日本統計学会誌
第 39 巻, 第 2 号, 2010 年 3 月
299 頁 ～ 314 頁

ダイレクトモンテカルロ法による Seemingly unrelated regression モデルのベイズ推定

アーノルド・ゼルナー*, 安道　知寛†

A Direct Monte Carlo Approach for Bayesian Analysis of the Seemingly Unrelated Regression Model with Informative Prior

Arnold Zellner* and Tomohiro Ando†

Seemingly unrelated regression (SUR) モデルのベイズ推定について考察する．一般的にベイズ推定においてはマルコフ連鎖モンテカルロ (Markov Chain Monte Carlo (MCMC)) 法が利用されているが，収束判定，サンプリングの効率性等の問題が常に MCMC 法には付随する．本稿では，ダイレクトモンテカルロ (direct Monte Carlo (DMC)) 法に基づく SUR モデルのベイズ推定法を提案する．Zellner and Ando (2008a) では，Jeffreys's prior (Jeffreys (1946, 1961)) に基づく SUR モデルのベイズ推定をおこなっているのに対し，本稿では正規分布を係数パラメータに仮定して，DMC 法を構築する．DMC 法と MCMC 法を数値実験に基づき比較した結果，アルゴリズムの効率性など様々な利点が DMC 法に確認された．

Computationally efficient methods for Bayesian analysis of seemingly unrelated regression (SUR) models are described and applied that involve use of a direct Monte Carlo (DMC) approach to calculate Bayesian estimation and prediction results. In contrast to Zellner and Ando (2008a), where a diffuse prior density is utililized, herelin we develop a DMC approach using an informative prior for the regression parameters. This DMC approach is employed to compute Bayesian marginal posterior densities, moments, intervals and other quantities, using data simulated from known models. The results obtained by the DMC approach are compared to those yielded by use of a Markov Chain Monte Carlo (MCMC) approach. It is concluded from these comparisons that the DMC approach is very worthwhile and applicable to many SUR and other problems.

キーワード: ベイズ推定，ダイレクトモンテカルロ，マルコフ連鎖モンテカルロ

1. はじめに

Seemingly unrelated regression (SUR) モデル (Zellner (1962)) は，計量経済学をはじめとする様々な学術分野で利用されている．SUR モデルの導入以降，推定，検定，予測等の

---

* Booth School of Business, University of Chicago, 5807 S. Woodlawn Avenue, Chicago, IL 60637, U.S.A. (E-mail: arnold.zellner@chicagogsb.edu.).
† 慶應義塾大学大学院経営管理研究科，〒 223-8526 神奈川県横浜市港北区日吉 4-1-1, (E-mail: an-doh@kbs.keio.ac.jp).

様々な観点から研究がなされている (例えば，Zellner (1962, 1963), Gallant (1975), Rocke (1989), Neudecker and Windmeijer (1991), Mandy and Martins-Hilho (1993), Kurata (1999), Liu (2002), Ng (2002), Carroll *et al.* (2006))．経済学分野で頻繁に利用される同時方程式モデル (Simultaneous equation model) も SUR モデルで定式化可能で，その応用範囲の広さから，教科書 (例えば，Greene (2002), Lancaster (2004), Geweke (2005), Rossi *et al.* (2005)) にも頻繁に取り上げられている．

SUR モデルの推定においては，Generalized least squares approach (Zellner (1962, 1963))，Likelihood distributional approach (Frasera *et al.* (2005))，Bayesian method of moments (van der Merwve and Viljoen (1988)) など様々な手法が提案されている．特に，現在，高速計算機の利用環境が整備されたことにより，MCMC 法を援用したベイズ推定が急速に広まっている (Percy (1992), Chib and Greenberg (1995), Percy (1996), Smith and Kohn (2000))．MCMC 法の一般論については，Carlin and Louis (1996)，Tierney (1994) などを参照されたい．

MCMC 法は，様々な統計モデルに対するベイズ推定の適用可能領域を広げた．しかし，MCMC 法には，現在においても完全には解決されていない問題が存在することも事実である．それは，MCMC 法に不慣れなユーザー側にとっては非常に厄介な問題でもある．例えば，Burn in period の長さ，提案分布の選択，収束の判定などの問題である．収束の判定には Schruben (1982), Heidelberger and Welch (1983), Gelman and Rubin (1992), Geweke (1992), Raftery and Lewis (1992), Zellner and Min (1995), Brooks and Gelman (1997) などの研究が提案されているものの，統計的検定等に基づいており，完璧な解決策ではないことは自明であろう．

そこで，本稿においては，MCMC 法に付随する問題を回避するため，DMC 法に基づいた SUR モデルのベイズ推定法を提案する．DMC 法の一般論については，Geweke (2005, p. 106) などを参照されたい．表 1 は MCMC 法と DMC 法の特徴に関する比較をおこなっている．表 1 からわかるように，提案する手法は事後分布からのサンプリングを非常に容易とし，ユーザー側にとっては非常に有用なツールとなる．つまり，ユーザー側は事後サンプリングの回数さえ指定すれば，後は事後サンプリングが自動的に実行されるのである．DMC 法に基づいた事後分布からのサンプルを利用することにより，パラメータの同時事後密度関数のみならず，条件付き事後密度関数，周辺事後密度関数，予測分布など様々な計算が容易となる．

一般的には，適用できない或いは事後分布が明示的に解けない場合，DMC 法が実行できないが，MCMC 法では推定可能な場合がある．そのため，DMC 法の適用範囲は限定的であることも指摘しておく．

本稿の構成は以下の通りである．2 節では，SUR モデルの概略と MCMC 法に基づくベ

表 1　MCMC 法と DMC 法の特徴に関する比較.

	DMC 法	MCMC 法
事後サンプリングの回数指定が必要	Yes	Yes
事後サンプルの独立性	Yes	No
100% で発生させたサンプルが採択される	Yes	No[a]
事後サンプリングにおいて初期値が必要	No	Yes
事後サンプリングにおいて burn-in period が必要	No	Yes
事後サンプリングにおいて収束判定が必要	No	Yes
事後サンプリングにおいて提案分布が必要	No	Yes
適用範囲	限定的	広い

[a] ギブスサンプリングの場合，100% で発生させたサンプルが採択される.

イズ推定について解説する．3 節では，DMC 法に基づいた SUR モデルのベイズ推定法を提案する．4 節では，数値実験をおこない提案する DMC 法を MCMC 法と比較し，5 節で結論をまとめる．

2.　SUR モデルとそのベイズ推定

2.1　SUR モデル

SUR モデルは m 本の回帰分析モデルで定義される．

$$\boldsymbol{y}_j = \boldsymbol{X}_j \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j,\ j = 1, \ldots, m, \tag{2.1}$$

ここで，$\boldsymbol{y}_j$，$\boldsymbol{\varepsilon}_j$ は $n \times 1$-次元ベクトル $\boldsymbol{X}_j$ は $n \times p_j$ 次元行列，$\boldsymbol{\beta}_j$ は p_j-次元ベクトルである．誤差項間に独立性を仮定した場合には，一般によく利用されている回帰モデルであるが，SUR モデルにおいては，誤差項間に相関

$$E[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j'] = \begin{cases} \sigma_{ij} I_n, & (i \neq j) \\ \sigma_i^2 I, & (i = j) \end{cases}$$

がある．ここで，I_n は n-次元単位行列とする．(2.1) 式から明らかなように，それぞれの方程式は異なる説明変数行列 $\boldsymbol{X}_j$，分散パラメータ σ_j^2 $(j = 1, \ldots, m)$ を持っており，誤差項は相関をもっているという特徴をもつ．

(2.1) 式の線形 SUR モデルにおいて，$n \times p_j$ 次元行列 $\boldsymbol{X}_j$ は様々な形式をとることができる.，最も標準的な形式としては，p_j-次元説明変数 $\boldsymbol{x}_j = (x_1, \ldots, x_{p_j})$ からなるものであろう．無論，定数項を含むこともでき，その非線形化も可能であり，多項式モデル，加法モデル (Hastie and Tibshirani (1990))，基底関数展開（Hastie *et al.* (2001)) などを説明変数行列 $\boldsymbol{X}_j$ に利用できる．

加法モデルは，各説明変数 x_k に基づく非線形関数 $h_k(x_k)$ の線形和 $u(\boldsymbol{x}) = \sum_k h_k(x_k)$ で定式化される．非線形関数 $h_k(x_k)$ の例としては，B-スプライン関数などの基底関数の線形和 $h_k(x_k) = \sum_{\ell=1}^{b} \gamma_{k\ell}\phi_{k\ell}(x_k) = \boldsymbol{\gamma}_k'\boldsymbol{\phi}_k(x_k)$ 等がある．ここで，$\phi_{k\ell}(x_k)$ は基底関数，$\gamma_{k\ell}$ はパラメータ，b は基底関数の個数である．基底関数の個数が各非線形関数 $h_k(x_k)$ で違っていてもよいが，表記の複雑さを避けるために，基底関数の個数を b で統一している．(2.1) 式において加法モデルを利用する場合，以下のようになる．

$$\begin{aligned}
\begin{pmatrix} y_{j,1} \\ \vdots \\ y_{j,n} \end{pmatrix} &= \begin{pmatrix} \phi_{j,11}(x_{j,11}) & \cdots & \phi_{j,1b}(x_{j,11}) \\ & \vdots & \\ \phi_{j,11}(x_{j,1n}) & \cdots & \phi_{j,1b}(x_{j,1n}) \end{pmatrix} \begin{pmatrix} \gamma_{j,11} \\ \vdots \\ \gamma_{j,1b} \end{pmatrix} + \cdots \\
&\quad + \begin{pmatrix} \phi_{j,p_j1}(x_{j,p_j1}) & \cdots & \phi_{j,p_jb}(x_{j,p_j1}) \\ & \vdots & \\ \phi_{j,p_j1}(x_{j,p_jn}) & \cdots & \phi_{j,p_jb}(x_{j,p_jn}) \end{pmatrix} \begin{pmatrix} \gamma_{j,p_j1} \\ \vdots \\ \gamma_{j,p_jb} \end{pmatrix} + \begin{pmatrix} \varepsilon_{j,1} \\ \vdots \\ \varepsilon_{j,n} \end{pmatrix} \\
&= B_{j1}\boldsymbol{\gamma}_{j1} + \cdots + B_{jp_j}\boldsymbol{\gamma}_{jp_j} + \boldsymbol{\varepsilon}_j
\end{aligned}$$

となる．ここで，

$$\begin{pmatrix} B_{j1} \vdots B_{j2} \vdots \cdots \vdots B_{jp_j} \end{pmatrix} \to \boldsymbol{X}_j, \qquad (\boldsymbol{\gamma}_{j1}', \ldots, \boldsymbol{\gamma}_{jp_j}')' \to \boldsymbol{\beta}_j$$

と対応させると，加法モデルを利用した場合にも (2.1) 式に帰着することが分かる．また，基底関数展開（Hastie *et al.* (2001)）等の非線形関数を利用した場合においても，(2.1) 式に帰着することが容易に類推できよう．説明変数行列 $\boldsymbol{X}_j$ の定式化は，特定の問題に依存するのでその議論はここまでにとどめておく．

(2.1) 式を行列で表現すると

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1 & O & \cdots & O \\ O & \boldsymbol{X}_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \boldsymbol{X}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_m \end{pmatrix},$$

もしくは

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \Sigma \otimes I),$$

と表現できる．ここで $\otimes$ はテンソル積，Σ は $m \times m$-次元共分散行列で，その対角成分は $(\Sigma)_{jj} = \sigma_j^2$，非対角成分は $(\Sigma)_{ij} = \sigma_{ij}$ で与えられる．SUR モデルに含まれるパラメータは $\boldsymbol{\beta}$，Σ であり，その最尤推定量は尤度関数

$$f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\beta}, \Sigma) = \frac{1}{(2\pi)^{nm/2}|\Sigma|^{n/2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{R\Sigma^{-1}\right\}\right],$$

の最大化による．ここで "tr" は行列の対角和 $|\Sigma| = \det(\Sigma)$ は Σ の行列式，$m \times m$-次元行列 R の ij 成分 $R = (r_{ij})$ は $r_{ij} = (\boldsymbol{y}_{ni} - \boldsymbol{X}_{ni}\boldsymbol{\beta}_i)'(\boldsymbol{y}_{nj} - \boldsymbol{X}_{nj}\boldsymbol{\beta}_j)$ で与えられる．次節では，SUR モデルのベイズ推定について触れる．

2.2 SUR モデルのベイズ推定

Zellner (1971)，Box and Tiao (1973)，Percy (1992) は，SUR モデルのベイズ推定について考察している．いま，パラメータに関する事前知識が皆無であるとする．このとき，自然な事前分布の設定の一つとして，Jeffreys's prior (Jeffreys (1946, 1961))

$$\pi_1(\boldsymbol{\beta}, \Sigma) = \pi_1(\boldsymbol{\beta})\pi_1(\Sigma) \propto |\Sigma|^{-\frac{m+1}{2}} \tag{2.2}$$

が考えられる．このとき，パラメータの同時事後分布は

$$\pi_1(\boldsymbol{\beta}, \Sigma|\boldsymbol{Y}, \boldsymbol{X}) \propto |\Sigma|^{-(n+m+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{R\Sigma^{-1}\right\}\right]$$

と表現される．残念ながら $\boldsymbol{\beta}$，Σ の同時事後分布は解析的に求められないことが知られているものの，$\boldsymbol{\beta}$，Σ の条件付き事後分布 $\pi_1(\boldsymbol{\beta}|\boldsymbol{Y}_n, \boldsymbol{X}_n, \Sigma)$，$\pi(\Sigma|\boldsymbol{Y}_n, \boldsymbol{X}_n, \boldsymbol{\beta})$ は解析的に求められる．

$$\pi_1\left(\boldsymbol{\beta}|\boldsymbol{Y}, \boldsymbol{X}, \Sigma\right) = N\left(\hat{\boldsymbol{\beta}}, \hat{\Omega}\right), \quad \pi_1\left(\Sigma|\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\beta}\right) = IW\left(R, n\right), \tag{2.3}$$

ただし，

$$\hat{\boldsymbol{\beta}} = \left\{\boldsymbol{X}'\left(\Sigma^{-1} \otimes I\right)\boldsymbol{X}\right\}^{-1}\boldsymbol{X}'\left(\Sigma^{-1} \otimes I\right)\boldsymbol{y},$$
$$\hat{\Omega} = \left(\boldsymbol{X}'\left(\Sigma^{-1} \otimes I\right)\boldsymbol{X}\right)^{-1}$$

とする．ここで，$IW(\cdot,\cdot)$ は Inverse Wishart 分布である．いま，$\boldsymbol{\beta}$，Σ の条件付き事後分布 $\pi_1(\boldsymbol{\beta}|\boldsymbol{Y}, \boldsymbol{X}, \Sigma)$，$\pi_1(\Sigma|\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\beta})$ は解析的に与えられているので，適当なパラメータの値（例えば，最尤推定量など）を初期値 $\boldsymbol{\beta}^{(0)}$，及び $\Sigma^{(0)}$ とし，ギブスサンプリング法により事後サンプリングをおこなえばよいこととなる．

また，パラメータ $\boldsymbol{\beta}$ に関して事前に何らかの情報を持っている場合，$\boldsymbol{\beta}$ の事前分布として正規分布 $\pi(\boldsymbol{\beta}) = N(\boldsymbol{\beta}_0, A^{-1})$ を利用することがある．事前分布は正規分布である必要はないが，解析的な利便性がある．$\boldsymbol{\beta}$，及び Σ の同時事前分布が $\pi_2(\boldsymbol{\beta}, \Sigma) = \pi_2(\boldsymbol{\beta})\pi_2(\Sigma)$ $\pi_2(\boldsymbol{\beta}) = N(\boldsymbol{\beta}_0, A^{-1})$，$\pi_2(\Sigma) = |\Sigma|^{-(m+1)/2}$ と設定される場合にも，$\boldsymbol{\beta}$，及び Σ の条件付き事後分布は解析的に求められる．それらは

$$\pi_2\left(\boldsymbol{\beta}|\Sigma, \boldsymbol{Y}, \boldsymbol{X}\right) = N\left(\hat{\boldsymbol{\beta}}_A, \hat{\Omega}_A\right), \quad \pi_2\left(\Sigma|\boldsymbol{\beta}, \boldsymbol{Y}, \boldsymbol{X}\right) = IW\left(R, n\right),$$

である．ただし，

$$\hat{\boldsymbol{\beta}}_A = (\boldsymbol{X}'(\Sigma^{-1} \otimes I)\boldsymbol{X} + A)^{-1}(\boldsymbol{X}'(\Sigma^{-1} \otimes I)\boldsymbol{X}\hat{\boldsymbol{\beta}} + A\boldsymbol{\beta}_0),$$
$$\hat{\Omega}_A = (\boldsymbol{X}'(\Omega^{-1} \otimes I)\boldsymbol{X}' + A)^{-1}.$$

先ほどと同様，$\boldsymbol{\beta}$，Σ の条件付き事後分布 $\pi_2(\boldsymbol{\beta}|\boldsymbol{Y},\boldsymbol{X},\Sigma)$，$\pi_2(\Sigma|\boldsymbol{Y},\boldsymbol{X},\boldsymbol{\beta})$ は解析的に与えられているので，ある初期値を設定してギブスサンプリング法から事後サンプリングを行えばよい．

しかし，1 節で指摘したとおり，MCMC 法に基づくベイズ推定には様々な注意点がある．それらの例については，4 節の数値実験でも触れているので参照されたい．次節では，MCMC 法に付随する問題を回避するため，DMC 法に基づいた SUR モデルのベイズ推定法を提案する．

3. DMC 法による SUR モデルのベイズ推定

3.1 SUR モデルの変換

ここでは，DMC 法による事後サンプリングをおこなうために以下の変換をおこなう．

$$\begin{cases} \boldsymbol{y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}_1, \quad \boldsymbol{e}_1 = \boldsymbol{\varepsilon}_1 \\ \boldsymbol{y}_2 = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \rho_{21}\boldsymbol{\varepsilon}_1 + \boldsymbol{e}_2, \\ \qquad\qquad \vdots \\ \boldsymbol{y}_m = \boldsymbol{X}_m\boldsymbol{\beta}_m + \sum_{j=1}^{m-1}\rho_{mj}\boldsymbol{\varepsilon}_j + \boldsymbol{e}_m. \end{cases}$$

ただし，

$$E[\boldsymbol{e}_i\boldsymbol{e}_j'] = \begin{cases} O_n, & (i \neq j) \\ \omega_i^2 I, & (i = j) \end{cases}, \quad \Omega = \mathrm{diag}\{\omega_1^2,\ldots,\omega_m^2\},$$

である．ここで，O_n は n-次元ゼロ行列とする．また，行列 Ω は変換された SUR モデルの誤差項の分散として定義した．無論，3.2 節で述べるように行列 Ω はオリジナルの SUR モデルのパラメータと対応関係が明示的にある．この変換を利用して，Zellner and Ando (2008a) は SUR モデルのベイズ推定法を提案している．Zellner *et al.* (1988)，Zellner and Chen (2002) は，この変換を同時方程式モデル (Simultaneous equation model) のベイズ推定に利用している．

このとき，変換された SUR モデルは以下で与えられる．

$$\begin{cases} \boldsymbol{y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}_1 \equiv \boldsymbol{Z}_1\boldsymbol{b}_1 + \boldsymbol{e}_1, \\ \boldsymbol{y}_j = \boldsymbol{X}_j\boldsymbol{\beta}_j + \sum_{l=1}^{j-1}\rho_{jl}(\boldsymbol{y}_l - \boldsymbol{X}_l\boldsymbol{\beta}_l) + \boldsymbol{e}_j \equiv \boldsymbol{Z}_j\boldsymbol{b}_j + \boldsymbol{e}_j, \quad j = 2,\ldots,m, \end{cases} \tag{3.1}$$

$n \times (p_j + j - 1)$-次元行列 $\boldsymbol{Z}_j$ は $\boldsymbol{\beta}_{j-1},\ldots,\boldsymbol{\beta}_1$ を含んでいるが，j-番目の方程式を変換する際には，$\boldsymbol{\beta}_{j-1},\ldots,\boldsymbol{\beta}_1$ は既に固定された値であることに注意されたい．また，この変換は誤差項 $\boldsymbol{e}_i$ を互いに独立にするため，パラメータ $\boldsymbol{b}_j$ $(j = 1,..,m)$ の推定が容易となる利点がある．すなわち，変換された SUR モデルの尤度関数は

$$f(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{b},\Omega) = \prod_{j=1}^{m}\frac{1}{(2\pi\omega_j^2)^{n/2}}\exp\left[-\frac{(\boldsymbol{y}_j - Z_j\boldsymbol{b}_j)'(\boldsymbol{y}_j - Z_j\boldsymbol{b}_j)}{2\omega_j^2}\right].$$

となる．ここで，$\boldsymbol{b} = (\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_m)'$ は係数パラメータとする．オリジナルの SUR モデル (2.1) 式と比較すると分かるように，$E[\boldsymbol{e}_i \boldsymbol{e}'_j] = O, (i \neq j)$ の理由から m-本の回帰モデルの尤度関数の積に分解されている．

3.2 パラメータの対応関係，及び事前分布の設定

Zellner and Ando (2008a) は，(2.1) 式で与えられるオリジナルの SUR モデルのパラメータ $\{\boldsymbol{\beta}, \Sigma\}$ と変換された SUR モデルのパラメータ $\{\boldsymbol{b}, \Omega\}$ には 1 対 1 の対応関係があることを示した．Σ と Ω の関係は

$$\begin{aligned} \sigma_1^2 &= \omega_1^2, \\ \sigma_j^2 &= \sum_{k=1}^{j-1} \rho_{jk}^2 \sigma_k^2 + \sum_{k,l=1,\ k<l}^{j-1} \rho_{jk}\rho_{jl}\sigma_{lk} + \omega_j^2, \quad (j \neq 1), \\ \sigma_{ji} &= \sum_{k=1, k\neq i}^{j-1} \rho_{jk}\sigma_{ki} + \rho_{ji}\sigma_i^2, \quad (j \neq 1). \end{aligned} \tag{3.2}$$

である．また，回帰係数については

$$\begin{aligned} \boldsymbol{b}_1 &= \boldsymbol{\beta}_1, \\ \boldsymbol{b}_j &= (\boldsymbol{\beta}'_j, \rho_{j1}, \ldots, \rho_{j,j-1})', \quad j = 2, \ldots, m \end{aligned}$$

の対応関係がある．

Zellner and Ando (2008a) は，パラメータに関する事前知識が皆無である場合を考え，事前分布として，

$$\pi(\boldsymbol{b}, \Omega) \propto |\Omega|^{-(m+1)/2} = \prod_{j=1}^{m} (\omega_j^2)^{-1/2} \tag{3.3}$$

を利用している．それに対し，本稿では，$\boldsymbol{\beta}$，及び Σ の事前分布は以下の定式化による．

$$\pi(\boldsymbol{b}, \Omega) = \pi(\boldsymbol{b}|\Omega)\pi(\Omega), \quad \pi(\boldsymbol{b}) = \prod_{j=1}^{m} N(\boldsymbol{b}_{j0}, \omega_j^2 D_j^{-1}), \quad \pi(\Omega) \propto \prod_{j=1}^{m} (\omega_j^2)^{-1/2}. \tag{3.4}$$

すなわち係数パラメータについてなんらかの事前情報がある場合に対応している．次節では，パラメータ $\{\boldsymbol{b}, \Omega\}$ の事後分布を導出し，DMC 法を構築する．

3.3 事後分布と DMC 法

変換された SUR モデルの尤度関数 $f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{b}, \Omega)$，及び定式化した事前分布より，パラメータ $\{\boldsymbol{b}, \Omega\}$ の同時事後分布は

$$\begin{aligned} \pi(\boldsymbol{b}, \Omega|\boldsymbol{Y}, \boldsymbol{X}) \quad &\propto \quad \prod_{j=1}^{m} \frac{1}{(2\pi\omega_j^2)^{(n+1)/2}} \exp\left[-\frac{(\boldsymbol{y}_j - Z_j\boldsymbol{b}_j)'(\boldsymbol{y}_j - Z_j\boldsymbol{b}_j)}{2\omega_j^2}\right] \\ &\quad \times \prod_{j=1}^{m} \frac{|D_j|^{1/2}}{(2\pi\omega_j^2)^{(p_j+j-1)/2}} \exp\left[-\frac{(\boldsymbol{b}_j - \boldsymbol{b}_{j0})' D_j (\boldsymbol{b}_j - \boldsymbol{b}_{j0})}{2\omega_j^2}\right]. \end{aligned}$$

となる．さらに同時事後分布は，以下の条件付き事後分布の積に分解される．

$$\pi(\boldsymbol{b}_j|\boldsymbol{b}_{j-1},\ldots,\boldsymbol{b}_1,\omega_j^2,\boldsymbol{Y},\boldsymbol{X}) = N\left(\hat{\boldsymbol{b}}_j,\omega_j^2(Z_j'Z_j+D_j)^{-1}\right), \tag{3.5}$$

$$\pi(\omega_j^2|\boldsymbol{b}_{j-1},\ldots,\boldsymbol{b}_1,\boldsymbol{Y},\boldsymbol{X}) = IG\left(\hat{\nu}_j/2,\hat{\lambda}_j/2\right), \tag{3.6}$$

$j=1,\ldots,m$．ただし，

$$\begin{aligned}
\hat{\boldsymbol{b}}_j &= \left(Z_j'Z_j+D_j\right)^{-1}\left(Z_j'\boldsymbol{y}_j+D_j\boldsymbol{b}_{j0}\right),\\
\bar{\boldsymbol{b}}_j &= \left(Z_j'Z_j\right)^{-1}Z_j'\boldsymbol{y}_j,\\
\hat{\lambda}_j &= \left(\boldsymbol{y}_j-Z_j\bar{\boldsymbol{b}}_j\right)'\left(\boldsymbol{y}_j-Z_j\bar{\boldsymbol{b}}_j\right)+\hat{\boldsymbol{b}}_j'\left[\left(Z_j'Z_j\right)^{-1}+D_j^{-1}\right]^{-1}\hat{\boldsymbol{b}}_j,\\
\hat{\nu}_j &= n-2
\end{aligned}$$

である．この分解を利用すると，DMC 法による事後サンプリングが以下のアルゴリズムで可能となる．

DMC 法による事後サンプリング

Step 1 初期化 m 本の方程式の順番，及び事後サンプリングの回数 N を与える．いま $j=1$ とし，$\omega_1^{2(k)}$, $k=1,\ldots,N$ を (3.6) 式の事後分布から発生させ，(3.5) 式の $\boldsymbol{b}_1$ の条件付き事後分布 $\pi(\boldsymbol{b}_1|\omega_1^2,\boldsymbol{Y},\boldsymbol{X})$ に代入する．次に条件付き事後分布 $\pi(\boldsymbol{b}_1|\omega_1^{2(k)},\boldsymbol{Y},\boldsymbol{X})$ から $\boldsymbol{b}_1^{(k)}$, $k=1,\ldots,N$ を発生させる．

Step 2 j を $j+1\leftarrow j$ とする．分散パラメータ $\omega_j{}^{(k)}$, $k=1,\ldots,N$ を (3.6) 式の条件付き事後分布 $\pi(\omega_j^2|\boldsymbol{b}_{j-1}^{(k)},\ldots,\boldsymbol{b}_1^{(k)},\boldsymbol{Y},\boldsymbol{X})$ から発生させ，(3.5) 式の $\boldsymbol{b}_j$ の条件付き事後分布に代入する．Step 1 と同様に，$\boldsymbol{b}_j^{(k)}$ を条件付き事後分布 $\pi(\boldsymbol{b}_j|\boldsymbol{b}_{j-1}^{(k)},\ldots,\boldsymbol{b}_1^{(k)},\omega_j^{2(k)},\boldsymbol{Y}_n,\boldsymbol{X}_n)$, $k=1,\ldots,N$ から発生させる．

Step 3 Step 2 を $j=m$ となるまで逐次的におこなう．

このDMC法により，変換された SUR モデルのパラメータ $\{\boldsymbol{b},\Omega\}$ の事後サンプルを得る．(2.1) 式で与えられるオリジナルの SUR モデルのベイズ推定には，3.2 節で与えたパラメータ $\{\boldsymbol{\beta},\Sigma\}$ との対応関係を利用して，$\{\boldsymbol{b},\Omega\}$ の事後サンプルを $\{\boldsymbol{\beta},\Sigma\}$ に変換すればよい．

3.4 Remark

SUR モデルのベイズ推定においては，(2.3) 式のように $\boldsymbol{\beta}$，Σ の条件付き事後分布は解析的に求められているため，ギブスサンプリング法により事後サンプリングをおこなうことも可能である．しかし，$\boldsymbol{\beta}$ の次元が非常に大きい場合には，計算機上での逆行列の計算ができないこともある．それとは対照的に，DMC ではパラメータ $\boldsymbol{b}$ を分割して逐次的に

事後サンプリングを行っているので，$\boldsymbol{\beta}$ の次元が非常に大きい場合にも計算負荷が小さいという利点がある．

任意の地点 $\boldsymbol{x}$ において予測をおこないたい場合には，$\boldsymbol{y}$ の予測分布を利用すればよい．SUR モデルにおいて $\boldsymbol{y}$ の予測分布は，事後サンプル $\{\beta^{(j)}, \Omega^{(j)}; j=1,\ldots,N\}$ を利用すると以下のように近似できる．

$$\int f\left(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\beta},\Sigma\right)\pi_1\left(\boldsymbol{\beta},\Sigma|\boldsymbol{Y},\boldsymbol{X}\right)d\boldsymbol{\beta}d\Sigma \approx \frac{1}{N}\sum_{k=1}^{N} f\left(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\beta}^{(k)},\Sigma^{(k)}\right).$$

また注意すべき点として，生成されたサンプルは SUR モデルを変換する際に与えた m 本の方程式の順番に依存することを注意しておく．仮に m 本の方程式の適切な順番が事前にわかっている場合には問題は生じないが，m 本の方程式を同等に扱いたい場合には注意が必要である．そのような場合には，Zellner and Ando (2008a) の順番に依存しない方法を利用することも考えられる．

適切なモデルの選択をおこないたい場合には，偏差情報量規準 (Spiegelhalter *et al.* (2002))，ベイズ予測型情報量規準 (Ando (2007))，予測型情報量規準 (Ando (2009b)) などを利用すればよい．これらの基準は

$$\mathrm{IC} = -2\int \log f\left(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\beta},\Sigma\right)\pi(\boldsymbol{\beta},\Sigma|\boldsymbol{Y},\boldsymbol{X})d\boldsymbol{\beta}d\Sigma + 2\times \mathrm{Penalty}$$

の形式をしている．例えば，偏差情報量規準の場合，Penalty 項は

$$\mathrm{Penalty} = \log f\left(\boldsymbol{Y}|\boldsymbol{X},\hat{\boldsymbol{\beta}},\hat{\Sigma}\right) - \int \log f\left(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\beta},\Sigma\right)\pi(\boldsymbol{\beta},\Sigma|\boldsymbol{Y},\boldsymbol{X})d\boldsymbol{\beta}d\Sigma$$

となる．ここで，$\hat{\boldsymbol{\beta}}$，$\hat{\Sigma}$ は事後平均とする．

また，ベイズファクター (Bayes factor (Kass and Raftery (1995))) に基づいたモデル選択も考えられる．2 つのモデル M_k，M_j を比較するとき，ベイズファクター (Bayes factor (Kass and Raftery (1995))) は，モデル M_k と M_j の周辺尤度の比

$$\text{Bayes factor}(M_k, M_j) \equiv \frac{P(\boldsymbol{Y}|\boldsymbol{X},M_k)}{P(\boldsymbol{Y}|\boldsymbol{X},M_j)}$$

もしくは

$$\text{Bayes factor}(M_k, M_j) = \frac{\text{Posterior odds}(M_k, M_j)}{\text{Prior odds}(M_k, M_j)}$$

で定義される．Proper な事前分布を利用する場合，ベイズファクターは適切に定義される．

一般にモデル M_k の事前確率 $P(M_k)$ は固定された値であり，例えば，2 つのモデル M_k，M_j の事前確率 $P(M_k)$ が等しいとき，Prior odds$(M_k, M_j)=1$ となる．しかし，事前確率にも確率分布をもたせることが可能である．いま，モデル M_k，M_j の事前確

率 $P(M_k)$ が等しいと考えている場合，モデルの事前確率に関する不確実性を，例えば Prior odds(M_k, M_j) の期待値が1となるような分布を利用して反映させればよい．モデル M_k と M_j の周辺尤度の比は固定されており，Prior odds(M_k, M_j) は確率的に分布するので，Posterior odds(M_k, M_j) の分布が得られることとなる．このように，様々な不確実性を考慮しつつモデルを構築できる利点がベイズアプローチにはある．

4. 数値実験

4.1 DMC 法の適用例

いま，$m=2$，$p_j=2$，$j=1,2$ として，(2.1) 式からデータを発生させる．

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_1 & O \\ O & \boldsymbol{X}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

ここで，$\boldsymbol{X}_j$ は $n \times 2$-次元行列，$\boldsymbol{\beta}_j$ は 2-次元ベクトル，共分散行列 Σ は

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 0.1 & -0.05 \\ -0.05 & 0.2 \end{pmatrix}.$$

とする．真の係数パラメータを $\boldsymbol{\beta}_1 = (3, -2)'$，$\boldsymbol{\beta}_2 = (-1, 1)'$ とし，行列 $\boldsymbol{X}_j$ $(j=1,2)$ のそれぞれ成分は $[-1, 1]$ の一様乱数とした．ここでは $n=50$ 個のデータを発生させ，提案する DMC 法によるベイズ推定をおこなう．

サンプリング回数を $N=1000$，行列 D_1，D_2 を対角成分が 10^{-10} の対角行列として，事後サンプリングをおこなった．行列 D_1，D_2 の対角成分が大きい場合，一様分布，或いは無情報事前分布のような事前分布を利用している場合とあまり変わらない．つまり，係数パラメータに関する情報がないような状況である．一般に，無情報事前分布のように Improper な事前分布を利用すると，ベイズファクターは適切に定義されない．しかし，ここで設定した事前分布は，限りなく無情報事前分布に近いものの，Proper な事前分布であり，ベイズファクターが適切に定義されるという利点がある．

表2は，パラメータの事後平均，事後モード，事後標準偏差，事後 95% 信頼区間である．表2からわかるように，パラメータの事後平均，及び事後モードは真のパラメータ値に近く，また 事後 95% 信頼区間は真のパラメータ値を含んでいる．提案する DMC 法が真のモデルを適切に推定していることが確認できる．

任意の地点 $\boldsymbol{x}$ における $\boldsymbol{y}$ の予測分布は，事後サンプルを利用して近似できることを3.4節で述べた．ここでは，観測地点を $\boldsymbol{x}_1 = (0.1, -0.4)'$，$\boldsymbol{x}_2 = (0.2, -0.3)'$ とし，予測分布を構築する．解析的な予測分布は求められないので，$\boldsymbol{y} = (y_1, y_2)'$ の確率密度関数と比較することとする．図1は，予測分布の推定結果である．図1からわかるように，DMC 法により構成された予測分布が $\boldsymbol{y}$ の確率密度関数を適切に推定していることがわかる．ここで

表 2 パラメータの事後平均，事後モード，事後標準偏差，事後 95% 信頼区間.

	平均	モード	標準偏差	95% 信頼区間	
β_{11}	3.106	3.085	0.083	2.909	3.239
β_{12}	−2.094	−2.090	0.082	−2.255	−1.935
β_{21}	−0.978	−0.938	0.101	−1.144	−0.740
β_{22}	1.098	1.095	0.117	0.868	1.319
ω_1^2	0.104	0.112	0.023	0.075	0.168
ω_{12}	−0.061	−0.071	0.025	−0.127	−0.027
ω_2^2	0.178	0.183	0.039	0.121	0.274

はサンプルサイズを $n = 50$ としたが，サンプルサイズが大きくなるにつれて推定の精度が向上したことも付記しておく.

4.2 DMC 法と MCMC 法の比較

ここでは，MCMC 法との比較をおこなう．$p_j = 25$, $j = 1, 2$ として前節で利用した SUR モデルからデータを発生させることとする．DMC 法は，前節と同様の設定で事後サンプリングをおこない，MCMC 法にはギブスサンプリング法を利用した．MCMC 法においては，$\boldsymbol{\beta}$ の事前分布として正規分布 $\pi(\boldsymbol{\beta}) = N(\boldsymbol{\beta}_0, A^{-1})$, Σ の事前分布として $\pi(\Sigma) = |\Sigma|^{-(m+1)/2}$ を利用した．ここで，$\boldsymbol{\beta}_0 = \mathbf{0}$，$A^{-1}$ は対角成分が 10^{10} の対角行列である．DMC 法で利用した限りなく無情報事前分布に近い Proper な事前分布として，このように設定した．初期値は最尤推定量である．事後サンプリングは 11,000 回行い，最初の 1,000 回を burn-in period とみなして残りの 10,000 回を事後分布からのサンプルとして利用する．定常分布に収束しているかの判定は，Geweke (1992) の検定を有意水準 5% でおこないその収束を確認する.

ここでは，100 組の観測データを発生させて，上記の事後サンプリングをそれぞれのデータセットについておこなった．無論，DMC 法は事後分布から直接サンプリングしているので収束判定の問題等は生じない．一方，MCMC 法については 100 回の繰り返しのうち，有意水準 5% の Geweke's の検定は，88 回収束していないという判定を下した．MCMC 法が収束していないと判定された場合，事後サンプリングの回数が非常に多くなってしまうのは自明であろう.

図 2 は MCMC 法により発生させた事後サンプルの自己相関関数である．図 2 からわかるように自己相関がみられる．モンテカルロ積分の定義に戻ると，独立なサンプルの平均によりある量の期待値を計算していることを注意しておく．たとえば，3.4 節の $\boldsymbol{y}$ の予測分布は，独立な事後サンプルに基づき近似されている．すなわち，事後サンプルに自己相関がある場合には，自己相関を考慮した期待値の計算が必要となる．一方，DMC により

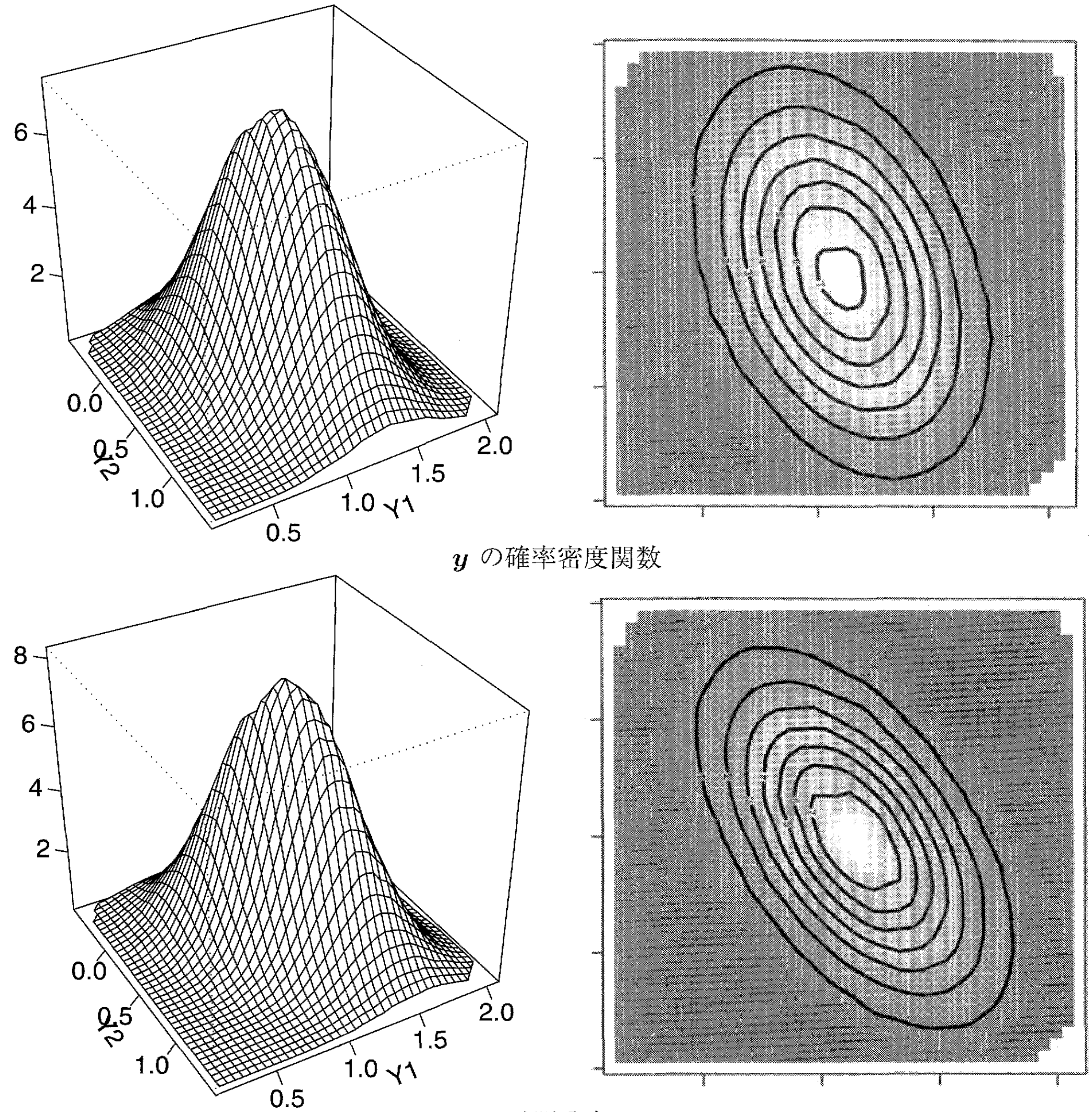
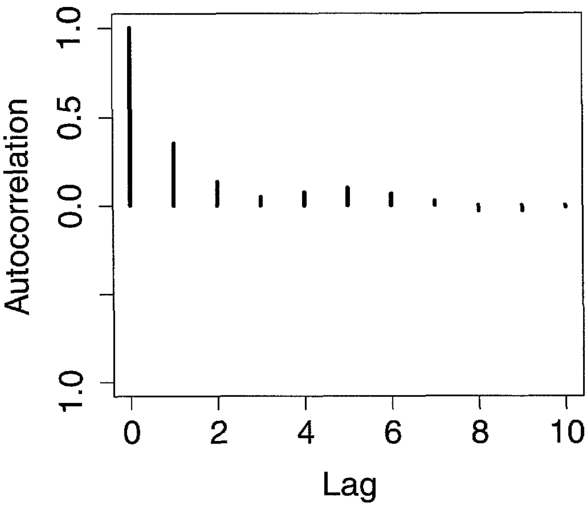


図 1　観測地点を $\boldsymbol{x}_1 = (0.1, -0.4)'$，$\boldsymbol{x}_2 = (0.2, -0.3)'$ における $\boldsymbol{y}$ の予測分布，及び $\boldsymbol{y}$ の確率密度関数．

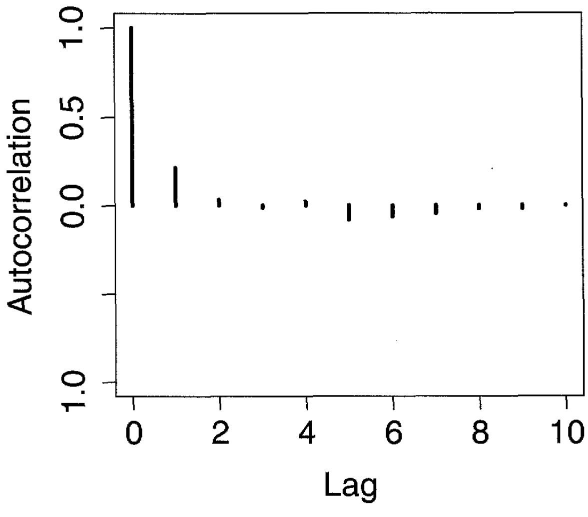
発生させた事後サンプルは完全に独立であるため，自己相関関数を考慮する必要は全くないという利点がある．事後サンプルの自己相関に対処する場合，一定間隔をおいて MCMC サンプリングを実行する方法がよく利用されている．例えば，5 回の MCMC サンプリング中，1 回のみ事後サンプルとして採用し，残りの 4 回については MCMC サンプリングを実行しているものの事後サンプルとして利用しないなどである．計算機環境が発展しているとはいえ，このような操作は計算負荷をかけることとなろう．

5.　おわりに

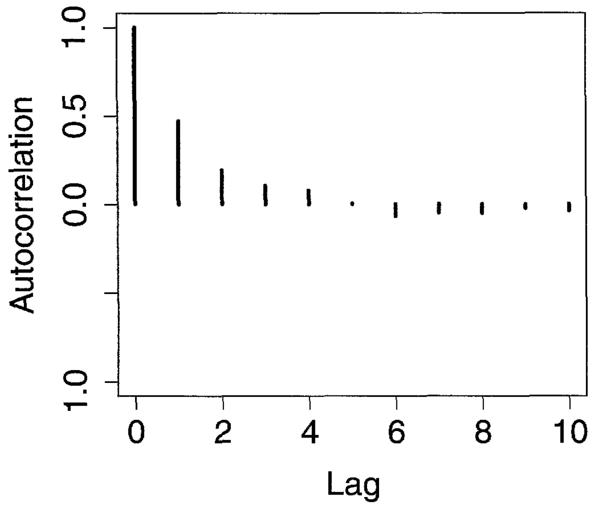
統計モデルのベイズ推定においては，マルコフ連鎖モンテカルロ (Markov Chain Monte Carlo) 法が一般に利用されている．しかし，収束判定，サンプリングの効率性等の問題が



σ_1^2



σ_{12}



σ_2^2

図 2　マルコフ連鎖モンテカルロ法により発生させた σ_1^2，σ_{12}，σ_2^2 に関する事後サンプルの自己相関関数.

常に MCMC 法には付随している．本稿では，SUR モデルのベイズ推定を取り上げ，ダイレクトモンテカルロ (direct Monte Carlo (DMC)) 法を提案した．さらに，DMC 法と MCMC 法の特徴を数値実験に基づき比較し，アルゴリズムの効率性などに関する DMC 法の利点を検証した．

数値実験においては，ギブスサンプリング法という，MCMC 法の中で非常にシンプルな方法を利用した．MCMC 法には，メトロポリス・ヘイスティングス法，もしくはメトロポリス・ヘイスティングス法とギブスサンプリング法を併用する方法等もある．仮に，ギブスサンプリング法の代わりに，これらの MCMC 法を数値実験で利用した場合，DMC 法の有用性がさらに鮮明となる．メトロポリス・ヘイスティングス法はギブスサンプリング法よりも適用範囲が広く，パラメータの完全条件付事後密度からの乱数生成が難しい時はギ

ブスサンプリング法が利用できず，その場合，メトロポリス・ヘイスティングス法が頻繁に利用される．しかし，その反面，メトロポリス・ヘイスティングス法は生成した乱数を確率的に採択するため，計算効率性の観点から，生成した乱数を100%の確率で採択するギブスサンプリング法に比べて非効率的である．査読者から有益な指摘があったように，メトロポリス・ヘイスティングス法を採用した場合，発生させたサンプルが100%の確率で採択されないため，DMC 法が好ましいことは自明である．さらに，メトロポリス・ヘイスティングス法においては提案分布を利用する必要がある．「どのような提案分布を用意すればいいのか？」という問題は，特に MCMC 法に不慣れなユーザーには面倒な作業でもある．無論，ギブスサンプリング法と同様，発生させたサンプルの自己相関の問題，burn-in period の設定，収束判定等の問題に対処する必要がある．

本稿に関連する研究としては以下のようなものが挙げられる．Zellner and Ando (2008b) は，Simultaneous equation model (SEM) を取り上げ，ダイレクトモンテカルロ法によるベイズ推定法を提案している．また，本稿では正規性に基づいた議論をおこなっているが，Zellner and Ando (2010) は，SUR モデルの誤差項に Student-t 分布を仮定して，ダイレクトモンテカルロ法の構築を試みている．さらに，Ando and Zellner (2009) では，SUR モデル，及び SEM の階層ベイズモデリング問題を考え，ダイレクトモンテカルロ法を提案している．Ando (2009a) では，因子分析モデルのベイズ分析を考え，ダイレクトモンテカルロ法を提案している．また，2 節で指摘したように SUR モデルの非線形化はスプライン関数などを利用すればよいので，本稿で構築したダイレクトモンテカルロ法，及び上記の研究は非常に幅広い SUR モデル，SEM モデルのベイズ推定に利用可能である．

ダイレクトモンテカルロ法はある条件下においてのみ構築できることも事実である．そのため，ダイレクトモンテカルロ法が利用できない場合には，自ずと MCMC 法に頼らざるを得ない．本稿では，MCMC 法の問題点を取り上げているが，無論，MCMC 法はベイズ的統計モデリングにおいて非常に有用な道具であり，その有用性を否定しているのではない．尤度関数，パラメータの事前分布などを設定すれば，事後分布からのサンプリングを可能としてくれる MCMC 法は，ベイズ的統計モデリングの発展に重要な役割をこれからも果たしていくであろう．

謝辞

西山慶彦教授 (京都大学)，及び査読者からは貴重なコメントを頂戴し，内容は大幅に改善されました．Herman K. van Dijk 教授 (Erasmus University Rotterdam), John F. Geweke 教授 (University of Iowa), Hedibert F. Lopes 教授 (University of Chicago) からのコメントも，論文執筆の上で非常に参考となりました．ここに記して感謝いたします．

参 考 文 献

Ando, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models, *Biometrika*, **94**, 443–458.

Ando, T. (2009a). Bayesian factor analysis with fat-tailed factors and its exact marginal likelihood, *J. Multivar. Anal.*, **100**, 1717–1726.

Ando, T. (2009b). Predictive Bayesian model selection, Working paper, Graduate School of Business Administration, Keio University.

Ando, T. and Zellner, A. (2009). Hierarchical Bayesian analysis of the seemingly unrelated regression and simultaneous equation models, *Bayesian Anal.* (to appear).

Box, G. E. P. and Tiao, G. C. (1973). Bayesian inference in statistical analysis, Reading, MA, Addison-Wesley.

Brooks, S. P. and Gelman, A. (1997). General methods for monitoring convergence of iterative simulations, *J. Comput. Graph. Stat.*, **7**, 434–455.

Carlin, B. and Louis, T. (1996). Bayes and empirical Bayes methods for data analysis, New York, Chapman and Hall.

Carroll, R. J., Doug M., Larry, F. and Victor, K. (2006). Seemingly unrelated measurement error models, with application to nutritional epidemiology, *Biometrics*, **62**, 75–84.

Chib, S. and Greenberg, E. (1995). Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models, *J. Econom.*, **68**, 339–360.

Frasera, D. A. S., Rekkasb, M. and Wong, A. (2005). Highly accurate likelihood analysis for the seemingly unrelated regression problem, *J. Econom.*, **127**, 17–33.

Gallant, R. (1975). Seemingly unrelated nonlinear regressions, *J. Econom.*, **3**, 35–50.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences, *Statist. Sci.*, **7**, 457–511.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments, in *Bayesian Statistics 4* (eds. J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. M. Smith), Oxford, Clarendon Press, pp. 169–193.

Geweke, J. (2005). *Contemporary bayesian econometrics and statistics*, New York, Wiley.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*, New York, Chapman & Hall.

Greene, W. H. (2002). *Econometric analysis* (5th ed.), New Jersey, Prentice-Hall.

Hastie, T. and Tibshirani, R. (1990). *Generalized additive models*, New York, Chapman & Hall/CRC.

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The elements of statistical learning*, New York, Springer.

Heidelberger, P. and Welch, P. D. (1983). Simulation run length control in the presence of an initial transient, *Oper. Res.*, **31**, 1109–1144.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems, *Proc. Roy. Soc. London, Ser. A*, **196**, 453–461.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.), Oxford, Oxford University Press.

Kass, R. E. and Raftery, A. 1995. Bayes factors, *J. Am. Statis.. Assoc.*, **90**, 773–795.

Kurata, H. (1999). On the efficiencies of several generalized least squares estimators in a seemingly unrelated regression model and a heteroscedastic model, *J. Multivar. Anal.*, **70**, 86–94.

Lancaster, T. (2004). *Introduction to modern Bayesian econometrics*, New Jersey, Cambridge University Press.

Liu, A. (2002). Efficient estimation of two seemingly unrelated regression equations, *J. Multivar. Anal.*, **82**, 445–456.

Mandy, D. M. and Martins-Filho, C. (1993). Seemingly unrelated regressions under additive heteroscedasticity: theory and share equation applications, *J. Econom.*, **58**, 315–346.

Neudecker, H. and Windmeijer, F. A. G. (1991). R^2 in seemingly unrelated regression equations, *Statistica Neerlandica*, **45**, 405–411.

Ng, V. M. (2002). Robust Bayesian inference for seemingly unrelated regressions with elliptical errors, *J. Multivar. Anal.*, **83**, 409–414.

Percy, D. (1992). Predictions for seemingly unrelated regressions, *J. Roy. Statist. Soc. B*, **54**, 243–252.

Percy, D. F. (1996). Zellner's influence on multivariate linear models, in *Bayesian analysis in statistics and econometrics: essays in honor of Arnold Zellner* (eds. D. A. Berry, K. M Chaloner and J. K. Geweke), New York, John Wiley and Sons, pp. 203–214.

Raftery, A. E. and Lewis, S. M. (1992). One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo, *Statist. Sci.*, **7**, 493–497.

Rocke, D. M. (1989). Bootstrap Bartlett adjustment in seemingly unrelated regression, *J. Am. Stat. Assoc.*, **84**, 598–601.

Rossi, P. E, Allenby, G. and McCulloch, R. (2005). *Bayesian statistics and marketing*, New Jersey, John Wiley and Sons.

Schruben L. W. (1982). Detecting initialization bias in simulation experiments, *Oper. Res.*, **30**, 569–590.

Smith, M. and Kohn, R. (2000). Nonparametric seemingly unrelated regression, *J. Econom.*, **98**, 257–282.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion), *J. Roy. Statist. Soc., Ser. B*, **64**, 583–639.

Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion), *Ann. Stat.*, **22**, 1701–1762.

van der Merwe, A. and Viljoen, C. (1988). Bayesian analysis of the seemingly unrelated regression model, Manuscript, University of the Free State, Department of Mathematical Statistics.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias, *J. Am. Stat. Assoc.*, **57**, 348–368.

Zellner, A. (1963). Estimators for seemingly unrelated regression equations: some exact finite sample results, *J. Am. Stat. Assoc.*, **58**, 977–992.

Zellner, A. 1971. *An introduction to Bayesian inference in econometrics*, Wiley, New York.

Zellner, A. and Ando, T. (2008a). A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model, Working paper, Graduate School of Business, University of Chicago.

Zellner, A. and Ando, T. (2008b). A direct Monte Carlo approach for Bayesian analysis of the simultaneous equation model, Working paper, Graduate School of Business, University of Chicago.

Zellner, A. and Ando, T. (2010). Bayesian and non-Bayesian analysis of the seemingly unrelated regression model with student-t errors and its application for forecasting, *Int. J. Forecasting* (in press).

Zellner, A. and Chen, B. (2002). Bayesian modeling of economies and data requirements, *Macroecon. Dyn.*, **5**, 673–700.

Zellner, A. and Min, C. K. (1995). Gibbs sampler convergence criteria, *J. Am. Stat. Assoc.*, **90**, 921–927.

Zellner, A., Bauwens, L. and Van Dijk, H. K. (1988). Bayesian specification analysis and estimation of simultaneous equation models using Monte Carlo Methods, *J. Econom.*, **38**, 39–72.