A Gibbs Sampler for Spatially-Varying Parameter Models with Simultaneous Spatial Effects

Levi John Wolf

August 18, 2016

1 Model

Stating a spatially-varying coefficient model with the most generic possible structure on the simultaneous spatial autoregressive effects will provide a Gibbs sampler that is valid for all restrictions of the complete model. A complete spatially-varying coefficient model with simultaneous spatial effects can be stated:

$$
\begin{aligned}
Y &= \rho \mathbf{W}_1 y + X_1 \beta + X_2 \Delta \alpha + \epsilon \\
\alpha &= \phi \mathbf{W}_2 \alpha + Z\gamma + \zeta \\
\epsilon &\sim \mathcal{N}(0, \Psi(\lambda)\sigma_e^2) \\
\zeta &\sim \mathcal{N}(0, \Psi(\theta)\sigma_\alpha^2)
\end{aligned}
\tag{1}
$$

where $\Psi(\theta), \Psi(\lambda)$ denote the spatial components of covariance due to spatial correlation parameters (θ, λ). For SMA or SAR-Error forms, the spatial component, $\Psi(.)$, is separable from the scale component, σ_*^2, at any level of a hierarchical spatial-autoregressive model. The model stated above is over-specified, and is not recommended for actual analysis. However, focusing on this model will provide a single sampler that is valid for any restrictions of the model, such as no endogenous lag ($\rho, \phi = 0$), or when no spatial effects exist on the upper-level ($\phi, \theta = 0$). In addition, a sampler for this specification will also hold for any restriction of X_1, X_2, and Z, so the sampler derived below is equivalent to the multilevel variance components Gibbs sampler under restrictions of X_1, X_2, and Z. It will also correctly sample a varying-intercept specification with appropriate restrictions of X_1, X_2, and X_1.

2 Likelihoods

The likelihood for the lower-level outcome Y conditional on *all parameters*, $\mathcal{L}(Y|\cdots)$, includes the α term in the linear predictor:

$$\begin{aligned} P(Y|\alpha,\rho,\beta,\Psi(\lambda),\sigma_e^2) =& |I-\rho\mathbf{W}_1| \times |\Psi(\lambda)|^{-\frac{1}{2}} \left(\sigma_e^2\right)^{-\frac{N}{2}} \\ & \times \exp\left[-\frac{1}{2}\left((Y-\rho\mathbf{W}_1 Y - X\beta - \Delta\alpha)' \frac{\Psi(\lambda)^{-1}}{\sigma_e^2}(Y-\rho\mathbf{W}_1 Y - X\beta - \Delta\alpha)\right)\right] \end{aligned} \tag{2}$$

By conditioning on α, we imply that upper-level parameters $(\gamma, \sigma_\alpha^2, \phi, \theta)$ are also contained in the conditioning. The lower-level model is made more tractable by exploiting this inherent structure. Since $\mathcal{L}(\alpha|\phi,\theta,\gamma,\sigma_\alpha^2)$ is similar in form as $\mathcal{L}(Y|\cdots)$, the upper-level likelihood is also at hand:

$$\begin{aligned} \mathcal{L}(\alpha|\phi,\theta,\gamma,\sigma_\alpha^2) =& |I-\phi\mathbf{W}_2| \times |\Psi(\theta)|^{-\frac{1}{2}} \left(\sigma_\alpha^2\right)^{-\frac{J}{2}} \\ & \times \exp\left[-\frac{1}{2}\left((\alpha-\phi\mathbf{W}_2\alpha - Z\gamma)' \frac{\Psi(\theta)^{-1}}{\sigma_\alpha^2}(\alpha-\phi\mathbf{W}_2\alpha - Z\gamma)\right)\right] \end{aligned} \tag{3}$$

Using these two likelihoods, deriving the full conditionals for the posterior distribution is possible.

3 Priors

Gibbs sampling strategies provide significant gains when conjugate priors are used. Therefore, common conjugate families for hierarchical normal models will be used. Specifically:

$$\begin{aligned} \sigma_e^2 &\sim IG(a_e, b_e) \\ \sigma_\alpha^2 &\sim IG(a_\alpha, b_\alpha) \\ \beta &\sim \mathcal{N}(\mu_\beta, \Sigma_{\beta 0}) \\ \gamma &\sim \mathcal{N}(\mu_\gamma, \Sigma_{\gamma 0}) \end{aligned} \tag{4}$$

4 Conditional Posteriors

4.1 Lower-Level Effects β

$$P(\beta|\cdots) \propto \mathcal{N}(\Sigma_\beta b, \Sigma_\beta) \tag{5}$$

$$\Sigma_\beta = \left(X'\Sigma_Y^{-1}X + \Sigma_{\beta 0}^{-1}\right)^{-1} \qquad b = \left(X'\Sigma_Y^{-1}(AY - \Delta\alpha) + \Sigma_{\beta 0}^{-1}\mu_\beta\right) \tag{6}$$

4.2 Upper-Level Effects γ

$$P(\gamma|\cdots) \propto \mathcal{N}(\Sigma_\gamma g_n, \Sigma_\gamma) \tag{7}$$

$$\Sigma_\gamma = \left(Z'\Sigma_\alpha^{-1}Z + \Sigma_{\gamma 0}^{-1}\right)^{-1} \qquad g_n = \left(Z'\Sigma_\alpha^{-1}B\alpha + \Sigma_{\gamma 0}^{-1}\mu_\gamma\right) \tag{8}$$

4.3 Upper-Level Outcomes α

$$P(\alpha|\cdots) \propto \mathcal{N}(\Sigma_a b_a, \Sigma_a) \tag{9}$$

$$b_a = \left(\Delta'\Sigma_Y^{-1}(AY - X\beta) + B'\Sigma_\alpha^{-1}Z\gamma\right) \qquad \Sigma_a = \left(\Delta'\Sigma_Y^{-1}\Delta + B'\Sigma_\alpha^{-1}B\right)^{-1} \tag{10}$$

4.4 Lower-Level Variance σ_e^2

Let $\eta_Y = AY - X\beta - \Delta\alpha$.

$$P(\sigma_e^2|\cdots) \propto IG\left(\frac{N}{2} + a_e, \frac{\eta_Y'\Psi(\lambda)^{-1}\eta_Y}{2} + b_e\right) \tag{11}$$

4.5 Upper-Level Variance σ_α^2

Let $\eta_\alpha = B\alpha - Z\gamma$.

$$P(\sigma_\alpha^2|\cdots) \propto IG\left(\frac{J}{2} + a_\alpha, \frac{\eta_\alpha'\Psi(\theta)^{-1}\eta_\alpha}{2} + b_\alpha\right) \tag{12}$$

4.6 Lower-Level Endogenous Autocorrelation ρ

$$P(\rho|\cdots) \propto |I - \rho\mathbf{W}_1| \exp\left[-\frac{1}{2\sigma_e^2}\eta_Y'\Psi(\lambda)^{-1}\eta_Y\right] \times P(\rho) \tag{13}$$

4.7 Lower-Level Error Autocorrelation λ

$$P(\lambda|\cdots) \propto |\Psi(\lambda)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma_e^2}\eta_Y' \Psi(\lambda)^{-1}\eta_Y\right] \times P(\lambda) \tag{14}$$

4.8 Upper-Level Endogenous Autocorrelation ϕ

$$P(\phi|\cdots) \propto |I - \phi \mathbf{W}_2| \exp\left[-\frac{1}{2\sigma_\alpha^2}\eta_\alpha' \Psi(\theta)^{-1}\eta_\alpha\right] \times P(\phi) \tag{15}$$

4.9 Upper-Level Error Autocorrelation θ

$$P(\theta|\cdots) \propto |\Psi(\theta)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma_\alpha^2}\eta_\alpha' \Psi(\theta)^{-1}\eta_\alpha\right] \times P(\theta) \tag{16}$$

5 Restrictions

- When X_2 = 0, single level.
- When X_2 = 1 and X_1 contains no constant, varying-intercept.
- When X_2 = 1 and $Z = 0$, variance components.
- Otherwise, varying slope/parameter.