Supplement Material for “Bayesian network–response regression”

S1 Posterior computation

Given the priors defined in equations (3)–(8) and based on the Pólya-Gamma data augmentation for Bayesian logistic regression, the Gibbs sampler for our network–response regression model in (1)–(2) alternates between the following steps.

- Update the Pólya-Gamma augmented data for each pair of brain regions \( l \) in every subject \( i \), from
  \[
  \omega_l^{(i)} | - \sim \text{PG} \left( 1, Z_l + \sum_{r=1}^R Y_{vr}^{(i)} Y_{ur}^{(i)} \right),
  \]
  for every \( l = 1, \ldots, V(V - 1)/2 \) and \( i = 1, \ldots, n \).

- Sample — for each subject \( i = 1, \ldots, n \) — the latent coordinates for her nodes comprising the \( V \times R \) matrix \( Y^{(i)} \). We accomplish this by block updating the elements in each row \( Y_v^{(i)} \), \( v = 1, \ldots, V \) of \( Y^{(i)} \) — representing the \( R \) coordinates of node \( v \) in subject \( i \) — given all the others \( u \neq v \). Recalling equations (1)–(2) we can obtain the full conditional posterior distribution for \( Y_v^{(i)} \), acting as a coefficient vector. In particular, let

  \[
  \mathcal{L}(A_i)_v \mid \pi_v^{(i)} \sim \text{Bern}\{\pi_v^{(i)}\},
  \]
  \[
  \text{logit}\{\pi_v^{(i)}\} = Z_v + Y_v^{(i)} Y_v^{(i)\top},
  \]

  where \( Y_v^{(i)} \) denotes the \( (V - 1) \times R \) matrix obtained by removing the \( v \)th row of \( Y^{(i)} \), while \( \mathcal{L}(A_i)_v \) and \( Z_v \) are \( (V - 1) \times 1 \) vectors obtained by stacking elements \( \mathcal{L}(A_i)_l \) and \( Z_l \) for all the \( l \) corresponding to pairs \( (u, v) \) such that \( u = v \) or \( w = v \), with \( u > w \) and ordered consistently with equation (11).

  Recalling equations (4)–(5), the prior for \( Y_v^{(i)\top} \) is \( \text{N}\{W(x_i)^\top G_v^\top, I_R\} \), with \( G \) the \( V \times K \) matrix of coefficients, \( W(x_i) \) the \( K \times R \) matrix containing the values of the basis functions at \( x_i \) and \( I_R \) the \( R \times R \) identity matrix. Hence, the Pólya-Gamma data augmentation for the model (11) ensures that the full conditional for each row of \( Y_v^{(i)} \) is

  \[
  Y_v^{(i)\top} \mid - \sim \text{N}_R\{\mu_v^{(i)}, \Sigma_v^{(i)}\}, \quad v = 1, \ldots, V,
  \]

  with

  \[
  \Sigma_v^{(i)} = \{I_R + Y_v^{(i)\top} (I_{v-v} Y_v^{(i)})^{-1} Y_v^{(i)}\}^{-1},
  \]
  \[
  \mu_v^{(i)} = \Sigma_v^{(i)} \{W(x_i)^\top G_v^\top + Y_v^{(i)\top} \psi_v^{(i)}\},
  \]
  \[
  \psi_v^{(i)} = \mathcal{L}(A_i)_v - 1/2I_{V-1} - \Omega_v^{(i)} Z_v,
  \]

  with \( \Omega_v^{(i)} \) the \((V-1)\times(V-1)\) diagonal matrix with entries obtained by stacking the Pólya-Gamma augmented data consistently with (11).

- Sample each shared similarity score \( Z_l \), \( l = 1, \ldots, V(V - 1)/2 \) from its Gaussian full conditional
  \[
  Z_l \mid - \sim \text{N}\left(\mu_l^Z, \sigma_l^2\right),
  \]
  where \( \sigma_l^2 = 1/\{\sigma_z^{-2} + \sum_{i=1}^n \omega_l^{(i)}\} \) and \( \mu_l^Z = \sigma_l^2 [\sigma_z^{-2} \mu_z + \sum_{i=1}^n \{\mathcal{L}(A_i)_l - 1/2 - \omega_l^{(i)} Y_v^{(i)\top} Y_v^{(i)\top}\}] \), with \( v \) and \( u \) the nodes corresponding to pair \( l \).
• Update each basis function $W_{kr}()$, $k = 1, \ldots, K$ and $r = 1, \ldots, R$ from its full conditional posterior. In particular, our Gaussian process prior (6) for the basis functions implies that

$$
\begin{pmatrix}
W_{kr}(x_1^*) \\
\vdots \\
W_{kr}(x_n^*)
\end{pmatrix}
\sim N_{n^*}
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix},
$$

where $(x_1^*, \ldots, x_n^*)$ are the unique values of $(x_1, \ldots, x_n)$ and $C$ is the Gaussian process covariance matrix with $C_{ij} = c(x_i^*, x_j^*)$. Hence, in updating $\{W_{kr}(x_1^*), \ldots, W_{kr}(x_n^*)\}^\top$, let $D = \text{diag}(\sum_i 1(x_i = x_1^*), \ldots, \sum_i 1(x_i = x_n^*))$, and

$$
\begin{pmatrix}
\sum_{i:x_i = x_1^*} Y_{1r}^{(i)} \\
\vdots \\
\sum_{i:x_i = x_n^*} Y_{1r}^{(i)} \\
\sum_{i:x_i = x_1^*} Y_{2r}^{(i)} \\
\vdots \\
\sum_{i:x_i = x_n^*} Y_{2r}^{(i)} \\
\vdots \\
\sum_{i:x_i = x_1^*} Y_{r}^{(i)} \\
\sum_{i:x_i = x_n^*} Y_{r}^{(i)}
\end{pmatrix}
, \quad
\begin{pmatrix}
W_{1r}(x_1^*) \\
\vdots \\
W_{1r}(x_n^*) \\
W_{2r}(x_1^*) \\
\vdots \\
W_{2r}(x_n^*) \\
\vdots \\
W_{Kr}(x_1^*) \\
\sum_{i:x_i = x_n^*} Y_{r}^{(i)}
\end{pmatrix}
$$

Standard conjugate posterior analysis provides the following full conditional

$$
\hat{W}_r | - \sim N_{K^{n^*}}(\mu^W_r, \Sigma^W_r),
$$

for each $r = 1, \ldots, R$, with

$$
\Sigma^W_r = (I_K \otimes C^{-1} + G^\top G \otimes D)^{-1},
\mu^W_r = \Sigma^W_r (G^\top \otimes I_{n^*}) \hat{Y}_r.
$$

• Conditioned on the hyperparameters $\tau_k$, the Gaussian prior on the elements of $G$ in equation (7) yields the following full conditional for each row of $G$:

$$
G^\top_v | - \sim N_K(\mu^G_v, \Sigma^G_v),
$$

for each $v = 1, \ldots, V$, with

$$
\Sigma^G_v = \{\tau + \sum_{i=1}^n W(x_i)W(x_i)\}^{-1},
\mu^G_v = \Sigma^G_v \sum_{i=1}^n W(x_i)Y_v \{i\}
$$

where $\tau = \text{diag}(\tau_1, \tau_2, \ldots, \tau_K)^\top$.

• The global shrinkage hyperparameters are updated as

$$
\tau_k | - \sim \text{Ga}\left(\alpha q^{3(k-1)} + \frac{V}{2} q^{2(k-1)} + \frac{1}{2} \sum_{v=1}^V G^2_{vk}\right),
$$

for each $k = 1, \ldots, K$.

• Update the subject-specific edge probabilities by applying equation

$$
\pi_l^{(i)} = \left\{1 + \exp(-Z_l - \sum_{r=1}^R Y_{vr}^{(i)} Y_{ur}^{(i)})\right\}^{-1},
$$

to the posterior samples of $Z_l$ and $Y^{(i)}$ for each $l = 1, \ldots, V(V - 1)/2$ and $i = 1, \ldots, n$. 

2