Spatial Bayesian Latent Factor Regression Modelling of Coordinate-Based Meta-Analysis Data

Silvia Montagna[1], Tor Wager[2], Lisa Feldman Barrett[3], Timothy D. Johnson[4], and Thomas E. Nichols[1]

[1]Department of Statistics, University of Warwick, Coventry, UK
[2]Department of Psychology and Neuroscience, University of Colorado at Boulder, Boulder, CO, USA
[3]Department of Psychology, Northeastern University, Boston, MA, USA
[4]Biostatistics Department, University of Michigan, Ann Arbor, MI, USA
Email: s.montagna@warwick.ac.uk,t.e.nichols@warwick.ac.uk

Introduction

Now 20 years old, functional MRI (fMRI) has a large and growing literature that is best synthesised with meta-analytic tools. As most authors do not share image data, only the peak activation coordinates (foci) reported in the paper are available for Coordinate Based Meta-Analysis (CBMA). Neuroimaging meta-analysis is used to

- Identify areas of consistent activation
- Build a predictive model of task type or cognitive process for new studies (reverse inference)

To address these aims, we propose a Bayesian hierarchical model that is both simpler and more flexible than previous Bayesian Point Process models for CBMA [1;2]. Within our framework, it is also possible to account for the effect of study-level covariates (meta-regression), significantly expanding the capabilities of the current neuroimaging MA methods available.

Methods

We model the foci from each study as a "doubly stochastic" Poison process (Cox process) (Eq. 1), where the study-specific log intensity function is characterised as a linear combination of a 3-dimensional basis set; here we use 3D isotropic Gaussian kernels (Eq. 2). The functional representation leads to computational speed-ups over traditional log-Gaussian Cox processes [3]. We adaptively drop unnecessary bases by imposing sparsity on their coefficients through a latent factor regression model [4] (Eq. 3), and information on covariates is incorporated through a simple linear regression model on the latent factors (Eq. 4). By interpreting the latent factors as coefficients vectors, our construction becomes analogous to a functional principal component analysis (fPCA) representation of the log intensities (Eq. 5), but bases $\{\tilde{\phi}_m\}$ are no longer mutually orthogonal functions [4]. The number of latent bases (k) is estimated via posterior computation as in [4].

The latent factorisation is used as a vehicle to link the intensity functions to a study-type as part of a scalar-on-image regression. Specifically, suppose the MA dataset can be split into studies of type A and of type B (i.e., emotion vs. executive control studies). We build a probit regression model to predict the study type, where the probability that study i is of type A is expressed as a function of the latent factors (Eq. 6). Along the same lines, we can easily accommodate prediction of two or more types of studies. Our fully Bayesian CBMA model permits explicit calculation of posterior predictive distribution for study type and, as a result, allows inference on the most likely domain for any new experiment by just using its foci.

Bayesian hierarchical model for point pattern data

Notation & method:

$\{\boldsymbol{\nu}_i, \mathbf{x}_i, Y_i\} = \{$foci in study $i = 1, \ldots, n$; vector of study-specific covariates; study type$\}$
$\mathcal{B}$ = analysis region, $\mathcal{B} \subset \mathbb{R}^3$, with volume $|\mathcal{B}|$
$\{\boldsymbol{\theta}_i, \boldsymbol{\Lambda}, \boldsymbol{\eta}_i, \boldsymbol{\Sigma}, \boldsymbol{\beta}, \alpha, \boldsymbol{\gamma}\}$ = model parameters

$$\pi(\{\boldsymbol{\nu}_i\}_{i=1}^n | \{\mu_i\}_{i=1}^n) \propto \exp\left\{-\sum_{i=1}^n \int_{\mathcal{B}} \mu_i(s)ds\right\} \prod_{i=1}^n \prod_{\boldsymbol{\nu}_{ij} \in \boldsymbol{\nu}_i} \mu_i(\boldsymbol{\nu}_{ij}), \tag{1}$$

$$\log \mu_i(\boldsymbol{\nu}) = \sum_{m=1}^p \theta_{im} b_m(\boldsymbol{\nu}) = \mathbf{b}(\boldsymbol{\nu})^\top \boldsymbol{\theta}_i \tag{2}$$

$b_m(\boldsymbol{\nu}) = \exp\{-\psi ||\boldsymbol{\nu} - \boldsymbol{\phi}_m||^2\}$, with $\{\psi, \boldsymbol{\phi}_m\}$ fixed

$$\boldsymbol{\theta}_i = \boldsymbol{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\zeta}_i, \quad \text{with} \quad \boldsymbol{\zeta}_i \sim N_p(0, \boldsymbol{\Sigma}), \quad \text{and} \quad \boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_p^2) \tag{3}$$

$$\boldsymbol{\eta}_i = \boldsymbol{\beta}^\top \mathbf{x}_i + \boldsymbol{\Delta}_i, \quad \boldsymbol{\Delta}_i \sim N_k(0, \mathbf{I}) \tag{4}$$

Equivalent (fPCA analogue) construction:

$$\log \mu_i(\boldsymbol{\nu}) = \sum_{m=1}^k \eta_{im} \tilde{\phi}_m(\boldsymbol{\nu}) + r_i(\boldsymbol{\nu}), \quad \text{with} \quad \tilde{\phi}_m(\boldsymbol{\nu}) = \sum_{l=1}^p \lambda_{lm} b_l(\boldsymbol{\nu}) \quad \text{and} \quad r_i(\boldsymbol{\nu}) = \sum_{l=1}^p \zeta_{il} b_l(\boldsymbol{\nu}) \tag{5}$$

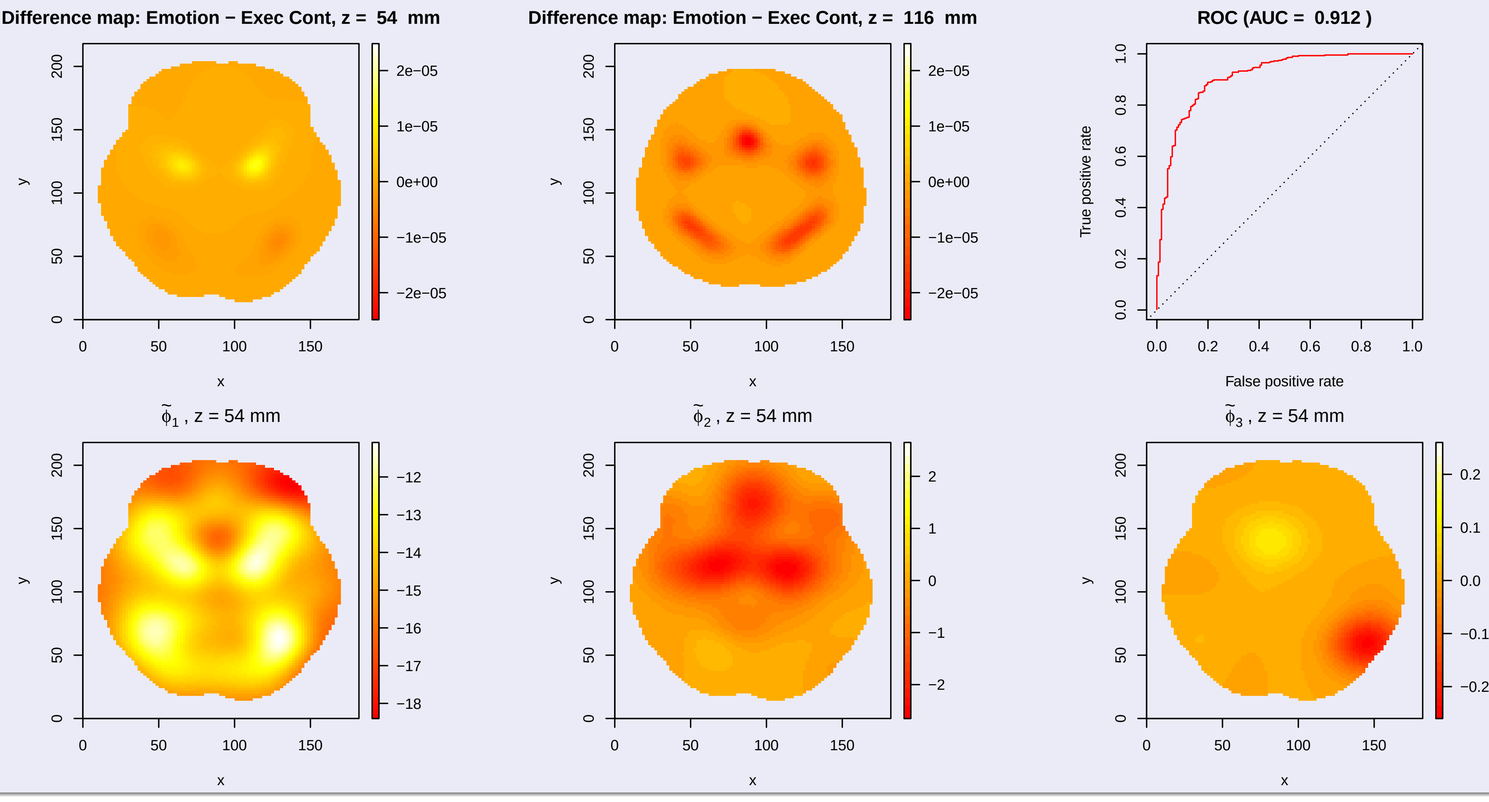
Inverse inference:

$$Y_i = \begin{cases} 1 & \text{if study } i \text{ is of type A} \\ 0 & \text{if study } i \text{ is of type B} \end{cases} \quad \text{with} \quad Pr(Y_i = 1 | \alpha, \boldsymbol{\gamma}, \boldsymbol{\eta}_i) = \Phi(\alpha + \boldsymbol{\gamma}^\top \boldsymbol{\eta}_i) \tag{6}$$

Data & results

The dataset used for our analysis consists of 1199 total contrasts categorised as either *emotion* (860 contrasts, 6481 foci) or *executive control* (339 contrasts, 4332 foci). We split the data into a training set, for which both foci and study type are retained for the analysis, and a testing test (50%). We used $p = 352$ bases for this analysis.

The first two images in the top row show the mean difference intensity map for (emotion - executive control) at $z = \{54, 116\}$ mm. These maps reflect the degree of consistency with which a region is activated by either task. The high-intensity regions at slice $z = 54$ are centred in the amygdalae, thus revealing stronger activation of emotion contrasts. Executive control tasks are more strongly associated with activation in more superior slices with a bilateral network. The ROC curve shows that our Bayesian classifier does an excellent job at predicting the study outcome. The bottom row shows 3 bases ($\tilde{\phi}_m$ in Eq. 5). Similar to fPCA, the variation in log intensity function explained by each basis decreases as the basis index increases.



Conclusions

The proposed model provides a unified approach for reverse inference by jointly modelling neuroimaging point pattern data with study types while explicitly accounting for the spatial structure of the data. Although the presented is for 2-class CBMA data, the proposed framework can be easily modified for joint modelling of data of many different types, e. g., the probit model for a binary outcome can be replaced by an appropriate predictive model for categorical, nominal, or continuous study features.

References

[1] Kang, J., et al., (2011), Meta-analysis of Functional Neuroimaging Data via Bayesian Spatial Point Processes, *Journal of the American Statistical Association*, vol. 106, no. 493, pp. 124–134
[2] Kang, J., et al., (2014), A Bayesian Hierarchical Spatial Point Process Model for Multitype Neuroimaging Meta-analysis, *The Annals of Applied Statistics*, vol. 8, no. 3, pp. 1800–1824
[3] Møller, J., et al., (1998), Log-Gaussian Cox Processes, *Scandinavian Journal of Statistics*, vol. 25, no. 3, pp. 451–482
[4] Montagna, S., et al. (2012), Bayesian Latent Factor Regression for Functional and Longitudinal Data, *Biometrics*, vol. 68, no. 4, pp. 1064–1073