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

Arnold Zellner
Tomohiro Ando
(Chicago GSB)

Theodore W. Anderson’s 90th Birthday Conference, Stanford U.
Outline

- Background, Motivation
- Bayesian Inference for the Seemingly Unrelated Regression Model Using a Direct Monte Carlo Procedure
- Numerical results
- Conclusion and future work
Outline

- Background, Motivation
- Bayesian Inference for the Seemingly Unrelated Regression Model Using a Direct Monte Carlo Procedure
- Numerical results
- Conclusion and future works
Background and motivation
SUR model: Overview

- Introduced by Zellner (1962)

- Many studies have contributed to the analysis of SUR models.

- Many textbooks and journal papers
SUR model: Overview Cont.

Generalized least squares
- Zellner (1962, 1963), GLS, Iterative GLS, finite sample
- Madansky (1964), Iterative GLS and ML

Bayesian Methods
- Zellner (1971), et al., approx. finite sample posteriors, Stein shrinkage, etc.
  - van der Merwve and Viljoen, (1998), Bayesian MOM

Bayesian Markov Chain Monte Carlo estimation
SUR model : Motivation

- MCMC methods are rather complicated and involve many decisions
  - Initial parameter value
  - Choice of an appropriate proposal density
  - The length of the burn in period
  - Check for convergence
SUR model : Motivation

- Several drawbacks of MCMC

- Can we estimate the Bayesian SUR model by using a direct Monte Carlo (DMC) procedure?
## Comparison of DMC and MCMC

<table>
<thead>
<tr>
<th>Items</th>
<th>DMC</th>
<th>MCMC (Gibbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need to fix the number of samples drawn</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>100% acceptance of draws</td>
<td>Yes</td>
<td>No (Yes)</td>
</tr>
<tr>
<td>Require initial parameters value</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Burn-in period setting</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Check for convergence</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Select convergence check criteria</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Selection of a proposal density</td>
<td>No</td>
<td>Yes (No)</td>
</tr>
<tr>
<td>Use of a proposal density</td>
<td>No</td>
<td>Yes (No)</td>
</tr>
</tbody>
</table>

User friendly
Outline

- Background, Motivation

- Bayesian Inference for the Seemingly Unrelated Regression Model Using a Direct Monte Carlo Procedure

- Numerical results

- Conclusion and future works
The standard SUR models
The standard SUR models (1/2)

- A set of m equations (E.g., Zellner (1971))

\[ y_j = X_j \beta_j + u_j, \quad j = 1, \ldots, m, \]

\[(n \times 1) \quad (n \times p_j) \quad (p_j \times 1) \quad (n \times 1)\]

\[
E[u_i u_j'] = \begin{cases} 
\omega_{ij} I \ (i \neq j) \\
\omega_i^2 I \ (i = j)
\end{cases}.
\]
The standard SUR models (1/2)

- The likelihood function

\[
L(D \mid \beta, Omega) = \frac{1}{(2\pi)^{nm/2} |Omega|^{n/2}} \exp \left[ -\frac{1}{2} \left( y - X\beta \right)' (Omega^{-1} \otimes I) \left( y - X\beta \right) \right] \\
= \frac{1}{(2\pi)^{nm/2} |Omega|^{n/2}} \exp \left[ -\frac{1}{2} \text{tr} \left\{ ROmega^{-1} \right\} \right].
\]

\[y' = (y'_1, ..., y'_m) \quad X = \text{diag}\{X_1, ..., X_m\}\]

\[R = (r_{ij}); \quad r_{ij} = (y_i - X_i\beta_i)' (y_j - X_j\beta_j)\]

\[\beta' = (\beta'_1, ..., \beta'_m)\]
Bayesian estimation

\{\beta, \Omega\}
Bayesian approach (1/3)

- Diffuse prior

\[ \pi_1(\beta, \Omega) = \pi_1(\beta)\pi_1(\Omega) \propto |\Omega|^{-\frac{m+1}{2}} \]

- Joint posterior distribution

\[ g_1(\beta, \Omega \mid D) \propto |\Omega|^{-(n+m+1)/2} \exp \left[ -\frac{1}{2} \text{tr} \left\{ R\Omega^{-1} \right\} \right] \]
Bayesian approach (2/3)

- Conditional posteriors

\[
\begin{align*}
    g_1(\beta \mid \Omega, D) &= N\left(\hat{\beta}, \hat{\Omega}_{\beta}\right) \\
    g_1(\Omega \mid \beta, D) &= IW(R, n)
\end{align*}
\]

with

\[
\begin{align*}
    \hat{\beta} &= \left\{X'(\Omega^{-1} \otimes I)X\right\}^{-1} X'(\Omega^{-1} \otimes I)y \\
    \hat{\Omega}_{\beta} &= (X'(\Omega^{-1} \otimes I)X)^{-1}
\end{align*}
\]
Bayesian approach (3/3)

- Normal/inverse Wishart priors

\[ \pi_2(\beta, \Omega) = \pi_2(\beta)\pi_2(\Omega) \]
\[ \pi_2(\beta) = N(\beta_0, A_\beta^{-1}), \quad \pi_2(\Omega) = IW(\Lambda_0, \nu_0) \]

- Conditional posteriors

\[ g_2(\beta \mid \Omega, D) = N(\overline{\beta}, \overline{\Omega}_\beta) \]
\[ g_2(\Omega \mid \beta, D) = IW(\Lambda_0 + R, n + \nu_0) \]
The DMC procedure for the transformed SUR model
Transformation (1/2)

- The original SUR model

\[ y_j = X_j \beta_j + u_j \quad j = 1, ..., m \]

- Transformation

\[
\begin{align*}
    u_1 &= e_1 \\
    u_2 &= \rho_{21} u_1 + e_2 \\
    &\vdots \\
    u_m &= \sum_{j=1}^{m-1} \rho_{mj} u_j + e_m
\end{align*}
\]

\[ E\left[ e_i e_j ' \right] = \begin{cases} 
    O & (i \neq j) \\
    \sigma_i^2 I & (i = j) 
\end{cases} \]
Transformation(2/2)

- The transformed model

\[
\begin{align*}
\mathbf{y}_1 &= \mathbf{X}_1 \mathbf{\beta}_1 + \mathbf{e}_1 = \mathbf{Z}_1 \mathbf{b}_1 + \mathbf{e}_1 \\
\mathbf{y}_j &= \mathbf{X}_j \mathbf{\beta}_j + \sum_{l=1}^{j-1} \rho_{jl} (\mathbf{y}_l - \mathbf{X}_l \mathbf{\beta}_l) + \mathbf{e}_j = \mathbf{Z}_j \mathbf{b}_j + \mathbf{e}_j, \quad j = 2, \ldots, m,
\end{align*}
\]

- The likelihood function

\[
L(D \mid \mathbf{b}, \Sigma) = \prod_{j=1}^{m} \frac{1}{(2\pi \sigma_j^2)^{n/2}} \exp \left[ - \frac{(\mathbf{y}_j - \mathbf{Z}_j \mathbf{b}_j)'(\mathbf{y}_j - \mathbf{Z}_j \mathbf{b}_j)}{2\sigma_j^2} \right].
\]
Bayesian estimation

\{ b, \Sigma \}
Bayesian analysis of the transformed model (1/3)

- Diffuse prior

\[ \pi_3(b, \Sigma) = \pi_3(b)\pi_3(\Sigma) \propto \prod_{j=1}^{m} (\sigma_j)^{-1}. \]

- Joint posteriors

\[ g(b, \Sigma | D) \propto \prod_{j=1}^{m} (\sigma_j)^{-(n+1)} \exp \left[ -\frac{(y_j - Z_j b_j)'(y_j - Z_j b_j)}{2\sigma_j^2} \right]. \]
Bayesian analysis of the transformed model (2/3)

- The conditional posteriors

**Normal**

\[
g(b_j | b_{j-1}, \ldots, b_1, \sigma_j^2, D) = N \left( \hat{b}_j, \sigma_j^2 \left( Z_j' Z_j \right)^{-1} \right),
\]

**Inverse gamma**

\[
g(\sigma_j^2 | b_{j-1}, \ldots, b_1, D) = IG \left( \frac{1}{2} (y_j - Z_j \hat{b}_j)' (y_j - Z_j \hat{b}_j), \frac{1}{2} (n - p_j - j + 1) \right),
\]

with \[
\hat{b}_j = \left( Z_j' Z_j \right)^{-1} Z_j' y_j.
\]
Bayesian analysis of the transformed model (3/3)

Graphical presentation of the conditional posteriors

\[
\begin{align*}
\sigma_1^2 &\sim g(\sigma_1^2 | D) \\
\sigma_2^2 &\sim g(\sigma_2^2 | b_1, D) \\
\sigma_m^2 &\sim g(\sigma_m^2 | b_{m-1}, \ldots, b_1, D)
\end{align*}
\]

\[
\begin{align*}
b_1 &\sim g(b_1 | \sigma_1^2, D) \\
b_2 &\sim g(b_2 | b_1, \sigma_2^2, D) \\
b_m &\sim g(b_m | b_{m-1}, \ldots, b_1, \sigma_j^2, D)
\end{align*}
\]
A direct Monte Carlo (DMC) sampling procedure:

1. Fix the order of a set of m equations. Set the number of samples N to be generated. Set $j = 1$

2. Generate $\sigma_{1}^{2(k)}$, $k = 1,\ldots, N$, and insert the drawn values in $\pi(b_{1} | \sigma_{1}^{2(k)}, D)$. Then make a draw $b_{1}^{(k)}$, from $\pi(b_{1} | \sigma_{1}^{2(k)}, D)$, for $k = 1,\ldots, N$.

3. Increase the index $j + 1 \leftarrow j$. Draw $\sigma_{j}^{(k)}$ from the conditional inverse gamma density $g(\sigma_{j}^{2} | b_{j-1}^{(k)}, \ldots, b_{1}^{(k)}, D)$, and then generate $b_{j}^{(k)}$ from $g(b_{j} | b_{j-1}^{(k)}, \ldots, b_{1}^{(k)}, \sigma_{j}^{(k)}, D)$ for $k = 1,\ldots, N$.

4. Repeat Step 3 sequentially until $j = m$. 

Some remarks (1/3)

- Only set the number of sampling

- Improper prior... Bayes factor?
  - BPIC (Ando, 2007)

- Informative prior?

- Inference on the original SUR model?

\[ \{\beta, \Omega\} \xleftarrow{\text{?}} \{b, \Sigma\} \]
Some remarks (2/3)

- **One to one relationship**

\[ \omega_i^2 = \sigma_i^2, \]
\[ \omega_j^2 = \sum_{k=1}^{j-1} \rho_{jk}^2 \omega_k^2 + \sum_{k,l=1, k < l}^{j-1} \rho_{jk} \rho_{jl} \omega_k \omega_l + \sigma_j^2, \quad (j \neq 1), \]
\[ \omega_{ji} = \sum_{k=1, k \neq i}^{j-1} \rho_{jk} \omega_{ki} + \rho_{ji} \omega_i^2, \quad (j \neq 1). \]

- **Transform a set of posterior samples**

\[ \{\sigma_1^{(k)}, \sigma_2^{(k)}, \ldots; k = 1, \ldots, N\} \quad \rightarrow \quad \{\hat{\Omega}_\beta^{(k)}; k = 1, \ldots, N\} \]

- **Generate \( \beta \) from the conditional posterior**

\[ \beta \sim N(\hat{\beta}^{(k)}, \hat{\Omega}_\beta^{(k)}) \]
Some remarks (3/3)

- The original model

\[ y_j = X_j \beta_j + u_j \]

\[ E[u_i u_j'] = \begin{cases} \omega_{ij} I (i \neq j) \\ \omega_i^2 I (i = j) \end{cases} \]

\[ \pi_1(\beta, \Omega) \propto |\Omega|^{-\frac{m+1}{2}} \]

- Transformed model

\[ y_j = Z_j b_j + e_j \]

\[ E[e_i e_j'] = \begin{cases} O (i \neq j) \\ \sigma_i^2 I (i = j) \end{cases} \]

\[ \pi_3(b, \Sigma) = \pi_3(b) \pi_3(\Sigma) \propto \prod_{j=1}^{m} (\sigma_j)^{-1} \]

\[ \pi_1(b, \Sigma) \propto |\Omega(b, \Sigma)|^{(m+1)/2} \times |J_m| = \prod_{j=1}^{m} (\sigma_j^2)^{\frac{m+1}{2}-j} \]
Inference on the original model

- The conditional posteriors

\[
g(b_j | b_{j-1}, ..., b_1, \sigma_j^2, D) = N\left(\hat{b}_j, \sigma_j^2 \left(Z_j' Z_j\right)^{-1}\right),
\]

\[
g(\sigma_j^2 | b_{j-1}, ..., b_1, D) = IG\left(\frac{1}{2}(y_j - Z_j \hat{b}_j)'(y_j - Z_j \hat{b}_j), \frac{1}{2}(n - p_j - j + 1)\right),
\]

\[
g(b_j | b_{j-1}, ..., b_1, \sigma_j^2, D) = N\left(\hat{b}_j, \sigma_j^2 \left(Z_j' Z_j\right)^{-1}\right),
\]

\[
g(\sigma_j^2 | b_{j-1}, ..., b_1, D) = IG\left(\frac{1}{2}(y_j - Z_j \hat{b}_j)'(y_j - Z_j \hat{b}_j), \frac{1}{2}(n - m - p_j - j + 1)\right),
\]
A direct Monte Carlo (DMC) sampling procedure:

1: Fix the order of a set of m equations. Set the number of samples N to be generated. Set \( j = 1 \)

2 Generate \( \sigma_1^{2(k)} \), \( k = 1, \ldots, N \), and insert the drawn values in \( \pi(b_1 \mid \sigma_1^2, D) \). Then make a draw \( b_1^{(k)} \), from \( \pi(b_1 \mid \sigma_1^{2(k)}, D) \), for \( k = 1, \ldots, N \).

3 Increase the index \( j + 1 \leftarrow j \). Draw \( \sigma_j^{(k)} \) from the conditional inverse gamma density \( g(\sigma_j^2 \mid b_{j-1}^{(k)}, \ldots, b_1^{(k)}, D) \), and then generate \( b_j^{(k)} \) from \( g(b_j \mid b_{j-1}^{(k)}, \ldots, b_1^{(k)}, \sigma_j^{(k)}, D) \) for \( k = 1, \ldots, N \).

4 Repeat Step 3 sequentially until completion.
Outline

- Overview of SUR model
- Bayesian Inference for the SUR Model Using a Direct Monte Carlo Procedure
- Numerical results
- Conclusion and future works
Numerical results
Simulation study
Simulation study

- MCMC approach

- Two DMC methods
  - DMC algorithm 1
    \[
    \pi_3(b, \Sigma) = \pi_3(b) \pi_3(\Sigma) \propto \prod_{j=1}^{m} (\sigma_j)^{-1}.
    \]
  - DMC algorithm 2
    \[
    \pi_1(b, \Sigma) \propto |\Omega(b, \Sigma)|^{(m+1)/2} |J|_m = \prod_{j=1}^{m} (\sigma_j^2)^{\frac{m-1}{2} - j}
    \]
Simulation study

- The true model

\[
\begin{pmatrix}
  y_1 \\
  y_2
\end{pmatrix} = \begin{pmatrix} X_1 & O \\ O & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}
\]

\[\beta_1 = (3, -2)' \quad \beta_2 = (2, 1)'
\]

\[\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} = \begin{pmatrix} 0.1 & -0.05 \\ -0.05 & 0.2 \end{pmatrix}.
\]

- Dimension of X=2
- n=100
- # of DMC sampling 10,000
- # of MCMC sampling 11,000 (1,000 burn in)
Predictive density

(a) True sampling density.

(b) Estimated predictive density based on the DMC 1.
## Summary of the parameter estimates for DMC1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>95% PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>3.0581</td>
<td>0.0597</td>
<td>2.9400</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-1.9675</td>
<td>0.0543</td>
<td>-2.0747</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>1.9887</td>
<td>0.0822</td>
<td>1.8279</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>0.9418</td>
<td>0.0816</td>
<td>0.7784</td>
</tr>
<tr>
<td>$\omega_1^2$</td>
<td>0.1109</td>
<td>0.0162</td>
<td>0.0834</td>
</tr>
<tr>
<td>$\omega_{12}$</td>
<td>-0.0345</td>
<td>0.0166</td>
<td>-0.0689</td>
</tr>
<tr>
<td>$\omega_2^2$</td>
<td>0.2196</td>
<td>0.0323</td>
<td>0.1641</td>
</tr>
</tbody>
</table>

**Note:** Result from 1 dataset
Summary of the computational time of each methods

<table>
<thead>
<tr>
<th></th>
<th>DMC1</th>
<th></th>
<th>DMC2</th>
<th></th>
<th>MCMC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
<td>SDs</td>
</tr>
<tr>
<td>n = 50</td>
<td>7.5881</td>
<td>0.0704</td>
<td>7.5736</td>
<td>0.0696</td>
<td>14.2223</td>
<td>0.2620</td>
</tr>
<tr>
<td>n = 100</td>
<td>12.3236</td>
<td>0.1266</td>
<td>12.3309</td>
<td>0.1214</td>
<td>15.1566</td>
<td>0.1609</td>
</tr>
</tbody>
</table>

The variation of the computational time for each method.

For MCMC, the computational times are measured from the initialization of parameters to the end of posterior sampling.

The computational times for our method are measured from the first posterior sampling to the end of posterior sampling.
Real data analysis 1
Incense product sales forecast

- In 2006, the size of the market for incense products in Japan was estimated to be about 30 billion yen.

- In Japan, traditional incense is used differently from lifestyle incense.

- Data consist of the daily sales figures for incense products from April, 2006 to June, 2006.

- The data were collected from two department stores, both located in Tokyo.
Incense product sales forecasting

- Model

\[
y_{jt} = \beta_{j0} + \beta_{j1}x_{jt} + \beta_{j2}x_{jt} + \beta_{j3}x_{jt} + \beta_{j4}x_{jt} + u_{jt},
\]

\[(t = 1, \ldots, 100, j = 1, 2)\]

- Holiday effect

\[
x_{j1t} = \begin{cases} 
1 & \text{(Sunday, Saturday, National holiday)} \\
0 & \text{(Otherwise)} 
\end{cases}
\]

- Weather effect

\[
x_{j3t} = \begin{cases} 
1 & \text{(Fine)} \\
0 & \text{(Cloudy)} \\
-1 & \text{(Rain)} 
\end{cases}
\]

- Promotion

\[
x_{j2t} = \begin{cases} 
1 & \text{(Execution)} \\
0 & \text{(Nonexecution)} 
\end{cases}
\]

- Event

\[
x_{j4t} = \begin{cases} 
1 & \text{(Holding)} \\
0 & \text{(Nonholding)} 
\end{cases}
\]
Estimation results

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>95%PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>54.5052</td>
<td>5.3536</td>
<td>43.9621</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-4.1032</td>
<td>5.3964</td>
<td>-14.6753</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>1.5103</td>
<td>5.3162</td>
<td>-8.8837</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-1.8303</td>
<td>3.4758</td>
<td>-8.698</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>17.6856</td>
<td>5.8354</td>
<td>6.2768</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>25.559</td>
<td>3.399</td>
<td>18.8576</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>9.623</td>
<td>4.2288</td>
<td>1.3021</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>11.9626</td>
<td>4.0891</td>
<td>3.8663</td>
</tr>
<tr>
<td>$\beta_8$</td>
<td>2.331</td>
<td>2.5327</td>
<td>-2.6756</td>
</tr>
<tr>
<td>$\beta_9$</td>
<td>8.4699</td>
<td>4.2123</td>
<td>0.0517</td>
</tr>
<tr>
<td>$\omega_1^2$</td>
<td>592.783</td>
<td>89.2884</td>
<td>443.5707</td>
</tr>
<tr>
<td>$\omega_2^2$</td>
<td>109.8778</td>
<td>48.2728</td>
<td>20.7837</td>
</tr>
<tr>
<td>$\omega_3^2$</td>
<td>331.8955</td>
<td>49.6295</td>
<td>248.3822</td>
</tr>
</tbody>
</table>

DMC1 algorithm is used.
Real data analysis 2
Forecasting economic growth

- We forecast the growth rates of real sales of several sectors of the Japanese economy
- Agriculture, Automobile and Service
- 1977~2004, Quarterly data.

\[
y_{j,t+1} = \beta_{j1}SR_t + \beta_{j2}M_t + \beta_{j3}GDP_t + \beta_{j4}GDP_{t-1} + \beta_{j5}GDP_{t-2} + e_{jt},
\]

- growth rate of the real monetary base (M2)
- growth rate of TOPIX
- the logarithm of quarterly real GDP
Forecasting economic growth

- Forecasting:
  - 1999, 2nd quarter ~2004 the 4th quarter.

- The mean of the predictive densities are used to forecast one quarter ahead growth rates.

- These sector growth rate forecasts are then transformed into sales forecasts for each sector.

\[
\hat{Y}_{j,t+1} = \hat{y}_{j,t+1} \times Y_{j,t}
\]

Forecast of one quarter ahead growth
Actual outcome at time t
Forecasting economic growth

Agriculture

Service

Automobile

Actual output
Predictive mean
95\% Posterior interval
## Comparison to AR model

The SUR model with DMC1 algorithm is used. An AIC score is used to select the AR model.

### Table

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUR with DMC</td>
<td>AR</td>
</tr>
<tr>
<td>Automobile</td>
<td>6.1469</td>
<td>7.0990</td>
</tr>
<tr>
<td>Agriculture</td>
<td>5.0991</td>
<td>5.7887</td>
</tr>
<tr>
<td>Service</td>
<td>6.6458</td>
<td>7.5656</td>
</tr>
</tbody>
</table>

### Formulas

\[
RMSE_j = \sum_{t=t_{\text{min}}}^{t_{\text{max}}} (Y_{j,t} - \hat{Y}_{j,t})^2,
\]

\[
MAE_j = \sum_{t=t_{\text{min}}}^{t_{\text{max}}} |Y_{j,t} - \hat{Y}_{j,t}|,
\]
Outline

- Overview of SUR model
- Bayesian Inference for the SUR Model Using a Direct Monte Carlo Procedure
- Numerical results
- Conclusion and future works
Conclusion
Conclusion

- A direct Monte Carlo (DMC) approach for Bayesian analysis of SUR models.

- Some advantages of DMC

- The method performed well in Monte Carlo experiments and applications using actual data.

- We can recommend our DMC approach for the analysis and use of the SUR model.
### Conclusion

<table>
<thead>
<tr>
<th>Items</th>
<th>DMC</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need to fix the number of samples drawn</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>100% acceptance of draws</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Doesn't require initial parameters value</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Doesn't need a burn-in period setting</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Doesn't need to check for convergence</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Doesn't need to select convergence check criteria</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Doesn't require selection of a proposal density</td>
<td>Yes</td>
<td>Yes / No</td>
</tr>
<tr>
<td>Doesn't require use of a proposal density</td>
<td>Yes</td>
<td>Yes / No</td>
</tr>
</tbody>
</table>

**User friendly**
Future work

- The DMC approach can be applied to more complex variants of the SUR model, for example those involving use of regression splines, wavelet bases, and so on.

- Fat-tailed error.

- Numb. of sampling

- Our method can be applied to the widely used simultaneous equation model.

- Forecasting Japanese economy by MMM model
References (1/5)

References (2/5)

References (3/5)

References (4/5)

References (5/5)

Appendix
Appendix 1:
MCMC sampling algorithm (1/3)

Sampling $\beta$

At the $k$-th iteration, we generate a candidate $\beta^{(k)}$ from the proposal density $g(\beta|\Omega_{MLE}, D)$ in (3), where $\Omega_{MLE}$ is set to be the maximum likelihood estimate. Then we accept the candidate draw with the probability $\alpha$ and reject the draw with the probability $1 - \alpha$. The probability $\alpha$ is defined as

$$\alpha = \min \left\{ 1, \frac{h_1(\beta^{(k)}, \Omega^{(k-1)})/g(\beta^{(k)}|\Omega_{MLE}, D)}{h_1(\beta^{(k-1)}, \Omega^{(k-1)})/g(\beta^{(k-1)}|\Omega_{MLE}, D)} \right\},$$

with

$$h_1(\beta, \Omega) = |\Omega|^{-(n+m+1)/2} \exp \left[ -\frac{1}{2} \text{tr} \left\{ R(\beta)\Omega^{-1} \right\} \right],$$

where the $ij$th elements of $m \times m$ matrix $R(\beta)$ is $(y_i - X_i\beta_i)'(y_j - X_j\beta_j)$. 

Appendix 1:
MCMC sampling algorithm (2/3)

Sampling $\Omega$

Generate a candidate $\Omega^{(k)}$ from the Wishart proposal density $g(\Omega|\beta_{MLE}, D)$ in (3), where $\beta_{MLE}$ is the maximum likelihood estimate. The candidate draw is accepted with the probability

$$\alpha = \min\left\{ 1, \frac{h_1(\beta^{(k)}, \Omega^{(k)})/g(\Omega^{(k)}|\beta_{MLE}, D)}{h_1(\beta^{(k)}, \Omega^{(k-1)})/g(\Omega^{(k-1)}|\beta_{MLE}, D)} \right\},$$

and reject the draw with the probability $1 - \alpha$. 
1. (Initialization). Initialize $\beta$ and $\Omega$ as the maximum likelihood estimates.

2. Sample the coefficient parameter $\beta$.

3. Sample the coefficient parameter $\Omega$.

4. Repeat Steps 2 and 3 for a sufficiently long time.
Summary of the parameter estimates for DMC2

<table>
<thead>
<tr>
<th>True values</th>
<th>Mean</th>
<th>SD</th>
<th>95%PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.00</td>
<td>3.0580</td>
<td>0.0597</td>
<td>2.9409</td>
</tr>
<tr>
<td>2.00</td>
<td>-1.9672</td>
<td>0.0549</td>
<td>-2.0728</td>
</tr>
<tr>
<td>2.00</td>
<td>1.9899</td>
<td>0.0821</td>
<td>1.8263</td>
</tr>
<tr>
<td>1.00</td>
<td>0.9438</td>
<td>0.0814</td>
<td>0.7818</td>
</tr>
<tr>
<td>0.10</td>
<td>0.1111</td>
<td>0.0161</td>
<td>0.0840</td>
</tr>
<tr>
<td>-0.05</td>
<td>-0.0347</td>
<td>0.0167</td>
<td>-0.0695</td>
</tr>
<tr>
<td>0.20</td>
<td>0.2224</td>
<td>0.0328</td>
<td>0.1676</td>
</tr>
</tbody>
</table>

Note: Result from 1 dataset
Summary of the parameter estimates for MCMC

<table>
<thead>
<tr>
<th>True values</th>
<th>Mean</th>
<th>SD</th>
<th>95%PI</th>
<th>INEF</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>3.0387</td>
<td>0.0517</td>
<td>2.9377</td>
<td>3.1418</td>
<td>1.4768</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-1.9790</td>
<td>0.0475</td>
<td>-2.0738</td>
<td>-1.8858</td>
<td>0.4317</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>1.9264</td>
<td>0.0711</td>
<td>1.7849</td>
<td>2.0660</td>
<td>1.2698</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>0.9542</td>
<td>0.0704</td>
<td>0.8160</td>
<td>1.0927</td>
<td>1.2881</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.1110</td>
<td>0.0160</td>
<td>0.0839</td>
<td>0.1472</td>
<td>0.5902</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>-0.0352</td>
<td>0.0165</td>
<td>-0.0701</td>
<td>-0.0051</td>
<td>0.6408</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>0.2176</td>
<td>0.0319</td>
<td>0.1636</td>
<td>0.2861</td>
<td>0.2162</td>
</tr>
</tbody>
</table>

Note: Result from 1 dataset
In Appendix, the MCMC details are described
Summary of the results
(100 Monte Carlo trials, n=50)

<table>
<thead>
<tr>
<th>True values</th>
<th>n=50</th>
<th>DMC1</th>
<th>DMC2</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
<td>SDs</td>
</tr>
<tr>
<td>3.00</td>
<td>3.0011</td>
<td>0.0804</td>
<td>3.0010</td>
<td>0.0803</td>
</tr>
<tr>
<td>-2.00</td>
<td>-1.9881</td>
<td>0.0782</td>
<td>-1.9880</td>
<td>0.0783</td>
</tr>
<tr>
<td>2.00</td>
<td>1.9945</td>
<td>0.1021</td>
<td>1.9944</td>
<td>0.1021</td>
</tr>
<tr>
<td>1.00</td>
<td>0.9911</td>
<td>0.1113</td>
<td>0.9910</td>
<td>0.1115</td>
</tr>
<tr>
<td>0.10</td>
<td>0.1053</td>
<td>0.0215</td>
<td>0.1052</td>
<td>0.0215</td>
</tr>
<tr>
<td>-0.05</td>
<td>-0.0507</td>
<td>0.0239</td>
<td>-0.0508</td>
<td>0.0239</td>
</tr>
<tr>
<td>0.20</td>
<td>0.2186</td>
<td>0.0484</td>
<td>0.2166</td>
<td>0.0482</td>
</tr>
</tbody>
</table>
Summary of the results
(100 Monte Carlo trials, n=100)

<table>
<thead>
<tr>
<th>True values</th>
<th>n=100</th>
<th>DMC1</th>
<th>DMC2</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
<td>SDs</td>
</tr>
<tr>
<td>3.00</td>
<td>2.9977</td>
<td>0.0484</td>
<td>2.9977</td>
<td>0.0485</td>
</tr>
<tr>
<td>-2.00</td>
<td>-2.0045</td>
<td>0.0520</td>
<td>-2.0045</td>
<td>0.0521</td>
</tr>
<tr>
<td>2.00</td>
<td>1.9939</td>
<td>0.0749</td>
<td>1.9940</td>
<td>0.0749</td>
</tr>
<tr>
<td>1.00</td>
<td>0.9919</td>
<td>0.0704</td>
<td>0.9920</td>
<td>0.0705</td>
</tr>
<tr>
<td>0.10</td>
<td>0.1067</td>
<td>0.0155</td>
<td>0.1067</td>
<td>0.0155</td>
</tr>
<tr>
<td>-0.05</td>
<td>-0.0527</td>
<td>0.0150</td>
<td>-0.0526</td>
<td>0.0150</td>
</tr>
<tr>
<td>0.20</td>
<td>0.2071</td>
<td>0.0253</td>
<td>0.2062</td>
<td>0.0252</td>
</tr>
</tbody>
</table>
### Summary of the results
(100 trials; the same data)

<table>
<thead>
<tr>
<th>n=50</th>
<th>DMC1</th>
<th>DMC2</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>2.9886</td>
<td>0.00076</td>
<td>2.9876</td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>-1.9118</td>
<td>0.00076</td>
<td>-1.9113</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>1.9800</td>
<td>0.00085</td>
<td>1.9800</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>0.9057</td>
<td>0.00078</td>
<td>0.9055</td>
</tr>
<tr>
<td>$\alpha_1^2$</td>
<td>0.1091</td>
<td>0.00024</td>
<td>0.1092</td>
</tr>
<tr>
<td>$\alpha_2^2$</td>
<td>-0.0667</td>
<td>0.00022</td>
<td>-0.0661</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>0.1934</td>
<td>0.00035</td>
<td>0.1930</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n=100</th>
<th>DMC1</th>
<th>DMC2</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SDs</td>
<td>Mean</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>3.0030</td>
<td>0.00056</td>
<td>3.0033</td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>-1.9578</td>
<td>0.00052</td>
<td>-1.9571</td>
</tr>
<tr>
<td>$\beta_{21}$</td>
<td>2.0045</td>
<td>0.00066</td>
<td>2.0043</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>1.0263</td>
<td>0.00059</td>
<td>1.0264</td>
</tr>
<tr>
<td>$\alpha_1^2$</td>
<td>0.1081</td>
<td>0.00014</td>
<td>0.1081</td>
</tr>
<tr>
<td>$\alpha_2^2$</td>
<td>-0.0352</td>
<td>0.00016</td>
<td>-0.0350</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>0.1945</td>
<td>0.00026</td>
<td>0.1926</td>
</tr>
</tbody>
</table>
Variable selection

- Improper prior $\rightarrow$ Bayes factor
- Information criteria

$$IC = -2 \int \log L(D \mid b, \Sigma) g(b, \Sigma \mid D) db d\Sigma + 2 \times s,$$

- BPIC (Ando, 2007) $s = \dim\{b\} + \dim\{\Sigma\}$
- DIC (Spiegelhalter et al. 2002)
  $$s = \log L(D \mid \bar{b}, \bar{\Sigma}) - \int \log L(D \mid b, \Sigma) g(b, \Sigma \mid D) db d\Sigma$$
Variable selection results (100 Monte Carlo trials)

| Model   | \( x_1 = x_{11}, x_2 = x_{21} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = x_{11}, x_2 = x_{22} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = x_{11}, x_2 = (x_{21}, x_{22})' \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = x_{21}, x_2 = x_{21} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = x_{21}, x_2 = x_{22} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = (x_{11}, x_{12})', x_2 = x_{21} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = (x_{11}, x_{12})', x_2 = x_{22} \) | \( k = 50 \) | \( k = 100 \) | \( x_1 = (x_{11}, x_{12})', x_2 = (x_{21}, x_{22})' \) | \( k = 50 \) | \( k = 100 \) |
|---------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|---------------------------------|-------------|-------------|
| BPIC    | 0 \( \uparrow \)               | 0 \( \uparrow \) | 0 \( \uparrow \) | 0 \( \uparrow \)               | 0 \( \uparrow \) | 0 \( \uparrow \) | 0 \( \uparrow \)               | 0 \( \uparrow \) | 0 \( \uparrow \) | 91 \( \uparrow \)               | 80 \( \uparrow \) | 94 \( \uparrow \) | 83 \( \uparrow \)               | 2 \( \uparrow \) | 4 \( \uparrow \) | 0 \( \uparrow \)               | 0 \( \uparrow \) | 0 \( \uparrow \) | 0 \( \uparrow \)               | 0 \( \uparrow \) | 2 \( \uparrow \) |
Summary of the parameter estimates from other MCMC setting

Meyer et al. (2003)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>95% Pls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>β₁</td>
<td>2.9342</td>
<td>0.0054</td>
<td>2.9236</td>
<td>2.9452</td>
</tr>
<tr>
<td>β₂</td>
<td>-1.9971</td>
<td>0.0050</td>
<td>-1.997</td>
<td>-1.9771</td>
</tr>
<tr>
<td>α₁</td>
<td>1.9994</td>
<td>0.0076</td>
<td>1.9276</td>
<td>2.0092</td>
</tr>
<tr>
<td>α₂</td>
<td>0.8994</td>
<td>0.0075</td>
<td>0.8845</td>
<td>0.9141</td>
</tr>
<tr>
<td>α₃</td>
<td>0.1041</td>
<td>0.0151</td>
<td>0.0787</td>
<td>0.1379</td>
</tr>
<tr>
<td>α₄</td>
<td>-0.0430</td>
<td>0.0159</td>
<td>-0.0789</td>
<td>-0.0133</td>
</tr>
<tr>
<td>α₅</td>
<td>0.2186</td>
<td>0.0316</td>
<td>0.1655</td>
<td>0.2887</td>
</tr>
</tbody>
</table>

Previous table

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>95% Pls</th>
<th>INEF</th>
<th>CD</th>
</tr>
</thead>
<tbody>
<tr>
<td>β₁</td>
<td>3.0387</td>
<td>0.0517</td>
<td>2.9377</td>
<td>3.1418</td>
<td>1.4768</td>
</tr>
<tr>
<td>β₂</td>
<td>-1.9790</td>
<td>0.0475</td>
<td>-2.0738</td>
<td>-1.8838</td>
<td>0.4317</td>
</tr>
<tr>
<td>α₁</td>
<td>1.9264</td>
<td>0.0711</td>
<td>1.7849</td>
<td>2.0660</td>
<td>1.2698</td>
</tr>
<tr>
<td>α₂</td>
<td>0.9542</td>
<td>0.0704</td>
<td>0.8160</td>
<td>1.0927</td>
<td>1.2881</td>
</tr>
<tr>
<td>α₃</td>
<td>0.1110</td>
<td>0.0160</td>
<td>0.0839</td>
<td>0.1472</td>
<td>0.5902</td>
</tr>
<tr>
<td>α₄</td>
<td>-0.0352</td>
<td>0.0165</td>
<td>-0.0701</td>
<td>-0.0051</td>
<td>0.6408</td>
</tr>
<tr>
<td>α₅</td>
<td>0.2176</td>
<td>0.0319</td>
<td>0.1636</td>
<td>0.2861</td>
<td>0.2162</td>
</tr>
</tbody>
</table>

Following the suggestion for the MCMC method of Meyer et al. (2003), we set the proposal density for beta as normal with mean equal to the posterior mode and the variance equal to 1/10 of the previous one.