

generativeMRF

Collection of functions to play around with different Markov random field models as generative processes for brain activity data.

Generative Potts Model

This is a generative Potts model of brain activity data. The main idea is that the brain consists of K regions, each with a specific activity profile \mathbf{v}_k for a specific task set. The model consists of an arrangement model that tells us how the K regions are arranged in a specific subject s , and an emission model that provides a probability of the measured data, given the individual arrangement of regions.

Arrangement model

The brain is sampled in P vertices (or voxels). Individual maps are aligned using anatomical normalization, such that each vertex refers to a (roughly) corresponding region in each individual brain. The assignment of each brain location to a specific parcel in subject s is expressed as the random variable $u_i^{(s)}$.

Across individual brains, we have the overall probability of a specific brain location being part of parcel k .

$$p(u_i = k) = \mu_{ki}$$

The spatial interdependence of brain locations is expressed as a Potts model. In this model, the overall probability of a specific assignment of brain locations to parcels (the vector \mathbf{u}) is expressed as the product of the overall prior and the product of all possible pairwise potentials (ψ_{ij}).

$$p(\mathbf{u}) \propto \prod_i \mu_{u_i, i} \prod_{i \neq j} \psi_{ij}(u_i, u_j)$$

Each local potential is defined by an exponential over all other that are connected to node i , i.e. nodes with connectivity weights of $w_{ji} = w_{ij} > 0$.

$$\psi_{ij} = \exp(\theta_w \mathbf{u}_i^T \mathbf{u}_j w_{ij})$$

Where we have introduced a one-hot encoding of u_i with a K vector of indicator variables \mathbf{u}_i , such that $\mathbf{u}_i^T \mathbf{u}_j = 1$ if $u_i = u_j$ and 0 otherwise.

The spatial co-dependence across the entire brain is therefore expressed with the pairwise weights w that encode how likely two nodes belong to the same parcel. The temperature parameter θ_w determines how strong this co-dependence overall influences the local probabilities (relative to the prior). We can use this notation to express local co-dependencies by using a graph, where we define

$$w_{ij} = \begin{cases} 1; & \text{if } i \text{ and } j \text{ are neighbours} \\ 0; & \text{otherwise} \end{cases}$$

This formulation would enforce local smoothness of the map. However, we could also express in these potential more medium range potentials (two specific parietal and premotor areas likely belong to the same parcel), as well as cross-hemispheric symmetry. Given this, the matrix \mathbf{W} could be simply derived from the underlying grid or be learned to reflect known brain-connectivity.

In summary, we can express the prior probability of a specific arrangement in terms of a set of conditional probabilities

$$p(u_i | u_{j \neq i}) \propto \prod_i \mu_{u_i, i} \prod_{i \neq j} \psi_{ij}(u_i, u_j)$$

and we the corresponding conditional log-probability

$$l(u_i | u_{j \neq i}) \propto \log \mu_{u_i, i} + \theta_w \sum_{i \neq j} \mathbf{u}_i^T \mathbf{u}_j \mathbf{w}_{ij}$$

Emission model

The emission model should depend on the type of data that is measured. A common application is that the data measured at location i are the task activation in N tasks, arranged in the $N \times 1$ data vector \mathbf{y}_i . The averaged expected response for each of the parcels is \mathbf{v}_k . One issue of the functional activation is that the signal-to-noise ratio (SNR) can be quite different across different participants, and voxels, with many voxels having relatively low SNR. We model this signal to noise for each brain location (and subject) as

$$s_i \sim \text{Gamma}(\theta_\alpha, \theta_{\beta s}).$$

Overall, the expected signal at each brain location is then

$$\mathbf{E}(\mathbf{y}_i) = \mathbf{u}_i^T \mathbf{V} \mathbf{s}_i$$

Finally, relative to the signal, we assume that the noise is distributed i.i.d Gaussian with:

$$\boldsymbol{\epsilon}_i \sim \text{Normal}(0, \mathbf{I}_K \theta_{\sigma s})$$

Thus, log likelihood emission model can be written as

$$l(\mathbf{y}_i | \mathbf{u}_i) = -\frac{K}{2} \log(2\pi\theta_\sigma) - \frac{1}{2\theta_\sigma} \|\mathbf{y}_i - \mathbf{u}_i^T \mathbf{V} \mathbf{s}_i\|^2$$

Sampling from the prior or posterior distribution

The problem with determining the overall prior or posterior distribution of the model (for purposes of data generation or inference) cannot be easily be computed. We can evaluate the prior probability of a parcellation $p(\mathbf{U})$ or the posterior distribution $p(\mathbf{U} | \mathbf{Y})$ up to a constant of proportionality, with for example

$$p(\mathbf{U} | \mathbf{Y}; \theta) = \frac{1}{Z(\theta)} \prod_i \mu_{u_i, i} \prod_{i \neq j} \psi_{ij}(u_i, u_j) \prod_i p(\mathbf{y}_i | u_i)$$

Calculating the normalization constant $Z(\theta)$ (partition function, Zustandssumme, or sum over states) would involve summing this probability over all possible states, which for P brain locations and K parcels is K^P , which is intractable.

However, the conditional probability for each node, given all the other nodes, can be easily computed. Here the normalization constant is just the sum of the potential functions over the K possible states for this node

$$p(u_i | u_{j \neq i}, \mathbf{y}_i; \theta) = \frac{1}{Z(\theta)} \mu_{u_i, i} p(\mathbf{y}_i | u_i) \prod_{i \neq j} \psi_{ij}(u_i, u_j)$$

With Gibbs sampling, we start with a pattern $\mathbf{u}^{(0)}$ and then update $u_1^{(1)}$ by sampling from $p(u_1 | u_2^{(0)} \dots u_P^{(0)})$. We then sample $u_2^{(1)}$ by sampling from $p(u_2 | u_1^{(1)}, u_3^{(0)} \dots u_P^{(0)})$ and so on, until we have sampled each node once. Then we return to the beginning and restart the process. After some burn-in period, the samples will come from desired overall distribution. If we want to sample from the prior, rather than from the posterior, we simply drop the $p(\mathbf{y}_i | u_i)$ term from the conditional probability above.