このように見ていくと、変数1の因果序列は1番目、変数2の序列は2番目、変数3の序列は5番目、変数3の序列は3番目、変数4の序列は4番目と推測される。すなわち、ブートストラップ法から次のような同時点内因果序列が推測される：マネタリーベース（1）⇒ 予想インフレ率（2）⇒ 消費者物価指数（5）⇒ 総合消費（3）⇒ 有効求人倍率（4）。（金融変数から実体経済への波及効果が見られる）

上に推定されたSVAR（5）の係数行列Bと Γ_{t-1} の推定値を使って因果序列図を作成することが出来る。図11は因果序列図の一例である。（なおこの例は、同じモデル、同じデータから1,000回のリサンプルによって得

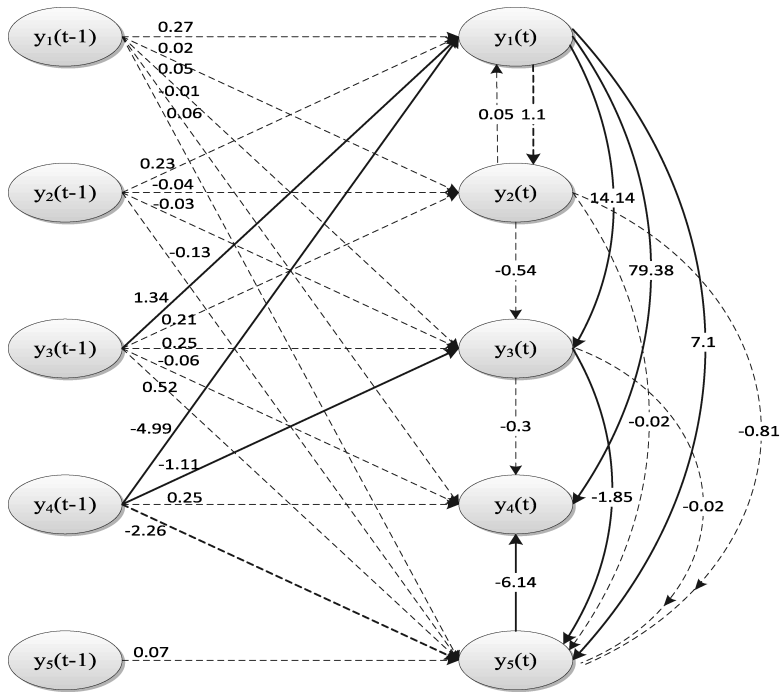


図11 同期及び1期遅れまでの因果序列図

られた推定値の平均に基づいている)。この図からも期待されるような因果序列がほぼ満されていることを読み取ることができる。

これに対して MKN 論文の VAR 分析ではマネタリーベース (1) ⇒ 予想インフレ率 (2) ⇒ 有効求人倍率 (4) ⇒ 総合消費 (3) ⇒ 消費者物価指数 (5) という因果序列が得られた。結論がこのように一致しない理由としては、MKN 論文で使用されたデータ期間は2009年1月～2014年12月であるのに対して、本研究で用いられたデータ期間は2012年1月～2014年12月であり、期間が大きく異なる点にあると思われる。

7. 宮尾モデル (2016) における因果序列

もう一つの LiNGAM の応用例として、この節では宮尾 (2016) 第3章

「非伝統的金融政策の効果はあるのか (II)」で採用されている日本の5変数マクロ金融モデルを借用して分析を試みた。ただしここでの分析と同書の分析に於いて以下の点で異なる。1) 本稿で用いたデータはネット上から採取されたものであるが、同書に採用されているデータと全く同じではない、2) 同書で用いられている四半期 GDP 時系列の月次変換法が不明であるので、本稿では単純に移動平均法を使用した、3) 同書では2014年4月以降は消費税増税による外生的な影響を除去するためにダミー変数を追加しているが、本稿ではダミー変数は用いなかった、4) 同書ではコア・コア消費者物価指数が採用されているがここではコア物価指数を採用した。

6 節と同様に ICA 分析に先立ち、誘導形残差の非正規性を Jarque - Bella 検定によってチェックした結果、図12に示されるように全ての残差の非正規性が確認された。

図13のグラフは、5変数の因果序列は $4 \Rightarrow 3 \Rightarrow 1 \Rightarrow 2 \Rightarrow 5$ すなわちマネタリベース \Rightarrow 長期金利 \Rightarrow 為替レート \Rightarrow インフレ率 \Rightarrow GDP である

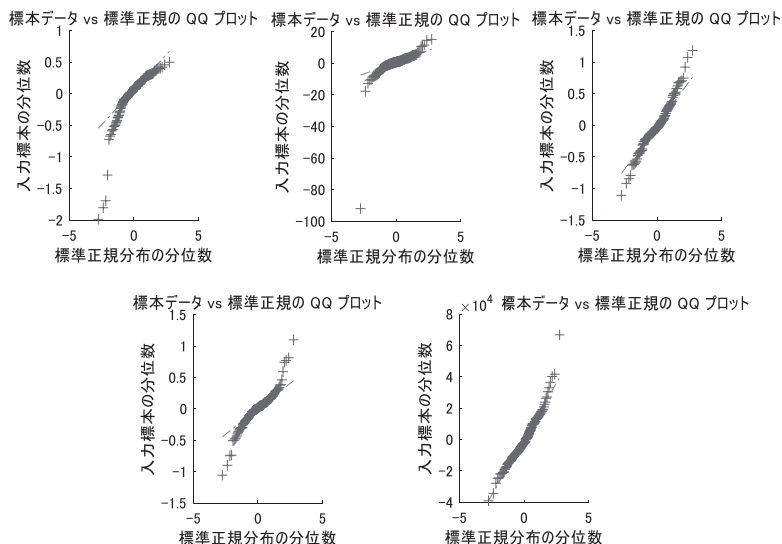
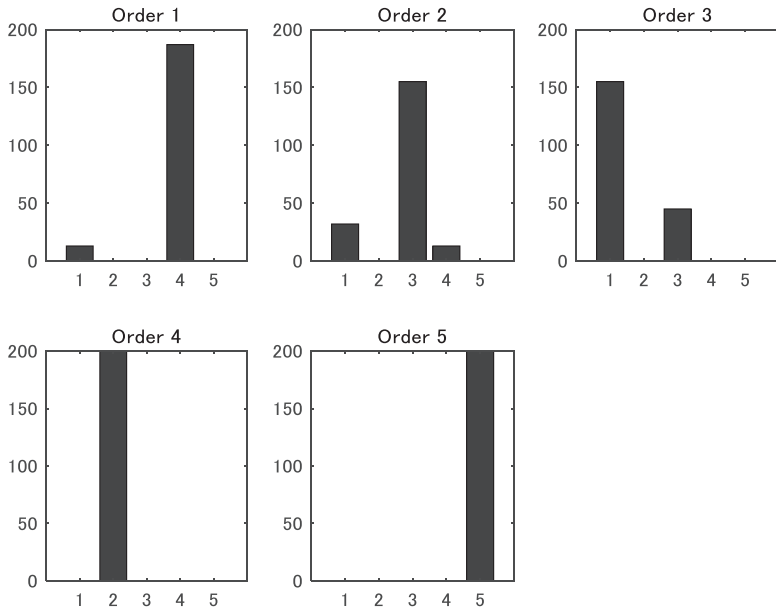


図12 Jarque - Bella 検定の結果



横軸：1 = 為替レート，2 = インフレ率（対前年同月比），3 = 長期金利，4 = マネタリーベース，5 = GDP

図13 宮尾マクロ5変数SVARモデルの因果序列の探索結果

ことを示している。金融変数から実体経済への波及経路が検出されている。上に述べたように宮尾（2016）と本稿は、用いたデータ、モデルの設定、因果序列決定方法等、全く同じではないので、両者の結果は当然異なるが、どちらもある程度経済的に自然な因果序列を検出している。しかしながらここで第5節に示した尤度比検定を用いて、宮尾モデルからLiNGAMによって推定された因果序列を帰無仮説とする尤度比検定を行った結果、帰無仮説は採択されなかった。

8. 残された課題

以上、限られた想定の下ではあるが、本稿で示されたモンテカルロ実験結果はICA-LiNGAMのSVAR分析への応用可能性及び有効性をかなりの

程度示しているといえる。しかしながら今後より一層、経済実証分析に耐えうる手法として確立するためには、次に示すような課題、特に統計学的な理論的課題を解決していく必要がある。

(1) SVAR モデルの係数行列 B は、同時点内での因果序列を表すが、上の実証例ではデータは月次データである。従て推定された係数行列 \hat{B} は同期内（1 か月内）での因果序列を表すものと解釈される。しかしここで注意すべき点は、このような解釈が可能であるのは、観測値の周期（観測データが採取される周期）が因果関係の周期（因果関係が現れるのに必要な期間）より長い場合である（Hyvärinen et al (2010) を見よ）。実証分析に於いては因果関係の周期は変数間で様ではないし、そもそもその周期自体が不明である。この問題に対する実際的な対応としては、同時的因果序列を表す係数の有意性から判断せざるを得ない。しかしながら、因果関係の周期についてはモデルを構築する前に何らかの検討を加えておくべきであろう。もう一点 SVAR モデルの解釈について述べておく。このモデルは係数行列 B を通して同時点内の因果序列のみを表しているように思われがちであるが、実は B は遅れのある変数の係数にも（ $\Gamma_0^{-1}\Gamma_1$ 等の形式を通して）影響を持つので、 B は遅れのある変数間の因果序列にも影響する。いずれにせよ LiNGAM が出力する因果序列を鵜呑みにするのではなく、様々な面から検討を加えて慎重に判断するべきであろう。例えば（本稿では行わなかったが）インパルス応答関数からも因果序列を推察してみるなどの多角的な検討が必要がある。これらの問題点に関しては引き続き検討する予定である。

(2) 本論で述べたように、ICA-LiNGAM はかなり複雑なアルゴリズムの集合体であり、その中には統計学的な推論になじまない面もある。本論ではその点を補足する意味でブートストラップ法を多用することによって、有意性の検定を試みた。しかし今後は、このような方法の理論的基礎を推測統計学的な立場から（少なくとも漸近理論的に）を固めておく必要がある。

(3) 5節でみたように、モンテカルロ実験のように真の DGP が分かっ

ている場合は、LiNGAM が検出した因果序列を、尤度比検定で追認することはできた。しかしながら実証分析のように真のモデル (DGP) が未知の場合は、LiNGAM が検出した因果序列を帰無仮説として (真の DGP が不明のまま) 尤度比検定を実施した場合の検定の精度は保証されない。このような場合の精度の高い検定方法を開発する必要がある。

(4) ファイナンス関係のデータでは、非正規性の仮定は現実的であるが、マクロデータの場合は必ずしも妥当しない場合もある。

以上述べたように、ICA に基づく変数間の因果序列の探索に関しては、様々な課題が残されているとはいえ、本稿に示された方法は第一次接近としては応用に耐えうとの感触が得られた。残された課題については研究を継続し、稿を改めて報告したい。

補論 1 Lanne *et al* (2017) の Identification Scheme に関する覚書

本論の第 1 節で述べたように、構造 VAR モデルには常に識別性の問題が存在する。計量経済学ではこの問題を解決するための一つの方法として、再帰的構造と言われる仮定が置かれる。再帰的構造とは、 $y_{1,t}$ に外生的なショックが加えられたとき、その効果は $y_{1,t} \rightarrow y_{2,t} \rightarrow \dots \rightarrow y_{n,t}$ の順に波及していくことを意味する。しかし事前に波及経路が正確に分かっているという場合は少ないと考えられるので、この仮定は強すぎると言わざるを得ない。これほど強い仮定を置くことなく識別性を確保するための様々な方法が提案されている。一つの方法はモデルの攪乱項が GARCH 過程に従う場合には、GARCH 構造を利用して識別性を獲得するという方向である。Lütkepohl and Netsunajev (2017) は、この分野の最近の動向をサーベイしている。識別性を獲得するもう一つの方向は、この補論で取り上げるモデルの攪乱項に非正規性を仮定する研究である。M. Lanne, M. Meitz and P. Saikkonen (2017) は、この方向の中でもっとも最新かつ包括的な論文である。この論文では、攪乱項に非正規性を仮定することによって識別可能なモデルが構成されることを数学的に厳密に証明すると同時に、識別性を

獲得するための手順（彼らは Identification Scheme と呼んでいる）を示している。ただしこの手順は数学的な構成法を示しているのみで、経済的な意味はまったくない（と同時にモデルの経済的な意味を損なうものでもない）。以下にその概要を、同論文が出版される前のインターネット・バージョン（2015年）に基づいて解説的に記述する。なお関連のある類似の定理として、Comon（1994）、Chan and Ho（2004）、Gourieroux and Zakaoian（2014）、Gourieroux and Monfort（2014）などが存在する。また Identification Scheme は LiNGAM の計算過程とも共通する部分が見られるが、両者の相違は前者は再帰的構造（B 行列が下側三角行列になること）を仮定しないのに対して、後者はそれを仮定する点である。（記号は M. Lanne, M. Meitz and P. Saikkonen（2015）に従う）

1. SVAR モデルの導入と識別性問題

以下で用いる記号は、M. Lanne, M. Meitz and P. Saikkonen（2015）に従う。時点 t における n 次元観測値ベクトルを $y_t = (y_{1,t}, \dots, y_{n,t})$ 、 k 時点前の観測値ベクトルを $y_{t-k} = (y_{1,t-k}, \dots, y_{n,t-k})$ 、 $k = 1, 2, \dots, p$ とする。この時 n 変量 SVAR モデルは、一般に以下のように定式化される。

$$\Gamma_0 y_t = c + \Gamma_1 y_{t-1} + \dots + \Gamma_p y_{t-p} + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (1)$$

ここに c は $n \times 1$ の定数項ベクトル、 Γ_0 は $n \times n$ の対角成分が 1 に等しい係数行列、 $\Gamma_1, \dots, \Gamma_p$ は $n \times n$ の係数行列である。また ε_t は $n \times 1$ の攪乱項ベクトル $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})$ で、一般的には ε_t の各要素は互いに無相関で平均 0、分散 $\sigma_{\varepsilon_i}^2 > 0$ を持つホワイトノイズ、すなわち $\varepsilon_t \sim WN(0, \Sigma_{\varepsilon_t})$ であると仮定される。このとき Σ_{ε_t} は対角行列である。(1) 式の両辺に左から Γ_0^{-1} を掛けると

$$y_t = \mu + A_1 y_{t-1} + \dots + A_p y_{t-p} + B \varepsilon_t \quad (2)$$

となる³⁾。ここに $\mu = \Gamma_0^{-1}c$, $B = \Gamma_0^{-1}$, $A_1 = \Gamma_0^{-1}\Gamma_1, \dots, \Gamma_0^{-1}\Gamma_p$ である。(2) 式は SVAR モデル (1) の誘導形と呼ばれる。この章では (2) 式を用いて SVAR を論じる。また以下においては、 y_t が定常な場合、すなわち

$$\det A(z) \stackrel{\text{def}}{=} \det(I_n - A_1 z - \cdots - A_p z^p) \neq 0, \quad |z| \leq 1 (z \in \mathbb{C}) \quad (3)$$

が成立する場合のみを扱う。

この章では後者の識別性の問題を扱う。識別性の問題とは次のような問題である。(2)式における攪乱項 $B\varepsilon_t$ を非特異行列 C を用いて, $BCC^{-1}\varepsilon_t$ と変形して $B^* = BC, \varepsilon_t^* = C^{-1}\varepsilon_t$ と置けば

$$y_t = \mu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B^* \varepsilon_t^* \quad (4)$$

となる。ここで(2)式と(4)式をそれぞれ MA 表現すると, 二つの MA 表現はまとめて

$$y_t = \mu + \sum_{j=0}^{\infty} \Psi_j B \varepsilon_{t-j} = \mu^* + \sum_{j=0}^{\infty} \Psi_j^* B^* \varepsilon_{t-j}^* \quad (5)$$

となる。(2)式と(4)式を比較すれば(あるいは(5)式をみれば)分かるように被説明変数 y_t は, 攪乱項 $B\varepsilon_t$ の下で生成されたのか $B^*\varepsilon_t^*$ の下で生成されたのかは, 一般的には識別できない。そのような場合, モデル(2)と(4)は識別不能, あるいは観測値上同等のモデルと呼ばれる。識別可能であるためにはモデルに現れる係数行列に関して何らかの先験的情報又は制約が必要となる。以下では部分的, または完全に識別可能であるための条件を検討していく。その検討の出発点となるのが, 上で導入した, ε_t の分散共分散行列 Σ_{ε_t} は対角行列という制約である。この制約の下では, 上に導入した非特異行列 C を全く自由に選ぶことはできず, $C=DO$ という形に限定される。ここに O は直交行列, D は対角行列である。さらに $\Sigma_{\varepsilon_t} = I$ であれば対角行列 D の対角要素は $+1$ または -1 に限定される。以下において識別性を獲得するためには, $C=DO$ という形をどこまで絞り込めばよいかを検討していく。

2. 非正規性の下での識別性問題

2.1 部分的識別可能性

ここで, $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})$ の要素に関して以下のような仮定を置く。

仮定 1

(i) 誤差過程 $\varepsilon_t = (\varepsilon_{1,t}, \dots, \varepsilon_{n,t})$ は定常な確率ベクトルであり, 各要素の平

均 0, 分散 $\sigma_i^2 > 0$ を持つ。

- (ii) 各要素 $\varepsilon_{i,t}, i=1, \dots, n$ は互いに独立で, それらの周辺分布のうち, ガウス分布に従うものは高々 1 個しかない。
- (iii) すべての $i=1, \dots, n$ に対して $\varepsilon_{1,t}$ は自己相関を持たない。すなわちすべての $k \neq 0$ に対して $\text{Cov}(\varepsilon_{1,t}, \varepsilon_{1,t+k}) = 0$ である。

仮定 (i) と (iii) は標準的な仮定であるが, 仮定 (ii) は完全な統計的独立性を仮定するものであり少し強い仮定である。これは (全ての $i \neq j$ に対して) $\varepsilon_{i,t}$ と $\varepsilon_{j,t}$ 独立という弱い仮定を含む。また仮定 (iii) では系列的に無相関ということだけしか仮定されていないので GARCH のよう不均一分散モデルが含まれる。

また仮定 (ii) は ε_t の非ガウス性を仮定しているが, ε_t の要素のどれか一つの周辺分布が正規分布であっても構わないことを含んでいる。前節では非特異行列 C を $C=DO$ という形に絞り込むことが出来ることを示したが, この仮定 1 を置くと以下の定理 1 に示されるように, 置換行列 P を使って $C=DO$ という形に限定することができる。

定理 1 (M. Lanne, M. Meitz and P. Saikkonen (2015, Proposition 1))

モデル (1) は定常性の条件 (3) と誤差過程 ε_t に関する仮定 1 を満たすとする。また (5) 式で示される二つの MA 表現が成立するとする。このとき以下の (i) ~ (iii) によって定められる, 定数項ベクトル μ^* , 係数行列 $\Psi_j^*, j=0, 1, \dots$, 及び誤差過程 $\varepsilon_t^* = (\varepsilon_{1,t}^*, \dots, \varepsilon_{n,t}^*)$ に対して, ある非ゼロな対角要素をもつ対角行列 $D = \text{diag}(d_1, \dots, d_n)$ と, ある $n \times n$ の置換行列 P を用いて全ての t に関して

$$B^* = BDP, \quad \varepsilon_t^* = P'D^{-1}\varepsilon_t, \quad \mu^* = \mu, \quad \text{及び} \quad \Psi_j^* = \Psi_j \quad (j=0, 1, \dots) \quad (6)$$

と表すことが出来る。ただしここにベクトル μ^* , 係数行列 $\Psi_j^*, j=0, 1, \dots$, 及び誤差過程 $\varepsilon_t^* = (\varepsilon_{1,t}^*, \dots, \varepsilon_{n,t}^*)$ は以下のように定義される:

- (i) μ^* は $(n \times 1)$ ベクトル。 B^* は $(n \times n)$ の非特異行列。

- (ii) $\Psi_j^*, j=0,1,\dots$ は, $\Psi^*(z) = A^*(z)^{-1} = \sum_{j=0}^{\infty} \Psi_j^* z^j$ によって決定される $(n \times n)$ 係数行列。ただし $A^*(z)$ は $A^*(z) = I_n - A_1^* z - \dots - A_p^* z^p$ で定義され, 条件 (3) の A を A^* に置き換えた式を満たすものとする。
- (iii) 誤差過程 $\varepsilon_t^* = (\varepsilon_{1,t}^*, \dots, \varepsilon_{n,t}^*)$ は仮定 1 を満たす $(n \times 1)$ ベクトル。

証明：省略。興味のある読者は原論文を参照されたい。

なお, 定理 1 は識別性の問題を完全に解決するものではなく, 行列 B の列の置換, 及びサイズ以外は識別されるというに過ぎない。すなわち平均ベクトル μ と係数行列 $\Psi_j, j=0,1,\dots$ は識別される (または同じことであるが, 自己回帰表現 (2) の定数項ベクトル μ と係数行列 A_1, \dots, A_p は識別される)。以上の結果は, (6) 式が満たされただけでは, (5) 式の二つの MA 表現は依然として観測値上同等なモデルであり, 完全には識別できないことを意味している。しかし SVAR モデルによる y_t の予測, インパルス応答分析, 分散分解を行うためには, 置換とスケールを除いて識別可能であるだけで十分である。

2.2 完全な識別可能性

前節の部分的識別可能性を示す定理 1 では, 置換とスケールまでは識別できなかったが, この節では, 定理 1 を拡張することによって完全に識別可能性を保証する条件を述べる。

先ず以下の説明に必要な次のような記号と用語を導入する。

1. $\mathcal{M}_n : (n \times n)$ の非特異行列の集合
2. $B_1 \sim B_2$: \mathcal{M}_n に含まれる二つの行列 B_1 と B_2 が $B_2 = B_1 D P$ の関係で結ばれているとき B_1 と B_2 は同等と呼び $B_1 \sim B_2$ であらわす。ここに D は非ゼロな対角要素をもつ対角行列 $D = \text{diag}(d_1, \dots, d_n)$, P は置換行列である。

同等関係 \sim は \mathcal{M}_n をいくつかの同等関係が保たれるグループに分割し, 各グループごとに観測値上同等な SVAR モデルが対応している。

以下では、先ず初めに完全識別性の枠組 (Identification Scheme) を示し、定理 2 によって、その枠内ではなぜ完全識別性が得られるかを示す。

完全識別性の枠組み

\mathcal{M}_n に含まれる要素 B (すなわち $B \in \mathcal{M}_n$) に対して以下の変換を順次施す

$$B \rightarrow BD_1 \rightarrow BD_1P \rightarrow BD_1PD_2$$

ここに $(n \times n)$ の行列 D_1, P, D_2 は (存在するときは常に) 以下の性質を満たす。

- (i) D_1 は正定符号の対角行列で、 BD_1 の各行のユークリッドノルムを 1 にするような行列である。
- (ii) P は置換行列で、置換後の行列 $C = BD_1P$ の要素 c_{ij} をとする。このときすべての $i < j$ 対して $|c_{ii}| > |c_{ij}|$ が成立する。
- (iii) D_2 は対角行列で、 BD_1PD_2 の対角要素はすべて 1 になる。

以上のような性質を持つ D_1, P, D_2 によって BD_1PD_2 のような変換が存在する B の集合を $I \in \mathcal{M}_n$ とする。この変換を $\Pi(B) = BD_1PD_2$ と表す。また \mathcal{M}_n における I の補集合を $\varepsilon = \mathcal{M}_n / I$ で表す。ここで、 I に含まれるある B に変換 $\Pi(B) = BD_1PD_2$ を施した結果を \tilde{B} と置く。当然 B も \tilde{B} も I に含まれるので⁴⁾、変換 $\Pi(\cdot)$ は I から I への変換であるから $\Pi(I)$ と書ける。以上を踏まえて I から I への変換 $\Pi(I)$ を定式化すれば以下のようにまとめることが出来る。

$$\Pi(\cdot) : I \rightarrow I \quad B \in \Pi(I) = \{ \tilde{B} \in \mathcal{M}_n \mid \tilde{B} = \Pi(B) \text{ for some } B \in I \}$$

この枠組みは B を識別するためにある特定の置換行列とスケールを決定する行列 D_1 と D_2 を選び出す手順を示している。したがってこの枠組みは次の定理で定式化された範囲で識別問題の解を与えるものである。

定理 2 (M. Lanne, M. Meitz and P. Saikkonen (2015, Proposition 2))

(a) 定理 1 の条件の下で、行列 B は完全識別性の枠組みで定義された集

合 \mathcal{B} の中で一意に識別される。

- (b)⁵⁾ 集合 $I \in \mathcal{M}_n$ に含まれる B_1 と B_1 が同等な関係 $B_1 \sim B_2$ にあるとき、
 $\Pi(\cdot)$ による変換 $\tilde{B}_1 = \Pi(B_1)$ 及び $\tilde{B}_2 = \Pi(B_2)$ は \mathcal{B} 含まれ、しかも
 $\tilde{B}_1 = \tilde{B}_2$ と一意 (unique) に定まる。また B_1 と B_1 が同等な関係にない
 $(B_1 \not\sim B_2)$ 場合は \tilde{B}_1 と \tilde{B}_2 は \mathcal{B} 含まれるが、両者は異なる (distinct),
すなわち $\tilde{B}_1 \neq \tilde{B}_2$ 。

- (c) 完全識別性の枠組から除外される集合 ε は、 $\mathbb{R}^{n \times n}$ 空間においてルベーグ測度は 0 である。また集合 I は \mathcal{M}_n の開かつ密な部分集合である

証明は省略するが、その含意は以下のように要約される。この定理の
(a) は、行列 B を $B = \Pi(I)$ に限定することによって識別性は一意に定まることを示している。(b) は、 B_1 と B_1 に対応する SVAR モデルの識別性の問題に読み替えれば、次のように述べる事が出来る。 I の中で二つの観測値上同等 (識別不能な) な SVAR モデルを変換 $\Pi(\cdot)$ を用いて \mathcal{B} に移行させると、 \mathcal{B} の中では一意 (unique) に識別されるモデルに変換される。また、 I の中で観測値上同等ではない (識別できる) 二つの SVAR モデルに変換 $\Pi(\cdot)$ を施すと、 \mathcal{B} の中でも等しくない (distinct な) 二つのモデルに変換されることを意味している。このことはさらに、(b) は、 \mathcal{B} に含まれるモデルは識別可能であり、また “ほとんど全て” の識別可能なモデルは \mathcal{B} に含まれることを意味している。

さらに簡潔に言えば、集合 \mathcal{B} には観測値上同等な SVAR モデルは含まれない (識別可能なモデルしか含まれていない)。またほとんど全ての観測値上同等でないモデル (識別可能なモデル) は集合 \mathcal{B} に含まれる。

(c) は “ほとんど全て” の意味を述べている。ここまでの議論では集合 I の中だけで SVAR モデルを考えてきたが、ほとんど全ての SVAR モデルは I に含まれている。しかし I の外側、すなわち上で定義した補集合

$\varepsilon = M_n / I$ の中にも非常に数は少ない SVAR モデルが存在している。ところが ε の測度は 0 であることが証明されるので、そのようなモデルは考察の対象から除外してもよい。

補論2 予想インフレ率について（前川・小村・永田（2015）より再掲）

予想インフレ率の算定方には物価連動国債の価格を用いる方法と消費者に対するアンケート調査結果を用いる方法がある。両者の一長一短については MKN 論文で論じられている。前者については本稿で採用したデータ期間全体をカバーすることが出来なかったので、本論では内閣府「消費動向調査」アンケート調査結果を用いた。

内閣府「消費動向調査」には、消費者に 1 年後の物価水準の予測を問うアンケート調査結果が掲載されている。そこには以下示されるような「～%未満～%以上」で区切られた区間（階級）が選択肢として用意されている。そして各選択肢に対して以下のような一つの数値（%）が割り当てられている（以下ではその値を階級値と呼ぶ）。階級値は次のように与えられている。すなわち、回答のうち、「-5%以下」（「-10%以下」と「-10%未満～-5%以下」の合計）に対して -5%，「-5%未満～-2%以下」に対して -3.5%，「-2%未満～」に対して -1%，「0%程度」に対して 0%，「～2%未満」に対して 1%，「2%以上～5%未満」に対して 3.5%，「5%以上」（「5%以上～10%未満」と「10%以上」の合計）に対して 5%をそれぞれ階級値として割り当てる。予想インフレ率は各区間の階級値をそれぞれの階級の回答者の比率をウェイトとする加重平均として計算される。

注

- 1) 本稿は、一橋大学、日本銀行、九州大学、Singapore Management University、ソウル国立大学などにおけるセミナー講演録に基づいて作成されている。これらのセミナーにおいて多くの方々から貴重なコメントを頂いた。また筆者の主催する ICA 研究会メンバーの片山直也関西大学教授、永田修一関西学院大学助教、

- Amirullah, S.H. インドネシアガジャマダ大学講師らとの刺激的な議論は本稿作成にとって非常に有益であった。これらの方々に感謝申し上げる。また本研究は日本学術振興会科学研究費（平成26年—28年基盤研究（C）（26380279））の補助を受けている。
- 2) そのほかの応用研究として、前川・Amirullah（2016）、永田（2016、2017）がある。また因果序列の探索が目的ではないが、Frank et al（2015）も ICA を応用している。
 - 3) このモデルに外生変数ベクトル x_t を追加したモデルが計量経済学における構造方程式モデルである。
 - 4) 当然 B も \tilde{B} も I に含まれる。
 - 5) (b) の原文は “The set B consists of unique and distinct representatives from each \sim -equivalence class of I ” であるが、簡潔すぎて分かりにくい。上の記述は、証明から読み取ることが出来る定理の内容を表現したものである。

参 考 文 献

欧文文献（本文で引用した文献及び本論文と関連の深い文献のみを掲載する。）

- Frank J. Fabozzi, Rosella Giacometti, and Naoshi Tsuchida (2015) The ICA-based Factor Decomposition of the Eurozone Sovereign CDS Spreads, IMES Discussion Paper N0. 2015-E-4, Bank of Japan
- Honda, Y and Y. Kuroki (2006), “Financial and Capital Markets’ Response to Changes in the Central Bank’s Target Interest Rate: The Case of Japan”, 2006, Economic Journal, Vol. 116, No. 513, July, pp. 812–842.
- 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）
- Hyvärinen, Zhang, Shimizu, and Hoyer (2010), Estimation of Structural Vector Autoregression Model Using Non-Gaussianity, J. Machine Learning Research 11, 1709–1731
- Lanne, M and Lütkepohl, H. (2006), Structural vector autoregressions with nonnormal residuals, CESIFO Working paper No. 1651, Category 10: Empirical and theoretical methods
- Lanne, M., Meitz, M., Saikkonen, P. (2017), Identification and estimation of non-Gaussian structural vector autoregressions. J. Econ. forthcoming.
- Lütkepohl (2012), Identifying Structural Vector Autoregressions via Changes in Volatil-

- ity, Department of Economics, Freie Universität Berlin and DIW Berlin
- Lütkepohl and Netšunajev (2017), Structural vector autoregressions with heteroskedasticity: A review of different volatility models, *Econometrics and Statistics* Vol. 1, pp. 2-18
- 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
- Shimizu, Hoyer, Hyvärinen, Kerminen (2006), A Linear non-Gaussian acyclic model for causal discovery, *Journal of Machine Learning Research*, Vol. 7

日本語文献 (量的緩和政策に関する夥しい数の文献が存在するが、ここでは本文で引用した文献及び本研究に関連の深いVAR分析に関連した最近の文献のみを掲載する)

- 岩田規久男・原田泰 (2013) 金融政策と生産：予想インフレ率の経路, 早稲田大学政治経済学術院ワーキングペーパー, No 1202
- 永田修一 (2016) 非ガウス型構造 VAR モデルの最尤推定－モンテカルロ実験による有限標本パフォーマンスの評価－, *商学論究* 64-1, pp. 97-115
- 永田修一 (2017) 非ガウス型構造 VAR モデルを用いた実証分析－我が国の金融政策の効果－*商学論究* 64-5, pp. 211-226
- 原田泰・増島稔 (2009) 「金融の量的緩和はどの経路で経済を改善したか」, 吉川洋 編『デフレ経済と金融政策第Ⅱ』慶應義塾大学出版会.
- 本多佑三・黒木祥弘・立花実 (2010) 「量的緩和政策—2001年から2006年にかけての日本の経験に基づく実証分析—」財務省財務総合政策研究所『フィナンシャル・レビュー』2010年第1号 (通巻第99号).
- 本多祐三 (2014) 「非伝統的金融政策の効果：日本の場合」, 『現代経済学の潮流 2014』東洋経済新報社
- 前川功一, 小村衆一, 永田修一 (2015), 「VAR モデルによる金融緩和政策効果の検証—2009年～2014年の期間について—」広島経済大学経済研究論集 第38巻第2号
- 前川功一 (2015) ICA 分析による因果性の検出—金融緩和政策の分析を事例として—, 2015年11月 日本銀行金融研究所及び一橋大学経済統計セミナー報告資料
- 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・ルビ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
- 宮尾龍蔵 (2016) 「非伝統的金融政策—政策当事者としての視点」有斐閣