Introduction to Bayesian Econometrics, 2nd Edition
Answers to Exercises, Chapter 11

1. For β, use the result in the text that

$$\hat{y}_t \sim N(\hat{x}_t'\beta, \sigma_u^2), \quad t = p+1, \ldots, T,$$

and $\pi(\beta) \sim N(\bar{\beta}_0, B_0)$ to find

$$\pi(\beta|y, \sigma_u^2, \phi) \propto \exp\left[-\frac{1}{2\sigma_u^2}\sum(\hat{y}_t - \hat{x}_t'\beta)'(\hat{y}_t - \hat{x}_t'\beta)\right] \exp\left[-\frac{1}{2}(\beta - \beta_0)'B_0^{-1}(\beta - \beta_0)\right];$$

the usual calculations yield

$$\beta|y, \sigma_u^2, \phi \sim N_K(\bar{\beta}, B_1),$$

where

$$B_1 = \left[\sigma_u^{-2}\sum_{p+1}^{T}\hat{x}_t\hat{x}_t' + B_0^{-1}\right]^{-1}$$
$$\bar{\beta} = B_1\left[\sigma_u^{-2}\sum_{p+1}^{T}\hat{x}_t\hat{y}_t + B_0^{-1}\bar{\beta}_0\right].$$

(Note that ϕ is needed to compute $\hat{y}_t$ and $\hat{x}_t$.)

The likelihood for $\hat{y}_t$ and the prior for $\sigma_u^2 = 1/h_u$ imply

$$\pi(h_u|y, \beta, \phi) \propto h_u^{(T-p)/2}\exp\left[\sum_{p+1}^{t}(\hat{y}_t - \hat{x}_t\beta)'(\hat{y}_t - \hat{x}_t\beta)\right]\, h_u^{\alpha_0/2-1}\, \exp\left[-\frac{\delta_0 h_u}{2}\right],$$

from which we find that

$$h_u|y, \beta, \phi \sim \mathrm{Ga}(\alpha_1/2, \delta_1/2),$$

where

$$\alpha_1 = \alpha_0 + T - p,$$
$$\delta_1 = \delta_0 + \sum_{p+1}^{t} (\hat{y}_t - \hat{x}_t\beta)'(\hat{y}_t - \hat{x}_t\beta).$$

Finally, for ϕ, from the definitions of y_t^* and Y_{t-1} in the text and from the prior for ϕ, we have

$$\phi|y, \sigma_u^2, \beta \propto \exp\left[-\frac{1}{2\sigma_u^2}\sum_{p+1}^{T}(y_t^* - Y_{t-1}^*\phi)'(y_t^* - Y_{t-1}^*\phi)\right]$$
$$\times \exp\left[-\frac{1}{2}(\phi - \phi_0)'\Phi_0^{-1}(\phi - \phi_0)\right] 1(\phi \in S_\phi).$$

Accordingly, we find

$$\phi|y, \beta, \sigma_u^2 \sim N_p(\hat{\phi}, \hat{\Phi})\, 1(\phi \in S_\phi),$$

where

$$\hat{\Phi} = \left[\sigma_u^{-2}\sum_{p+1}^{T} Y_{t+1}^{*'} Y_{t-1}^* + \Phi_0^{-1}\right]^{-1}$$
$$\hat{\phi} = \hat{\Phi}\left[\sigma_u^{-2}\sum_{p+1}^{T} Y_{t-1}^{*'} y_{t-1}^* + \Phi_0^{-1}\phi_0\right].$$

2. We take the case $p = 3$ as an example; the general case follows readily. With $e_1' = (1, 0, 0)$ and

$$E_t = \begin{pmatrix} \epsilon_t \\ \epsilon_{t-1} \\ \epsilon_{t-2} \end{pmatrix},$$

we have $e_1' E_t = \epsilon_t$, which verifies that $y_t = x_t'\beta + \epsilon_t$. And from $E_t = GE_{t-1} +$

$e_1 u_t$, we have

$$\begin{pmatrix} \epsilon_t \\ \epsilon_{t-1} \\ \epsilon_{t-2} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \epsilon_{t-1} \\ \epsilon_{t-2} \\ \epsilon_{t-3} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} u_t.$$

The first row shows that

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \phi_3 \epsilon_{t-3} + u_t.$$

What do the second and third rows show?

3. The predictive distribution is defined by

$$y_{T+1}|Y_T = \int f(y_{t+1}|Y_T, \beta, \phi, \sigma_u^2)\, \pi(\beta, \phi, \sigma_u^2|Y_T)\, d\beta\, d\phi\, d\sigma_u^2.$$

To approximate the right-hand side (note that this calculation requires a value for x_{T+1}), write

$$\begin{aligned} y_{T+1} &= x'_{T+1}\beta + \epsilon_{T+1} \\ &= x'_{T+1}\beta + \phi_1 \epsilon_T + \cdots + \phi_p \epsilon_{T-p+1} + u_{T+1} \\ &= \left[x'_{T+1}\beta + \phi_1(y_T - x'_T\beta) + \cdots + \phi_p(y_{T-p+1} - x'_{T-p+1}\beta)\right] + u_{T+1} \\ &\equiv E(y_{T+1}) + u_{T+1}. \end{aligned}$$

Now define

$$E(y_{T+1}^{(g)}) = x'_{T+1}\beta^{(g)} + \phi_1^{(g)}(y_T - x'_T\beta^{(g)}) + \cdots + \phi_p^{(g)}(y_{T-p+1} - x'_{T-p+1}\beta^{(g)}).$$

Finally, we draw a sample

$$y_{T+1}^{(g)} \sim N\big(E(y_{T+1}^{(g)}), \sigma_u^{2(g)}\big).$$

The sample values are an approximation to the predictive distribution.

4. From (10.11), the likelihood function for the model of (11.13), and the prior for ρ, we have

$$\begin{aligned}
\pi(\beta, b, h_u, D, \rho|y) \propto \big(h_u(1-\rho^2)\big)^{T/2} & \frac{1}{|\Omega|^{T/2}} \\
& \times \exp\left[-\frac{h_u(1-\rho^2)}{2}\sum_i (y_i - X_i\beta - W_i b_i)'\Omega^{-1}(y_i - X_i\beta - W_i b_i)\right] \\
& \times \exp\left[-\frac{1}{2}(\beta-\beta_0)'B_0^{-1}(\beta-\beta_0)\right] h_u^{\alpha_0/2-1}\exp\left[-\frac{\delta_0 h_u}{2}\right] \\
& \times |D|^{-n/2}\exp\left[-\frac{1}{2}\sum b_i' D^{-1} b_i\right] \\
& \times |D|^{-(\nu_0-K_2-1)/2}\exp\left[-\frac{1}{2}\operatorname{tr}(D_0^{-1}D^{-1})\right] \\
& \times \mathrm{TN}_{(-1,1)}(\rho_0, R_0).
\end{aligned}$$

Given ρ, algorithm 10.3 is easily modified. (Note that I'm using $h_u = 1/\sigma^2$ and a gamma distribution prior, $h_u \sim \mathrm{Ga}(\alpha_0/2, \delta_0/2)$.) The distributions needed to construct an algorithm are as follows:

$$h_u|y, \beta, b, D, \rho \sim \mathrm{Ga}(\alpha_1/2, \delta_1/2),$$

where

$$\begin{aligned}
\alpha_1 &= \alpha_0 + T \\
\delta_1 &= \delta_0 + (1-\rho)^2 \sum_i (y_i - X_i\beta - W_i b_i)'\Omega^{-1}(y_i - X_i\beta - W_i b_i);
\end{aligned}$$

$$D^{-1}|b \sim W_{K_2}(\nu_1, D_1),$$

where

$$\begin{aligned} \nu_1 &= \nu_0 + n, \\ D_1 &= \left[D_0^{-1} + \sum b_i b_i'\right]^{-1}; \end{aligned}$$

$$b_i|y, \beta, h_u, D, \rho \sim N_{K_2}(\bar{b}_i, D_{1i}),$$

where

$$\begin{aligned} D_{1i} &= [h_u(1-\rho)^2 W_i' \Omega^{-1} W_i + D^{-1}]^{-1} \\ \bar{b}_i &= D_{1i} h_u (1-\rho)^2 W_i' \Omega^{-1} \tilde{y}_i. \end{aligned}$$

For β, start with

$$\begin{aligned} \mathrm{Cov}(y_i) &= E[(W_i b_i + \epsilon_i)(W_i b_i + \epsilon_i)'] \\ &= W_i D W_i' + \frac{1}{h_u(1-\rho^2)}\Omega \\ &\equiv B_{1i}. \end{aligned}$$

Then

$$\beta|y, h_u, D \sim N_{K_1}(\bar{\beta}, B_1),$$

where

$$\begin{aligned} B_1 &= \left[\sum X_i' B_{1i}^{-1} X_i + B_0^{-1}\right]^{-1} \\ \bar{\beta} &= B_1 \left[\sum X_i' B_{1i}^{-1} y_i + B_0^{-1} \beta_0\right]. \end{aligned}$$

The sampling for ρ requires an MH step. We have

$$\pi(\rho|y,\beta,h_u) \propto \left(\frac{1-\rho^2}{|\Omega|}\right)^{T/2} \exp\left[-\frac{h_u(1-\rho^2)}{2}\sum_i (y_i - X_i\beta - W_ib_i)'\Omega^{-1}(y_i - X_i\beta - W_ib_i)\right]$$
$$\times \exp\left[-\frac{1}{2}(\rho-\rho_0)'R_0^{-1}(\rho-\rho_0)\right] 1\big(\rho \in (-1,1)\big).$$

5. This model is considered in Chib and Greenberg (1995a, sec. 3). I sketch out the general idea here and refer to that paper for details. (The paper also includes a hierarchical structure for β, which I have ignored in the following discussion.) The model is written as

$$y_j = X_j\beta_j + \epsilon_j,$$

where

$$\epsilon_j = \Phi\epsilon_{j-1} + u_j, \quad \Phi \in 1_S(\Phi),$$

where $1_S(\Phi) = 1$ if all roots of Φ line inside the unit circle and 0 otherwise, and

$$u_j \sim N_S(0,\Omega).$$

The assumption on the ϵ_j is a first order vector autoregressive process; the assumption on u_j embodies a SUR specification. We continue with the priors assumed in the text for β and Σ^{-1}, and assume Φ is distributed uniformly on the stationary region. For simplicity, we condition on y_1.

We have, for $j = 2,\ldots,J$,

$$\begin{aligned} y_j &= X_j\beta + \epsilon_j \\ &= X_j\beta + \Phi\epsilon_{j-1} + u_j \\ &= X_j\beta + \Phi(y_{j-1} - X_{j-1}\beta) + u_j, \end{aligned}$$

from which we can write

$$y_j - \Phi y_{j-1} = (X_j - \Phi X_{j-1})\beta + u_j,$$
$$y_j^* = X_j^* \beta + u_j,$$

where $y_j^* = y_j - \Phi y_{j-1}$ and $X_j^* = X_j - \Phi X_{j-1}$. It follows immediately that

$$\beta | y, \Sigma, \Phi \sim N_K(\bar{\beta}, B_1),$$

where

$$B_1 = \left[\sum_{j=2}^{j=J} X_j^{*'} \Sigma^{-1} X_j^* + B_1^{-1} \right]^{-1},$$
$$\bar{\beta} = B_1 \left[\sum_{j=2}^{j=J} X_j^{*'} \Sigma^{-1} y_j + B_1^{-1} \beta_0 \right].$$

An argument similar to that in the text implies

$$\Sigma^{-1} \sim W_S(\nu_1, R_1),$$

where

$$\nu_1 = \nu_0 + J - 1,$$
$$R_1 = \left[R_0^{-1} + \sum_{j=2} (y_j^* - X_j^* \beta)(y_j^* - X_j^* \beta)' \right]^{-1}.$$

Finally, for Φ, the expression above can be rewritten as

$$y_j - X_j \beta = \Phi(y_{j-1} - X_{j-1}\beta) + u_j,$$

or

$$e_j = \Phi e_{j-1} + u_j.$$

By Bayes theorem,

$$\begin{aligned}
\pi(\Phi|y,\beta,\Sigma) &\propto f(e_2|e_1,\beta,\Sigma,\Phi)\cdots f(e_J|e_{J-1},\beta,\Sigma,\Phi)\,\pi(\Phi) \\
&\propto \exp\left[-\frac{1}{2}\sum_{j=2}(e_j-\Phi e_{j-1})'\Sigma^{-1}(e_j-\Phi e_{j-1})\right]\,1_S(\Phi) \\
&= \exp\left[-\frac{1}{2}\,\mathrm{tr}\{\Sigma^{-1}\sum_{j=2}(e_j-\Phi e_{j-1})(e_j-\Phi e_{j-1})'\}\right]\,1_S(\Phi) \\
&\propto \exp\left[-\frac{1}{2}\,\mathrm{tr}\{\Sigma^{-1}\,\Phi\sum_{j=2}e_{j-1}e_{j-1}'\Phi' - \Phi\sum_{j=2}e_{j-1}e_j' - \sum_{j=2}e_j e_{j-1}'\Phi'\}\right]\,1_S(\Phi).
\end{aligned}$$

Comparing this last expression with that for the matricvariate normal distribution of section A.1.15, with $X = \Phi'$ and $\Omega = \Sigma$, you can verify that

$$\Phi' \sim \mathrm{MN}_{S\times S}(\bar{\Phi}', \Sigma\otimes P),$$

where $P = (E_J'E_J)^{-1}$ and

$$E_J = (e_1,\ldots,e_{J-1})',$$

so that $\sum_{j=2} e_{j-1}e_{j-1}' = E_J'E_J$, and

$$\bar{\Phi}' = (E_J'E_J)^{-1}\sum_{j=2}e_{j-1}e_j'.$$

Note that the notation for the matricvariate normal is slightly different from that used in Chib and Greenberg (1995a).

To sample from the full conditional distribution of Φ, we use the theorem that

$X \sim \mathrm{MN}_{p\times q}(\mu, \Sigma \otimes P)$ if and only if $x \equiv \mathrm{vec}(X) \sim N_{pq}(\mathrm{vec}(\mu), \Sigma \otimes P)$. Then let $AA' = P$ and $BB' = \Sigma$ (A and B are Cholesky decompositions of their respective matrices) and let

$$X = \mu + AZB',$$

where Z is a $p \times q$ matrix of standard normal variates. Then

$$\begin{aligned} \mathrm{vec}(X) &= \mathrm{vec}(\mu) + \mathrm{vec}(AZB') \\ &= \mathrm{vec}(\mu) + (B \otimes A)\,\mathrm{vec}(Z), \end{aligned}$$

from which it follows that $E\big(\,\mathrm{vec}(X)\big) = \mathrm{vec}(\mu)$ and

$$\mathrm{Cov}\,\big(\,\mathrm{vec}(X)\big) = (B\otimes A)E\big(\,\mathrm{vec}(Z)\,\mathrm{vec}(Z)'\big)(B\otimes A)' = BB'\otimes AA' = \Sigma\otimes P,$$

where we have used a theorem from Section A.2. Accordingly, we can draw Z and form A and B from the current values of Σ and P, and then add AZB to the current value of $\bar{\Phi}'$ to obtain a candidate draw that is accepted if the roots of Φ are less than one.