Introduction to Bayesian Econometrics, 2nd Edition
Answers to Exercises, Chapter 10

1. For the first part of the exercise, we have

$$
\begin{aligned}
f(y_j|X_j,\beta,\Sigma) &\propto \int_0^\infty \lambda_j^{S/2}\exp\left[-\frac{\lambda_j}{2}(y_j - X_j\beta)'\Sigma^{-1}(y_j - X_j\beta)\right]\lambda_j^{\nu/2-1}\exp\left[-\frac{\nu\lambda_j}{2}\right]d\lambda_j \\
&= \int_0^\infty \lambda_j^{(\nu+S)/2-1}\exp\left[-\lambda_j\left(\frac{\nu + (y_j - X_j\beta)'\Sigma^{-1}(y_j - X_j\beta)}{2}\right)\right]d\lambda_j.
\end{aligned}
$$

The integral is over the kernel of a gamma distribution, which implies

$$
f(y_j|X_j,\beta,\Sigma) \propto \left(\frac{\nu + (y_j - X_j\beta)'\Sigma^{-1}(y_j - X_j\beta)}{2}\right)^{-(\nu+S)/2},
$$

from which we have

$$
f(y_j|X_j,\beta,\Sigma) \propto \left(1 + \frac{1}{\nu}\big((y_j - X_j\beta)'\Sigma^{-1}(y_j - X_j\beta)\big)\right)^{-(\nu+S)/2}.
$$

This establishes the result.

A Gibbs algorithm proceeds as follows.

(a) Sample

$$
\lambda_j^{(g)} \sim \mathrm{Ga}\left(\frac{\nu+1}{2}, \frac{\nu + (y_j - X_j\beta^{(g-1)})'\Sigma^{-1(g-1)}(y_j - X_j\beta^{(g-1)})}{2}\right).
$$

(b) Sample $\beta^{(g)} \sim N_K(\bar\beta^{(g)}, B_1^{(g)}$ where

$$
\begin{aligned}
B_1^{(g)} &= \left[\sum_j \lambda_j^{(g)} X_j'\Sigma^{-1(g-1)}X_j + B_0^{-1}\right]^{-1}, \\
\bar\beta^{(g)} &= B_1^{(g)}\left[\sum_j \lambda_j^{(g)} X_j'\Sigma^{-1(g-1)}y_j + B_0^{-1}\beta_0\right].
\end{aligned}
$$

(c) Sample $\Sigma^{(g)} \sim W_S(\nu_1, R_1^{(g)})$, where

$$\nu_1 = \nu_0 + J,$$
$$R_1^{(g)} = \left[R_0^{-1} + \sum_j \lambda_j^{(g)} (y_j - X_j\beta^{(g)})(y_j - X_j\beta^{(g)})' \right]^{-1}.$$

2. Take as prior distributions $\sigma_j^2 \sim \mathrm{IG}(\alpha_{j0}/2, \delta_{j0}/2)$. Then

$$\begin{aligned}\pi(\sigma_j^2|y,\beta) &\propto \left(\frac{1}{\sigma_j^2}\right)^{S/2} \exp\left[-\frac{1}{2\sigma_j^2}(y_j - X_j\beta)'(y_j - X_j\beta)\right] \left(\frac{1}{\sigma_j^2}\right)^{\alpha_{j0}/2+1} \exp\left[-\frac{\delta_{j0}}{2\sigma_j^2}\right] \\ &= \left(\frac{1}{\sigma_j^2}\right)^{(S+\alpha_{j0})/2+1} \exp\left[-\frac{1}{\sigma_j^2}\left(\frac{\delta_{j0} + (y_j - X_j\beta)'(y_j - X_j\beta)}{2}\right)\right],\end{aligned}$$

which implies that σ_j^2 should be drawn from

$$\mathrm{IG}\left(\frac{S + \alpha_{j0}}{2}, \frac{\delta_{j0} + (y_j - X_j\beta)'(y_j - X_j\beta)}{2}\right).$$

Finally, it is easy to see that

$$\beta|\sigma_1^2, \ldots, \sigma_S^2 \sim N(\bar{\beta}, B_1),$$

where

$$\begin{aligned} B_1 &= \left[\sum \frac{1}{\sigma_j^2} X_j' X_j + B_0^{-1}\right]^{-1} \\ \bar{\beta} &= B_1 \left[\sum \frac{1}{\sigma_j^2} X_j' y_j + B_0^{-1}\beta_0\right]. \end{aligned}$$

3. The most straightforward way to design an algorithm is to add y_{it}^* to the sampler and set $h_u = 1$ for the usual identification reason. As usual, y_{it}^* is drawn from a truncated normal distribution, with the truncation depending on y_{it}.

Carlin and Chib (1999) suggest marginalizing out the b_i from the full conditional for β and from the full conditional for the y^*_{it}, which they call z_{it}. From $y^*_i = X_i\beta + W_i b_i + u_i$, we can think of $W_i b_i + u_i$ as an unobserved error term, from which it follows that

$$\begin{aligned} \mathrm{Var}(y^*_i) &= E(W_i b_i + u_i)(W_i b_i + u_i)' \\ &= W_i D W_i' + I. \end{aligned}$$

In this formulation, we have

$$\begin{aligned} \pi(\beta|y, y^*, D) \propto \exp &\left[-\frac{1}{2} \sum_i \big(y^*_i - X_i\beta\big)'(I + W_i D W_i')^{-1}\big(y^*_i - X_i\beta\big) \right] \\ &\times \exp\left[-\frac{1}{2}(\beta - \beta_0)' B_1^{-1}(\beta - \beta_0) \right]. \end{aligned}$$

A standard computation now follows to find the distribution of $\beta|y^*, D$. The same setup shows that we can sample for $y^*|y, \beta, D$ marginalized over the b_i. Ignoring for a moment the truncation of the y^*_{it} induced by y_{it}, we have

$$f(y^*_i|y_i, \beta) \propto \exp\left[-\frac{1}{2} \big(y^*_i - X_i\beta\big)'(I + W_i D W_i')^{-1}\big(y^*_i - X_i\beta\big) \right],$$

which implies that we can draw y^*_i from a multivariate Gaussian distribution with mean $X_i\beta$ and covariance matrix $(I + W_i D W_i')$, truncated according to the values of y_i; that is, y^*_{it} is restricted to negative values if $y_{it} = 0$ and positive values otherwise.

4. We derive the full conditional distributions, from which a Gibbs algorithm can easily be assembled. Start from

$$y_i = X_i\beta + W_i b_i + u_i,$$

and let $\Lambda_i = \text{diag}(\lambda_1, \ldots, \lambda_T)$. We can, as in the text, integrate out the b_i from $b_i \sim N_T(0, D)$ to obtain

$$y_i \sim N(X_i\beta, V_i),$$

where

$$V_i = (h_u\Lambda_i)^{-1} + W_i D W_i'.$$

It follows that

$$\beta | y, h_u, D, \{\Lambda_i\} \sim N(\bar{\beta}, B_1),$$

where

$$B_1 = \left(\sum_i X_i' V_i^{-1} X_i + B_0^{-1} \right)^{-1}$$

$$\bar{\beta} = B_1 \left(\sum_i X_i' V_i^{-1} y_i + B_0^{-1}\beta_0 \right).$$

For b_i, take $\bar{y}_i$ as defined in the text and write

$$\pi(b_i | y_i, \beta, h_u, D, \Lambda_i) \propto \exp\left[-\frac{h_u}{2} (\bar{y}_i - W_i b_i)' \Lambda_i (\bar{y}_i - W_i b_i) \right] \exp\left[-\frac{1}{2}\, b_i D b_i \right],$$

from which standard computations imply

$$b_i | y_i, \beta, h_u, D, \Lambda_i \sim N(\bar{b}_i, D_{1i}),$$

where

$$D_{1i} = \left(h_u W_i' \Lambda_i W_i + D^{-1} \right)^{-1},$$

$$\bar{b}_i = D_{1i}\big(h_u W_i' \Lambda_i \bar{y}_i \big).$$

For h_u, the only change from the text is to redefine δ_1 as

$$\delta_1 = \delta_0 + \sum (y_i - X_i\beta - W_i b_i)' \Lambda_i (y_i - X_i\beta - W_i b_i).$$

Finally, the full conditional distribution of D^{-1} is unchanged.

5. For panel data, start from the definition of y^*_{it},

$$y^*_{it} = \begin{cases} y^*_{Cit}, & \text{if } it \in C, \\ y_{it}, & \text{if } it \in C^c, \end{cases}.$$

Here, C is the set of y_{it} for which $y_{it} = 0$. Our model is

$$y^*_{it} = x'_{it}\beta + w'_{it} b_i + u_{it},$$

where y^*_{it} represents observed or latent data, depending on whether y_{it} is observed. Algorithm 10.3 can now be applied with y^*_{it} in place of y_{it}, and the sampling for y^*_{it} is from a truncated Gaussian distribution if $y_{it} = 0$.

For SUR data, we again add latent data to the sampler. We now write

$$y^*_{sj} = x^*_{sj}\beta_s + u_{sj},$$

where $y^*_{sj} = y_{sj}$ if the latter is observed and is drawn from a Gaussian distribution truncated at 0 if $y_{sj} = 0$. Algorithm 10.1 still applies, with y^*_j in place of y_j.

6. This is a rather difficult problem because there is no natural multivariate distribution for this case as there is for the probit case, where the multivariate normal is associated with normal marginal distributions.

This model is discussed from the Bayesian viewpoint in O'Brien, S. M. and

Dunson, David, B. (2004). Bayesian multivariate logistic regression. *Biometrics 60*, 3 (September), 739-746.

They propose a joint distribution that has the correlation structure of a scale mixture of normals and univariate logistic marginal distributions. The paper also proposes an importance sampling algorithm estimate the posterior distribution of the parameters.

7. To find A_i, we find the normalizing constant in $\pi(\lambda_i|y_i, \beta, h_u)$. By the usual calculations, we have

$$\begin{aligned}\pi(\lambda_i|y_i, \beta, h_u) &\propto f(y_i|\beta, h_u, \lambda_i)\,\pi(\lambda_i) \\ &\propto \lambda_i^{T/2} \exp\left[-\frac{h_u\lambda_i}{2}(y_i - X_i\beta)'(y_i - X_i\beta)\right] \lambda_i^{\nu/2-1} \exp\left[-\frac{\nu\lambda_i}{2}\right] \\ &= \lambda_i^{(\nu+T)/2-1} \exp\left[-\lambda_i \frac{\nu + h_u(y_i - X_i\beta)'(y_i - X_i\beta)}{2}\right].\end{aligned}$$

The last expression is recognized as a

$$\mathrm{Ga}\left(\frac{\nu+T}{2}, \frac{\nu + h_u(y_i - X_i\beta)'(y_i - X_i\beta)}{2}\right)$$

distribution, for which the normalizing constant is

$$K_i = \Gamma\left(\frac{\nu+T}{2}, \frac{\nu + h_u(y_i - X_i\beta)'(y_i - X_i\beta)}{2}\right),$$

and $A_i = \alpha K_i$. This result and the distribution for the λ_i in the text provides an algorithm to sample $\lambda_i|y_i, \beta, h_u, \lambda_{-i}$.

To specify an algorithm for the other parameters, start with the joint posterior distribution that appears below (10.11) modified to include $\lambda = (\lambda_1, \ldots, \lambda_n)$:

$$
\begin{aligned}
\pi(\beta, b, h_u, D, \lambda | y) \propto h_u^{nT/2} & \left[\prod \lambda_i^{T/2}\right] \exp\left[-\frac{h_u}{2} \sum \lambda_i (y_i - X_i\beta - W_i b_i)'(y_i - X_i\beta - W_i b_i)\right] \\
& \times \exp\left[-\frac{1}{2}(\beta - \beta_0)' B_0^{-1}(\beta - \beta_0)\right] h_u^{\alpha_0/2 - 1} \exp\left[-\frac{\delta_0 h_u}{2}\right] \\
& \times |D|^{-n/2} \exp\left[-\frac{1}{2} \sum b_i' D^{-1} b_i\right] \\
& \times |D|^{-(\nu_0 - K_2 - 1)/2} \exp\left[-\frac{1}{2}\operatorname{tr}(D_0^{-1} D^{-1})\right] \left[\prod \pi(\lambda_i)\right].
\end{aligned}
$$

The distribution $\pi(D|b)$ is unchanged.

To find $h_u | y, \beta, b, D, \lambda$, we have

$$
\begin{aligned}
h_u | y, \beta, b, D, \lambda \propto\ & h_u^{nT/2} \exp\left[-\frac{h_u}{2} \sum \lambda_i (y_i - X_i\beta - W_i b_i)'(y_i - X_i\beta - W_i b_i)\right] \\
& \times h_u^{\alpha_0/2 - 1} \exp\left[-\frac{\delta_0 h_u}{2}\right] \\
=\ & h_u^{(nT + \alpha_0)/2 - 1} \exp\left[-h_u \frac{\delta_0 + \sum \lambda_i (y_i - X_i\beta - W_i b_i)'(y_i - X_i\beta - W_i b_i)}{2}\right],
\end{aligned}
$$

which is proportional to a gamma distribution.

β and the b_i can be handled as in the text:

$$
\pi(b_i | y, \beta, h_u, D, \lambda_i) \propto \exp\left[-\frac{h_u \lambda_i}{2}(y_i - X_i\beta - W_i b_i)'(y_i - X_i\beta - W_i b_i)\right] \exp\left[-\frac{1}{2} b_i' D^{-1} b_i\right].
$$

Complete the square in b_i to obtain that $b_i | y_i, h_u, \beta, D \sim N_{K_2}(\bar{b}_i, D_{1i})$, where

$$
\begin{aligned}
D_{1i} &= \left[h_u \lambda_i W_i' W_i + D^{-1}\right]^{-1}, \\
\bar{b}_i &= D_{1i}\, h_u \lambda_i W_i' \bar{y}_i.
\end{aligned}
$$

Finally, for β, we write $y_i = X_i\beta + W_i b_i + u_i$ to obtain $y_i \sim N_T\big(X_i\beta, B_{1i}\big)$, where $B_{1i} = W_i D W_i' + (h_u \lambda_i)^{-1} I_T$. This is combined with the prior for β to obtain

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

where

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