Gibbs' Sampler: Dynamic Spatio-Temporal Multivariate Model

Elliot Shannon

shann125@msu.edu

Michigan State University

Setting

We consider multivariate data observed over a spatial domain which is partitioned into a fixed number of small areal units. A fixed number of sampling units are present within each areal unit, where we measure some (or all) outcomes variables from a subset of the sampling units at each time step. We do not have the exact locations of the sampling units; we only know that they are associated with a specific areal unit. The quantity of interest is then the latent mean for each outcome within each areal unit at each time step. We aim to model this latent mean using spatially and temporally explicit covariates through spatially-varying regression coefficients, as well as a spatially-varying intercept term which is allowed to evolve dynamically over time. We will employ conditional autoregressive (CAR) spatial structures for these random effects, specifically multivariate CAR structures where correlation among outcomes is induced through a linear model of coregionalization.

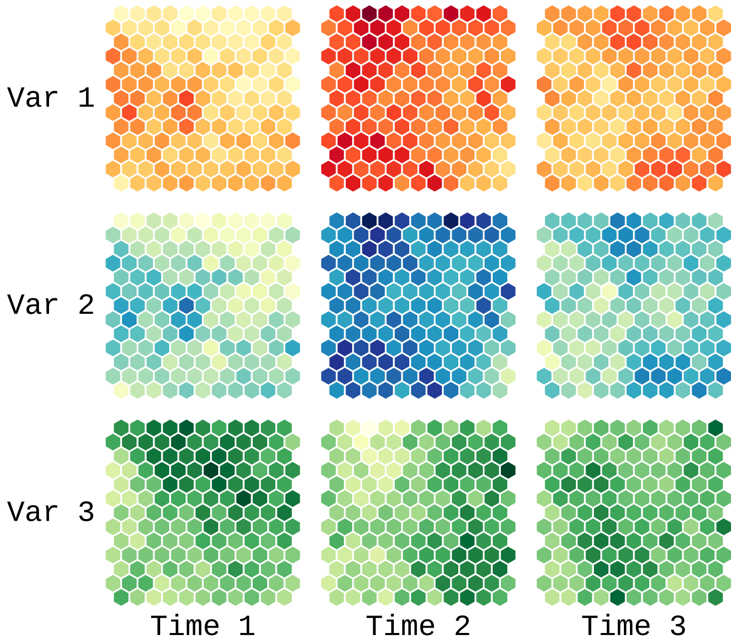


Figure 1: A dynamic 3-dimensional multivariate CAR spatial effect evolving over time. In this example, `Var 1` and `Var 2` are positively correlated, while `Var 1` and `Var 2` are negatively correlated with `Var 3`.

Notation

- Let $m = 1, \dots, M$ index outcome variables (for the multivariate setting, $M > 1$).
- Let $j = 1, \dots, J$ index areal units.
- Let $t = 1, \dots, T$ index discrete time steps.

- Let $i = 1, \ldots, n_{j,t}$ index sampling units where some (or all) of the M outcome variables may be observed within areal unit j at time t.

For areal unit j at time t, let

- $y_{i,m,j,t}$ be the measured value for outcome m from the i^{th} sampling unit $i = 1, \ldots, n_{j,t}$, $m = 1, \ldots, M$. Note that we may have $y_{i,m,j,t} = \texttt{NA}$.
- $\mathbf{y}_{i,j,t} = (y_{i,1,j,t}, \ldots, y_{i,M,j,t})^\top \in \mathbb{R}^M$, $i = 1, \ldots, n_{j,t}$.
- $\boldsymbol{\mu}_{j,t} = (\mu_{1,j,t}, \ldots, \mu_{M,j,t})^\top \in \mathbb{R}^M$ be the vector of latent means (variable of interest).

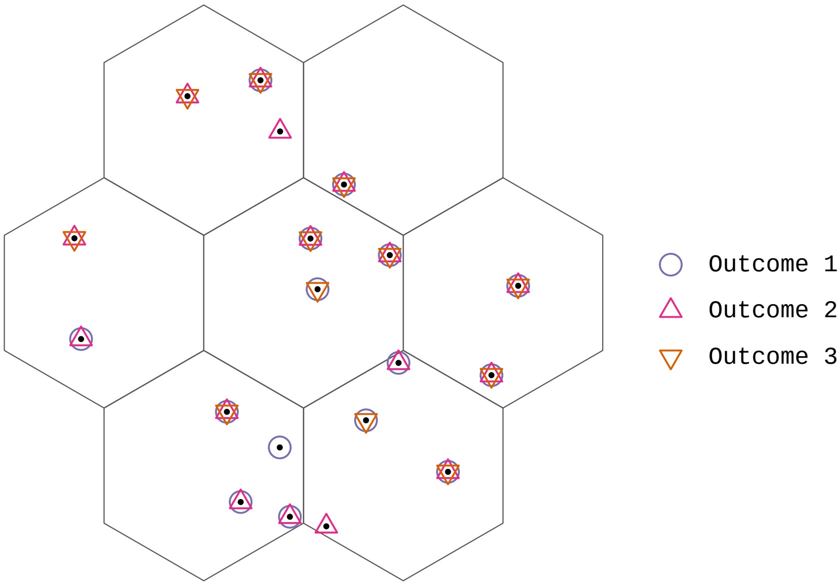


Figure 2: Example areal units (hexagons) and sampling units (points) for a single time point. At each sampling unit, as many as $M = 3$ outcome variables may be measured. However, we do not have measurements for all outcome variables at all sampling units. We also do not know the exact location of each sampling unit, only that it is located in a given areal unit.

Distributions

Let $N(\mu, \sigma^2)$ be the normal distribution with mean $\mu \in \mathbb{R}$, variance $\sigma^2 > 0$, and pdf

$$p(x \mid \mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right). \tag{1}$$

Let $MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ be the Multivariate Normal distribution with mean vector $\boldsymbol{\mu} \in \mathbb{R}^N$, postivie semi-definite covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$, and pdf

$$p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-N/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right). \tag{2}$$

Let $IW(\mathbf{H}, \nu)$ be the Inverse Wishart distribution with (positive definite) scale matrix $\mathbf{H} \in \mathbb{R}^{N \times N}$, degrees of freedom $\nu > N - 1$, and pdf

$$p(\boldsymbol{\Sigma} \mid \mathbf{H}, \nu) = \frac{|\mathbf{H}|^{\nu/2}}{2^{\nu N/2} \Gamma_N\left(\frac{\nu}{2}\right)} |\boldsymbol{\Sigma}|^{-(\nu+N+1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{H}\boldsymbol{\Sigma}^{-1}\right)\right). \tag{3}$$

Let $IG(a, b)$ be the Inverse Gamma distribution with shape parameter $a > 0$, scale parameter $b > 0$, and pdf

$$p(\sigma^2 \mid a, b) = \frac{b^a}{\Gamma(a)}(\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right). \tag{4}$$

Let $U(c, d)$ be the continuous Uniform distribution with support $[c, d]$ and pdf

$$p(\rho \mid c, d) = \begin{cases} \frac{1}{d-c} & \text{if } c \leq \rho \leq d \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

CAR Spatial Structure

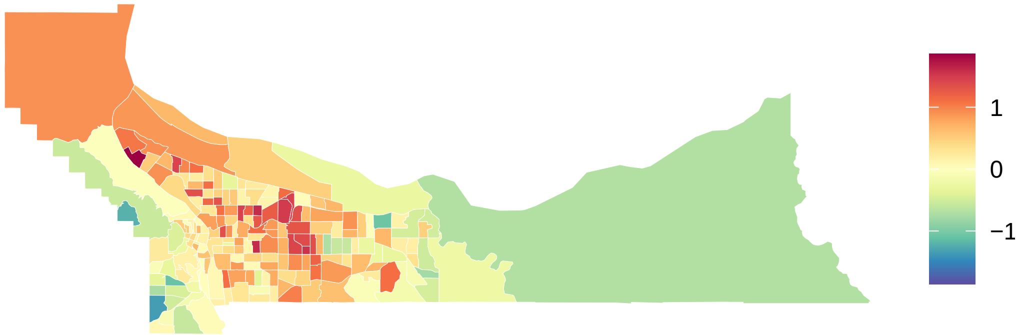
A CAR model is a special form of the multivariate normal (2) distribution that allows areal units to be correlated through a neighborhood structure. Specifically, a CAR model for J many areal units may be specified as a multivariate normal distribution of the form

$$MVN(\boldsymbol{\mu}, \tau^2 \mathbf{Q}(\rho)), \tag{6}$$

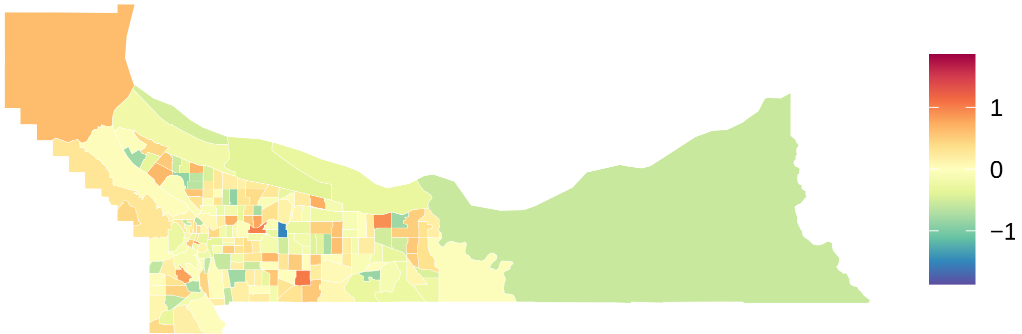
where $\boldsymbol{\mu}$ is a length J mean vector, $\tau^2 > 0$ is a scalar variance parameter, ρ is a spatial correlation parameter, and $\mathbf{Q}(\rho)$ is a $J \times J$ correlation matrix of the form

$$\mathbf{Q}(\rho) = (\mathbf{D} - \rho\mathbf{W})^{-1}, \tag{7}$$

where $\mathbf{D}$ is a $J \times J$ diagonal matrix whose j^{th} diagonal element is the number of neighbors of areal unit j, and $\mathbf{W}$ is a binary $J \times J$ spatial adjacency matrix with diagonal elements set equal to 0. In practice, $\mathbf{W}$ may follow from a variety of nieghborhood definitions. Here, we consider areal units to be neighbors if they share a common border. This covariance structure yields correlations between neighbors, where the strength of this association is dictated by ρ. Specifically, when ρ is very close to 1, we have very high correlation between neighbors (Figure 3a). When ρ is very close to 0, there is very little correlation among neighbors (Figure 3b).



(a) High spatial correlation CAR



(b) Low spatial correlation CAR

Figure 3: Examples of simulated CAR models for 2020 census tracts in Multnomah County, OR with high (a) and low (b) spatial correlation.

In the case of large J, it may be less attractive computationally to work with the Multivariate Normal specification of the CAR model. One advantage of the CAR model structure is that it can be equivalently represented as the joint distribution of J many univariate normal random variables, where the distribution of each variable is conditional on the remaining $J-1$ variables. Better still, the distribution of each variable only depends on it's neighbors (hence the name, 'conditional autoregressive model'), so we can recursively update each variable more efficiently.

To illustrate this, let $\mathbf{u}$ be a length J vector following a CAR Multivariate Normal distribution (6). Let $\mathcal{N}_j$ be the set of neighbor indices of areal unit j. Similarly, let d_j be the number of neighbors associated with areal unit j. We can then write

$$u_j | u_{k,k\neq j} \sim N\left(\mu_j + \rho \sum_{k\in\mathcal{N}_j} \frac{u_k - \mu_k}{d_j}, \frac{\tau^2}{d_j}\right), \tag{8}$$

where u_j is the j^{th} element of $\mathbf{u}$ and μ_j is the j^{th} element of $\boldsymbol{\mu}$.

The intuition behind equation (8) is that the distribution of u_j given all other $u_{k,k\neq t}$ is normally distributed with a mean that is a weighted average of its neighbors centered on μ_j, and with scaled variance which decreases as d_j increases. Using this form, we can derive the full conditional distribution of each variable in turn, leading to J many 1-dimensional distributions rather than a single J-dimensional distribution. In practice, this may be slower for small J, but when J is large, the computational savings can be substantial.

Regions

We consider the setting where areal units $j = 1, \ldots, J$ are each uniquely assigned to a region. These regions are ecologically meaningful aggregates of areal units that partition the spatial domain and capture regional relationships among the outcomes (after accounting for the covariates)

(Figure 4). We will use these regions to capture (space-varying) differences in the cross-covariance among outcomes' random effects, as well as regional pool-specific variance terms in the model. Specifically, let $r = 1, \dots, R$ index regions, where $R < J$, and let $\mathcal{J}_r$ be the set of areal unit indices associated with region r. Finally, let $n_r = |\mathcal{J}_r|$ be the number of areal units comprising region r.

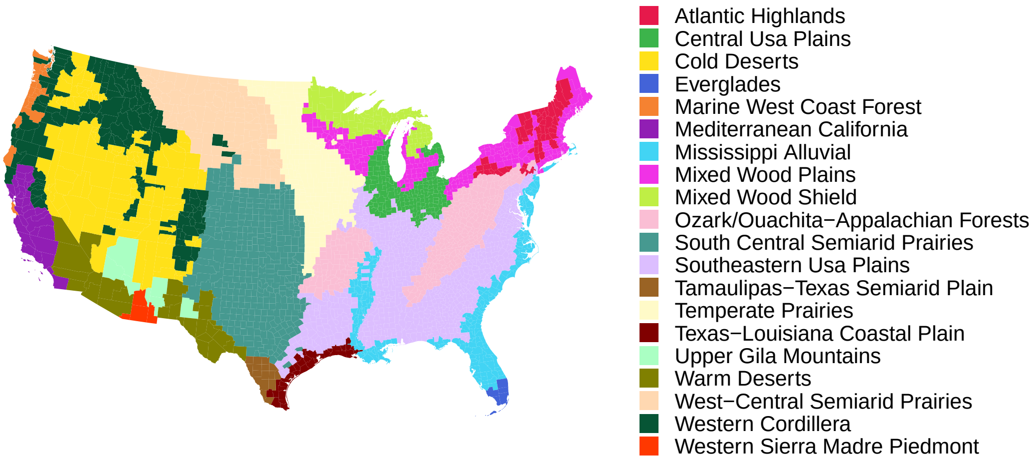


Figure 4: U.S. counties assigned to 20 ecoregions.

Model

For each areal unit $j = 1, \dots, J$ and time step $t = 1, \dots T$, our goal is to model $\boldsymbol{\mu}_{j,t}$ using sampling unit measurements $\mathbf{y}_{i,j,t}$, $i = 1, \dots, n_{j,t}$, as well as covariates which are available for all areal units and time steps. Here, we assume the same set of P many covariates are used for each of the M many outcomes. However, the proposed model can be easily extended to accommodate different covariates for each outcome variable. The regression coefficients for these covariates are fixed, while a subset of covariates are also endowed spatially-varying regression coefficients, where we assume no correlation between the spatial processes for regression coefficients between outcome variables. Finally, a dynamically-evolving spatially-varying intercept term is employed to capture correlations between outcome variables, and is modeled through a linear model of coregionalization.

For each areal unit $j = 1, \dots, J$ and time step $t = 1, \dots T$, the proposed model is

$$\mathbf{y}_{i,j,t} = \boldsymbol{\mu}_{j,t} + \boldsymbol{\delta}_{i,j,t}, \quad i = 1, \dots, n_{j,t} \tag{9}$$
$$\boldsymbol{\mu}_{j,t} = \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j + \boldsymbol{\varepsilon}_{j,t}, \tag{10}$$

where $j \in \mathcal{J}_r$ and

$$\boldsymbol{\delta}_{i,j,t} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_{\delta,j}), \quad \boldsymbol{\Sigma}_{\delta,j} \sim IW(\mathbf{H}_\delta, \nu_\delta) \tag{11}$$
$$\boldsymbol{\varepsilon}_{j,t} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad \boldsymbol{\Sigma}_{\varepsilon,r} \sim IW(\mathbf{H}_\varepsilon, \nu_\varepsilon). \tag{12}$$

In (10), we define $\mathbf{u}_{j,t} = (u_{1,j,t}, \dots, u_{M,j,t})^\top$ as a spatio-temporally varying intercept term, which we model dynamically as

$$\mathbf{u}_{j,t} = \mathbf{u}_{j,t-1} + \mathbf{w}_{j,t}, \quad \mathbf{u}_{\cdot,t} \equiv \mathbf{0}, \tag{13}$$

$$\mathbf{w}_{j,t} = \mathbf{A}_r \mathbf{v}_{j,t}, \qquad j \in \mathcal{J}_r \tag{14}$$

where $\mathbf{A}_r$ is a $M \times M$ lower triangular matrix with $\mathbf{\Sigma}_{u,r} = \mathbf{A}_r \mathbf{A}_r^\top \succcurlyeq 0$. We give $\mathbf{\Sigma}_{u,r}$ an Inverse Wishart prior of the form $IW(\mathbf{H}_u, \nu_u)$ for $r = 1, \ldots, R$.

Further, the elements of $\mathbf{v}_{j,t} = (v_{1,j,t}, \ldots, v_{M,j,t})^\top$ are modeled via CAR spatial structures for each $m = 1, \ldots, M$ and $t = 1, \ldots, T$ as

$$\begin{pmatrix} v_{m,1,t} \\ \vdots \\ v_{m,J,t} \end{pmatrix} \sim MVN\left(\mathbf{0}, \mathbf{Q}(\rho_{v,m})\right),$$

where $\mathbf{Q}(\rho_{v,m})$ is as defined in (7). We assign Uniform priors for each $\rho_{v,\cdot}$.

To directly specify the distribution of $\mathbf{u}_{j,t}$, we can plug (14) into (13) as

$$\mathbf{u}_{j,t} = \mathbf{u}_{j,t-1} + \mathbf{A}_r \mathbf{v}_{j,t} \quad \text{where } j \in \mathcal{J}_r. \tag{15}$$

Following (8), we can write the distribution of $\mathbf{v}_{j,t}$ conditional on all other $\mathbf{v}_{k,t,k\neq j}$ as $N(\boldsymbol{\mu}_{v,j,t}, \mathbf{\Sigma}_{v,j,t})$ where

$$\boldsymbol{\mu}_{v,j,t} = \begin{pmatrix} \rho_{v,1} \sum_{k \in \mathcal{N}_j} \frac{v_{1,k,t}}{d_j} \\ \rho_{v,2} \sum_{k \in \mathcal{N}_j} \frac{v_{2,k,t}}{d_j} \\ \vdots \\ \rho_{v,M} \sum_{k \in \mathcal{N}_j} \frac{v_{M,k,t}}{d_j} \end{pmatrix}, \quad \mathbf{\Sigma}_{v,j,t} = \begin{pmatrix} \frac{1}{d_j} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{d_j} \end{pmatrix}.$$

It is well known that for $\mathbf{Y} \sim MVN(\boldsymbol{\mu}, \mathbf{\Sigma}) \in \mathbb{R}^n$, multiplying by matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ yields another Multivariate Normal distribution of the form $\mathbf{AY} \sim MVN\left(\mathbf{A}\boldsymbol{\mu}, \mathbf{A\Sigma A}^\top\right)$. Therefore, we can write the distribution of $\mathbf{A}_r \mathbf{v}_{j,t}$ as $MVN\left(\mathbf{A}_r \boldsymbol{\mu}_{v,j,t}, \mathbf{A}_r \mathbf{\Sigma}_{v,j,t} \mathbf{A}_r^\top\right)$. And since $\mathbf{\Sigma}_{v,j,t} = \frac{1}{d_j}\mathbf{I}_M$ where $\mathbf{I}_M$ is the $M \times M$ identity matrix, we can write the distribution of $\mathbf{A}_r \mathbf{v}_{j,t}$ as $MVN\left(\mathbf{A}_r \boldsymbol{\mu}_{v,j,t}, \frac{1}{d_j}\mathbf{\Sigma}_{u,r}\right)$. Then from (15), we have

$$\mathbf{u}_{j,t} \sim MVN\left(\mathbf{u}_{j,t-1} + \mathbf{A}_r \boldsymbol{\mu}_{v,j,t}, \frac{1}{d_j}\mathbf{\Sigma}_{u,r}\right). \tag{16}$$

Writing this distribution in terms of $\mathbf{u}$ only, we first note that rearranging (15) gives

$$\mathbf{A}_r \mathbf{v}_{j,t} = \mathbf{u}_{j,t} - \mathbf{u}_{j,t-1},$$

and since $\mathbf{A}_r$ is invertible, we may then write

$$\mathbf{v}_{j,t} = \mathbf{A}_r^{-1}\mathbf{u}_{j,t} - \mathbf{A}_r^{-1}\mathbf{u}_{j,t-1}.$$

Rewriting $\boldsymbol{\mu}_{v,j,t}$ gives

$$\boldsymbol{\mu}_{v,j,t} = \begin{pmatrix} \rho_{v,1} \sum_{k \in \mathcal{N}_j} \frac{v_{1,k,t}}{d_j} \\ \rho_{v,2} \sum_{k \in \mathcal{N}_j} \frac{v_{2,k,t}}{d_j} \\ \vdots \\ \rho_{v,M} \sum_{k \in \mathcal{N}_j} \frac{v_{M,k,t}}{d_j} \end{pmatrix} = \sum_{k \in \mathcal{N}_j} \frac{1}{d_j} \text{diag}(\rho_{v,1}, \ldots, \rho_{v,M}) \mathbf{v}_{k,t}$$
$$= \sum_{k \in \mathcal{N}_j} \frac{1}{d_j} \text{diag}(\rho_{v,1}, \ldots, \rho_{v,M}) (\mathbf{A}_r^{-1} \mathbf{u}_{k,t} - \mathbf{A}_r^{-1} \mathbf{u}_{k,t-1}),$$

and hence,

$$\mathbf{A}_r \boldsymbol{\mu}_{v,j,t} = \sum_{k \in \mathcal{N}_j} \frac{1}{d_j} \text{diag}(\rho_{v,1}, \ldots, \rho_{v,M}) (\mathbf{u}_{k,t} - \mathbf{u}_{k,t-1})$$

which we denote as $\boldsymbol{\mu}_{u,j,t}$. Finally, we can write (16) more simply as

$$\mathbf{u}_{j,t} \sim MVN\left(\mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}, \frac{1}{d_j} \boldsymbol{\Sigma}_{u,r}\right). \tag{17}$$

We include two regression components in (10); a fixed effects regression term $\mathbf{X}_{j,t}\boldsymbol{\beta}$ and a spatially-varying regression term $\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j$. For the fixed term, the design matrix $\mathbf{X}_{j,t} \in \mathbb{R}^{M \times (M \times P)}$ is given as

$$\mathbf{X}_{j,t} = \mathbf{I}_M \otimes \mathbf{x}_{j,t},$$
$$\mathbf{x}_{j,t} = (x_{1,j,t}, \ldots, x_{P,j,t}),$$

where $\otimes$ indicates the Kronecker product and $\mathbf{x}_{j,t}$ is the vector of P many covariates observed at areal unit j at time t. Accordingly, $\boldsymbol{\beta} \in \mathbb{R}^{(M \times P)}$ may be written as

$$\boldsymbol{\beta} = (\beta_{1,1}, \ldots, \beta_{1,P}, \ldots, \beta_{M,1}, \ldots, \beta_{M,P})^\top,$$

so we have

$$\mathbf{X}_{j,t}\boldsymbol{\beta} = \begin{pmatrix} x_{1,j,t} & \cdots & x_{P,j,t} & & \cdots & & 0 \\ & & & \ddots & & & \\ 0 & & \cdots & & x_{1,j,t} & \cdots & x_{P,j,t} \end{pmatrix} \begin{pmatrix} \beta_{1,1} \\ \vdots \\ \beta_{1,P} \\ \vdots \\ \beta_{M,1} \\ \vdots \\ \beta_{M,P} \end{pmatrix},$$

where we assign $\boldsymbol{\beta} \sim MVN(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$. For the spatially-varying term $\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j$, the design matrix $\tilde{\mathbf{X}}_{j,t} \in \mathbb{R}^{M \times (M \times Q)}$ is composed of a subset of Q many predictor variables from $\mathbf{X}_{j,t}$ ($Q \leq P$). In this way, we allow the effect for some (or all when $Q = P$) predictors to vary spatially across the domain. Specifically, we have

$$
\begin{aligned}
\tilde{\mathbf{X}}_{j,t} &= \mathbf{I}_M \otimes \tilde{\mathbf{x}}_{j,t}, \\
\tilde{\mathbf{x}}_{j,t} &= (\tilde{x}_{1,j,t}, \dots, \tilde{x}_{Q,j,t}),
\end{aligned}
$$

and $\boldsymbol{\eta}_j \in \mathbb{R}^{(M \times Q)}$ may be written as

$$
\boldsymbol{\eta}_j = (\eta_{1,1,j}, \dots, \eta_{1,Q,j}, \dots, \eta_{M,1,j}, \dots, \eta_{M,Q,j})^\top,
$$

so

$$
\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j = \begin{pmatrix} \tilde{x}_{1,j,t} & \cdots & \tilde{x}_{Q,j,t} & & \cdots & & 0 \\ & & & \ddots & & & \\ 0 & & \cdots & & \tilde{x}_{1,j,t} & \cdots & \tilde{x}_{Q,j,t} \end{pmatrix} \begin{pmatrix} \eta_{1,1,j} \\ \vdots \\ \eta_{1,Q,j} \\ \vdots \\ \eta_{M,1,j} \\ \vdots \\ \eta_{M,Q,j} \end{pmatrix}.
$$

To model the elements of $\boldsymbol{\eta}_j, j = 1, \dots, J$, we again employ CAR spatial structures, where for response m and predictor q, we have

$$
\boldsymbol{\eta}_{m,q} = \begin{pmatrix} \eta_{m,q,1} \\ \vdots \\ \eta_{m,q,J} \end{pmatrix} \sim MVN\left(\mathbf{0}, \tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})\right).
$$

Following (8), we can then write the distribution of each $\eta_{m,q,j}$ conditional all other $\eta_{m,q,k}, k \neq j$ as

$$
\eta_{m,q,j} | \eta_{m,q,k,k\neq j} \sim N\left(\rho_{\eta,m,q} \sum_{k \in \mathcal{N}_j} \frac{\eta_{m,q,k}}{d_j}, \frac{\tau^2_{\eta,m,q}}{d_j}\right).
$$

We then have that $\boldsymbol{\eta}_j$ conditional on all other $\boldsymbol{\eta}_k, k \neq j$ is distributed as

$$
\boldsymbol{\eta}_j | \boldsymbol{\eta}_{k,k\neq j} \sim MVN\left(\boldsymbol{\mu}_{\eta,j}, \boldsymbol{\Sigma}_{\eta,j}\right),
$$

where

$$
\begin{aligned}
\boldsymbol{\mu}_{\eta,j} &= \frac{1}{d_j}\text{diag}(\rho_{\eta,1,1}, \dots, \rho_{\eta,1,Q}, \dots, \rho_{\eta,M,1}, \dots, \rho_{\eta,M,Q}) \sum_{k \in \mathcal{N}_j} \boldsymbol{\eta}_k, \\
\boldsymbol{\Sigma}_{\eta,j} &= \frac{1}{d_j}\text{diag}(\tau^2_{\eta,1,1}, \dots, \tau^2_{\eta,1,Q}, \dots, \tau^2_{\eta,M,1}, \dots, \tau^2_{\eta,M,Q}).
\end{aligned}
$$

Finally, we assign Inverse Gamma priors to each $\tau^2_{\eta,\cdot,\cdot}$ and Uniform priors to each $\rho_{\eta,\cdot,\cdot}$.

Posterior

We know that the posterior distribution in Bayesian inference is proportional to the likelihood (data distribution) times the prior. Here, our likelihood is the distribution of the data $\mathbf{y}_{i,j,t}, i = 1, \ldots, n_{j,t}, \quad j = 1, \ldots, J, \quad t = 1, \ldots, T$. Since the distribution of the residuals is Multivariate Normal, we will have a Multivariate Normal likelihood, written as

$$\prod_{t=1}^{T} \prod_{j=1}^{J} \prod_{i=1}^{n_{j,t}} MVN(\mathbf{y}_{i,j,t} | \boldsymbol{\mu}_{j,t}, \boldsymbol{\Sigma}_{\delta,j}). \tag{18}$$

We will also have priors (some of which are induced) for each parameter for which we seek inference. Specifically, we have prior distributions of the form

$$\begin{aligned}
&\prod_{t=1}^{T} \prod_{j=1}^{J} MVN\left(\boldsymbol{\mu}_{j,t} | \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}\right), \quad j \in \mathcal{J}_r \\
&\times \prod_{t=1}^{T} \prod_{j=1}^{J} MVN\left(\mathbf{u}_{j,t} | \mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right) \times MVN(\boldsymbol{\beta} \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \\
&\times \prod_{j=1}^{J} MVN\left(\boldsymbol{\eta}_j | \boldsymbol{\mu}_{\eta,j}, \boldsymbol{\Sigma}_{\eta,j}\right) \times \prod_{m=1}^{M} \prod_{q=1}^{Q} IG(\tau^2_{\eta,m,q} | c_\eta, d_\eta) \\
&\times \prod_{j=1}^{J} IW(\boldsymbol{\Sigma}_{\delta,j} | \mathbf{H}_\delta, \nu_\delta) \times \prod_{r=1}^{R} IW\left(\boldsymbol{\Sigma}_{u,r} | \mathbf{H}_u, \nu_u\right) \times \prod_{r=1}^{R} IW\left(\boldsymbol{\Sigma}_{\varepsilon,r} | \mathbf{H}_\varepsilon, \nu_\varepsilon\right) \\
&\times \prod_{m=1}^{M} U(\rho_{v,m} | a_v, b_v) \times \prod_{m=1}^{M} \prod_{q=1}^{Q} U(\rho_{\eta,m,q} | a_\eta, b_\eta).
\end{aligned} \tag{19}$$

Given (18) and (19), their product will be proportional to the posterior distribution, and moving forward, we will work with this proportional posterior.

Working it out

We will now go through each of the parameters in model and derive their full conditional distributions. We will use different colors to help keep track of different prior and likelihood terms as we work through the derivations. This will help us see where each component of the full conditional distribution is coming from. We have specifically chosen prior forms to induce conjugacy, which allows us to more easily determine the form of each full conditional distribution. In the cases of Normal-Normal conjugacy, we will use the stretegy of **incompleting the square**, which is outlined in Section 4.2 of James Clark's book 'Models for Ecological Data', and is described here briefly.

Incompleting the square

Following (2), we know that the kernel of a Multivariate Normal distribution is of the form

$$p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$

where $\propto$ indicates "proportional to". This structure gives us some insight into the placement of the parameters in the pdf. To see this further, we can focus on the terms inside the exponential, ignoring the $-\frac{1}{2}$ coefficient. Expanding this term gives us

$$(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = \mathbf{x}^\top \boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{20}$$

Writing the mean as a function of the covariance matrix, we can define

$$\begin{aligned}\boldsymbol{\mu} &= \mathbf{V}\mathbf{v}\\ \boldsymbol{\Sigma} &= \mathbf{V},\end{aligned}$$

where $\mathbf{V}$ is an $n \times n$ matrix and $\mathbf{v}$ is a length n vector.

This allows us to rewrite (20) as

$$\begin{aligned}&\mathbf{x}^\top \mathbf{V}^{-1}\mathbf{x} - \mathbf{x}^\top \mathbf{V}^{-1}\mathbf{V}\mathbf{v} - \mathbf{v}^\top \mathbf{V}^\top \mathbf{V}^{-1}\mathbf{x} + \mathbf{v}^\top \mathbf{V}^\top \mathbf{V}^{-1}\mathbf{V}\mathbf{v}\\ &= \mathbf{x}^\top \mathbf{V}^{-1}\mathbf{x} - \mathbf{x}^\top \mathbf{v} - \mathbf{v}^\top \mathbf{x} + \mathbf{v}^\top \mathbf{V}\mathbf{v}\\ &= \mathbf{x}^\top \mathbf{V}^{-1}\mathbf{x} - \mathbf{x}^\top \mathbf{v} \cdots\end{aligned}$$

This is super helpful, because when we are multiplying together normal priors and likelihoods, we can combine like terms and look for $\mathbf{V}^{-1}$ and $\mathbf{v}$ to get the mean and variance of the full conditional Multivariate Normal. They will always be in the same place! $\mathbf{V}^{-1}$ will be sandwiched between the variable of interest (i.e., the $\mathbf{x}^\top \mathbf{V}^{-1}\mathbf{x}$ term), and $\mathbf{v}$ will be attached to the negative transpose of the variable of interest (i.e., $-\mathbf{x}^\top \mathbf{v}$) (in this example, $\mathbf{x}$ is the variable of interest). Once we identify $\mathbf{V}^{-1}$ and $\mathbf{v}$, we can sample from the full conditional distribution as $MVN(\mathbf{V}\mathbf{v}, \mathbf{V})$. The case for the univariate Normal distribution follows in the same way.

Update missing elements of $\mathbf{y}_{i,j,t}, i = 1, \ldots, n_{j,t}, j = 1, \ldots, J, t = 1, \ldots, T$

From (18), we know that $\mathbf{y}_{i,j,t} \sim MVN(\boldsymbol{\mu}_{j,t}, \boldsymbol{\Sigma}_{\delta,j})$. If any elements of $\mathbf{y}_{i,j,t}$ are `NA`, we update them from their posterior predictive distribution. Suppose $K < M$ many elements of $\mathbf{y}_{i,j,t}$ are missing ($=$ `NA`). Rewrite terms as

$$\mathbf{y}_{i,j,t} \to \begin{pmatrix}\mathbf{y}_{1,i,j,t}\\ \mathbf{y}_{2,i,j,t}\end{pmatrix}, \quad \boldsymbol{\mu}_{j,t} \to \begin{pmatrix}\boldsymbol{\mu}_{1,j,t}\\ \boldsymbol{\mu}_{2,j,t}\end{pmatrix}, \quad \boldsymbol{\Sigma}_{\delta,j} \to \begin{pmatrix}\boldsymbol{\Sigma}_{1,\delta,j} & \boldsymbol{\Sigma}_{2,\delta,j}\\ \boldsymbol{\Sigma}_{3,\delta,j} & \boldsymbol{\Sigma}_{4,\delta,j}\end{pmatrix},$$

where $\mathbf{y}_{1,i,j,t}$ are the K many elements of $\mathbf{y}_{i,j,t}$ that are *missing*, and $\mathbf{y}_{2,i,j,t}$ are the remaining $M-K$ many elements of $\mathbf{y}_{i,j,t}$ that are measured. Then, $\boldsymbol{\mu}_{1,j,t}$ and $\boldsymbol{\mu}_{2,j,t}$ are composed of elements of $\boldsymbol{\mu}_{j,t}$ that are associated with $(\mathbf{y}_{1,i,j,t}, \mathbf{y}_{2,i,j,t})^\top$. Finally, $\boldsymbol{\Sigma}_{\delta,j}$ is reordered into four submatrices corresponding to the stacked ordering of $(\mathbf{y}_{1,i,j,t}, \mathbf{y}_{2,i,j,t})^\top$ and $(\boldsymbol{\mu}_{1,j,t}, \boldsymbol{\mu}_{2,j,t})^\top$. We then know

$$\mathbf{y}_{1,i,j,t} | \mathbf{y}_{2,i,j,t} \sim MVN\left(\tilde{\boldsymbol{\mu}}_{j,t}, \tilde{\boldsymbol{\Sigma}}_{\delta,j}\right),$$

where

$$\begin{aligned}\tilde{\boldsymbol{\mu}}_{j,t} &= \boldsymbol{\mu}_{1,j,t} + \boldsymbol{\Sigma}_{2,\delta,j}\boldsymbol{\Sigma}_{4,\delta,j}^{-1}(\mathbf{y}_{2,i,j,t} - \boldsymbol{\mu}_{2,j,t}),\\ \tilde{\boldsymbol{\Sigma}}_{\delta,j} &= \boldsymbol{\Sigma}_{1,\delta,j} - \boldsymbol{\Sigma}_{2,\delta,j}\boldsymbol{\Sigma}_{4,\delta,j}^{-1}\boldsymbol{\Sigma}_{3,\delta,j},\end{aligned}$$

and we can sample these missing values from this Multivariate Normal distribution.

Update latent mean $\boldsymbol{\mu}_{j,t}, j = 1, \ldots, J, t = 1, \ldots T$

Looking at (18) and (19), we see that $\boldsymbol{\mu}_{j,t}$ appears in two terms, namely in the likelihood of $\mathbf{y}_{i,j,t}$ and in its own prior distribution. The product of these terms looks like

$$\prod_{i=1}^{n_{j,t}} MVN(\mathbf{y}_{i,j,t}|\boldsymbol{\mu}_{j,t}, \boldsymbol{\Sigma}_{\delta,j}) \times MVN(\boldsymbol{\mu}_{j,t}|\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad j \in \mathcal{J}_r,$$

which is the product of two Multivariate Normal random variables, hence their product will also be distributed as Multivariate Normal. To identify the parameters of this new Multivariate Normal distribution, we can simply incomplete the square. Specifically, we have

$$\begin{aligned}
&\prod_{i=1}^{n_{j,t}} MVN(\mathbf{y}_{i,j,t}|\boldsymbol{\mu}_{j,t}, \boldsymbol{\Sigma}_{\delta,j}) \times MVN(\boldsymbol{\mu}_{j,t}|\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad j \in \mathcal{J}_r,\\
&\propto \prod_{i=1}^{n_{j,t}} \exp\left((\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}(\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})\right)\\
&\quad \times \exp\left((\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right)\\
&= \exp\Bigg(\sum_{i=1}^{n_{j,t}} \left((\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}(\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})\right)\\
&\qquad + (\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\Bigg)\\
&= \exp\Bigg(\sum_{i=1}^{n_{j,t}} \left(\mathbf{y}_{i,j,t}^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}\mathbf{y}_{i,j,t} - \mathbf{y}_{i,j,t}^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}\boldsymbol{\mu}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}\mathbf{y}_{i,j,t} + \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\delta,j}^{-1}\boldsymbol{\mu}_{j,t} + \cdots\right)\\
&\qquad + \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) - \cdots\Bigg)\\
&= \exp\left(\boldsymbol{\mu}_{j,t}^\top \left(n_{j,t}\boldsymbol{\Sigma}_{\delta,j}^{-1} + \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\right) - \boldsymbol{\mu}_{j,t}^\top \left(\sum_{i=1}^{n_{j,t}} \boldsymbol{\Sigma}_{\delta,j}^{-1}\mathbf{y}_{i,j,t} + \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right) + \cdots\right)
\end{aligned}$$

where $j \in \mathcal{J}_r$. Then we can identify

$$\begin{aligned}
\mathbf{V}^{-1} &= n_{j,t}\boldsymbol{\Sigma}_{\delta,j}^{-1} + \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\\
\mathbf{v} &= \sum_{i=1}^{n_{j,t}} \boldsymbol{\Sigma}_{\delta,j}^{-1}\mathbf{y}_{i,j,t} + \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j),
\end{aligned}$$

and we sample $\boldsymbol{\mu}_{j,t}$ from $MVN(\mathbf{Vv}, \mathbf{V})$.

Update $\boldsymbol{\beta}$

Looking at (18) and (19), we see that $\boldsymbol{\beta}$ appears in two terms, namely in the prior for $\boldsymbol{\mu}_{j,t}, j = 1, \ldots, J, t = 1, \ldots T$ and in its own prior distribution. The product of these terms is

$$\prod_{t=1}^{T}\prod_{j=1}^{J} MVN(\boldsymbol{\mu}_{j,t} \mid \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}) \times MVN(\boldsymbol{\beta} \mid \boldsymbol{\mu}_{\beta}, \boldsymbol{\Sigma}_{\beta}), \quad j \in \mathcal{J}_r,$$

which is the product of Multivariate Normal distributions. Hence, the posterior will also be distributed as Multivariate Normal. We can identify the parameters of this Multivariate Normal distribution by incompleting the square. We have

$$
\begin{aligned}
&\prod_{t=1}^{T}\prod_{j=1}^{J} MVN(\boldsymbol{\mu}_{j,t} \mid \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}) \times MVN(\boldsymbol{\beta} \mid \boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) \\
&\propto \prod_{t=1}^{T}\prod_{j=1}^{J} \exp\left((\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right) \\
&\quad \times \exp\left((\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^\top \boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)\right) \\
&= \exp\left(\sum_{t=1}^{T}\sum_{j=1}^{J}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right. \\
&\qquad \left. + (\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)^\top \boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)\right) \\
&= \exp\left(\sum_{t=1}^{T}\sum_{j=1}^{J}\left(\boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\right.\right. \\
&\qquad -\mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad -\boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad \left. -\boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\right) \\
&\qquad \left. + \boldsymbol{\beta}^\top \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}^\top \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta - \boldsymbol{\mu}_\beta^\top \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\beta} + \boldsymbol{\mu}_\beta^\top \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta\right) \\
&\propto \exp\left[\boldsymbol{\beta}^\top\left(\sum_{t=1}^{T}\sum_{j=1}^{J}(\mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}) + \boldsymbol{\Sigma}_\beta^{-1}\right)\boldsymbol{\beta}\right. \\
&\qquad \left. -\boldsymbol{\beta}^\top\left(\sum_{t=1}^{T}\sum_{j=1}^{J}(\mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} - \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} - \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) + \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta\right) - \cdots\right],
\end{aligned}
$$

where r is such that $j \in \mathcal{J}_r$. So, we can identify

$$
\begin{aligned}
\mathbf{V}^{-1} &= \sum_{t=1}^{T}\sum_{j=1}^{J}(\mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}) + \boldsymbol{\Sigma}_\beta^{-1} \\
\mathbf{v} &= \sum_{t=1}^{T}\sum_{j=1}^{J}(\mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} - \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} - \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) + \boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta,
\end{aligned}
$$

and we can update $\boldsymbol{\beta}$ from its full conditional as $MVN(\mathbf{Vv}, \mathbf{V})$.

Update $\mathbf{u}_{j,t}, j = 1, \ldots, J, t = 1, \ldots, T-1$

Looking at (18) and (19), we see that $\mathbf{u}_{j,t}$ appears in three terms, namely in the prior for $\boldsymbol{\mu}_{j,t}$, its own prior distribution, and in the prior for $\mathbf{u}_{j,t+1}$. The product of these terms looks like

$$MVN(\boldsymbol{\mu}_{j,t} \mid \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\beta + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad j \in \mathcal{J}_r,$$
$$\times MVN\left(\mathbf{u}_{j,t}|\mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right) \times MVN\left(\mathbf{u}_{j,t+1}|\mathbf{u}_{j,t} + \boldsymbol{\mu}_{u,j,t+1}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)$$

which is the product of Multivariate Normal distributions. Hence, the posterior will also be distributed as Multivariate Normal. We can identify the parameters of this Multivariate Normal distribution by incompleting the square. We have

$$
\begin{aligned}
&MVN(\boldsymbol{\mu}_{j,t} \mid \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad j \in \mathcal{J}_r, \\
&\quad \times MVN\left(\mathbf{u}_{j,t}|\mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right) \times MVN\left(\mathbf{u}_{j,t+1}|\mathbf{u}_{j,t} + \boldsymbol{\mu}_{u,j,t+1}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right) \\
&\propto \exp\left((\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right) \\
&\quad \times \exp\left((\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})\right) \\
&\quad \times \exp\left((\mathbf{u}_{j,t+1} - \mathbf{u}_{j,t} - \boldsymbol{\mu}_{u,j,t+1})^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t+1} - \mathbf{u}_{j,t} - \boldsymbol{\mu}_{u,j,t+1})\right) \\
&= \exp\Bigg((\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) \\
&\qquad + (\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t}) \\
&\qquad + (\mathbf{u}_{j,t+1} - \mathbf{u}_{j,t} - \boldsymbol{\mu}_{u,j,t+1})^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t+1} - \mathbf{u}_{j,t} - \boldsymbol{\mu}_{u,j,t+1})\Bigg) \\
&= \exp\Bigg(\boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} - \boldsymbol{\mu}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad - \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \mathbf{u}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad - \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \mathbf{X}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad - \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\boldsymbol{\beta} + \boldsymbol{\eta}_j^\top \tilde{\mathbf{X}}_{j,t}^\top \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j \\
&\qquad + \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} - \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t-1} - \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t} \\
&\qquad - \mathbf{u}_{j,t-1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} + \mathbf{u}_{j,t-1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t-1} + \mathbf{u}_{j,t-1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t} \\
&\qquad - \boldsymbol{\mu}_{u,j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} + \boldsymbol{\mu}_{u,j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t} \\
&\qquad + \mathbf{u}_{j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t+1} - \mathbf{u}_{j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} - \mathbf{u}_{j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t+1} \\
&\qquad - \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t+1} + \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} + \mathbf{u}_{j,t}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t+1} \\
&\qquad - \boldsymbol{\mu}_{u,j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t+1} + \boldsymbol{\mu}_{u,j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \mathbf{u}_{j,t} + \boldsymbol{\mu}_{u,j,t+1}^\top \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \boldsymbol{\mu}_{u,j,t+1}\Bigg) \\
&\propto \exp\Bigg[\mathbf{u}_{j,t}^\top \left(\boldsymbol{\Sigma}_{\varepsilon,r}^{-1} + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1}\right)\mathbf{u}_{j,t} \\
&- \mathbf{u}_{j,t}^\top \left(\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}) + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} (\mathbf{u}_{j,t+1} - \boldsymbol{\mu}_{u,j,t+1})\right) - \cdots\Bigg]
\end{aligned}
$$

so we can identify

$$
\begin{aligned}
\mathbf{V}^{-1} &= \boldsymbol{\Sigma}_{\varepsilon,r}^{-1} + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \\
\mathbf{v} &= \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j) + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1}(\mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}) + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1}(\mathbf{u}_{j,t+1} - \boldsymbol{\mu}_{u,j,t+1}),
\end{aligned}
$$

and we can then update $\mathbf{u}_{j,t}$ from its full conditional as $MVN(\mathbf{Vv}, \mathbf{V})$.

Update $\mathbf{u}_{j,t}, j = 1, \ldots, J, t = T$

Looking at (18) and (19), we see that $\mathbf{u}_{j,T}$ appears in two terms, namely in the prior for $\boldsymbol{\mu}_{j,T}$ and its own prior distribution. The product of these terms looks like

$$
\begin{aligned}
&MVN(\boldsymbol{\mu}_{j,t} \mid \mathbf{u}_{j,T} + \mathbf{X}_{j,T}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,T}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}), \quad j \in \mathcal{J}_r, \\
&\times MVN\left(\mathbf{u}_{j,T} | \mathbf{u}_{j,T-1} + \boldsymbol{\mu}_{u,j,T}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)
\end{aligned}
$$

which is the product of Multivariate Normal distributions. Hence, the posterior will also be distributed as Multivariate Normal. We can identify the parameters of this Multivariate Normal distribution by incompleting the square. We can identify

$$
\begin{aligned}
\mathbf{V}^{-1} &= \boldsymbol{\Sigma}_{\varepsilon,r}^{-1} + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1} \\
\mathbf{v} &= \boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,T} - \mathbf{X}_{j,T}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,T}\boldsymbol{\eta}_j) + \left(\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)^{-1}(\mathbf{u}_{j,T-1} + \boldsymbol{\mu}_{u,j,T}),
\end{aligned}
$$

and we can then update $\mathbf{u}_{j,T}$ from its full conditional as $MVN(\mathbf{Vv}, \mathbf{V})$.

Update $\boldsymbol{\eta}_j, j = 1, \ldots, J$

Looking at (18) and (19), we see that $\boldsymbol{\eta}_j$ appears in two terms, namely in the prior for $\boldsymbol{\mu}_{j,t}, t = 1, \ldots, T$ and in its own prior distribution. The product of these terms looks like

$$
\prod_{t=1}^{T} MVN\left(\boldsymbol{\mu}_{j,t} | \mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}\right) \times MVN\left(\boldsymbol{\eta}_j | \boldsymbol{\mu}_{\eta,j}, \boldsymbol{\Sigma}_{\eta,j}\right), \quad j \in \mathcal{J}_r
$$

which is the product of Multivariate Normal distributions. Hence, the posterior will also be distributed as Multivariate Normal. We can identify the parameters of this Multivariate Normal distribution by incompleting the square. We have

$$
\begin{aligned}
&\prod_{t=1}^{T} MVN\left(\boldsymbol{\mu}_{j,t}|\mathbf{u}_{j,t}+\mathbf{X}_{j,t}\beta+\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j,\boldsymbol{\Sigma}_{\varepsilon,r}\right)\times MVN\left(\boldsymbol{\eta}_j|\boldsymbol{\mu}_{\eta,j},\boldsymbol{\Sigma}_{\eta,j}\right),\quad j\in\mathcal{J}_r\\
&\propto\prod_{t=1}^{T}\exp\left((\boldsymbol{\mu}_{j,t}-\mathbf{u}_{j,t}-\mathbf{X}_{j,t}\beta-\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t}-\mathbf{u}_{j,t}-\mathbf{X}_{j,t}\beta-\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right)\\
&\qquad\times\exp\left((\boldsymbol{\eta}_j-\boldsymbol{\mu}_{\eta,j})^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}(\boldsymbol{\eta}_j-\boldsymbol{\mu}_{\eta,j})\right)\\
&=\exp\left(\sum_{t=1}^{T}(\boldsymbol{\mu}_{j,t}-\mathbf{u}_{j,t}-\mathbf{X}_{j,t}\beta-\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t}-\mathbf{u}_{j,t}-\mathbf{X}_{j,t}\beta-\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right.\\
&\qquad\left.+(\boldsymbol{\eta}_j-\boldsymbol{\mu}_{\eta,j})^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}(\boldsymbol{\eta}_j-\boldsymbol{\mu}_{\eta,j})\right)\\
&=\exp\left(\sum_{t=1}^{T}\left(\boldsymbol{\mu}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}-\boldsymbol{\mu}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}-\boldsymbol{\mu}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta-\boldsymbol{\mu}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\right.\right.\\
&\qquad-\mathbf{u}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}+\mathbf{u}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}+\mathbf{u}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta+\mathbf{u}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\\
&\qquad-\beta^\top\mathbf{X}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}+\beta^\top\mathbf{X}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}+\beta^\top\mathbf{X}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta+\beta^\top\mathbf{X}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\\
&\qquad\left.-\boldsymbol{\eta}_j^\top\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}+\boldsymbol{\eta}_j^\top\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}+\boldsymbol{\eta}_j^\top\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta+\boldsymbol{\eta}_j^\top\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j\right)\\
&\qquad\left.+\boldsymbol{\eta}_j^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\eta}_j-\boldsymbol{\eta}_j^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\mu}_{\eta,j}-\boldsymbol{\mu}_{\eta,j}^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\eta}_j+\boldsymbol{\mu}_{\eta,j}^\top\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\mu}_{\eta,j}\right)\\
&\propto\exp\left[\boldsymbol{\eta}_j^\top\left(\sum_{t=1}^{T}\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}+\boldsymbol{\Sigma}_{\eta,j}^{-1}\right)\boldsymbol{\eta}_j\right.\\
&\qquad\left.-\boldsymbol{\eta}_{m,j}^\top\left(\sum_{t=1}^{T}\left(\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}-\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}-\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta\right)+\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\mu}_{\eta,j}\right)-\cdots\right],
\end{aligned}
$$

so we can identify

$$
\begin{aligned}
\mathbf{V}^{-1}&=\sum_{t=1}^{T}\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\tilde{\mathbf{X}}_{j,t}+\boldsymbol{\Sigma}_{\eta,j}^{-1}\\
\mathbf{v}&=\sum_{t=1}^{T}\left(\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\boldsymbol{\mu}_{j,t}-\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{u}_{j,t}-\tilde{\mathbf{X}}_{j,t}^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\mathbf{X}_{j,t}\beta\right)+\boldsymbol{\Sigma}_{\eta,j}^{-1}\boldsymbol{\mu}_{\eta,j},
\end{aligned}
$$

and we can then update $\boldsymbol{\eta}_j$ from its full conditional as $MVN(\mathbf{V}\mathbf{v},\mathbf{V})$.

Update $\boldsymbol{\Sigma}_{u,r}, r=1,\ldots,R$

Looking at (18) and (19), we see that $\boldsymbol{\Sigma}_{u,r}$ appears in two terms, namely in the prior for $\mathbf{u}_{j,t}$, for $j\in\mathcal{J}_r$ and $t=1,\ldots,T$ and in its own prior distribution. The product of these terms looks like

$$
\prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r}MVN\left(\mathbf{u}_{j,t}|\mathbf{u}_{j,t-1}+\boldsymbol{\mu}_{u,j,t},\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)\times IW\left(\boldsymbol{\Sigma}_{u,r}|\mathbf{H}_u,\nu_u\right),
$$

which is the product of Multivariate Normal distributions with an Inverse Wishart distribution. Since the Inverse Wishart is a conjugate prior for the covariance matrix of a Multivariate Normal, we know that the posterior will also follow an Inverse Wishart distribution (3). Specifcally we have

$$
\begin{aligned}
&\prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r} MVN\left(\mathbf{u}_{j,t}|\mathbf{u}_{j,t-1}+\boldsymbol{\mu}_{u,j,t},\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}\right)\times IW\left(\boldsymbol{\Sigma}_{u,r}|\mathbf{H}_u,\nu_u\right)\\
&=\prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r}(2\pi)^{-M/2}|\frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}|^{-1/2}\exp\left(-\frac{d_j}{2}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})\right)\\
&\quad\times\frac{|\mathbf{H}_u|^{\nu_u/2}}{2^{\nu_u M/2}\Gamma_M\left(\frac{\nu_u}{2}\right)}|\boldsymbol{\Sigma}_{u,r}|^{-(\nu_u+M+1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{H}_u\boldsymbol{\Sigma}_{u,r}^{-1}\right)\right)\\
&\propto|\boldsymbol{\Sigma}_{u,r}|^{-(T\times n_r+\nu_u+M+1)/2}\\
&\quad\times\exp\left(-\frac{1}{2}\left[\underbrace{\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})}_{I}+\mathrm{tr}(\mathbf{H}_u\boldsymbol{\Sigma}_{u,r}^{-1})\right]\right).
\end{aligned}
$$

We will now use a trick to rearrange the expression into the form of the IW distribution. Since I is evaluated as a scalar (sum of quadratic forms), we can take the *trace* of it without changing the expression (the trace of a scalar is the scalar itself). So we can write

$$
\begin{aligned}
I&=\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})\\
&=\mathrm{tr}\left(\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})\right).
\end{aligned}
$$

We also know that the trace of a sum is equal to the sum of a trace, so we can move the trace operator inside the sum.

$$
=\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}\mathrm{tr}\left(d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})\right)
$$

Another useful property of the trace operator is the following

$$
\mathrm{tr}(\mathbf{ABC})=\mathrm{tr}(\mathbf{BCA})=\mathrm{tr}(\mathbf{CAB})
$$

Applying this yields

$$
=\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}\mathrm{tr}\left(d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}\right)
$$

and plugging I back in gives

$$
\begin{aligned}
&|\boldsymbol{\Sigma}_{u,r}|^{-(T\times n_r+\nu_u+M+1)/2} \\
&\quad\times \exp\left(-\frac{1}{2}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}\mathrm{tr}\left(d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top\boldsymbol{\Sigma}_{u,r}^{-1}\right)+\mathrm{tr}(\mathbf{H}_u\boldsymbol{\Sigma}_{u,r}^{-1})\right]\right) \\
&=|\boldsymbol{\Sigma}_{u,r}|^{-(T\times n_r+\nu_u+M+1)/2} \\
&\quad\times \exp\left(-\frac{1}{2}\mathrm{tr}\left[\left(\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top+\mathbf{H}_u\right)\boldsymbol{\Sigma}_{u,r}^{-1}\right]\right)
\end{aligned}
$$

which is in the same form as (3), with parameter values

$$
\begin{aligned}
\tilde{\nu}_{u,r}&=\nu_u+T\times n_r \\
\tilde{\mathbf{H}}_{u,r}&=\mathbf{H}_u+\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}d_j(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})(\mathbf{u}_{j,t}-\mathbf{u}_{j,t-1}-\boldsymbol{\mu}_{u,j,t})^\top
\end{aligned}
$$

and we can update $\boldsymbol{\Sigma}_{u,r}$ from its full conditional distribution as $IW(\tilde{\mathbf{H}}_{u,r},\tilde{\nu}_{u,r})$.

Update $\boldsymbol{\Sigma}_{\delta,j}, j=1,\dots,J$

Looking at (18) and (19), we see that $\boldsymbol{\Sigma}_{\delta,j}$ appears in two terms, namely in the likelihood for $\mathbf{y}_{i,j,t}$, for $i=1,\dots,n_{j,t}, t=1,\dots,T$ and in its own prior distribution. The product of these terms looks like

$$
\prod_{t=1}^{T}\prod_{i=1}^{n_{j,t}}MVN\left(\mathbf{y}_{i,j,t}|\boldsymbol{\mu}_{j,t},\boldsymbol{\Sigma}_{\delta,j}\right)\times IW\left(\boldsymbol{\Sigma}_{\delta,j}|\mathbf{H}_\delta,\nu_\delta\right),
$$

which is the product of Multivariate Normal distributions with an Inverse Wishart distribution. Hence, the posterior will also follow an Inverse Wishart distribution (3). Using the same trick as before, we have

$$
\begin{aligned}
&\prod_{t=1}^{T}\prod_{i=1}^{n_{j,t}}MVN\left(\mathbf{y}_{i,j,t}|\boldsymbol{\mu}_{j,t},\boldsymbol{\Sigma}_{\delta,j}\right)\times IW\left(\boldsymbol{\Sigma}_{\delta,j}|\mathbf{H}_\delta,\nu_\delta\right) \\
&=\prod_{t=1}^{T}\prod_{i=1}^{n_{j,t}}(2\pi)^{-M/2}|\boldsymbol{\Sigma}_{\delta,j}|^{-1/2}\exp\left(-\frac{1}{2}(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})^\top\boldsymbol{\Sigma}_{\delta,j}^{-1}(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})\right) \\
&\quad\times\frac{|\mathbf{H}_\delta|^{\nu_\delta/2}}{2^{\nu_\delta M/2}\Gamma_M\left(\frac{\nu_\delta}{2}\right)}|\boldsymbol{\Sigma}_{\delta,j}|^{-(\nu_\delta+M+1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{H}_\delta\boldsymbol{\Sigma}_{\delta,j}^{-1}\right)\right) \\
&\propto|\boldsymbol{\Sigma}_{\delta,j}|^{-(\sum_{t=1}^{T}n_{j,t}+\nu_\delta+M+1)/2} \\
&\quad\times\exp\left(-\frac{1}{2}\left[\sum_{t=1}^{T}\sum_{i=1}^{n_j}(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})^\top\boldsymbol{\Sigma}_{\delta,j}^{-1}(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})+\mathrm{tr}(\mathbf{H}_\delta\boldsymbol{\Sigma}_{\delta,j}^{-1})\right]\right) \\
&=|\boldsymbol{\Sigma}_{\delta,j}|^{-(\sum_{t=1}^{T}n_{j,t}+\nu_\delta+M+1)/2} \\
&\quad\times\exp\left(-\frac{1}{2}\mathrm{tr}\left[\left(\sum_{t=1}^{T}\sum_{i=1}^{n_{j,t}}(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})(\mathbf{y}_{i,j,t}-\boldsymbol{\mu}_{j,t})^\top+\mathbf{H}_\delta\right)\boldsymbol{\Sigma}_{\delta,j}^{-1}\right]\right)
\end{aligned}
$$

which is in the same form as (3), with parameter values

$$\tilde{\nu}_{\delta,j} = \nu_\delta + \sum_{t=1}^{T} n_{j,t}$$
$$\tilde{\mathbf{H}}_{\delta,j} = \mathbf{H}_\delta + \sum_{t=1}^{T}\sum_{i=1}^{n_{j,t}} (\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})(\mathbf{y}_{i,j,t} - \boldsymbol{\mu}_{j,t})^\top$$

and we can update $\boldsymbol{\Sigma}_{\delta,j}$ from its full conditional distribution as $IW(\tilde{\mathbf{H}}_{\delta,j}, \tilde{\nu}_{\delta,j})$.

Update $\boldsymbol{\Sigma}_{\varepsilon,r}, r = 1, \ldots, R$

Looking at (18) and (19), we see that $\boldsymbol{\Sigma}_{\varepsilon,r}$ appears in two terms, namely in the prior for $\boldsymbol{\mu}_{j,t}$, for $j \in \mathcal{J}_r$ and $t = 1, \ldots, T$ and in its own prior distribution. The product of these terms looks like

$$\prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r} MVN\left(\boldsymbol{\mu}_{j,t}|\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}\right) \times IW\left(\boldsymbol{\Sigma}_{\varepsilon,r}|\mathbf{H}_\varepsilon, \nu_\varepsilon\right),$$

which is the product of Multivariate Normal distributions with an Inverse Wishart distribution. Hence, the posterior will also follow an Inverse Wishart distribution (3). Applying the same trick as before, we have

$$\begin{aligned}
&\prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r} MVN\left(\boldsymbol{\mu}_{j,t}|\mathbf{u}_{j,t} + \mathbf{X}_{j,t}\boldsymbol{\beta} + \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j, \boldsymbol{\Sigma}_{\varepsilon,r}\right) \times IW\left(\boldsymbol{\Sigma}_{\varepsilon,r}|\mathbf{H}_\varepsilon, \nu_\varepsilon\right)\\
&= \prod_{t=1}^{T}\prod_{j\in\mathcal{J}_r} (2\pi)^{-M/2}|\boldsymbol{\Sigma}_{\varepsilon,r}|^{-1/2}\exp\left(-\frac{1}{2}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)\right)\\
&\quad\times \frac{|\mathbf{H}_\varepsilon|^{\nu_\varepsilon/2}}{2^{\nu_\varepsilon M/2}\Gamma_M\left(\frac{\nu_\varepsilon}{2}\right)}|\boldsymbol{\Sigma}_{\varepsilon,r}|^{-(\nu_\varepsilon+M+1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{H}_\varepsilon\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\right)\right)\\
&\propto |\boldsymbol{\Sigma}_{\varepsilon,r}|^{-(T\times n_r+\nu_\varepsilon+M+1)/2}\\
&\quad\times\exp\left(-\frac{1}{2}\left[\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j + \mathrm{tr}(\mathbf{H}_\varepsilon\boldsymbol{\Sigma}_{\varepsilon,r}^{-1})\right]\right)\\
&= |\boldsymbol{\Sigma}_{\varepsilon,r}|^{-(T\times n_r+\nu_\varepsilon+M+1)/2}\\
&\quad\times\exp\left(-\frac{1}{2}\mathrm{tr}\left[\left(\sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top + \mathbf{H}_\varepsilon\right)\boldsymbol{\Sigma}_{\varepsilon,r}^{-1}\right]\right)
\end{aligned}$$

which is in the same form as (3), with parameter values

$$\tilde{\nu}_{\varepsilon,r} = \nu_\varepsilon + T \times n_r$$
$$\tilde{\mathbf{H}}_{\varepsilon,r} = \mathbf{H}_\varepsilon + \sum_{t=1}^{T}\sum_{j\in\mathcal{J}_r}(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)(\boldsymbol{\mu}_{j,t} - \mathbf{u}_{j,t} - \mathbf{X}_{j,t}\boldsymbol{\beta} - \tilde{\mathbf{X}}_{j,t}\boldsymbol{\eta}_j)^\top$$

and we can update $\boldsymbol{\Sigma}_{\varepsilon,r}$ from its full conditional distribution as $IW(\tilde{\mathbf{H}}_{\varepsilon,r}, \tilde{\nu}_{\varepsilon,r})$.

Update $\tau^2_{\eta,m,q}, m = 1, \ldots, M, \quad q = 1, \ldots, Q$

In considering the full model, we see that $\tau^2_{\eta,m,q}$ appears in two terms, namely in the prior for $\boldsymbol{\eta}_{m,q}$ and in its own prior distribution. The product of these terms is

$$MVN\left(\boldsymbol{\eta}_{m,q}|\mathbf{0}, \tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})\right) \times IG\left(\tau^2_{\eta,m,q}|a_\eta, b_\eta\right),$$

which is the product of a Multivariate Normal distribution with an Inverse Gamma distribution. Since the Inverse Gamma is a conjugate prior for the variance of a Normal distribution, we know that the posterior will also follow an Inverse Gamma distribution (4). We have

$$\begin{aligned}
&MVN\left(\boldsymbol{\eta}_{m,q}|\mathbf{0}, \tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})\right) \times IG\left(\tau^2_{\eta,m,q}|a_\eta, b_\eta\right) \\
&= (2\pi)^{-J/2}|\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^\top_{m,q}(\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q}))^{-1}\boldsymbol{\eta}_{m,q}\right) \\
&\quad \times \frac{b_\eta^{a_\eta}}{\Gamma(a_\eta)}(\tau^2_{\eta,m,q})^{-a_\eta-1} \exp\left(\frac{-b_\eta}{\tau^2_{\eta,m,q}}\right) \\
&\propto (\tau^2_{\eta,m,q})^{-\frac{J}{2}-a_\eta-1} \exp\left(\frac{-\frac{1}{2}\boldsymbol{\eta}^\top_{m,q}\mathbf{Q}(\rho_{\eta,m,q})^{-1}\boldsymbol{\eta}_{m,q} - b_\eta}{\tau^2_{\eta,m,q}}\right)
\end{aligned}$$

which we can identify as in the form of an Inverse Gamma pdf (4), with parameters

$$\begin{aligned}
\tilde{a_\eta} &= a_\eta + \frac{J}{2} \\
\tilde{b_\eta} &= b_\eta + \frac{1}{2}\boldsymbol{\eta}^\top_{m,q}\mathbf{Q}(\rho_{\eta,m,q})^{-1}\boldsymbol{\eta}_{m,q}
\end{aligned}$$

and we can then update $\tau^2_{\eta,m,q}$ via it's full conditional as $IG(\tilde{a_\eta}, \tilde{b_\eta})$.

Update $\rho_{v,m}, m = 1, \ldots, M$

Until now we have worked directly with full conditional distributions for model parameters via conjugate priors. This has made things easier, as we could directly determine the form of the full conditional distribution. Now, we introduce a *Metropolis step* to update $\rho_{v,m}, m = 1, \ldots, M$. The metropolis step is essentially a process whereby a proposed value for the parameter is either accepted or rejected based on a proposal distribution. Although we are using a Uniform distribution with bounded support, our proposal distribution (which is typically Guassian) may propose values outside the support of $\rho_{v,m}$. Hence, we will introduce a transformation for $\rho_{v,m}$, which will give it support on the real line $\mathbb{R}$. This way, we can use a Gaussian proposal distribution, evaluate the acceptance probability of this transformed variable value, and then back transform our sample upon acceptance.

When considering the full model, we identify the terms in which $\rho_{v,m}$ appears, namely in the prior for $\mathbf{u}_{j,t}, j = 1, \ldots, J, t = 1, \ldots, T$ and in its own prior distribution, which is Uniform on $[a_v, b_v]$. The product of these densities is

$$
\begin{aligned}
&\prod_{t=1}^{T}\prod_{j=1}^{J} MVN(\mathbf{u}_{j,t} \mid \mathbf{u}_{j,t-1} + \boldsymbol{\mu}_{u,j,t}, \frac{1}{d_j}\boldsymbol{\Sigma}_{u,r}) \times U(\rho_{v,m} \mid a_v, b_v), \quad j \in \mathcal{J}_r \\
&= \prod_{t=1}^{T}\prod_{j=1}^{J} (2\pi)^{-M/2}|\boldsymbol{\Sigma}_{u,r}|^{-1/2} \exp\left(-\frac{d_j}{2}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})^{\top}(\boldsymbol{\Sigma}_{u,r})^{-1}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})\right) \\
&\quad \times \frac{1}{b_v - a_v}\mathbf{1}_{a_v<\rho_{v,m,t}<b_v}
\end{aligned}
$$

where $\mathbf{1}$ is the indicator function. Next, we employ transformations on the Uniform variable $\rho_{v,m}$ of the form

$$
g(\rho_{v,m}) = \log\left(\frac{\rho_{v,m} - a_v}{b_v - \rho_{v,m}}\right).
$$

To get the distribution of each newly transformed variable $g(\rho_{v,m})$, we require the Jacobian. First, we find the inverse function of the transformation, which is

$$
g^{-1}(g(\rho_{v,m})) = \frac{b_v e^{g(\rho_{v,m})} + a_v}{e^{g(\rho_{v,m})} + 1}.
$$

Next, we take the derivative of this function with respect to each $g(\rho_{v,m,t}), m = 1, \ldots, M$, which gives

$$
\frac{e^{g(\rho_{v,m})}(b_v - a_v)}{(e^{g(\rho_{v,m})} + 1)^2},
$$

and plugging the expression for $g(\rho_{v,m})$ into this equation yields

$$
\frac{\frac{\rho_{v,m}-a_v}{b_v-\rho_{v,m}}(b_v - a_v)}{\left(\frac{\rho_{v,m}-a_v}{b_v-\rho_{v,m}} + 1\right)^2} \propto \frac{\frac{\rho_{v,m}-a_v}{b_v-\rho_{v,m}}}{\left(\frac{b_v-\rho_{v,m,t}+\rho_{v,m}-a_v}{b_v-\rho_{v,m}}\right)^2} \propto \frac{\frac{\rho_{v,m}-a_v}{b_v-\rho_{v,m}}}{\frac{1}{(b_v-\rho_{v,m})^2}} = (\rho_{v,m} - a_v)(b_v - \rho_{v,m}).
$$

With this jacobian adjustment, we can now determine the *log target density* to which we will add the Jacobian adjustment. We have

$$
\begin{aligned}
&\log\Bigg[\prod_{t=1}^{T}\prod_{j=1}^{J} (2\pi)^{-M/2}|\boldsymbol{\Sigma}_{u,r}|^{-1/2} \exp\left(-\frac{d_j}{2}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})^{\top}(\boldsymbol{\Sigma}_{u,r})^{-1}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})\right) \\
&\quad \times \frac{1}{b_v - a_v}\mathbf{1}_{a_v<\rho_{v,m,t}<b_v} \times (\rho_{v,m} - a_v)(b_v - \rho_{v,m})\Bigg] \\
&\propto \sum_{t=1}^{T}\sum_{j=1}^{J}\left(-\frac{d_j}{2}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})^{\top}(\boldsymbol{\Sigma}_{u,r})^{-1}(\mathbf{u}_{j,t} - \mathbf{u}_{j,t-1} - \boldsymbol{\mu}_{u,j,t})\right) \\
&\quad + \log(\rho_{v,m} - a_v) + \log(b_v - \rho_{v,m})
\end{aligned}
$$

We can then proceed by generating proposed values of $g(\rho_{v,m})$ from some proposal distribution (typically Gaussian), evaluating the log target density plus Jacobian adjustment, and then accepting or rejecting the back-transformed parameter value based on this acceptance probability.

Update $\rho_{\eta,m,q}, \quad m = 1, \ldots, M, q = 1, \ldots, Q$

We again use Metropolis steps to update $\rho_{\eta,m,q}, m = 1, \ldots, M, q = 1, \ldots, Q$. When considering the full model, we identify the terms in which $\rho_{\eta,m,q}$ appears, namely in the prior for $\boldsymbol{\eta}_{m,q}$ and in it's own prior distribution, which is Uniform on $[a_\eta, b_\eta]$. The product of these densities is

$$
\begin{aligned}
&MVN(\boldsymbol{\eta}_{m,q} \mid \mathbf{0}, \tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})) \times U(\rho_{\eta,m,q} \mid a_\eta, b_\eta) \\
&= (2\pi)^{-J/2} |\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^\top_{m,q}(\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q}))^{-1}\boldsymbol{\eta}_{m,q}\right) \\
&\quad \times \frac{1}{b_\eta - a_\eta}\mathbf{1}_{a_\eta < \rho_{\eta,m,q} < b_\eta}
\end{aligned}
$$

where $\mathbf{1}$ is the indicator function. Next, we employ a transformation on the Uniform variable $\rho_{\eta,m,q}$ of the form

$$
g(\rho_{\eta,m,q}) = \log\left(\frac{\rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}}\right).
$$

To get the distribution of this newly transformed variable $g(\rho_{\eta,m,q})$, we require the Jacobian. First, we find the inverse function of the transformation, which is

$$
g^{-1}(g(\rho_{\eta,m,q})) = \frac{b_\eta e^{g(\rho_{\eta,m,q})} + a_\eta}{e^{g(\rho_{\eta,m,q})} + 1}.
$$

Next, we take the derivative of this function with respect to $g(\rho_{\eta,m,q})$, which gives

$$
\frac{e^{g(\rho_{\eta,m,q})}(b_\eta - a_\eta)}{(e^{g(\rho_{\eta,m,q})} + 1)^2},
$$

and plugging the expression for $g(\rho_{\eta,m,q})$ into this equation yields

$$
\frac{\frac{\rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}}(b_\eta - a_\eta)}{\left(\frac{\rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}} + 1\right)^2} \propto \frac{\frac{\rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}}}{\left(\frac{b_\eta - \rho_{\eta,m,q} + \rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}}\right)^2} \propto \frac{\frac{\rho_{\eta,m,q} - a_\eta}{b_\eta - \rho_{\eta,m,q}}}{\frac{1}{(b_\eta - \rho_{\eta,m,q})^2}} = (\rho_{\eta,m,q} - a_\eta)(b_\eta - \rho_{\eta,m,q}).
$$

With this jacobian adjustment, we can now determine the log target density to which we will add the Jacobian adjustment. We have

$$
\begin{aligned}
&\log\Bigg[(2\pi)^{-J/2} |\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^\top_{m,q}(\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q}))^{-1}\boldsymbol{\eta}_{m,q}\right) \\
&\quad \times \frac{1}{b_\eta - a_\eta}\mathbf{1}_{a_\eta < \rho_{\eta,m,q} < b_\eta} \times (\rho_{\eta,m,q} - a_\eta)(b_\eta - \rho_{\eta,m,q})\Bigg] \\
&\propto -\frac{1}{2}\log|\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})| - \frac{1}{2}\boldsymbol{\eta}^\top_{m,q}(\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q}))^{-1}\boldsymbol{\eta}_{m,q} + \log(\rho_{\eta,m,q} - a_\eta) + \log(b_\eta - \rho_{\eta,m,q})
\end{aligned}
$$

We can then proceed by generating a proposed value of $g(\rho_{\eta,m,q})$ from some proposal distribution (usually Gaussian), evaluating the log target density plus Jacobian adjustment, and then accepting or rejecting the back-transformed $g^{-1}(g(\rho_{\eta,m,q}))$ based on this acceptance probability.

Efficiently evaluate $\log|\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})|$ **and** $(\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q}))^{-1}$

First, define $\mathbf{Q}(\rho_{\eta,m,q})^{-1} = (\mathbf{D} - \rho_{\eta,m,q}\mathbf{W}) = \mathbf{D}^{1/2}(\mathbf{I} - \rho_{\eta,m,q}\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2})\mathbf{D}^{1/2}$, where $\mathbf{W}$ and $\mathbf{D}$ are as defined in (7) and $\mathbf{I}$ is the $J \times J$ identity matrix. Next, let $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}^\top$, where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues and the columns of $\mathbf{P}$ are the eigenvectors of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. This allows $\mathbf{Q}(\rho_{\eta,m,q})^{-1}$ to be expressed as

$$\mathbf{Q}(\rho_{\eta,m,q})^{-1} = \sum_{j=1}^{J}(1 - \rho_{\eta,m,q}\lambda_j)\mathbf{z}_j\mathbf{z}_j^\top = \sum_{j=1}^{J}\mathbf{z}_j\mathbf{z}_j^\top - \rho_{\eta,m,q}\left(\sum_{j=1}^{J}\lambda_j\mathbf{z}_j\mathbf{z}_j^\top\right)$$

where $\mathbf{z}_j$ are the columns of $\mathbf{D}^{1/2}\mathbf{P}$ and λ_j are the diagonal elements of $\boldsymbol{\Lambda}$. Expressing the precision matrix in this way removes the need for costly matrix formation and Cholesky decomposition in each MCMC iteration. Rather, the first term, i.e., $\sum_{j=1}^{J}\mathbf{z}_j\mathbf{z}_j^\top$, remains the same across MCMC iterations and the second term, i.e., $\rho_{\eta,m,q}\left(\sum_{j=1}^{J}\lambda_j\mathbf{z}_j\mathbf{z}_j^\top\right)$, only varies by a multiplicative constant $\rho_{\eta,m,q}$.

To evaluate $\log|\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})|$, we use the fact that $|\mathbf{Q}(\rho_{\eta,m,q})^{-1}| = \prod_{j=1}^{J}\mathbf{D}_{jj}(1 - \rho_{\eta,m,q}\lambda_j)$ and obtain

$$\begin{aligned}
\log|\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})| &= -\log|1/\tau^2_{\eta,m,q}\mathbf{Q}(\rho_{\eta,m,q})^{-1}| \\
&= -\log\left[\left(\frac{1}{\tau^2_{\eta,m,q}}\right)^J|\mathbf{Q}(\rho_{\eta,m,q})^{-1}|\right] \\
&= -J\log\frac{1}{\tau^2_{\eta,m,q}} - \log|\mathbf{Q}(\rho_{\eta,m,q})^{-1}| \\
&= -J\log\frac{1}{\tau^2_{\eta,m,q}} - \log\left[\prod_{j=1}^{J}\mathbf{D}_{jj}(1 - \rho_{\eta,m,q}\lambda_j)\right] \\
&= -J\log\frac{1}{\tau^2_{\eta,m,q}} - \sum_{j=1}^{J}\left[\log\mathbf{D}_{jj}(1 - \rho_{\eta,m,q}\lambda_j)\right]
\end{aligned}$$