NeuroImage 86 (2014) 294–305

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Structurally-informed Bayesian functional connectivity analysis

Max Hinne [a,b,*], Luca Ambrogioni [a], Ronald J. Janssen [a], Tom Heskes [b], Marcel A.J. van Gerven [a]

[a] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands
[b] Radboud University Nijmegen, Institute for Computing and Information Sciences, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 30 September 2013
Available online 10 October 2013

Keywords:
Functional connectivity
Structural connectivity
G-Wishart prior
Bayesian inference

ABSTRACT

Functional connectivity refers to covarying activity between spatially segregated brain regions and can be studied by measuring correlation between functional magnetic resonance imaging (fMRI) time series. These correlations can be caused either by direct communication via active axonal pathways or indirectly via the interaction with other regions. It is not possible to discriminate between these two kinds of functional interaction simply by considering the covariance matrix. However, the non-diagonal elements of its inverse, the precision matrix, can be naturally related to direct communication between brain areas and interpreted in terms of partial correlations. In this paper, we propose a Bayesian model for functional connectivity analysis which allows estimation of a posterior density over precision matrices, and, consequently, allows one to quantify the uncertainty about estimated partial correlations. In order to make model estimation feasible it is assumed that the sparseness structure of the precision matrices is given by an estimate of structural connectivity obtained using diffusion imaging data. The model was tested on simulated data as well as resting-state fMRI data and compared with a graphical lasso analysis. The presented approach provides a theoretically solid foundation for quantifying functional connectivity in the presence of uncertainty.

Introduction

One of the oldest and most influential ideas in cognitive neuroscience is that the brain, and in particular the cortex, can be divided into specialized functional regions (Friston, 2011). In recent times, the neuroscience community has become increasingly interested in determining how these regions are organized as large functional networks and how their modulation reflects ongoing cognitive processing (Bullmore and Sporns, 2009). The organization of these functional networks can be described using the umbrella term 'functional connectivity', defined as the deviations from statistical independence between distributed and often spatially remote neuronal units (Craddock et al., 2013; Friston, 1994). Despite the indirect nature of the blood oxygenation level dependent (BOLD) signal, functional magnetic resonance imaging (fMRI) has proven to be able to extract patterns of co-activation between clusters of voxels (Lowe et al., 2000).

The easiest way to operationalize the notion of functional connectivity is to calculate a covariance matrix which, in case of standardized variables, is equivalent to the correlation structure between brain regions. However, this approach is not able to identify direct (monosynaptic) functional connections as it is also sensitive to indirect (polysynaptic) functional interactions. For example, if regions A and B as well as regions B and C display correlated activity, then A and C will also show correlated activity even if they are not directly connected (Smith, 2012; Varoquaux and Craddock, 2013).

In contrast, the precision matrix, defined as the inverse of the covariance matrix, captures conditional independence between brain regions (Lauritzen, 1996; Whittaker, 2009). That is, elements of the precision matrix are related to partial rather than full correlations and zero elements of the precision matrix imply an absence of direct functional connectivity. Therefore, sparse precision matrices provide us with valuable information about how different regions interact, though the estimates need to be interpreted with care (Friston, 2011; Hutchison et al., 2013; Marrelec and Benali, 2009; Woolrich and Stephan, 2013).

A common approach to obtain a point estimate for a sparse precision matrix is by means of the graphical lasso (Friedman et al., 2008; Smith et al., 2011; Varoquaux et al., 2010), which achieves sparseness through ℓ_1 regularization. Although the graphical lasso provides a reasonable point estimate, it is biased due to the induced shrinkage of the partial correlations. Furthermore, it does not directly provide a measure of uncertainty regarding the partial correlation estimates. This could lead to possibly erroneous conclusions about functional connectivity.

From a Bayesian perspective we are interested in the posterior density of the precision matrix given observed data. Ultimately, this should lead to more reliable inferences about a subject's cognitive state. In order to facilitate the estimation problem, we will not resort to shrinkage, as in the graphical lasso. Rather, we assume that the conditional independence structure between brain regions is given by an independent estimate of structural connectivity.

Structural connectivity refers to the presence of white matter tracts between spatially segregated brain regions (Hagmann et al., 2008). In

* Corresponding author at: Radboud University Nijmegen, Faculty of Science, Institute for Computing and Information Sciences, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands. Fax: +31 24 365 27 28.
E-mail address: mhinne@cs.ru.nl (M. Hinne).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.09.075

humans, these tracts can be estimated in vivo by diffusion weighted imaging (DWI) which measures the anisotropy in the diffusion of water molecules (Le Bihan et al., 2001). The final result is usually a binary undirected graph which reports whether or not two areas are structurally connected. Clearly two brain regions can be directly functionally coupled only if they are physically connected, therefore the concepts of functional and structural connectivity are intimately related (Damoiseaux and Greicius, 2009). The idea is to infer structural connectivity from DWI data and use it as an additional constraint in our Bayesian model. The validity of this approach is supported by several recent experimental studies which found a substantial overlap between structural and functional networks both inside specific cortical areas (Koch et al., 2002) and on a whole brain scale (Cabral et al., 2012; Damoiseaux and Greicius, 2009; Greicius et al., 2009; Hagmann et al., 2008; Honey et al., 2007, 2009). Related approaches have been used before in the context of functional and effective connectivity analysis (Deligianni et al., 2011; Ng et al., 2012; Stephan et al., 2009).

In the following we present a new Bayesian framework for estimating functional connectivity. The framework, which we refer to as Bayesian functional connectivity (BFC) analysis, makes use of a *G*-Wishart prior (Roverato, 2002). This prior allows the sparseness structure of estimated precision matrices to be determined by a graph *G*, corresponding to structural connectivity. BFC analysis then amounts to computing a posterior density over sparse precision matrices. This posterior may then be used to compute marginal densities for partial correlations of interest. Our approach is compared with existing approaches using both simulated data and empirical data. We show that our approach provides robust partial correlation estimates while at the same time quantifying the uncertainty about functional connectivity.

Materials and methods

Conventional functional connectivity estimation

Traditionally, functional connectivity estimation has relied on estimating covariance structure between p brain regions from time series data $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$. Each vector $\mathbf{x}_n = (x_{n1}, \ldots, x_{np})$ reflects neuronal activity (e.g. BOLD responses) for p brain regions. Without loss of generality, we assume that data is standardized to have zero mean and unit standard deviation such that covariance coincides with correlation. It is assumed that the data are generated according to a zero-mean multivariate Gaussian density

$$p(\mathbf{X}|\mathbf{\Omega}) = \prod_n^N \mathcal{N}(\mathbf{x}_n|0,\mathbf{\Omega}) \propto |\mathbf{\Omega}|^{N/2} \exp\left[-\frac{1}{2}\langle \mathbf{S\Omega}\rangle\right] \tag{1}$$

with precision (inverse covariance) matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$, scatter matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ and trace operator $\langle \cdot \rangle$. The choice of this distribution is justified by the fact that it is the maximum entropy distribution among all distributions with a specified mean and covariance (Cover and Thomas, 2006). Alternatively, the likelihood may be characterized in terms of the scatter matrix $\mathbf{S}$ which follows a Wishart distribution $\mathcal{W}_p(\mathbf{\Sigma}, N)$ if its density is

$$p(\mathbf{S}|\mathbf{\Sigma}, N) = \frac{|\mathbf{S}|^{N/2}}{Z(N,\mathbf{\Sigma})} \exp\left[-\frac{1}{2}\left\langle \mathbf{S\Sigma}^{-1}\right\rangle\right], \tag{2}$$

with $Z(N,\mathbf{\Sigma})$ the normalizing constant. This perspective can be applied more easily for distributions with a mean different from zero (Anderson, 1984).

We focus on estimating the precision matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ rather than the covariance matrix. As mentioned before, zero elements in $\mathbf{\Omega}$ reflect the absence of direct interactions. More formally, the sparseness structure of $\mathbf{\Omega}$, represented in terms of an undirected graph G where $V(G)$ is a set of nodes and $E(G)$ is a set of undirected edges between nodes, is equivalent to the conditional independence structure of a Gaussian Markov random field (Lauritzen, 1996; Whittaker, 2009). In other words, in the context of connectivity analysis, $\omega_{ij} = 0$ corresponds to the absence of structural connectivity between brain regions i and j.

In order to estimate the precision matrix $\mathbf{\Omega}$ of a zero-mean multivariate Gaussian density from data $\mathbf{X}$ one may maximize the log likelihood

$$\log p(\mathbf{X}|\mathbf{\Omega}) = \frac{1}{2}[N \log|\mathbf{\Omega}| - \langle \mathbf{S\Omega}\rangle]$$

which gives the maximum likelihood estimate (MLE):

$$\hat{\mathbf{\Omega}} = \arg\max_{\mathbf{\Omega}\in M^+} [N \log|\mathbf{\Omega}| - \langle \mathbf{S\Omega}\rangle] = N\mathbf{S}^{-1} \tag{3}$$

where the maximization is constrained to precision matrices in the family of $p \times p$ positive definite matrices M^+.

In practice, however, this empirical estimate does not contain zero elements. Furthermore, in case $N < p$, the maximum likelihood solution does not exist since $\mathbf{S}/N$ becomes singular. Even in case $N > p$, the MLE is often poorly behaved, and regularization is called for (Pourahmadi, 2011). The graphical lasso (Friedman et al., 2008) regularizes the preceding MLE through sparsification by solving

$$\hat{\mathbf{\Omega}} = \arg\max_{\mathbf{\Omega}\in M^+} \left[\log|\mathbf{\Omega}| - \frac{1}{N}\langle \mathbf{S\Omega}\rangle - \lambda\|\mathbf{\Omega}\|_1\right]. \tag{4}$$

The employed ℓ_1 regularizer encourages sparse precision matrices as determined by the regularization parameter λ. This maximization problem can be solved using established coordinate descent methods (Friedman et al., 2008). The graphical lasso has been proposed as the method of choice for functional connectivity estimation (Smith et al., 2011; Varoquaux and Craddock, 2013; Varoquaux et al., 2010).

Even though the graphical lasso is commonly used to estimate sparse precision matrices, it suffers from two issues. First, since the graphical lasso employs shrinkage, pushing precision values towards zero, the resulting functional connectivity estimate is biased. Second, the graphical lasso produces a point estimate which does not directly allow inferences to be drawn about the uncertainty in our estimates arising from sampling noise and finite sample size.

Bayesian functional connectivity estimation

In order to tackle the aforementioned issues, we developed a Bayesian framework for inferring functional connectivity which does not rely on shrinkage but rather assumes that the sparseness structure G of $\mathbf{\Omega}$ is given (Dempster, 1972). Specifically, we assume that the graph G is given by the structural connectivity as estimated from DWI data.

We start by assuming a *G*-Wishart distribution as the conjugate prior on precision matrices $\mathbf{\Omega}$. The *G*-Wishart is defined for the cone $M^+(G)$ of positive-definite symmetric matrices with off-diagonal elements $\omega_{ij} = 0$ whenever $(i,j) \notin E(G)$. A zero-constrained random matrix $\mathbf{\Omega}$ has the *G*-Wishart distribution $\mathcal{W}_G(\delta_0, \mathbf{D})$ if its density is (Wang and Li, 2012):

$$p(\mathbf{\Omega}|G) = \frac{|\mathbf{\Omega}|^{(\delta_0-2)/2}}{Z_G(\delta_0,\mathbf{D})} \exp\left(-\frac{1}{2}\langle \mathbf{D\Omega}\rangle\right) 1_{\{\mathbf{\Omega}\in M^+(G)\}}$$

where δ are the prior degrees of freedom, $\mathbf{D}$ a symmetric positive definite prior scatter matrix, and $Z_G(\delta,\mathbf{D})$ the normalizing constant. The indicator function 1_x evaluates to 1 if its argument x is true and to 0 if its argument is false. In our experiments, we set $\delta_0 = 3$ and choose $\mathbf{D} = \mathbf{I}_{p\times p}$ (Moghaddam et al., 2009). This amounts to a vague prior for the precision matrix in Eq. (1), except that its support is restricted by *G*. We may now use Bayes' rule to obtain the posterior density for $\mathbf{\Omega}$ according to

$$p(\mathbf{\Omega}|\mathbf{X},G) \propto p(\mathbf{X}|\mathbf{\Omega})p(\mathbf{\Omega}|G) = \frac{|\mathbf{\Omega}|^{(\delta_n-2)/2}}{Z_G(\delta_n,\mathbf{B})} \exp\left(-\frac{1}{2}\langle \mathbf{B\Omega}\rangle\right) 1_{\{\mathbf{\Omega}\in M^+(G)\}}. \tag{5}$$