

研究ノート

非ガウス型構造 VAR モデルにおける 尤度比及び Wald 検定について*

——シミュレーション分析——

前 川 功 一**

1. 構造 VAR モデルと誘導形

1.1 構造 VAR モデルと ICA モデル

K 変数を含む時系列ベクトル $y_t = (y_{1t}, \dots, y_{Kt})'$ を考えよう。 y_t に対する標準的な K 次元構造 VAR (Structural Vector Auto-regression: SVAR) モデルは

$$B_0 y_t = B_1 y_{t-1} + \dots + B_p y_{t-p} + c + w_t, \quad (1)$$
$$1 \leq t \leq T$$

と表される。 w_t は平均ベクトル $\mathbf{0}$ 、 $K \times K$ の分散共分散行列 Σ_w を持つ非正規分布に従う誤差項である。 w_t に含まれる K 個の要素 w_{1t}, \dots, w_{Kt} は互いに独立で全ての標準偏差は 1 になるように標準化されていると仮定する。従って $\Sigma_w = I_K$ である。経済時系列では w_t の経済的含意を示したいときに、これを構造ショックと呼ぶ。 B_0 はモデルに含まれる K 変数相互間の同時的な関係を表す非特異なパラメータ行列である。この SVAR モデルには通常定数項ベクトル c が含まれるのであるが、定数項 c を省略しても以下の議論の一般性を失わないので、省略する。(1) 式はラグオペレータ L を用いて簡潔に

$$B(L)y_t = w_t$$

と書ける。ここに $B(L) = B_0 - B_1 L - B_2 L^2 - \dots - B_p L^p$ である。

次に (1) 式の両辺の左から B_0^{-1} を掛けると

$$B_0^{-1} B_0 y_t = B_0^{-1} B_1 y_{t-1} + \dots + B_0^{-1} B_p y_{t-p} + B_0^{-1} w_t$$

を得る。ここで $A_i = B_0^{-1} B_i$, $i = 1, \dots, p$, $u_t = B_0^{-1} w_t$ と置けば

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

と VAR モデルの形になる。この式を SVAR の誘導形という。

ここで、誘導形の誤差項と構造形の誤差項を結びつける関係式

$$u_t = B_0^{-1} w_t \quad (3)$$

は、機械学習や信号理論における独立成分分析 (Independent Component Analysis: ICA) における基本的な関係式と同じ構造であることに着目する。ICA を簡単に説明するためによく使われる例として、カクテルパーティ効果というものがある。カクテルパーティーで多数の人が会話をしているとき、個々の会話が混合されたものは、全体としては雑然とした騒音にしか聞こえない。この時、会場に K 個のマイクロフォンを設置しこの騒音を収録する状況を考えて、(3) 式右辺の w_t が K 人の音声 (源信号) に相当し、それらの音声が空中で混合され (雑音の混ざった) た音声を、K 個のマイクロフォンによって受信された記録が u_t である。行列 A は

* 本研究は平成30年度～令和2年度の科学研究費基盤研究 C (一般) (課題番号: 18K01555, 研究課題「非ガウス型構造 VAR モデルの統計理論と応用」) の助成を受けて行われた。

** 広島経済大学・広島大学名誉教授

K 個の音源を混合する働きを表しているので、A を混合行列または *mixing matrix* という。このモデルにおける、最も重要な仮定は、 w_t に含まれる K 個の要素は互いに独立な非正規分布に従うという点である（但し一つは正規分布であってもよい）。ICA とは、合成された信号 u_t から源信号 w_t の構成要素 w_{1t}, \dots, w_{Kt} を分離するための理論と計算法の体系である。

1.2 ICA のアルゴリズム

u_t から w_t を分離するアルゴリズムとして、最も普及していると思われる方法は *fastICA* と呼ばれる方法である。情報理論では、非正規性の尺度としてネグントロピー (*ngentropy*) という概念がある。その定義は

$$J(x) = H(x_{gauss}) - H(x)$$

で与えられる ($J(x)$ は非負であることが保障されている)。ここには $H(x_{gauss})$ は正規分布に従う確率ベクトル x_{gauss} のエントロピー、 $H(x)$ は非ガウス分布に従う確率ベクトル x のエントロピーである。この値が大きいほど x は正規分布から遠いと言える。このネグントロピーを最大化することによって、音源分離を行うというのが ICA の一つの基本的考え方である。 $J(x)$ 計算のためのアルゴリズムとして *fastICA* というアルゴリズムがある。これは R ライブラリーの中に実装されており、広く使用されている。*fastICA* 及び ICA への応用については文献 [1], [2] 及びそれらに掲げられている文献を参照されたい。さらに *fastICA* に依拠して作られた *LiNGAM* というパッケージがある。これは構造 VAR モデルの推定と変数間の因果序列の検出用に開発されている。言い換えると (3) 式の係数行列 B_0^{-1} が下側三角行列である場合は変数間に因果序列があることを意味するので、*LiNGAM* は可能な限り B_0^{-1} が下側三角行列に近づくように計算を進めるプログラムで

ある。この方法の理論と応用例は文献 [5]~[8] を参考されたい。

もう一つの ICA のアルゴリズムは本稿が以下で取り上げる最尤法である。ICA における最尤法は R のライブラリーの中に “*svars*” というパッケージの中の関数 “*id.ngml*” に実装されている。ただし “*id.ngml*” では、t 分布を前提にする尤度関数に基づく最尤法が使われている。ICA の分野では、雑音が多い音声観測データから音源を分離することが主たる関心事であるためか、*mixing matrix* の各要素の推定の精度や仮説検定ということにはあまり関心が払われない傾向がある。しかし計量経済学では B_0^{-1} の推定及び検定も重要課題である。そこで次節では B_0^{-1} に関する検定に重点を置いて、簡単なモデル設定の下で尤度比検定と Wald 検定について *id.ngml* 関数を使ってシミュレーションを行った結果を報告し、その結果について考察を加える。

2. シミュレーション・モデルの設定

以下のシミュレーションにおいてデータの発生装置 (*Data Generating Process: DGP*) として 3 変数構造 VAR モデル

$$B_0 y_t = B_1 y_{t-1} + w_t, \quad t = 1, \dots, T,$$

を想定する。その誘導形は

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

と表わされる。ここに $A_1 = B_0^{-1} B_1$ 及び $u_t = B_0^{-1} w_t$ である。

(4) 式がシミュレーションにおける DGP である。(4) 式において w_t が、平均ベクトル 0、分散行列が $\Sigma = I$ であれば、 u_t の分散行列は

$$E(u_t u_t') = E(B_0^{-1} w_t w_t' B_0^{-1'}) = B_0^{-1} B_0^{-1'} \quad (5)$$

となる。したがって *id.ngml* によって B_0^{-1} を

推定するためには w_t は平均ベクトル 0, 分散行列が $\Sigma = I$ でなければならない。

A と B_0^{-1} に数値を与え, w_t に特定の非正規分布に従う乱数を与えれば, 人工的観測値 y_t データを生成することができる。本稿では 3 変数 SVAR モデルを使ってシミュレーションを行う。すなわち係数行列 A_1 と B_0^{-1} は 3×3 の行列

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, B_0^{-1} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

である。さらに誤差項 $w_t = \begin{pmatrix} w_{1t} \\ w_{2t} \\ w_{3t} \end{pmatrix}$ に対しては,

次の二通りの乱数を与える。

- 1) 独立に t 分布に従う 3 つのランダムな確率変数 $w_{it}, i=1,2,3$ を与える。以下では自由度 5 の t 分布に従う乱数を 3 系統 w_{1t}, w_{2t}, w_{3t} 発生させ, 上に述べた理由により, それぞれの w_{it} を平均 0, 標準偏差 1 となるように標準化 = $\frac{w_{it} - \text{mean}(w_{it})}{\text{sd}(w_{it})}$ しておく。

- 2) 独立に Laplace 分布に従う乱数を与える

本稿では, B_0^{-1} に与える数値と w_t の分布を色々変えてその効果を調べる。以下では, B_0^{-1} が下側三角行列の場合を考える。すなわち

$$B_0^{-1} = \begin{pmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

以下のシミュレーションでは

$$A_1 = \begin{pmatrix} 0.941 & 0.005 & -0.026 \\ 0.073 & 0.859 & 0.084 \\ 0.081 & 0.109 & 0.931 \end{pmatrix},$$

$$B_0^{-1} = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ -0.5 & 1.0 & 0.0 \\ 0.15 & -0.5 & 1.0 \end{pmatrix},$$

とした。以上の基本設定の下で, 以下のステップに従ってシミュレーションを行う。

2.1 w_{1t}, w_{2t}, w_{3t} が独立な t 分布に従う場合 ステップ 1

自由度 5 の t 分布から互いに独立な 3 つの乱数 $w_t = (w_{1t}, w_{2t}, w_{3t})'$ を発生させ, それらを標準化した値を使って, (4) 式によって y_t を計算する。

ステップ 2

ステップ 1 で生成された y_t を観測データとみなし VAR (1) モデルを推定する。

ステップ 3

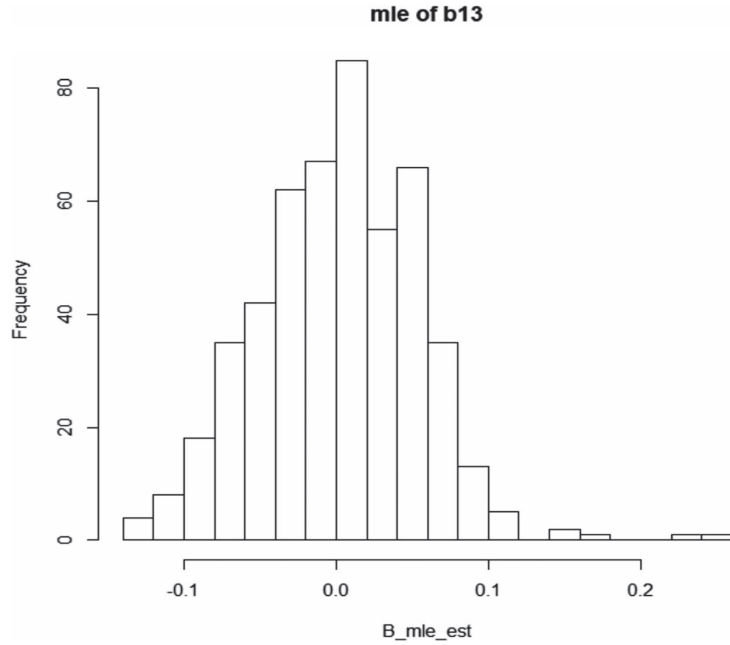
ステップ 2 で得られる VAR (1) の残差系列を関数 id.ngml に引き渡し,

- 1) B_0^{-1} の最尤推定値を計算する
- 2) 検定仮説を次のように設定する:
帰無仮説 $H_0: b_{13} = 0$, 対立仮説 $H_1: b_{13} \neq 0$
- 3) 尤度比検定と Wald 検定を計算する

ステップ 4

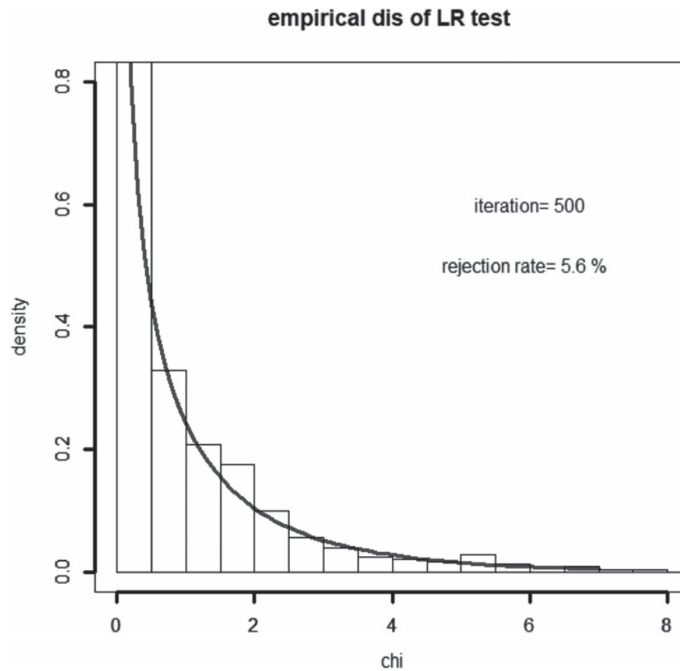
ステップ 1 ~ 3 を 500 回繰り返し, 得られた 500 個の尤度比検定統計量の値と Wald 検定統計量の値が自由度 1 のカイ 2 乗分布の 5 % 臨界値 3.84 を超える回数を調べる。またそれぞれの検定統計量の経験分布 (ヒストグラム) を作成する。以下シミュレーション結果を示す。

以上の尤度比検定と Wald 検定のシミュレーションにおいて, DGP の真値としてでは $b_{13} = 0$ が割り当てられているので, b_{13} の推定値は 0 に近い値とる確率が高い。従って, 尤度比検定と Wald 検定の値は, 理論的には自由度 1 のカイ 2 乗分布の臨界値 3.84 を超えるほど大きな値を取る確率は 5 % そこそこであるはずである。シミュレーションの結果この臨界値を超える比率がほぼ 5 % 程度であるので, これらの検定法は妥当であると判断することができる。



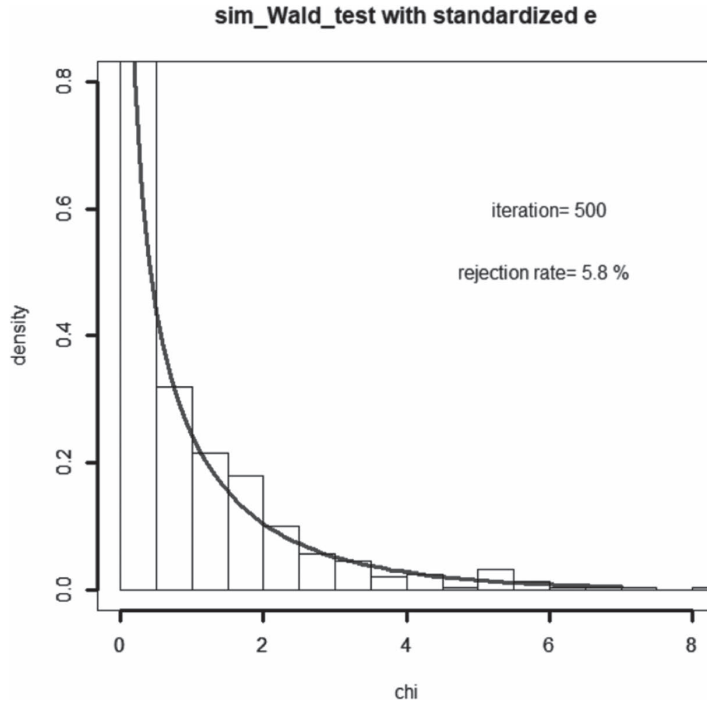
繰り返し = 500, b_{13} の真値 = 0.0, 最尤推定値 \hat{b}_{13} の平均値 = 0.001859311, 標準偏差 = 0.05273108, 右裾が若干長い³⁾, 推定は概ねうまくいっている。

グラフ 1 最尤推定値 \hat{b}_{13} の経験分布 (t-分布の下で)



$H_0: b_{13} = 0, H_1: b_{13} \neq 0$ の下で尤度比検定値の経験分布。黒の曲線は自由度 1 のカイ 2 乗分布 (5% 臨界値 = 3.84)。シミュレーションの結果, 500 回繰り返したとき, 5.6% が棄却された。

グラフ 2 尤度比検定 (t-分布の下で) の経験分布



$H_0: b_{13} = 0, H_1: b_{13} \neq 0$ の下で Wald 検定値の経験分布。黒の曲線は自由度 1 のカイ 2 乗分布（5% 臨界値 = 3.84）。シミュレーションの結果、500 回繰り返したとき、5.6% が棄却された。

グラフ 3 Wald 検定 (t 分布の下で) の経験分布

2.2 w_{1t}, w_{2t}, w_{3t} が独立な Laplace 分布に従う場合

次に非正規性の強い Laplace 分布の下でのシミュレーションを実施する。このような場合を想定する理由は、t 分布に基づく尤度を使用している id.ngml が t 分布以外の確率分布から生成されたデータ y_t にどの程度対応できるか等について考察するためである。言い換えれば、これは、id.ngml が、t 分布以外の非正規分布から生成された観測値に対して、疑似尤度としてどの程度の有効性と頑健性を持つかをシミュレーションによって確かめようとする試みである。シミュレーションのステップは、上記のステップ 1~4 において、誤差項の分布を Laplace 分布に変えて実行する。Laplace 分布の密度関数、期待値、分散は以下に示すとおりである。

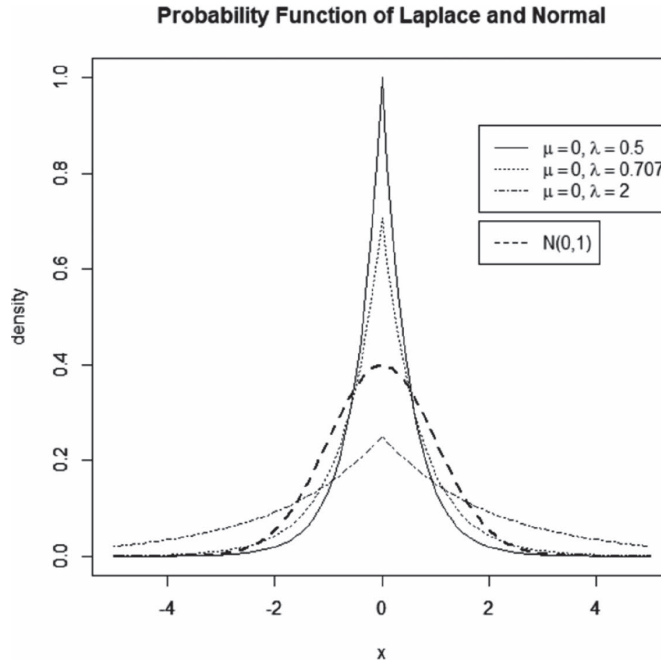
$$\text{密度関数: } f(x) = \frac{1}{2\lambda} \exp\left(-\frac{|x-\mu|}{\lambda}\right)$$

$$\text{期待値: } E(x) = \mu, \text{ 分散: } \text{var}(x) = 2\lambda^2$$

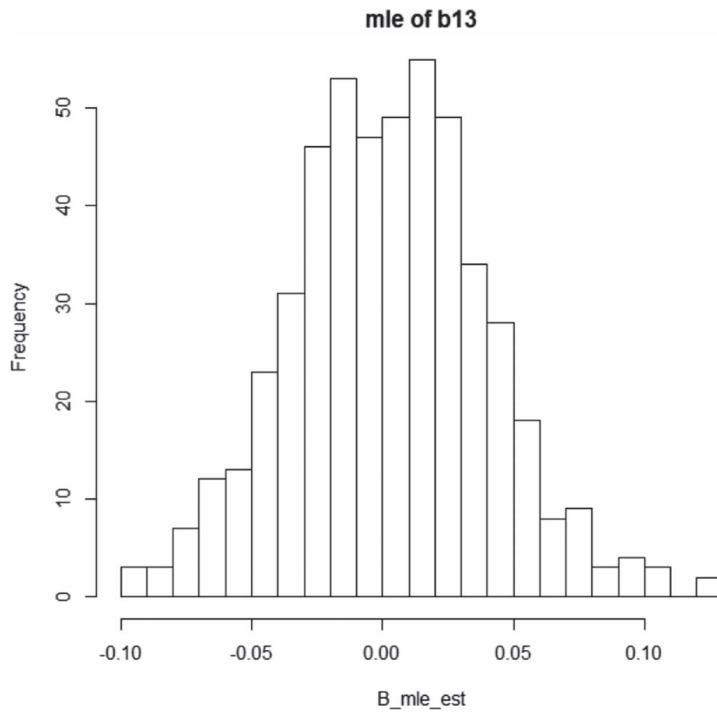
したがって $\lambda \cong 0.707$ のとき、Laplace 分布の分散は 1 となる。異なる λ ごとの Laplace 分布の形状はグラフ 8 の通りである。なお、比較のために標準正規分布を鎖線で併記しておく。

2.3 w_{1t}, w_{2t}, w_{3t} が独立でない場合

id.ngml 関数は構造形の誤差項 w_t の要素は独立に平均 0、分散 1 の非正規分布に従うと仮定すれば、誘導形誤差項の共分散行列が $\Sigma_u = B_0^{-1} B_0^{-1'}$ と表されることを前提にして B_0^{-1} を最尤推定する構造になっている。したがってこの仮定が満たされない場合（誤差項 w_t の要素間に相関がある場合）は、id.ngml 関数が返して

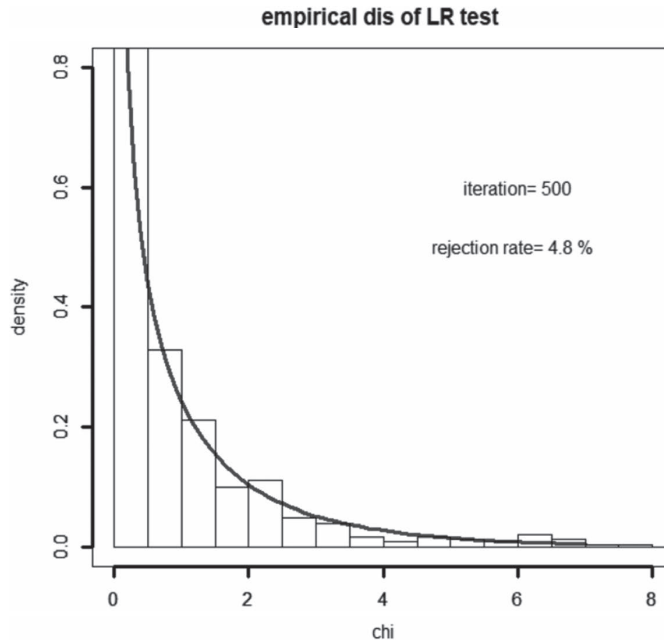


グラフ4 Laplace 分布と標準正規分布



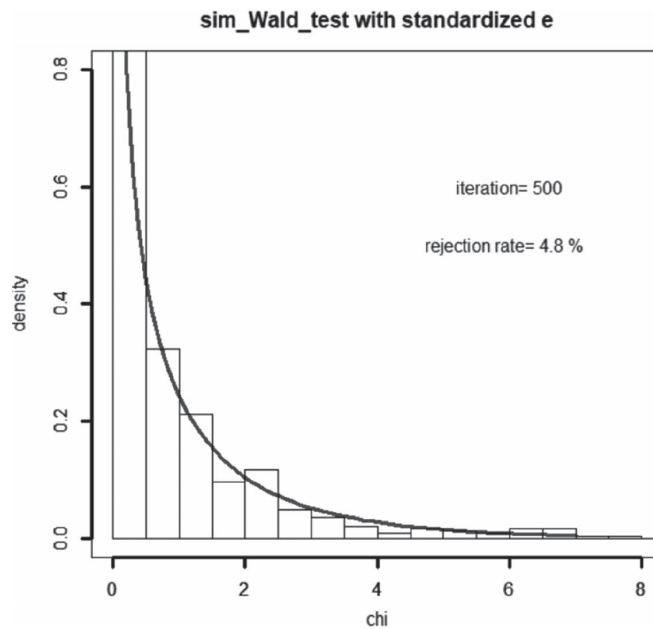
繰り返し = 500, b_{13} の真値 = 0.0, 最尤推定値 \hat{b}_{13} の平均値 = 0.0025528, 標準偏差 = 0.03772447. 最尤推定は概ね上手くいっている。

グラフ5 誤差項が Laplace 分布 ($\mu = 0, \lambda \cong 0.707$) のときの最尤推定値 \hat{b}_{13} の経験分布



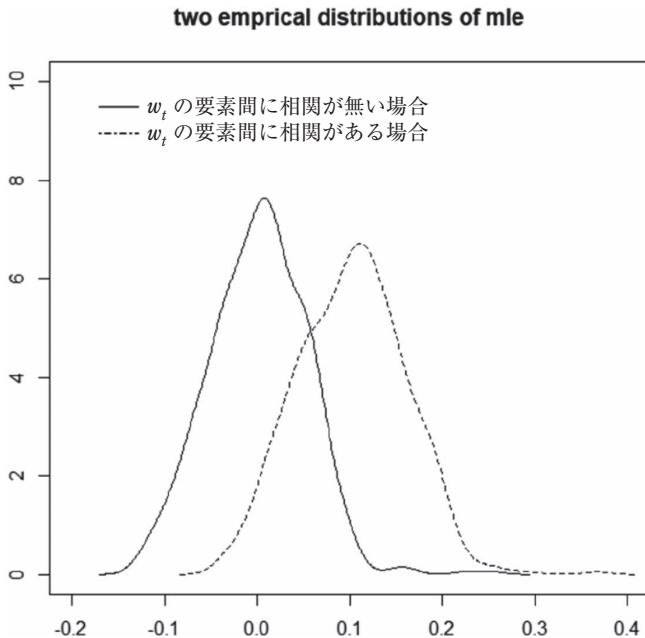
尤度比検定の棄却サイズがグラフ 3 (t-分布の場合) に比べて約 1 % 小さい。

グラフ 6 誤差項が Laplace 分布 ($\mu=0, \lambda \cong 0.707$) のときの尤度比検定値の経験分布。



グラフ 6 の場合と同様の傾向。尤度比検定の棄却サイズがグラフ 3 (t-分布の場合) に比べて約 1 % 小さい。

グラフ 7 誤差項が Laplace 分布 ($\mu=0, \lambda \cong 0.707$) のときの Wald test の経験分布。



実線の分布の平均と分散：M = 0.001859311, SD = 0.05273108,
 点線の分布の平均と分散：M = 0.1022904, SD = 0.0583169

グラフ 8 w_{1t}, w_{2t}, w_{3t} の間に相関がある場合とない場合とにおける、最尤推定値 \hat{b}_{13} の経験分布の比較

くる最尤推定値は、この相関が高いほど、偏りが大きくなることが予想される（ただし A_1 の推定は影響を受けない）。以下に、そのような偏りを示す一例として、誤差項 w_t の共分散行列が w_t

$$\Sigma_w = \begin{pmatrix} 1.0 & 0.5 & 0.1 \\ 0.5 & .0 & 0.5 \\ 0.1 & 0.5 & 1.0 \end{pmatrix} \quad (6)$$

の場合（ w_t の要素間に相関がある場合）、 B_0^{-1} の (1,3) 要素 b_{13} の最尤推定のシミュレーション結果をグラフ 8 に示す。このグラフの実線は、 Σ_w が単位行列の場合の、点線は Σ_w が (6) 式の場合の、id.ngml によって計算された最尤推定値の経験分布である。明らかに後者の場合に推定の大きな偏りがみられる。

3. ま と め

本稿は (4) 式を DGP とし、非ガウス型構造 VAR (1) モデルにおける最尤推定法、尤度比検定、及び Wald 検定のパフォーマンスをシミュレーションによって示そうという試みである。非ガウス確率分布として、自由度 5 の t 分布と平均 0、分散 1 の Laplace 分布から乱数を生成し、両者が推定と検定に及ぼす効果を R に実装されている id.ngml 関数を使って調べた。その結果、構造係数 B_0^{-1} に関する推定、検定に対して上述の二つの乱数における相異は予想外に小さいことが分かった。しかし誤差項の要素間の同時点間での相関の有無は B_0^{-1} の推定には非常に大きい影響があり、他方検定のサイズに関しては大きな影響はないことが分かった。

経済モデルの実証分析においては、ほとんどの場合モデルの誤差項に慣例的に正規分布が仮

定される。しかし現実のデータを見ると正規分布ではないと思われる場合が少なくない（特にファイナンスデータの場合）。したがって非正規性を前提とする ICA を経済モデル分析に導入することは、十分意味がある。しかしどんな非正規分布であるかが特定できなければ、尤度関数を導くことができず、したがって最尤法も使えないことになる。そこでやむを得ず疑似最尤法を使うことになるが、疑似尤度として正規分布のもとでの尤度関数を使うことは自己矛盾をきたす。id.ngml 関数を使うということは、見方を変えれば疑似尤度として t 分布に基づく尤度関数を使うことと解釈することができる。本稿で示されたシミュレーション結果は、もし現実の分布が Laplace 分布のように極度に先鋒度の高い分布であっても、 t 分布に基づく尤度関数を疑似尤度として使用しても、推定、検定のパフォーマンスはさほど悪くないことを示している。その理由はとしては、 t 分布は高次のモーメントまで有するので、幅広い分布に対応できるからではないかと思われる。したがって t 分布を前提とする id.ngml の尤度関数を疑似尤度として利用して推定、検定を行っても、シミュレーションでかなり良好な結果が得られたことも納得ができる。

今回のシミュレーション結果から以上のような傾向が判明したが、それがどの程度普遍的妥当性を持つかについては、今後さらに分析が必要である。また今回の経験から、より柔軟性が

ある疑似尤度として Peason 分布族を導入すればいいのではないかという予想が浮かび上がってくる。

参 考 文 献

- [1] Hyvärinen, A., and Erkki Oja (2000), Independent Component Analysis: algorithms and applications, *Neural Networks* 13, 411–430
- [2] 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)
- [3] Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A. (2006), A Linear non-Gaussian acyclic model for causal discovery, *Journal of Machine Learning Research*, Vol. 7
- [4] Hyvärinen, A., Zhang, K., Shimizu, S., and Hoyer, P. O. (2010), Estimation of Structural Vector Autoregression Model Using Non-Gaussianity, *J. Machine Learning Research* 11, 1709–1731
- [5] 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
- [6] 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刷「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
- [7] Lane, M., Meitz, M., and Saikonen, P. (2017), Identification and estimation of non-Gaussian structural vector autoregressions, *Journal of Econometrics* 196, 288–304
- [8] 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と政策の分析を事例として— 広島経済大学創立五十周年記念論文集上巻, 25–62
- [9] 前川功一 (2020), 構造 VAR モデルの識別性と推定 (I) 広島経済大学 経済研究論集 第43巻第2号