

Relating brain functional connectivity to anatomical connections: Model Selection

Fani Deligianni¹, Gael Varoquaux^{2,3}, Bertrand Thirion³, Emma Robinson¹,
David J. Sharp⁴, A. David Edwards⁵, Daniel Rueckert¹

¹ Department of Computing, Imperial College London, UK

² INSERM U922, Neurospin, CEA Saclay, France

³ Parietal project team, INRIA, Saclay, France

⁴ C3NL, Division of Experimental Medicine, Imperial College London, UK

⁵ Institute of Clinical Sciences, Imperial College London, UK

Abstract. We aim to learn across several subjects a mapping from brain anatomical connectivity to functional connectivity. Following [1], we formulate this problem as estimating a multivariate autoregressive (MAR) model with sparse linear regression. We introduce a model selection framework based on cross-validation. We select the appropriate sparsity of the connectivity matrices and demonstrate that choosing an ordering for the MAR that lends to sparser models is more appropriate than a random. Finally, we suggest randomized LASSO in order to identify relevant anatomo-functional links with better recovery of ground truth.

1 Introduction

Disruption of brain connectivity have been implicated in a number of diseases [2] including schizophrenia, ADHD, autism and brain trauma. This has spurred interest in network organization and dynamics, studied with diffusion weighted MRI (DWI) or resting-state functional MRI (rs-fMRI) [3]. Multi-modal imaging, integrating both functional and anatomical descriptions can yield a more detailed picture of brain architecture. Model-based studies have shown links between anatomical and functional whole-brain connectivity [4]. Here we investigate data-driven predictive modeling from anatomical to functional networks to uncover mechanisms ignored in studies with a-priori defined interactions.

We set the problem as the prediction of whole-brain resting-state functional connectivity from anatomical brain connectivity. In a population of S subjects, we consider the connectivity between a set of N regions of interest (ROIs). Our supervised learning task is to link across subjects the anatomical connections between these brain regions, as estimated by tractography, to functional connections, *i.e.* the observed synchronization in the brain activity observed via rs-fMRI. We represent the set of anatomical connections for a subject s as a symmetric connectivity matrix $\mathbf{A}^s \in \text{Sym}_N$. The corresponding brain activity in the ROIs is summarized by N time series of length t , $\mathbf{X}^s \in \mathbb{R}^{N \times t}$. We want to

explain the correlation structure of the observed fMRI time-series \mathbf{X}^s from the subject’s anatomical connectivity matrix \mathbf{A}^s .

In [1] we introduced a learning framework using multiple penalized regression to link \mathbf{A}^s and the covariance matrix of the fMRI data $\mathbf{\Sigma}^s \in \mathbb{R}^{N \times N}$. The generative model of the functional signal relies on autoregressive Gaussian processes spanning a graphical model, the Markov structure of which is restricted by the anatomical connectivity. Estimating the fMRI model is then a covariance selection problem, computing the maximum likelihood estimate of the precision–inverse covariance– matrix, \mathbf{K} parameterize a multivariate Gaussian model, subject to the conditional independence constraints. The output space of the regression from anatomical to functional connectivity, is then the set of symmetric positive definite (SPD) matrices. We take an unconstrained parametrization of SPD matrices using a matrix square-root of the precision matrix \mathbf{K} to be predicted. To do that we use a Cholesky decomposition of \mathbf{K} , which is not invariant by a permutation of the rows and columns of \mathbf{K} .

Here, we address the selection of the optimal ordering in this framework. Using cross-validation we show that a permutation of \mathbf{K} that favors a sparser Cholesky decomposition results in higher prediction performance than a random ordering. Subsequently, we select the optimal sparsity of the joint anatomical and functional graphs. Finally, we introduce the randomized Lasso for model identification: recovery of the underlying anatomo-functional links.

2 Generative model and learning strategy

Generative model. We use multivariate autoregressive models (MAR) as a generative model for fMRI time series [1]. They imply a Markov structure between the different time series: conditional independences forming an undirected graphical model. Each ROI is represented as a node. Nodes that are not connected have mutually independent time-series, conditionally on their neighbors. MAR models of stationary process are parametrized by a transition matrix, \mathbf{T} , the square of which is directly related to the precision matrix of the signal: $\mathbf{K} = \mathbf{B}^T \mathbf{B}$, where $\mathbf{B} = \mathbf{I} - \mathbf{T}$. Note that \mathbf{B} is a matrix square root of the precision matrix.

Markov structure. Conditional independence between variables is given by the zeros in the precision matrix. Estimating from the fMRI data a sparse precision matrix reduces the small-sample estimation error present in the empirical covariance matrix [5]. We use anatomical connections that fail to be significantly positive across subjects to impose the Markov structure of the fMRI model [1]. The maximal likelihood estimate of \mathbf{K} is then computed using the iterative proportional scaling algorithm [6].

Imposing definite positive predictions. To parameterize a valid Gaussian model, the predicted precision matrix must be positive definite. This condition on the output space corresponds to constraints on the coefficients of \mathbf{K} : we cannot predict independently these coefficients and ensure that the resulting matrix will be in Sym_N^+ . Thus, rather than predicting \mathbf{K} , we predict a square root, \mathbf{B} , that is not constrained. Here, we use the Cholesky decomposition of the precision matrix \mathbf{K}

to estimate \mathbf{B} . The Cholesky decomposition is not invariant to a permutation of the rows and columns of the input matrix \mathbf{K} . We use the approximate minimum degree (AMD) ordering [7]: a permutation of \mathbf{K} favoring a sparser Cholesky decomposition. As it only depends on the support of the matrix \mathbf{K} , it is identical across subjects. Note that we are interested in predicting correlation and not covariance, thus, we rescale the diagonal of \mathbf{B} to ones: $\tilde{\mathbf{B}} = \mathbf{B} \text{diag}(\mathbf{B})^{-1}$. This amounts to setting the innovation terms of the MAR to one.

Statistical learning. We use a multivariate linear model between all the anatomical connectivity values $\mathbf{a} = \{\mathbf{A}_{i,j}\}$ and the coefficients of the matrix $\tilde{\mathbf{B}}$, $\mathbf{b} = \{\tilde{\mathbf{B}}_{i,j}\}$. The learning problem takes the form of multiple regressions:

$$\mathbf{b}_k = \beta_{k,0} + \sum_{s=1}^S \beta_{s,j} \mathbf{a}_j, \quad (1)$$

where β_0 is the intercept and β_k are the coefficients relating the functional connection \mathbf{b}_k to the whole-brain anatomical connectivity, \mathbf{a} . There are many more candidate anatomical connections than subjects: we are in high-dimensional settings. Therefore we resort to the Lasso: sparse penalized regression. We use the LARS algorithm, implemented in the scikit-learn toolbox [8].

3 Model selection and identification

Model selection metric. We work on the space of sparse Gaussian models, parametrized by correlation or precision matrices, symmetric definite positive matrices. This space, Sym_N^+ , is not a vector space. The standard Euclidean distance on matrices, the Frobenius norm, does not account for the geometry of this space. Thus, it is ill suited to quantify prediction errors. However, Sym_N^+ can be parametrized as a Riemannian manifold using an intrinsic metric [1, 9]:

$$d_{AR}(\mathbf{C}, \mathbf{D})^2 = \text{tr}(\log \mathbf{C}^{-\frac{1}{2}} \mathbf{D} \mathbf{C}^{-\frac{1}{2}})^2 \quad (2)$$

This metric is invariant under affine scaling and inversion of the matrices, and is thus independent of the parametrization of the Gaussian models.

Choice of Markov structure. We set the same Markov structure across individuals, testing with a T-statistic for each anatomical connection whether it is non zero in the group. The threshold controls the sparsity of the MAR model for fMRI. We set it using random hold-out cross validation with 10% held out.

Randomized Lasso. We use the randomized Lasso [10] to identify non-zero coefficients β_k . It is a generalization of Lasso with better recovery properties. The randomized Lasso estimate is computed by solving the Lasso problem with weights W_j random within specified bounds:

$$\hat{\beta} = \underset{\beta}{\text{argmin}} \left(\sum_{s=1}^S (\mathbf{b}_s^k - \beta_0 - \sum_j \beta_j \mathbf{a}_{s,j})^2 + \lambda \sum_{j=1}^N \frac{|\beta_j|}{W_j} \right) \quad (3)$$

Markov model sparsity		80%	70%	60%	50%	40%	30%	20%	10%
Random order	d_{AI}	25.0	26.0	27.8	29.5	34.1	39.2	50.6	67.2
	std	32.4	8.0	8.2	8.4	8.2	8.1	8.1	8.0
AMD order	d_{AI}	31.3	24.5	24.0	25.3	26.1	27.3	42.5	63.2
	std	49.3	8.2	8.4	8.4	9.0	9.5	9.6	9.8

Table 1: Prediction performance, measured with the d_{AI} metric, for different sparsity of the Markov structure and under different ordering of \mathbf{K} .

This randomized penalization regression is solved many times. The probability that a functional connection is related to a anatomical connection is then given by the fraction of times the coefficient is selected during the repetitions.

4 Experimental results

Brain connectivity analysis was performed in 26 normal adults, with the imaging parameters, pre-processing, definition of ROIs and extraction of structural networks described previously [1]. To set the level of support that gives the optimum prediction performance, we vary the threshold on the t-test so that the percentage of connections selected ranges from 10% to 80%. We also compare two choices of ordering for the Cholesky decomposition of the precision matrix K : random ordering and model averaging, or AMD ordering for sparser Cholesky factors. Cross-validation results (Table 1) show that the optimum sparsity for the Markov structure corresponds to selecting 60% of the connections, and that AMD ordering is to be preferred to averaging on random orderings. Fig. 1 shows structural connections associated with the default mode network. The diameter of the tubes is associated with the probability of the connection to be selected.

5 Conclusions

We use the probabilistic framework introduced in [1] to learn a predictive model from anatomical to functional brain connectivity. Here, we use cross-validation to set the properties of the Markov model of fMRI: filling factor, and ordering of the nodes in the MAR. As we are interested in recovering the relevant anatomic-functional links, we also introduce randomized Lasso that can identify better non-zero coefficients in linear models with correlated designs.

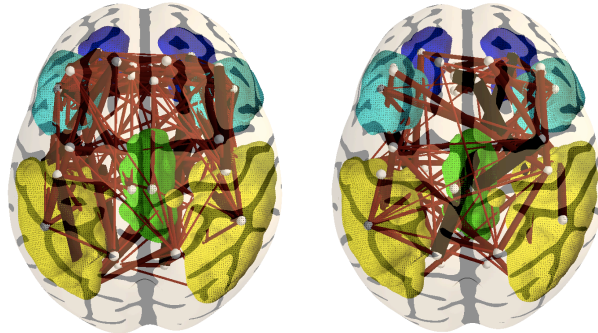


Fig. 1: Structural connections associated with the default mode network: **left** random order; **right** AMD.

References

1. F. Deligianni, et al. : A probabilistic framework to infer brain functional connectivity from anatomical connections. *IPMI* (2011) 296–307
2. O. Sporns: The non-random brain: efficiency, economy, and complex dynamics. *Front Comput Neurosc* **5** (2011) 5
3. E. Bullmore and O. Sporns: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* **10** (2009) 186–198
4. C. Honey, et al. : Predicting human resting-state functional connectivity from structural connectivity. *P Natl Acad Sci Usa* **106**(6) (2009) 2035–2040
5. G. Varoquaux, et al. : Brain covariance selection: better individual functional connectivity models using population prior. *NIPS* (2010)
6. S. Lauritzen: *Graphical models*. Oxford University Press, USA (1996)
7. P. Amestoy, et al. : An approximate minimum degree ordering algorithm. *SIAM Journal on Matrix Analysis and Applications* **17**(4) (1996) 886–905
8. F. Pedregosa, et al. : Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research* **to appear** (2011) pub ahead of print
9. G. Varoquaux, et al. : Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In: *MICCAI*. (2010)
10. N. Meinshausen and P. Bühlmann: Stability selection. *Journal of the Royal Statistical Society: Series B* **27** (2010) 417–473