Generalised Pareto Distribution with Predictive Processes – MCMC updates

John O'Sullivan

May 7, 2018

Algorithm details

(Filled in below):

Algorithm 1: Gaussian Process Generalised Pareto Distribution Markov chain Monte Carlo

Data: y_{ij}, the declustered threshold excesses at locations $i = 1 \ldots n$, with $j = 1 \ldots n_i$ excesses at location i;

X_ϕ, X_ξ, $n \times (p+1)$ and $n \times (q+1)$ matrices of p and q covariates at locations $i = 1 \ldots n$.

Result: Samples from the posterior distributions of $\phi = \log(\sigma)$ and ξ (the unknown parameters of interest), which can then be used to calculate return level estimates

Initialisation;

Random starting values of $\tilde{\phi}$ and $\tilde{\xi}$;

Projection of ϕ and ξ from $\tilde{\phi}$ and $\tilde{\xi}$ respectively;

Hyper-parameter values of α_ϕ, β_ϕ, ς^2_ϕ, τ^2_ϕ, ν_ϕ, α_ξ, β_ξ, ς^2_ξ, τ^2_ξ, and ν_ξ;

Number of iterations N;

for *iterations i from 1 to N* **do**

    Generate $u \sim U(0, 1)$;

    for *sub-grid locations k from 1 to m* **do**

        Simulate $\tilde{\phi}_{new,k}$;

        Project ϕ_{new} from $\tilde{\phi}_{new}$;

        Set l_{new} = log full conditional of new vector $\tilde{\phi}_{new}$;

        Set l_{old} = log full conditional of old vector $\tilde{\phi}$;

        Set $a = \exp(l_{new} - l_{old})$, that is, evaluate equation (1);

        if $a > u$ **then**

            Set $\tilde{\phi} = \tilde{\phi}_{new}$;

        end

        Simulate $\tilde{\xi}_{new,k}$;

        Project ξ_{new} from $\tilde{\xi}_{new}$;

        Set l_{new} = log full conditional of new vector $\tilde{\xi}_{new}$;

        Set l_{old} = log full conditional of old vector $\tilde{\xi}$;

        Set $a = \exp(l_{new} - l_{old})$, that is, evaluate equation (2);

        if $a > u$ **then**

            Set $\tilde{\xi} = \tilde{\xi}_{new}$;

        end

    end

end

for *iterations i from 1 to N* **do**

  for *each element c of the vector* α_ϕ **do**

    Simulate $\alpha_{\phi_{new,c}}$;

    Set $a =$ result of equation (3);

    If $a > u$, set $\alpha_\phi = \alpha_{\phi_{new}}$;

  end

  for *each element d of the vector* α_ξ **do**

    As above, but set $a =$ result of equation (4);

  end

  for *elements e in the lower-triangle of the matrix* β_ϕ **do**

    Simulate $\beta_{\phi_{new,e}}$;

    Set $a =$ result of equation (5);

    If $a > u$, set $\beta_\phi = \beta_{\phi_{new}}$;

  end

  for *elements f in the lower-triangle of the matrix* β_ξ **do**

    As above, but set $a =$ result of equation (6);

  end

  Simulate $\nu_{\phi_{new}}$;

  Set $a =$ result of equation (7);

  If $a > u$, set $\nu_\phi = \nu_{\phi_{new}}$;

  Repeat for ν_ξ with equation (8) ;

  Simulate $\varsigma^2_{\phi_{new}}$;

  Set $a =$ result of equation (9);

  If $a > u$, set $\varsigma^2_\phi = \varsigma^2_{\phi_{new}}$;

  Repeat for ς^2_ξ with equation (10) ;

  Simulate $\tau^2_{\phi_{new}}$;

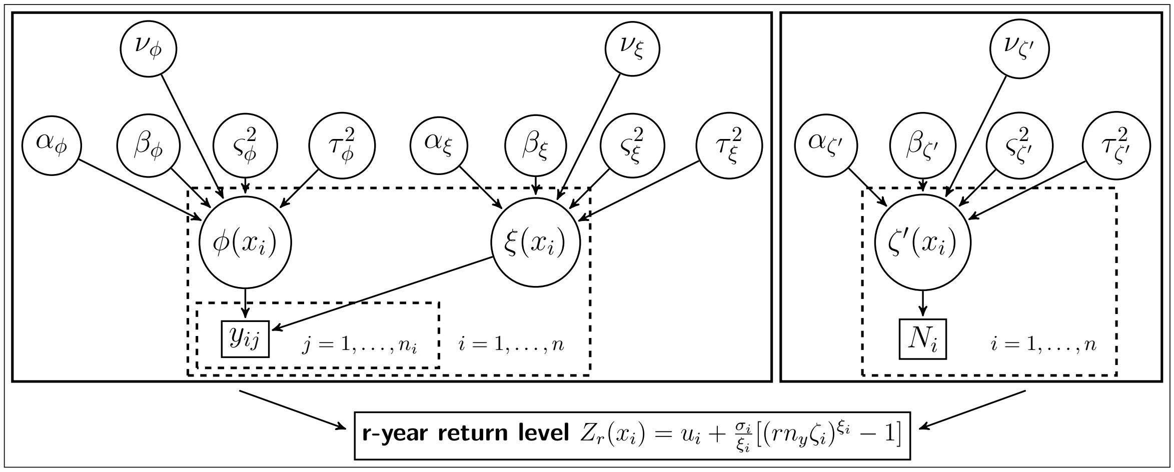
  Set $a =$ result of equation (11);

  If $a > u$, set $\tau^2_\phi = \tau^2_{\phi_{new}}$;

  Repeat for τ^2_ξ with equation (12) ;

end

1 DAG



2 Notation

- x_i Multivariate (bivariate or tri-variate) location values for location i, $i = 1, \ldots, n$. Write the matrix of all locations as just x
- y_{ij} Excess j for observation i, $j = 1, \ldots, n_i$ where n_i is the number of excesses at location i
- $\sigma(x_i)$ Scale parameter for location x_i
- $\phi(x_i) = \log(\sigma(x_i))$ The re-parameterised scale parameter
- $\xi(x_i)$ Shape parameter for location x_i
- z_k Sub-grid locations $k = 1, \ldots, m$. Together written as z
- A, B Projection matrices of dimension $n \times m$
- $\tilde{\phi}, \tilde{\xi}$ Gaussian processes of ϕ and ξ defined on sub-grid z
- $\mu_\phi(z), \mu_\xi(z)$ Means for the Gaussian processes
- Σ, Ψ Auto-covariance matrices for the Gaussian processes
- τ^2_ϕ, τ^2_ξ Nugget standard deviation parameters
- α_ϕ, α_ξ vectors of coefficients including intercept and slope parameters for Gaussian process means

- $X_\phi, X_\xi, Z_\sigma, Z_\xi$ Matrices of covariates (including column for intercept term) on the x grid and the z sub-grid

- ν_ϕ, ν_ξ Matern smoothness parameters

- β_ϕ, β_ξ Matern length scale matrices

- $\varsigma_\phi^2, \varsigma_\xi^2$ Variance parameters for Gaussian process

3 Model outline

In hierarchical notation:

$$
\begin{aligned}
y_{ij} &\sim GPD(\sigma(x_i), \xi(x_i)) \\
\log(\sigma(x)) = \phi(x) &= A(x,z)\Sigma^{-1}(z,z)\tilde{\phi}(z) \\
\xi(x) &= B(x,z)\Psi^{-1}(z,z)\tilde{\xi}(z) \\
A(x_i, z_k) &= \varsigma_\phi^2 \frac{2^{1-\nu_\phi}}{\Gamma(\nu_\phi)} \left(\sqrt{2\nu_\phi} \frac{\|x_i - z_k\|}{\beta_\phi} \right)^{\nu_\phi} K_{\nu_\phi} \left(\sqrt{2\nu_\phi} \frac{\|x_i - z_k\|}{\beta_\phi} \right) \\
B(x_i, z_k) &= \varsigma_\xi^2 \frac{2^{1-\nu_\xi}}{\Gamma(\nu_\xi)} \left(\sqrt{2\nu_\xi} \frac{\|x_i - z_k\|}{\beta_\xi} \right)^{\nu_\xi} K_{\nu_\xi} \left(\sqrt{2\nu_\xi} \frac{\|x_i - z_k\|}{\beta_\xi} \right) \\
\tilde{\phi} &\sim MVN_m(\mu_\phi(z), \tau_\phi^2 I_m + \Sigma(z,z)) \\
\tilde{\xi} &\sim MVN_m(\mu_\xi(z), \tau_\xi^2 I_m + \Psi(z,z)) \\
\mu_\phi(z) &= Z_\phi \alpha_\phi \\
\mu_\xi(z) &= Z_\xi \alpha_\xi \\
\mu_\phi(x) &= X_\phi \alpha_\phi \\
\mu_\xi(x) &= X_\xi \alpha_\xi \\
\Sigma(z_k, z_l) &= \varsigma_\phi^2 \frac{2^{1-\nu_\phi}}{\Gamma(\nu_\phi)} \left(\sqrt{2\nu_\phi} \frac{\|z_k - z_l\|}{\beta_\phi} \right)^{\nu_\phi} K_{\nu_\phi} \left(\sqrt{2\nu_\phi} \frac{\|z_k - z_l\|}{\beta_\phi} \right) \\
\Psi(z_k, z_l) &= \varsigma_\xi^2 \frac{2^{1-\nu_\xi}}{\Gamma(\nu_\xi)} \left(\sqrt{2\nu_\xi} \frac{\|z_k - z_l\|}{\beta_\xi} \right)^{\nu_\xi} K_{\nu_\xi} \left(\sqrt{2\nu_\xi} \frac{\|z_k - z_l\|}{\beta_\xi} \right)
\end{aligned}
$$

Hyper-parameter prior distributions (subject to change):

$$\begin{aligned}
\log(\varsigma_\phi^2), \log(\varsigma_\xi^2), \log(\tau_\phi^2), \log(\tau_\xi^2) &\sim N(0, 10) \\
\beta_\phi, \beta_\xi, \nu_\phi, \nu_\xi &\sim DU(0.001, 0.01, 0.1, 1, 10, 100, 1000) \\
\alpha_\phi &\sim MVN(\vec{\eta}_\phi, H_\phi) \\
\alpha_\xi &\sim MVN(\vec{\eta}_\xi, H_\xi)
\end{aligned}$$

To begin, $\vec{\eta}_\phi$ and $\vec{\eta}_\xi$ are vectors of the appropriate length with all entries equal to 0. H_ϕ and H_ξ are the relevant covariance matrices - to begin, these are diagonal matrices with 10 on the diagonal. The code is designed to allow the relationship between the coefficients to be modelled (if desired) by adjusting this covariance matrix.

4 Posterior distribution

The full posterior distribution is:

$$\begin{aligned}
p(\alpha_\phi, \alpha_\xi, \beta_\phi, \beta_\xi, \varsigma_\phi^2, \varsigma_\xi^2, \tau_\phi^2, \tau_\xi^2, \nu_\phi, \nu_\xi, \tilde{\phi}, \tilde{\xi} | y, x, z, X_\phi, X_\xi, Z_\phi, Z_\xi) \propto & \left[\prod_{i=1}^{n} \prod_{j=1}^{n_i} p(y_{ij} | \sigma(x_i) = \exp(\phi(x_i)), \xi(x_i)) \right] \times \\
& p(\tilde{\phi}(x) | \mu_\phi, \tau_\phi^2, \Sigma) p(\tilde{\xi}(x) | \mu_\xi, \tau_\xi^2, \Psi) \times \\
& p(\alpha_\phi) p(\alpha_\xi) p(\beta_\phi) p(\beta_\xi) p(\nu_\phi) p(\nu_\xi) \times \\
& p(\tau_\phi^2) p(\tau_\xi^2) p(\varsigma_\phi^2) p(\varsigma_\xi^2)
\end{aligned}$$

5 Conditional posterior distributions: Layer 1

Updating the first layer of the DAG - that is, parameters $\tilde{\phi} = \log(\tilde{\sigma})$ and $\tilde{\xi}$. The conditional posterior distribution of $\tilde{\phi}$ is given by:

$$\pi(\tilde{\phi}|y, x, z, X_\phi, X_\xi, Z_\phi, Z_\xi, \alpha_\phi, \alpha_\xi, \beta_\phi, \beta_\xi, \varsigma_\phi^2, \varsigma_\xi^2, \tau_\phi^2, \tau_\xi^2, \nu_\phi, \nu_\xi, \tilde{\xi}) \propto$$

$$p(y|\tilde{\phi}, x, z, X_\phi, X_\xi, Z_\phi, Z_\xi, \alpha_\phi, \alpha_\xi, \beta_\phi, \beta_\xi, \varsigma_\phi^2, \varsigma_\xi^2, \tau_\phi^2, \tau_\xi^2, \nu_\phi, \nu_\xi, \tilde{\xi})\times$$

$$p(\tilde{\phi}|x, z, X_\phi, X_\xi, Z_\phi, Z_\xi, \alpha_\phi, \alpha_\xi, \beta_\phi, \beta_\xi, \varsigma_\phi^2, \varsigma_\xi^2, \tau_\phi^2, \tau_\xi^2, \nu_\phi, \nu_\xi, \tilde{\xi})$$

$$\propto p(y|\sigma, \xi)p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)$$

That is, the conditional posterior of $\tilde{\phi}$ is proportional to the product of the likelihood of the data y given $\tilde{\phi}$ and all other parameters, and the prior probability density of $\tilde{\phi}$ given all of the other parameters. In the final line, most parameters have dropped out from the right-hand side, as the densities are independent of these, given the remaining terms (see the DAG for this). Of the remaining terms σ, ξ, μ_ϕ and Σ are all deterministic given the other parameters. These will have been calculated using the formulae in section 3.

In briefer notation (to be used from now on):

$$\begin{aligned}
\pi(\tilde{\phi}|\dots) &\propto p(y|\phi, \xi)p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma) \\
&= \left[\prod_{i=1}^{n}\prod_{j=1}^{n_i} p(y_{ij}|\sigma(x_i), \xi(x_i))\right] p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma) \\
&= \left[\prod_{i=1}^{n}\prod_{j=1}^{n_i} \frac{1}{\sigma(x_i)}\left(1 + \xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\right)^{-(1/\xi(x_i)+1)}\right] \times \\
&\quad \frac{1}{\sqrt{\det(2\pi(\Sigma + \tau_\phi^2 I_m))}} e^{-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)}
\end{aligned}$$

Similarly:

$$
\begin{aligned}
\pi(\tilde{\xi}|\dots) \propto & p(y|\phi,\xi)p(\tilde{\xi}|\mu_\xi,\tau_\xi^2,\Psi) \\
= & \left[\prod_{i=1}^{n}\prod_{j=1}^{n_i} p(y_{ij}|\sigma(x_i),\xi(x_i))\right] p(\tilde{\xi}|\mu_\xi,\tau_\xi^2,\Psi) \\
= & \left[\prod_{i=1}^{n}\prod_{j=1}^{n_i} \frac{1}{\sigma(x_i)}\left(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\right)^{-(1/\xi(x_i)+1)}\right] \times \\
& \frac{1}{\sqrt{\det(2\pi(\Psi+\tau_\xi^2 I_m))}} e^{-\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi)}
\end{aligned}
$$

5.1 Metropolis-Hastings Markov chain Monte Carlo sampling

In order to sample from these conditional posterior distributions, we use the Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm. New samples are accepted or rejected at random according to the algorithm outlined below.

A new value of $\tilde{\phi}_i$ is suggested: $\tilde{\phi}_i{}'$. The new vector ϕ' is then calculated using the projection formula.

Though $\tilde{\phi}'$ differs from $\tilde{\phi}$ in just one location i, ϕ and ϕ' can be different from each other in many locations i due to this projection.

Suggested updates are drawn from a Normal distribution centred on the old value and with a variance of a manually set tuning parameter used to control the size of the proposed steps.

We calculate:

$$
\rho(\tilde{\phi}_i,\tilde{\phi}_i') = \min\left(1, \frac{\pi(\tilde{\phi}'|\dots)q_t(\tilde{\phi}_i'\to\tilde{\phi}_i)}{\pi(\tilde{\phi}|\dots)q_t(\tilde{\phi}_i\to\tilde{\phi}_i')}\right)
$$

where $\pi(\tilde{\phi}|\dots)$ is as defined above, and $q_t(a \to b)$ is the transition probability of proposing value b given value a. Since updates are proposed using a Normal distribution, these transition probabilites above and below the line will always cancel, so the above simplifies to:

$$\rho(\tilde{\phi}_i, \tilde{\phi}'_i) = \min\left(1, \frac{\pi(\tilde{\phi}'|\dots)}{\pi(\tilde{\phi}|\dots)}\right)$$

Following this calculation, we always accept proposed value $\tilde{\phi}'_i$ when $\rho(\tilde{\phi}_i, \tilde{\phi}'_i)$ is bigger than 1 and we reject accordingly when the ratio is smaller than 1 by simulating a random variable $u \sim U[0,1]$ and accepting proposed value $\tilde{\phi}'_i$ when $u \leq \rho(\tilde{\phi}_i, \tilde{\phi}'_i)$.

Evaluating $\rho(\tilde{\phi}_i, \tilde{\phi}'_i)$ typically involves products and quotients of many terms which may be close to 0. In order to work with something far more stable, we use the property that $x = \exp(\log(x))$.

Following this observation we need to evaluate:

$$\begin{aligned}
&\exp\left(\log\left(\frac{\pi(\tilde{\phi}'|\dots)}{\pi(\tilde{\phi}|\dots)}\right)\right) \\
&= \exp\left[\log(\pi(\tilde{\phi}'|\dots)) - \log(\pi(\tilde{\phi}|\dots))\right] \\
&= \exp\Bigg[\log\Bigg(\Bigg[\prod_{i=1}^{n}\prod_{j=1}^{n_i} p(y_{ij}|\sigma'(x_i), \xi(x_i))\Bigg] p(\tilde{\phi}'|\mu_\phi, \tau_\phi^2, \Sigma)\Bigg) - \\
&\qquad \log\Bigg(\Bigg[\prod_{i=1}^{n}\prod_{j=1}^{n_i} p(y_{ij}|\sigma(x_i), \xi(x_i))\Bigg] p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)\Bigg)\Bigg]
\end{aligned}$$

Filling in the specific distribution from above, this becomes:

$$
\begin{aligned}
&\exp\Bigg[\log\Big(\Big[\prod_{i=1}^{n}\prod_{j=1}^{n_i}\frac{1}{\sigma'(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\times \\
&\qquad\frac{1}{\sqrt{\det(2\pi(\Sigma+\tau_\phi^2 I_m))}}e^{-\frac{1}{2}(\tilde{\phi}'-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}'-\mu_\phi)}\Big) \\
&\quad-\log\Big(\Big[\prod_{i=1}^{n}\prod_{j=1}^{n_i}\frac{1}{\sigma(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\times \\
&\qquad\frac{1}{\sqrt{\det(2\pi(\Sigma+\tau_\phi^2 I_m))}}e^{-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)}\Big)\Bigg] \\
&=\exp\Bigg[\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\Big(\Big[\frac{1}{\sigma'(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\Big)+ \\
&\qquad\log\left(\frac{1}{\sqrt{\det(2\pi(\Sigma+\tau_\phi^2 I_m))}}\right)-\frac{1}{2}(\tilde{\phi}'-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}'-\mu_\phi) \\
&\quad-\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\Big(\Big[\frac{1}{\sigma(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\Big)- \\
&\qquad\log\left(\frac{1}{\sqrt{\det(2\pi(\Sigma+\tau_\phi^2 I_m))}}\right)+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)\Bigg] \\
&=\exp\Bigg[\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\Big(\Big[\frac{1}{\sigma'(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\Big)-\frac{1}{2}(\tilde{\phi}'-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}'-\mu_\phi) \\
&\quad-\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\Big(\Big[\frac{1}{\sigma(x_i)}\Big(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\Big)^{-(1/\xi(x_i)+1)}\Big]\Big)+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)\Bigg]
\end{aligned}
$$

The common term in both numerator and denominator above dropped out as it will be equal in both (it has no dependence on $\tilde{\phi}$).

The GPD can clearly be simplified further using the properties of logs. Just taking one of the functions on its own for clarity:

$$\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\left(\left[\frac{1}{\sigma'(x_i)}\left(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\right)^{-(1/\xi(x_i)+1)}\right]\right)=$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n_i}\left[\log\Big(\frac{1}{\sigma'(x_i)}\Big)+\log\left(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\right)^{-(1/\xi(x_i)+1)}\right]=$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n_i}\left[-\log(\sigma'(x_i))-\left(\frac{1}{\xi(x_i)}+1\right)\log\left(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\right)\right]$$

Following this, the full term we need to evaluate in order to update $\tilde{\phi}$ is:

$$\begin{aligned}=\exp\Bigg[&\sum_{i=1}^{n}\sum_{j=1}^{n_i}\Big[-\log(\sigma'(x_i))-\left(\frac{1}{\xi(x_i)}+1\right)\log\left(1+\xi(x_i)\frac{y_{ij}}{\sigma'(x_i)}\right)\Big]-\frac{1}{2}(\tilde{\phi}'-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}'-\mu_\phi)\\&-\sum_{i=1}^{n}\sum_{j=1}^{n_i}\Big[-\log(\sigma(x_i))-\left(\frac{1}{\xi(x_i)}+1\right)\log\left(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\right)\Big]+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)\Bigg]\end{aligned}\tag{1}$$

Similar reasoning leads to the update for $\tilde{\xi}$. We need to evaluate:

$$\rho(\tilde{\xi}_i,\tilde{\xi}_i')=\min\left(1,\frac{\pi(\tilde{\xi}'|\dots)q_t(\tilde{\xi}_i'\to\tilde{\xi}_i)}{\pi(\tilde{\xi}|\dots)q_t(\tilde{\xi}_i\to\tilde{\xi}_i')}\right)$$

The full term we need to evaluate is:

$$\begin{aligned}=\exp\Bigg[&\sum_{i=1}^{n}\sum_{j=1}^{n_i}\Big[-\log(\sigma(x_i))-\left(\frac{1}{\xi'(x_i)}+1\right)\log\left(1+\xi'(x_i)\frac{y_{ij}}{\sigma(x_i)}\right)\Big]-\frac{1}{2}(\tilde{\xi}'-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}'-\mu_\xi)\\&-\sum_{i=1}^{n}\sum_{j=1}^{n_i}\Big[-\log(\sigma(x_i))-\left(\frac{1}{\xi(x_i)}+1\right)\log\left(1+\xi(x_i)\frac{y_{ij}}{\sigma(x_i)}\right)\Big]+\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi)\Bigg]\end{aligned}\tag{2}$$

6 Conditional posterior distributions: Layer 2

6.1 α

Updating the second layer of the DAG - that is, all hyper-parameters - starting with α_ϕ:

$$\begin{aligned}\pi(\alpha_\phi|y,\dots) &\propto p(y|\alpha_\phi,\dots)p(\alpha_\phi|\dots)\\ &\propto p(y|\sigma,\xi)p(\tilde{\phi}|\mu_\phi,\tau_\phi^2,\Sigma)p(\alpha_\phi)\\ &\propto p(\tilde{\phi}|\mu_\phi,\tau_\phi^2,\Sigma)p(\alpha_\phi)\end{aligned}$$

Essentially the GPD component is independent of α_ϕ once the other parameters are known, and so can be absorbed into the constant of proportionality. What we're left with is the MVN piece (since α_ϕ features in the calculation of μ_ϕ) and the prior on α_ϕ.

Then we have:

$$\begin{aligned}\pi(\alpha_\phi|\dots) &\propto p(\tilde{\phi}|\mu_\phi,\tau_\phi^2,\Sigma)p(\alpha_\phi)\\ &\propto \frac{1}{\sqrt{\det(2\pi(\Sigma+\tau_\phi^2 I_m))}}e^{-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)}\times\\ &\quad \frac{1}{\sqrt{\det(2\pi H_\phi)}}e^{-\frac{1}{2}(\alpha_\phi-\eta_\phi)'H_\phi^{-1}(\alpha_\phi-\eta_\phi)}\end{aligned}$$

where H_ϕ is the covariance matrix for the prior distribution of α_ϕ and η_ϕ is the prior mean.
Similarly:

$$\begin{aligned}\pi(\alpha_\xi|\dots) &\propto p(\tilde{\xi}|\mu_\xi,\tau_\xi^2,\Psi)p(\alpha_\xi)\\ &\propto \frac{1}{\sqrt{\det(2\pi(\Psi+\tau_\xi^2 I_m))}}e^{-\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi)}\times\\ &\quad \frac{1}{\sqrt{\det(2\pi H_\xi)}}e^{-\frac{1}{2}(\alpha_\xi-\eta_\xi)'H_\xi^{-1}(\alpha_\xi-\eta_\xi)}\end{aligned}$$

where H_ξ is the covariance matrix for the prior distribution of α_ξ and η_ξ is the prior mean.

As before, updates are suggested for alpha element-wise: $\alpha_{\phi,k} \to \alpha'_{\phi,k}$. The ratio $\rho(\alpha_{\phi,k}, \alpha'_{\phi,k})$ is then calculated. Similar manipulations to those used in the previous section lead to the following calculation needed to update α_ϕ:

$$\begin{aligned}
\exp\Big[&-\frac{1}{2}\log(\det(2\pi(\Sigma+\tau_\phi^2 I_m))) - \frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)\\
&-\frac{1}{2}\log(\det(2\pi H_\phi)) - \frac{1}{2}(\alpha'_\phi-\eta_\phi)'H_\phi^{-1}(\alpha'_\phi-\eta_\phi))\\
&+\frac{1}{2}\log(\det(2\pi(\Sigma+\tau_\phi^2 I_m))) + \frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)\\
&+\frac{1}{2}\log(\det(2\pi H_\phi)) + \frac{1}{2}(\alpha_\phi-\eta_\phi)'H_\phi^{-1}(\alpha_\phi-\eta_\phi))\Big] =\\
\exp\Big[&-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi) - \frac{1}{2}(\alpha'_\phi-\eta_\phi)'H_\xi^{-1}(\alpha'_\phi-\eta_\phi))\\
&+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi) + \frac{1}{2}(\alpha_\phi-\eta_\phi)'H_\phi^{-1}(\alpha_\phi-\eta_\phi))\Big] \qquad (3)
\end{aligned}$$

And for α_ξ:

$$\begin{aligned}
\exp\Big[&-\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi) - \frac{1}{2}(\alpha'_\xi-\eta_\xi)'H_\xi^{-1}(\alpha'_\xi-\eta_\xi))\\
&+\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi) + \frac{1}{2}(\alpha_\xi-\eta_\xi)'H_\xi^{-1}(\alpha_\xi-\eta_\xi))\Big] \qquad (4)
\end{aligned}$$

Updates for the other hyper-parameters are very similar - though none feature the MVN prior that α_ϕ and α_ξ do, the part involving the log determinant can't be dropped from the m-dim MVN, since the matrix changes with any change in β, ν, ς^2 or τ^2. The remaining hyper-parameters either have a discrete update (in which case the prior probability will be $^1/_{\#discrete.values}$ and will cancel above and below the line, or have a univariate Normal prior on them (or their log).

6.2 β

For β_ϕ we have:

$$\pi(\beta_\phi|\dots) \propto p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)p(\beta_\phi)$$

Since β_ϕ has a discrete prior, to update β_ϕ we need to evaluate:

$$\begin{aligned}\exp\Big[&-\frac{1}{2}\log(\det(2\pi(\Sigma' + \tau_\phi^2 I_m))) - \frac{1}{2}(\tilde{\phi} - \mu_\phi)'(\Sigma' + \tau_\phi^2 I_m)^{-1}(\tilde{\phi} - \mu_\phi) \\ &+ \frac{1}{2}\log(\det(2\pi(\Sigma + \tau_\phi^2 I_m))) + \frac{1}{2}(\tilde{\phi} - \mu_\phi)'(\Sigma + \tau_\phi^2 I_m)^{-1}(\tilde{\phi} - \mu_\phi)\Big] \end{aligned} \tag{5}$$

where Σ' has been formed using β'_ϕ, the new proposal value.

Then for β_ξ we have:

$$\pi(\beta_\xi|\dots) \propto p(\tilde{\xi}|\mu_\xi, \tau_\xi^2, \Psi)p(\beta_\xi)$$

To update β_ξ we need to evaluate:

$$\begin{aligned}\exp\Big[&-\frac{1}{2}\log(\det(2\pi(\Psi' + \tau_\xi^2 I_m))) - \frac{1}{2}(\tilde{\xi} - \mu_\xi)'(\Psi' + \tau_\xi^2 I_m)^{-1}(\tilde{\xi} - \mu_\xi) \\ &+ \frac{1}{2}\log(\det(2\pi(\Psi + \tau_\xi^2 I_m))) + \frac{1}{2}(\tilde{\xi} - \mu_\xi)'(\Psi + \tau_\xi^2 I_m)^{-1}(\tilde{\xi} - \mu_\xi)\Big] \end{aligned} \tag{6}$$

where Ψ' has been formed using β'_ξ, the new proposal value.

6.3 ν

ν_ϕ also has a discrete update and so will look identical to the update for β_ϕ. It has conditional posterior distribution of:

$$\pi(\nu_\phi|\dots) \propto p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)p(\nu_\phi)$$

We need to evaluate:

$$\exp\Big[-\frac{1}{2}\log(\det(2\pi(\Sigma' + \tau_\phi^2 I_m))) - \frac{1}{2}(\tilde{\phi} - \mu_\phi)'(\Sigma' + \tau_\phi^2 I_m)^{-1}(\tilde{\phi} - \mu_\phi)$$
$$+ \frac{1}{2}\log(\det(2\pi(\Sigma + \tau_\phi^2 I_m))) + \frac{1}{2}(\tilde{\phi} - \mu_\phi)'(\Sigma + \tau_\phi^2 I_m)^{-1}(\tilde{\phi} - \mu_\phi)\Big] \quad (7)$$

where Σ' has been formed using ν'_ϕ, the new proposal value.

In order to update ν_ξ we have:

$$\pi(\nu_\xi | \dots) \propto p(\tilde{\xi}|\mu_\xi, \tau_\xi^2, \Psi)p(\nu_\xi)$$

We therefore need to evaluate:

$$\exp\Big[-\frac{1}{2}\log(\det(2\pi(\Psi' + \tau_\xi^2 I_m))) - \frac{1}{2}(\tilde{\xi} - \mu_\xi)'(\Psi' + \tau_\xi^2 I_m)^{-1}(\tilde{\xi} - \mu_\xi)$$
$$+ \frac{1}{2}\log(\det(2\pi(\Psi + \tau_\xi^2 I_m))) + \frac{1}{2}(\tilde{\xi} - \mu_\xi)'(\Psi + \tau_\xi^2 I_m)^{-1}(\tilde{\xi} - \mu_\xi)\Big] \quad (8)$$

where Ψ' has been formed using ν'_ξ, the new proposal value.

6.4 ς^2

ς_ϕ^2 has a conditional posterior distribution of:

$$\pi(\varsigma_\phi^2 | \dots) \propto p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)p(\varsigma_\phi^2)$$

The prior distribution of the log of ς_ϕ^2 is a univariate Normal. So to update ς_ϕ^2 we need to evaluate:

$$\exp\Bigg[-\frac{1}{2}\log(\det(2\pi(\Sigma'+\tau_\phi^2 I_m)))-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma'+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)-\frac{(\log(\varsigma_\phi^{2'})-m)^2}{2s^2}$$
$$+\frac{1}{2}\log(\det(2\pi(\Sigma+\tau_\phi^2 I_m)))+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^2 I_m)^{-1}(\tilde{\phi}-\mu_\phi)+\frac{(\log(\varsigma_\phi^{2})-m)^2}{2s^2}\Bigg] \quad (9)$$

where Σ' has been formed using $\varsigma_\phi^{2'}$, the new proposal value, and m and s are the prior mean and standard deviation repectively.

ς_ξ^2 has a conditional posterior distribution of:

$$\pi(\varsigma_\xi^2|\dots) \propto p(\tilde{\xi}|\mu_\xi, \tau_\xi^2, \Psi)p(\varsigma_\xi^2)$$

To update ς_ξ^2 we need to evaluate:

$$\exp\Bigg[-\frac{1}{2}\log(\det(2\pi(\Psi'+\tau_\xi^2 I_m)))-\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi'+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi)-\frac{(\log(\varsigma_\xi^{2'})-m)^2}{2s^2}$$
$$+\frac{1}{2}\log(\det(2\pi(\Psi+\tau_\xi^2 I_m)))+\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^2 I_m)^{-1}(\tilde{\xi}-\mu_\xi)+\frac{(\log(\varsigma_\xi^{2})-m)^2}{2s^2}\Bigg] \quad (10)$$

where Ψ' has been formed using $\varsigma_\xi^{2'}$, the new proposal value, and m and s are the prior mean and standard deviation repectively.

6.5 τ^2

τ_ϕ^2 has a conditional posterior distribution of:

$$\pi(\tau_\phi^2|\dots) \propto p(\tilde{\phi}|\mu_\phi, \tau_\phi^2, \Sigma)p(\tau_\phi^2)$$

The prior distribution of the log of τ_ϕ^2 is a univariate Normal. So to update τ_ϕ^2 we need to evaluate:

$$\exp\Bigg[-\frac{1}{2}\log(\det(2\pi(\Sigma+\tau_\phi^{2'}I_m)))-\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^{2'}I_m)^{-1}(\tilde{\phi}-\mu_\phi)-\frac{(\log(\tau_\phi^{2'})-m)^2}{2s^2}$$
$$+\frac{1}{2}\log(\det(2\pi(\Sigma+\tau_\phi^{2}I_m)))+\frac{1}{2}(\tilde{\phi}-\mu_\phi)'(\Sigma+\tau_\phi^{2}I_m)^{-1}(\tilde{\phi}-\mu_\phi)+\frac{(\log(\tau_\phi^{2})-m)^2}{2s^2}\Bigg] \quad (11)$$

where m and s are the prior mean and standard deviation repectively.

τ_ξ^2 has a conditional posterior distribution of:

$$\pi(\tau_\xi^2|\dots) \propto p(\tilde{\xi}|\mu_\xi,\tau_\xi^2,\Psi)p(\tau_\xi^2)$$

To update τ_ξ^2 we need to evaluate:

$$\exp\Bigg[-\frac{1}{2}\log(\det(2\pi(\Psi+\tau_\xi^{2'}I_m)))-\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^{2'}I_m)^{-1}(\tilde{\xi}-\mu_\xi)-\frac{(\log(\tau_\xi^{2'})-m)^2}{2s^2}$$
$$+\frac{1}{2}\log(\det(2\pi(\Psi+\tau_\xi^{2}I_m)))+\frac{1}{2}(\tilde{\xi}-\mu_\xi)'(\Psi+\tau_\xi^{2}I_m)^{-1}(\tilde{\xi}-\mu_\xi)+\frac{(\log(\tau_\xi^{2})-m)^2}{2s^2}\Bigg] \quad (12)$$

where m and s are the prior mean and standard deviation repectively.