広島経済大学経済研究論集
第44巻第２号　2021年11月
http://dx.doi.org/10.18996/keizai2021440202

構造 VAR モデルの識別性と推定（Ⅱ）*

——非正規誤差項の場合——

前　　川　　功　　一**

要　　約

前回では誤差項が正規分布に従う場合の構造 VAR モデルの識別性と推定問題を考察したが，今回は非正規誤差の場合について論じる。伝統的な計量経済学に於いては，誤差項は正規分に従う仮定されてきた。この仮定が正しければ，伝統的な統計的推測理論が有効に機能し，構造 VAR モデルの実証分析に対して，統計的推測はほぼ機械的に進められる。しかし現実のデータ分析においては，誤差項の分布は正規分布に従わない場合が少ない。しかもどのような非正規分布に従っているかに関する情報が得られる場合は限られている。このような状況の下では，尤度関数を構成することができないため最尤法を適用することができない。そこで本稿では，観察された誤差のみを使って非ガウス分布を推定し，推定された非ガウス分布に基づいて疑似尤度を構成するという手続き提案し，さらにこの方法の有効性をモンテカルロ実験によって検証する。

1.　は　じ　め　に

伝統的な構造 VAR モデルに於いては，ほとんどの場合，モデルの誤差項は正規分布に従うと仮定される。この仮定の下では古典的な数理統計学の手法を機械的に応用することによって推定，検定，予測等を実施することは（少なくとも近似的には）許容されている。しかし近年ファイナンス計量分析に見られるように，正規性の仮定が満たされないモデルが使用される場合が増えたためであろうか，誤差項に正規性を仮定しない構造 VAR モデルの研究が徐々にではあるが活発化しつつある。誤差項に非正規性を仮定する構造 VAR モデルを，本稿では，非ガウス型構造 VAR モデルと呼ぶ。（その英訳として non-Gaussian Structural VAR model，略して NG-SVAR モデルと表すことにする）。この系列に属する論文のいくつかを挙げれば，Shimizu et al.（2006），Moneta et al.（2013），Gourieroux et al.（2017），Lanne et al.（2017），前川（2017, 2020），Maekawa et al.（2021）などがある。NG-SVAR モデルが注目される原因としては，上に挙げたファイナンス計量分析における非ガウス分布の重要性に加えて，次の２点を付加することができる。第１点は，構造 VAR モデルには識別性の問題が内在するが，非正規性の仮定のもとでは，識別性の問題が解決される場合がありうるという点である。この問題に関しては，最近 Lanne et al.（2017）は，かなり広い枠組みの中で識別性の必要十分条件を示した。これについては前川（2017）に解説がある。第２点として，モデルに含まれる変数間に因果序列が存在する場合には，NG-SVAR モデルはその因果序列の検出が容易に行われるという点を挙げることができる。R のパッケー

* 本研究は JSPS 科研費 JP18K01555の助成を受けて行われた。
** 広島経済大学名誉教授。本稿は EcoSta2021をはじめとする多くの学会，セミナー等における一連の研究報告をもとに書かれている。それらの機会において多くの参加者の方々から貴重なご意見を頂いたことをここに記し感謝申し上げる。しかしもし誤りがあれば，それらすべて著者の責任である。

ジ“LiNGAM”には，この因果序列の検出に必要なアルゴリズムが実装されている。それを用いて前川（2017）は，日本銀行によるの非伝統的金融緩和政策の波及効果を実証的に分析した。以上挙げた NG-SVAR モデルによるモデル分析では，機械学習（または信号理論）で確立された独立成分分析（Independent Component Analysis ICA）の理論が重要な役割を果たしている。ICA の理論モデルと NG-SVAR model との類似性については次節以降で詳しく説明する。

最後に本稿の構成を示しておく。第２節では非ガウス型構造 VAR モデルの定式化と疑似尤度関数の構成法が示される。第３節では，この疑似尤度最尤法のパフォーマンスを検証するためのモンテカルロ実験に必要なデータ生成過程（Data Generating Process; DGP）を示し，第４節ではモンテカルロ実験結果が示される。最後に第４節において実験結果の総括をおこなう。

2.　非ガウス型構造 VAR モデル

2.1　モデルの表現形式

まず初めに，k 変数を持つ p 次元非ガウス型構造 VAR モデルは

$$y_t = v^* + A_0^* y_t + A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + \varepsilon_t \quad (1)$$
$$t = 1,\cdots,T$$

と表すことができる。ここに $y_t = (y_{1t},\cdots,y_{kt})'$ は k 個の変数に関する時点 t における $k \times 1$ の観測値ベクトル，$\varepsilon_t = (\varepsilon_{1t},\cdots,\varepsilon_{kt})'$ は $k \times 1$ の（観測されない）誤差項ベクトルである。NG-SVAR モデルでは，ε_t は非正規分布に従うと仮定されるが，ここでは Lanne et al. (2017) に倣って ε_t の仮定を次のように整理する。なお（1）式において両辺に y_t が含まれていることから，A_0^* の対角要素は 0 とすることが適当である。

誤差項 $\boldsymbol{\varepsilon_t}$ に関する仮定（Lanne et al.（2017））：

(i)　過程 $\varepsilon_t = (\varepsilon_{1t},\cdots,\varepsilon_{kt})', t = 1,\cdots,T$ は独立・同一の確率分布に従い，ε_t の各要素 $\varepsilon_{1t},\cdots,\varepsilon_{kt}$ は平均 0 と有限な分散 σ_i^2 を持つ。

(ii)　ε_t の各要素 $\varepsilon_{1t},\cdots,\varepsilon_{kt}$ は互いに独立で，これらの要素の内，最大１個の要素は正規分布に従っていてもよい。

(iii)　各々の $i = 1,\cdots,k$ に対して誤差項 ε_{it} はある密度関数 $f_{i,\sigma_i}(x;\lambda_i) = \sigma_i^{-1} f_i(\sigma_i^{-1} x;\lambda_i)$ を持つ。ここに λ_i は平均と分散以外の密度関数固有のパラメータである。なお $f_{i,\sigma_i}(x;\lambda_i)$ は（必ずしも必要ではないが）同じ分布族に属することができる。

これらの仮定に従う密度関数として，例えば１変数 t 分布が挙げられる。このとき λ_i は i 番目の密度関数の自由度が対応する。

NG-SVAR モデル（1）をさらに変形すれば

$$\left(I - A_0^*\right) y_t = v^* + A_1^* y_{t-1} + \cdots + A_p^* y_{t-p} + \varepsilon_t \quad (2)$$

となる。ここで $B_0^{-1} = \left(I - A_0^*\right)^{-1}$（対角要素は１になることに注意）と置けば（2）式は

$$\begin{aligned} y_t &= B_0^{-1} v^* + B_0^{-1} A_1^* y_{t-1} + \cdots + B_0^{-1} A_p^* y_{t-p} \\ &\quad + B_0^{-1} \varepsilon_t \\ &= \mu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \end{aligned} \quad (3)$$

となる（B_0^{-1} は Lanne et al. における B に対応する）。ここに $\mu = B_0^{-1} v^*, A_i = B_0^{-1} A_i^*, i = 1,\cdots,p$。また

$$u_t = B_0^{-1} \varepsilon_t \quad (4)$$

である。そして A_i は（3）式の y_t が定常であるための条件

$$\mathrm{detA}(z) \overset{\mathrm{def}}{=} \det\left(I_T - A_1 z - \cdots - A_p z^p\right) \neq 0,$$
$$|z \leq 1| \; (z \in \mathbb{C})$$

を満すものとする。ここで注意しなければならない点は，任意の非特異行列 C によって $u_t = B_0^{-1} C C^{-1} \varepsilon_t$ と変形できるが，このとき B_0^{-1} は $B_0^{-1} C$ に，ε_t は $C^{-1} \varepsilon_t$ に置き換えられた形になる。しかしこのように変形しても，誤差項

ε_t に関する仮定が損なわれること無く，u_t に対して二つの表現が成立する。すなわち u_t と ε_t とを関係づける係数が一意に決まらないという問題が発生する。これがいわゆる識別性の問題である。従って誘導形誤差 u_t 自体は，係数行列が B_0^{-1} であっても，$B_0^{-1}C$ であっても変わらない。

識別性の問題を回避する方法はいくつかある。

1，B_0^{-1} が下三角行列であると仮定する。この仮定は変数間に因果序列を仮定することに等しい。因果序列のパターンは順列の数だけ存在するが，それらの中から経済的に意味のあるパターンを推定するという方法が一般的である。因果序列が存在位する場合は，R のパッケージ“LiNGAM”が自動的に最適な因果序列を探し出してくれる。

2，Lenne et al.（2017）の提案。この論文は，かなり広い枠組みで B_0^{-1} の識別可能性の条件を導き，その条件を満たす B_0^{-1} を導出するアルゴリズムを示している。

3，B_0^{-1} に対して経済学的又は数学的な観点から妥当な何らかの制約を置くことによって識別可能性を獲得する。例えば B_0^{-1} の対角要素を 0 と置く，経済的に明らかにある係数は 0 としてよいなど。

2.2　尤度関数の構成

本稿では尤度関数の構成法を述べる。モデルの未知パラメータを尤度関数に引き渡すために以下のような記号を導入する。まず誘導形（3）式に現れる係数行列を $\pi_1=\mu$，$\pi_2=vec(\mathrm{A}_1,\cdots,A_p)$ と置く。また B_0^{-1}（対角要素は 1 ）の非対角要素をベクトル化したものを β と置く。また各密度関数 $f_{i,\sigma_i}\left(x;\lambda_i\right), i=1,2,\cdots,k$ に現れる未知パラメータ σ_i と λ_i をそれぞれベクトル σ，λ で表す。そしてすべての未知パラメータを一括したものを $\theta=\left(\pi,\beta,\sigma,\lambda\right)$ と置く。これらの記号を用いて Lanne et al. p. 293（2017）に倣って尤度関数を次のように書き表わす。使用するデータ時点は $t=1,\cdots,T$，最大ラグ $=p$ であるから時点 $t=1$ より前のデータ $y_{-p+1},\cdots,y_0$, 及び $y_1,\cdots,y_T$ は与えられているものとする。この時，尤度関数は一般的に下のように書くことができる。

$$L_T\left(\theta\right)=T^{-1}\sum_{t=1}^{T}l_t\left(\theta\right), \tag{5}$$

ここに

$$l_t\left(\theta\right)=\sum_{i=1}^{n}log\,f_i\left(\sigma_i^{-1}\iota_i'B\left(\beta\right)^{-1}u_t\left(\pi\right):\lambda_i\right) \\ -log\left|\det\left(B\left(\beta\right)\right)\right|-\sum_{i=1}^{n}log\sigma_i \tag{6}$$

である。この式に使われている ι_i は第 i 要素を 1，それ以外を 0 とする単位ベクトル $\iota_i'=\left(0,\cdots0,1,0,\cdots,0\right)$ である。また誘導形残差は $u_t\left(\pi\right)=y_t-\nu-A_1y_{t-1}-\cdots-A_py_{t-p}$ と書くことが出来る。（6）式の最後の二つの項は変数変換 $u_t=B_0^{-1}\varepsilon_t$ に伴うヤコビアンに由来する項である。

さて尤度関数が構成されたので，理論的には密度関数を特定した後で $L_T\left(\theta\right)$ を未知パラメータ θ に関して最大化することによって最尤推定量 $\tilde{\theta}$ が得られる。しかし未知パラメータの個数が多いので，すべてを一括して最適化を図ろうとすると数値計算に伴う，計算時間，計算精度，収束などの問題が生じやすい。このような問題を回避するために Lanne et al. p. 293（2017）は，第一段階として，未知パラメータ π に関しては誘導形（3）の最小 2 乗推定値 $\hat{\pi}$ を求め，第二段階として，$\hat{\pi}$ を所与として残りのパラメータを最尤推定するという 2 段階法を用いている。

3.　モンテカルロ実験

3.1　データ生成過程（DGP）

前節で示された尤度関数を使って 2 段階最尤法のパフォーマンスをモンテカルロ実験で確かめてみよう。実験用モデルとしては，以下に示

すような単純な，定数項を持たないラグ次数１の４変数 NG-SVAR モデルを使った。

データ生成過程（**Data Generating Process; DGP**）を以下のように設定する。この DGP を使って人工的にデータを発生させる。

●定数項のない NG-SVAR モデルを使用する。:

$$B_0 y_t = A_1^* y_{t-1} + \varepsilon_t \tag{7}$$

ここに $\varepsilon_t = (\varepsilon_{1t},\cdots,\varepsilon_{kt})'$ の各要素は平均０，分散１の独立同一の非正規確率分布に従う。したがって $E\left(\varepsilon_t \varepsilon_t'\right) = I$ である。

●誘導形（2）は次のように単純化される：

$$y_t = A_1 y_{t-1} + u_t, \tag{8}$$

ここに $A_1 = B_0^{-1} A_1^*$，$u_t = B_0^{-1}\varepsilon_t$ である。また ε_t の構成法より u_t の分散共分散行列は

$$\Sigma_{uu'} = E(u_t u_t') = B_0^{-1} E(\varepsilon_t \varepsilon_t')\, B_0^{-1\prime} = B_0^{-1} B_0^{-1\prime} \tag{9}$$

となる。

● A_1 と B_0^{-1} を次のように与える。

$$A_1 = \begin{pmatrix} .96 & .01 & .21 & -.03 \\ .19 & 1.24 & .62 & -.07 \\ -.12 & -.25 & .39 & .07 \\ -.22 & -.05 & -.37 & .95 \end{pmatrix},$$

$$B_0^{-1} = \begin{pmatrix} 1 & 0 & \mathbf{1.41} & \mathbf{1} \\ 0 & 1 & \mathbf{1} & 0 \\ \mathbf{.57} & 0 & 1 & 0 \\ 0 & \mathbf{.57} & 0 & 1 \end{pmatrix} \tag{10}$$

ここに B_0^{-1} は Lanne et al.（2017）に示されている識別可能な例から採用した。以下の実験では太字で示された係数を未知パラメータと見なし，それ以外のパラメータ値は所与とする。そして未知パラメータの最尤推定のパフォーマンスを検証することがここでの目的である。また A_1 は（8）が定常となるように，B. Pfaff（2008）のカナダのマクロモデルの推定値を採用した。

●さらに次のような単純化を行った。

- ・すべての $i = 1,\cdots,4$ に対して $\sigma_i = \sigma, f_i = f$，$\lambda_i = \lambda$ と置く
- ・β は，B_0^{-1} の太字で示された未知パラメータを要素とするベクトル
- ・$\iota_i' B(\beta)^{-1} u_t(\pi) = \varepsilon_i$

ここに ι_i' は，第 i 要素の値は１，それ以外の値は０となる 1×4 のベクトルである。

以上の設定の下で B_0^{-1} の中の未知パラメータに関する２段階最尤法の推定パフォーマンスを検証するモンテカルロ実験を次のような手順で行なう。まず DGP における非正規密度関数として，t 分布，ラプラス（*Laplace*）分布，双曲線正割（*Hyperbolic secant*）分布を取り上げる。それぞれの密度関数とそのグラフを下に示しておく。

● t 分布（自由度＝ν）

$$f(x) = \frac{\Gamma\left[(\nu+1)/2\right]}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}\left[1 + \left(x^2/\nu\right)\right]^{-(\nu+1)/2}$$

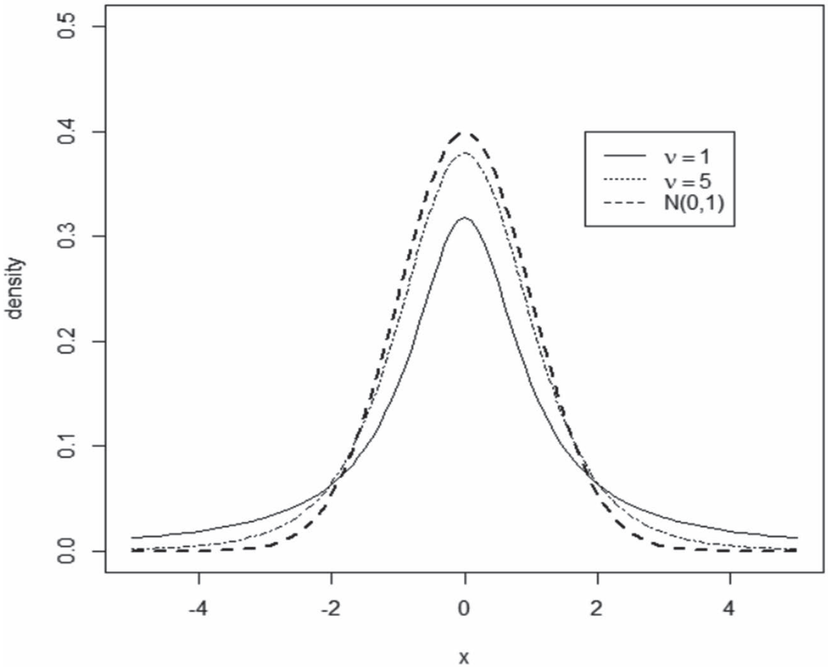
● *Laplace*:　$f(x) = \dfrac{1}{2\lambda} exp\left(-\dfrac{|x-\mu|}{\lambda}\right)$

● *Hyperbolic secant*:

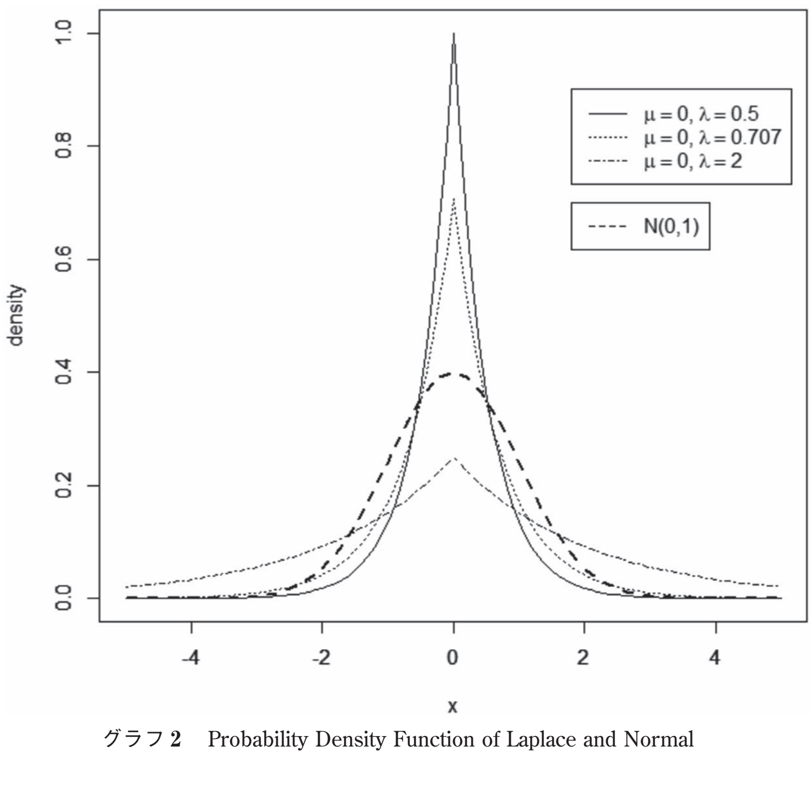
$$f(x) = \frac{1}{2\lambda} sech\left\{\frac{\pi(x-\mu)}{2\lambda}\right\} = \frac{1}{\theta\left[exp\left\{\frac{\pi(x-\mu)}{2\lambda}\right\} + exp\left\{-\frac{\pi(x-\mu)}{2\lambda}\right\}\right]}$$

ここに μ と λ は，範囲（$-\infty < x < \infty, -\infty < \mu < \infty$，$\theta > 0$）で定義される位置とスケールを表すパラメータである。

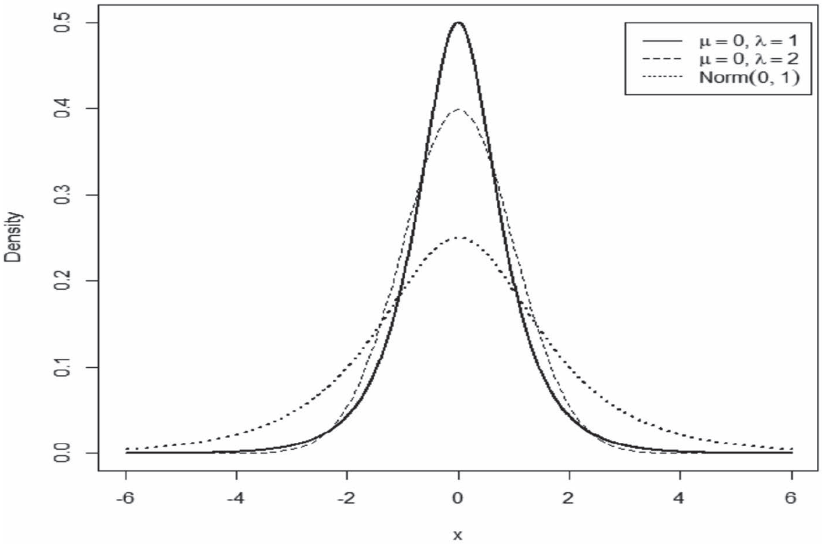
これらのグラフには，比較のために標準正規分布の密度関数が表示されている。例えばラプラス分布の場合，パラメータ λ が0.707のとき標準偏差が１であるので，標準正規分布に比べて尖度がどの程度大きいかがわかる。



グラフ **1**　Probability Function of t and Normal



グラフ **2**　Probability Density Function of Laplace and Normal



グラフ 3　Probability Density Function of Hyperbolic secant and N(0,1)

3.2　計算ステップ

以上でモンテカルロ実験を行う準備が整ったので，上の3つの分布ごとに DGP を特定化し実験を実行した結果を示そう。まず実験開始前に，DGP における真の密度関数 $f_{i,\sigma_i}(x;\lambda_i)$, $i=1,2,\cdots,k$ とパラメータ行列 A_1 と B_0^{-1} の値を与えておかなければならない。それらの真値を使って DGP から人工的にデータ $y_t, t=1,\cdots,T$ を発生させる。その後，（真値を伏して）真値を推定する。しかし真値を伏せると言っても，以下の実験では $\varepsilon_t=(\varepsilon_{1t},\cdots,\varepsilon_{4t})'$ の要素間の独立性（したがって無相関性）と B_0^{-1} の識別条件だけは既知とする。このような想定の下で実験を実施する。その具体的手順を以下に示す。

Step 1：3つの非正規分布の中から一つを選び，その分布から乱数 $\varepsilon_t=(\varepsilon_{1t},\cdots,\varepsilon_{4t})'$ を発生させる。

Step 2：発生させた ε_t から，あらかじめ与えられた A_1 と B_0^{-1} の値と $u_t=B_0^{-1}\varepsilon_t$ から計算される誘導形残差 u_t を使って $y_t, t=1,\cdots,T$ を発生させる。この y_t が（人工的に発生させた）観測値である。

Step 3：観測値 y_t から誘導形係数行列 A_1 の要素をベクトル化して格納した π の最小2乗推定値量 $\hat{\pi}$ を計算する。

Step 4：$\hat{\pi}$ を所与とし，y_t を以下に示す疑似尤度関数 $L_T(\theta)$ に引き渡し，未知パラメータ β と σ に関してこの尤度を最大化する。

ここで問題は，われわれにとって DGP に採用された密度関数 $f_{i,\sigma_i}(x;\lambda_i), i=1,2,\cdots,k$ は伏されているという前提なので，尤度関数もまた未知であるという点である。そこで真の尤度関数の代わりに疑似尤度関数 $L_T(\theta)$ で代用したいのであるが，それをどのように構成法すればいいのであろうか？この問題に対して，本稿では，以下のような疑似尤度の2通りの構成法を提案する。

疑似尤度関数の構成法

(i) 構造誤差が無相関の場合

上に述べたように DGP の仮定から $\varepsilon_{1t},\cdots,\varepsilon_{4t}$ は互いに独立であるから $\tilde{u}_t$ の分散共分散行列を $\Sigma_{\tilde{u}\tilde{u}'}=\tilde{B}_0^{-1}\tilde{B}_0^{-1\prime}$ と直交分解することによって $\tilde{B}_0^{-1}$ を B_0^{-1} の推定値とすることができる。その結果を使って $\tilde{\varepsilon}_t=\tilde{B}_0\tilde{u}_t$ から $\tilde{\varepsilon}_t$ を得ることができる。一般にはこの直交分解は一意ではないので，識別性の問題が生じるが，ここでは

DGP において B_0^{-1} の一意性が保障されているので，識別性の問題は生じない。このようにして求められた $\tilde{\varepsilon}_t$ に対して下に述べる方法によって，最適な $\tilde{\varepsilon}_t$ の密度関数を推定する。この推定された密度関数から構成される尤度関数を，ε_t の疑似尤度関数とみなことにする。本稿では，$\tilde{\varepsilon}_t$ の密度関数の推定に際して，R の関数 “PearsonMSC” を用いた。この関数に $\tilde{\varepsilon}_t$ を入力すると，AIC，BIC，等のいくつかの基準に基づいて，ピアソン分布族 typeI～VII の中から最適な密度関数形とその関数形を定めるパラメータ値が出力される。

（注）$\tilde{\varepsilon}_t$ に含まれる 4 つの誤差項ごとに，PearsonMSC を適用したが，4 つとも PearsonVII が選ばれた場合は，それら 4 つの結果における Pearson 分布の特性値の平均を使用した。稀に 4 つの誤差項の分布に PearsonIV と PearsonVII が混在する場合も発生したが，PearsonVII が多数を占めたので PearsonVII を採用した。理論的には二つの分布が混在する尤度関数を構成することは可能であるが，本稿では尤度関数を単純化するためにどちらか一方に依拠する尤度関数を用いた。

次に，選ばれたピアソン分布の密度関数 $f_{i,\sigma_i}(\varepsilon_{it};\lambda_i)$ に $\varepsilon_{it} = \iota_i' B(\beta)^{-1} u_t(\pi)$ を代入する。このとき ε_{it} から u_t への変数変換が施されるので，変換のヤコビアンが必要になる。このように構成された尤度関数を未知パラメータに関して最大化することによって β と σ の，すなわち $B(\beta)^{-1}$ の中の未知パラメータと分布のスケールパラメータ σ の疑似最尤推定量が得られる。

ピアソン分布族には Type I～VII があるが，本稿のモンテカルロ実験ではほとんどの場合，Type IV または Type VII のどちらかしか選ばれないので，以下の分析では，この二つの分布に焦点を絞った。

(ii) 構造誤差の無相関性が未知の場合
—fastICA に基づく疑似尤度関数の利用—

まず初めに ICA と fastICA の概略を述べる。いま n 個の独立な分布に従う潜在変数 $s_1, s_2, \cdots, s_n$ 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さた結果として，n の観測値 $x_1, x_2, \cdots, x_n$ が観測される状況を考える。この状況は，式を用いて

$$x = \mathrm{As}$$

と表すことが出来る。ここに $x = (x_1, x_2, \cdots, x_n)'$，$s = (s_1, s_2, \cdots, s_n)'$ である。また A は，フルランクを持つ $n \times n$ 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 と A をいかに分離抽出するかという問題である。ただし $\varepsilon_{1t}, \cdots, \varepsilon_{4t}$ は独立に非ガウス分布にしたがって分布する（したがって無相関である）と仮定される。ここで x を誘導形誤差 u_t に，s を構造誤差 ε_t に，A を B_0^{-1} に置き換えれば ICA モデルは（4）式の

$$u_t = B_0^{-1} \varepsilon_t$$

と同じ形になるので，ICA を応用すれば u_t を B_0^{-1} と ε_t に分離して推定することが可能になる。この計算の基本方針は，ε_t の非正規性の尺度を最大化することによって，B_0^{-1} と ε_t を推定するというものである。非正規性の尺度として情報理論における negantropy という概念が用いられる。fastICA においては，ε_t の要素間に相関があれば，それらを無相関化（白色化）するプロセスが組み込まれている。

この計算は R の fastICA によって実行することができる。計算の概略は，Hyvärinen, A., and Erkki Oja（2000）または Hyvärinen, A., Juha Karhunen, Erkki Oja（2001），Marchini et al.（2019）などを参照されたい。

このように，現実には観測できない $\varepsilon_{1t}, \cdots, \varepsilon_{4t}$ の代わりとして，fastICA を使って可能な限り独立（従って無相関）な $\hat{\varepsilon}_{1t}, \cdots, \hat{\varepsilon}_{4t}$（いわば $\varepsilon_{1t}, \cdots, \varepsilon_{4t}$ の推定値）を求めることができる。

R の fastICA のプログラムは，その名の通り

計算が非常に高速化された優れたアルゴリズムを実装している。しかし限界もある。その限界とは出力される混合行列の推定値は，変数の順序，係数値の符号とスケールに関しては一意性を持たないことである。係数推定が主目的である場合は，結果を見ながら経済的意味などの外部情報を用いて，これらの不定性を正さなければならない。この作業は変数の数が大きくなると容易ではない。しかし独立成分の推定値のみを見たい場合にはこの不定性は問題にならない。そこで本稿では fastICA を $\hat{\varepsilon}_{1t},\cdots,\hat{\varepsilon}_{4t}$ を取り出すためだけに利用した。

次に，このようにして得られた $\hat{\varepsilon}_{1t},\cdots,\hat{\varepsilon}_{4t}$ を pearsonMSC に引き渡し，最も適合度の高い確率密度関数を選択する。選択された密度関数を用いて構成される尤度関数を $\varepsilon_{1t},\cdots,\varepsilon_{4t}$ の疑似尤度関数として B_0^{-1} に含まれる未知パラメータの疑似最尤推定値を求めることができる。下記の実験11-2ではこの疑似最尤推定法を用いて計算した。

4.　モンテカルロ実験結果

以下の実験では，繰り返し回数 N は常に200とした。他方発生させる時系列の長さ T は200と1,000を用いた。乱数発生確率分布のタイプとしては t 分布，ラプラス（Laplace）分布，双曲線正割（Heperboloic Secant）分布を用いた。各分布に対して T = 200と1,000の場合を実行した。実験１～10では上述の疑似尤度関数構法（i）を用い，実験11では（i）と（ii）の比較を行っている。

実験結果１　グラフ４と表１
1）繰り返し回数＝200，時系列長＝200
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 0.5）分布
3）“pearsonMSC” の選択結果：PearsonVII（7.3, 0.0, 0.8）

実験結果２　表２
1）繰り返し回数＝200，時系列長＝1,000
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 0.5）分布
3）“pearsonMSC” の選択結果：PearsonVII（7.3, 0.0, 0.8）

実験結果３　表３
1）繰り返し回数＝200，時系列長＝200
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 1.0）分布
3）“pearsonMSC” の選択結果：PearsonVII（7.0, 0.0, 1.6）

実験結果４　表４
1）繰り返し回数＝200，時系列長＝1,000
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 1.0）分布
3）“pearsonMSC” の選択結果：PearsonVII（7.0, 0.0, 1.6）

実験結果５　表５
1）繰り返し回数＝200，時系列長＝200
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 2.0）分布
3）“pearsonMSC” の選択結果：PearsonVII（6.8, 0.0, 3.1）

実験結果６　表６
1）繰り返し回数＝200，時系列長＝1,000
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 2.0）分布
3）“pearsonMSC” の選択結果：PearsonVII（6.8, 0.0, 3.1）

実験結果７　表７
1）繰り返し回数＝200，時系列長＝200
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Hyperbolic secant

（mu = 0.0, sig = 2）分布
3）“pearsonMSC”の選択結果：PearsonVII（10.5, 0, 2.3）

実験結果 8　表 8
1）繰り返し回数＝200，時系列長＝1,000
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Hyperbolic secant（mu = 0.0, sig = 2）分布
3）“pearsonMSC”の選択結果：PearsonVII（10.5, 0, 2.3）

実験結果 9　表 9
1）繰り返し回数＝200，時系列長＝200
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：t（df = 5）分布
3）“pearsonMSC”の選択結果：PearsonVII（8.8, 0.0, 1.4）

実験結果10　表10
1）繰り返し回数＝200，時系列長＝1,000
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：t（df = 5）分布
3）“pearsonMSC”の選択結果：PearsonVII（8.8, 0.0, 1.4）

実験結果11-1　$\boldsymbol{E}\left(\boldsymbol{\varepsilon_t\varepsilon_t'}\right)\neq\boldsymbol{I}$ の場合（**before whitening**）　表11-1
1）繰り返し回数＝200，時系列長＝100
2）$\varepsilon_t=\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)'$ 生成過程：Laplace（r = 0.5）分布
3）ε_t の相関行列

$$\mathrm{Corr}(\varepsilon_t)=\begin{pmatrix}1.0 & 0.8 & 0.6 & 0.4\\ 0.8 & 1.0 & 0.8 & 0.6\\ 0.6 & 0.8 & 1.0 & 0.8\\ 0.4 & 0.6 & 0.8 & 1.0\end{pmatrix}$$

の下で，ε_t を生成する。
4）この相関を無視して，$\Sigma_{\tilde{u}\tilde{u}'}$ の直行変換を使ってこれまでと同じ疑似最尤法を用いる。“pearsonMSC”の選択結果：PearsonVII（7.7, 0.0, 0.78）
5）この PearsonVII に基づく尤度関数を使って，疑似最尤推定値を計算する。

実験11-2　$\boldsymbol{E}\left(\boldsymbol{\varepsilon_t\varepsilon_t'}\right)\neq\boldsymbol{I}$ の場合に **fastICA fast ICA** を適用した（**after whitening**）　表11-2
手順1），2），3）は実験11-1と同じ
4）誘導形残差を fastICA に渡し，無相関化の手続きを経た残差系列を取り出す。
5）無相関化された残差の分布を，“pearsonMSC”に選択させる。選択結果 PearsonVII（3.49, 0.0, 0.71）が選ばれた。
6）この PearsonVII に基づく尤度関数を使って，疑似最尤推定値を計算する。

実験結果に見られる傾向的特徴。

以上の実験に於いて，DGP においてラプラス分布，双曲線正割分布，t 分布の何れの分布に対しても，PearsonMSC は，少数の例外を除けば，PearsonVII 分布を選択したので，PearsonVII 分布に基づく疑似尤度最尤推定の結果のみを掲げた。そして推定精度に関しては，以下のような傾向的特徴が見られた。

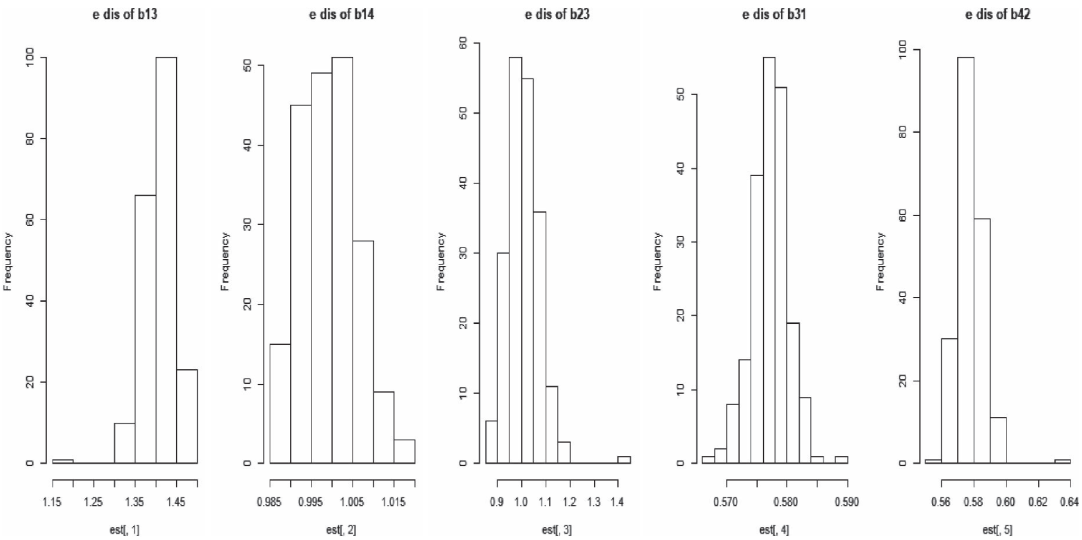
１．実験結果１－６：DGP にラプラス分布を使用した場合。分布のパラメター r は結果の推定精度にやや影響を与える。すなわち，r = 0.5 のとき，T の大きさに拘わらず推定の精度はよい。しかし r = 1.0のときは推定にやや偏りが見受けられる。

２．実験結果７－８：DGP に Hyperbolic secant（mu = 0.0, sig = 2）を使用した場合。推定精度が良くない場合が見られる（太字で示した個所）。それ以外については，推定精度は良い。T の大きさの違いは推定結果にほとんど影響がない。

３．実験結果９－10：DGP に t（df=5）分布を使用した場合。

推定精度は T の大きさに拘わらず，非常に良い。

４．実験結果11：　誤差項 ε_t に相関があるに



グラフ 4　B_0^{-1} の 5 つの未知パラメータ疑似最尤推定量の経験分布

表 1　標本実験 1 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4091	0.0368	1.375e-03
1.0000	0.9992	0.0067	4.527e-05
1.0000	1.0104	0.0688	4.828e-03
0.5733	0.5771	0.0030	9.425e-06
0.5733	0.5776	0.0082	6.697e-05

表 2　標本実験 2 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4144	0.0142	1.997e-04
1.0000	1.0000	0.0031	8.979e-06
1.0000	1.001	0.0271	7.331e-04
0.5733	0.5775	0.0014	2.158e-06
0.5733	0.5778	0.0029	8.382e-06

表 3　標本実験 3 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.5627	0.0610	0.0256
1.0000	0.9839	0.0144	0.0004
1.0000	**0.6798**	0.1098	0.1145
0.5733	0.5575	0.0089	0.0004
0.5733	0.5511	0.0185	0.0010

表 4　標本実験 4 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.5705	0.2904	0.1083
1.0000	0.9855	0.0324	0.0012
1.0000	**0.6761**	0.4978	0.3515
0.5733	0.5578	0.0128	0.0005
0.5733	0.5528	0.0286	0.0014

表 5　標本実験 5 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	**1.7741**	0.8170	0.7937
1.0000	0.9787	0.0133	0.0006
1.0000	**0.2891**	1.5014	2.7484
0.5733	0.5434	0.0086	0.0012
0.5733	0.5372	0.0223	0.0021

表 6　標本実験 6 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	**1.8131**	0.8311	0.8464
1.0000	0.9839	0.0264	0.0009
1.0000	**0.2269**	1.4671	2.7392
0.5733	0.5420	0.0098	0.0013
0.5733	0.5381	0.0324	0.0025

表 7 標本実験 7 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4689	0.0581	0.0063
1.0000	0.9914	0.0091	0.0001
1.0000	**0.8759**	0.1119	0.0278
0.5733	0.5669	0.0066	0.0001
0.5733	0.5269	0.0107	0.0003

表 8 標本実験 8 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4691	0.0495	5.4574e-03
1.0000	0.9924	0.0043	7.5807e-05
1.0000	**0.8769**	0.0983	2.4786e-02
0.5773	0.5677	0.0025	9.9076e-05
0.5773	0.5642	0.0057	2.0448e-04

表 9 標本実験 9 の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4506	0.0314	2.3050e-03
1.0000	0.9954	0.0076	7.8892e-05
1.0000	0.9159	0.0575	1.0364e-02
0.5773	0.5686	0.0027	8.3268e-05
0.5773	0.5676	0.0068	1.4184e-04

表10 標本実験10の結果の平均，標本標準偏差，真値に対する平均平方誤差

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4537	0.0142	1.7667e-03
1.0000	0.9953	0.0030	3.1542e-05
1.0000	0.9111	0.0270	8.6142e-03
0.5773	0.5687	0.0015	7.6368e-05
0.5773	0.5679	0.0031	9.8706e-05

表11-1 Before whitening の結果

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.3191703	0.133387644	0.16324079
1.0000	1.0120633	0.020894101	0.02403584
1.0000	0.9588554	0.273249085	0.27497507
0.5773	0.6008275	0.008121468	0.02482899
0.5773	0.4933481	0.014219820	0.08518532

表11-2 After whitening の結果

DGP の真値	平均	標準偏差	平均平方誤差
1.4142	1.4064665	0.037817728	0.036703951
1.0000	1.0062049	0.007510153	0.009447937
1.0000	1.0159151	0.069363040	0.067700800
0.5773	0.5758451	0.003730274	0.003845662
0.5773	0.5758451	0.005989417	0.005890742

も拘らず（$E\left(\varepsilon_t\varepsilon_t'\right) \neq I$ の場合），相関が無いという仮定に立って直交変換法を用いた疑似尤度最尤法を行うと，推定結果の精度はかなり劣化する。しかし fasstICA を用いて相関を無相関化した上で推定を行うと，推定精度は非常に改善されることが示された。

５． 最後に全体的な留意点を幾つか挙げておく。

1） “pearsonMSC” によって正規分布に非常に近い分布が選ばれた場合は疑似尤度最尤法は当然よくない結果に終わる。

2） 実験 1 － 4 において，たとえ $\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)$ の全てを同一のラプラス分布から発生させても，それぞれに対して “pearsonMSC” が異なる分布を選択する可能性がある。そのような場合に，同一の分布を当てはめると，推定値にバイアスが生じることがある。そのときは $\left(\varepsilon_{1t},\cdots,\varepsilon_{4t}\right)$ の個々の要素にそれぞれに対応する選ばれた分布を用いて尤度関数を構成する方がよい。

3） 実際のデータ分析においては，時系列の長さは十分長いとは言い難い場合も多い。その時は bootstrap 法を援用してみる価値はある。しかし bootstrap 法を試してみたが，常に改善されるとは限らなかった。

5. む す び

本稿では可能な限り先験的情報を用いず，「データに語らせる」という data-driven な立場に立つ分析法として，非ガウス型構造 VAR モデルの推定と識別性に関する理論的及び数値計算的立場から考察した。そして構造誤 ε_t が非ガウス分布に従うという仮定の下で，ある疑似最尤法を提案し，それに基づく疑似尤度最尤法のパフォーマンスをモンテカルロ実験で検証した。その結果から判断する限り，本稿で提案された疑似最尤法は，よいパフォーマンスを示した。今後は，この方法を実際の経済データを使って様々なモデルに応用し，有効性を確認する予定である。

英文参考文献

(1) Hyvärinen, A., and Erkki Oja (2000), Independent Component Analysis: Algorithms and applications, Neural Networks 13, 411–430
(2) Hyvärinen, A., Juha Karhunen, and Erkki Oja (2001), Independent Component Analysis, John wiley & sons, Inc.
(3) Shimizu, S., Hoyer, P. O., Hyvärinen, A., and Kerminen, A. (2006), A Linear non-Gaussian acyclic model for causal discovery, Journal of Machine Leaning Research, Vol. 7
(4) Hyvarinen, A., Zhang, K., Shimizu, S., and Hoyer, P. (2010), Estimation of Structural Vector Autoregression Model Using Non-Gaussianity, J. Machine Learning Research 11, 1709–1731
(5) Moneta, A., Entner, D., Hoyer, P. O., Coad, A. (2013), Causal Inference by Independent Component Analysis: Theory and Applications. Oxford Bulletin of Economics and Statistics, 75, 5
(7) Gourieroux, C., Monfort., A., and Renme, J. (2017), Statistical inference for independent component analysis: Application to structural VAR models, Journal of Econometrics 196, 111–126
(8) Lanne, M., Meitz, M., and Saikonen, P. (2017), Identification and estimation of non-Gausian structural vector autoregressions, Journal of Econometrics 196, 288–304
(9) Maekawa, K., and Nakanishi, T. (2021), Estimation of Non- Gaussian Structural VAR model—A Flexible Quasi Likelihood Function Approach—, (2021), presented at EcoSta 2021

日本語参考文献

(1) 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―，広島経済大学研究叢書　第４冊　東アジアの経済成長の持続可能性について，第９章　広島経済大学地域経済研究所　2016年
(2) 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―，広島経済大学創立五十周年記念論文集（上巻）pp. 25–62
(3) 前川功一（2020）構造 VAR モデルの識別性と推定（I）―正規性誤差の場合―　広島経済大学経済研究論集　第43巻第２号　pp. 23–40，2020年11月
(4) 前川功一（2021）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―，広島経済大学経済研究論集　第43巻第３号　pp. 119–127，2021年３月

本稿で使用した主な R パッケージ

(1) Marchini, J. L., C. Heaton and B. D. Ripley, Package ‘fastICA’ July 2019
(2) Martin Becker, and Stefan Klobner, Package ‘PearsonDS’, July 2017
(3) Alexander Lange, Bernhard Dalheimer, Helmut Herwartz, and Simone Maxand, Package ‘svars’, September, 2019