Learning Brain fMRI Structure Through Sparseness and Local Constancy

Jean Honorio[a,b] Luis Ortiz[a] Dimitris Samaras[a] Rita Goldstein[b]
[a] Computer Science Department, Stony Brook University, NY
[b] Medical Department, Brookhaven National Laboratory, NY

1 Objective

We propose sparse and locally constant Gaussian graphical models as well as structural equation models for learning the functional connectivity in the whole-brain.

fMRI datasets are typically under-sampled and high-dimensional, they often need to be represented with low-complexity statistical models, which are comprised of only the important probabilistic dependencies. Most methods attempt to reduce model complexity by enforcing structure sparseness. However, sparseness cannot describe inherent regularities in the structure. Locality information is one aspect of fMRI datasets that has not been modeled in the past. In this paper, we propose adding sparseness and local constancy priors in two classes of models: Gaussian graphical models (GGM) and structural equation models (SEM).

Different methods have been proposed for finding functionally connected networks. Most of these methods do not allow a fully exploratory analysis, i.e. without a priori set of regions of interest. Methods which focus on linear dependencies or correlations include seed-region [1], structural equation models [2, 3] and Granger causality maps [4]. Dynamic Bayesian networks was proposed in [5, 6] for finding non-linear interactions among brain regions.

2 Method

The problem of learning functional connectivity can be modeled as a maximum likelihood estimation of the inverse covariance matrix given a dataset [7]. We encourage finding probabilistic connectivities between two close or distant clusters of voxels, instead of between isolated voxels. This is done by adding sparseness and local constancy priors on the estimation process. It was shown that this method leads to better generalization performance in a cross-validation setting.

In this paper, we propose regularized regression as a method for learning the structure of a structural equation model. We impose constraints on the difference of entries in the coefficient matrix $\mathbf{\Theta} \in \mathbb{R}^{N \times N}$ for N variables, which correspond to spatially neighboring variables. Let $\mathbf{X} \in \mathbb{R}^{N \times S}$ be the dataset for S subjects and $\mathbf{D} \in \mathbb{R}^{M \times N}$ be the discrete derivative operator on the manifold, where $M \in O(N)$ is the number of spatial neighborhood relationships. For instance, in a 3D image, M is the number of voxel pairs that are spatial neighbors on the manifold. More specifically, if voxel n_1 and voxel n_2 are spatial neighbors, we include a row m in $\mathbf{D}$ such that $d_{mn_1} = 1$, $d_{mn_2} = -1$ and $d_{mn_3} = 0$ for $n_3 \notin \{n_1, n_2\}$. The following penalized regression is proposed:

$$\min_{\mathbf{diag}(\mathbf{\Theta})=\mathbf{0}} \left(\frac{1}{2} \left\| \mathbf{\Theta}^{\mathrm{T}} \mathbf{X} - \mathbf{X} \right\|_{\mathfrak{F}}^2 + \rho \|\mathbf{\Theta}\|_1 + \tau \|\mathbf{D}\mathbf{\Theta}\|_1 \right) \tag{1}$$

for some $\rho, \tau > 0$. The first term measures the quality of the regression, the second term $\rho\|\mathbf{\Theta}\|_1$ encourages sparseness, and the third term $\tau\|\mathbf{D}\mathbf{\Theta}\|_1$ encourages local constancy in the coefficient matrix by penalizing the differences of spatially neighboring variables. For optimizing eq.(1), we followed the *coordinate-direction descent algorithm* described in [7].

3 Results

An experiment was performed in order to discover differences in processing monetary rewards between cocaine addicted subjects versus healthy control subjects. The time series consists of 87 frames taken every 3.5 seconds. The dataset collected by [8] contains 28 subjects: 16 drug-addicted and 12 healthy non-drug-using control individuals. Preprocessing of the dataset was performed in SPM2 (**http://www.fil.ion.ucl.ac.uk/spm/**), and it included deforming all time series to the same spatial reference template, spatial smoothing, cropping and regular sampling. Each subject has a sequence of 87 images of $53 \times 63 \times 46$ voxels. After preprocessing,

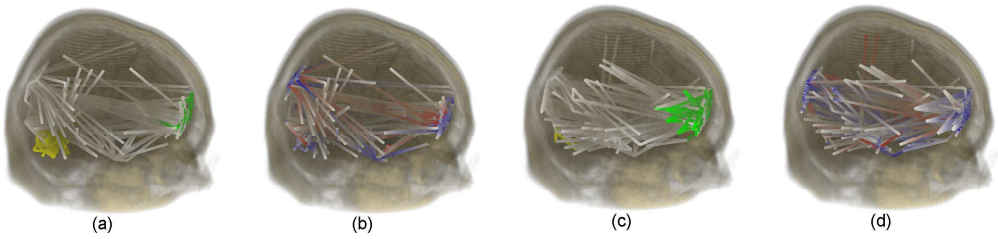

Figure 1: (a) GGM and (b) SEM for cocaine addicted subjects, (c) GGM and (d) SEM for control subjects. (a),(c): drug addicted subjects have more connections in the cerebellum (in yellow), control subjects have more connections in the prefrontal cortex (in green); (b),(d): blue lines correspond to positive interactions while red lines correspond to negative interactions. In order to represent the directionality, the cause is in darker colors while the effect is in lighter colors

the number of voxels was reduced to 869. We applied our algorithm with sparseness parameter $\rho = 0.15$, local constancy parameter $\tau = 0.01$ and $K = 10$ iterations.

Figure 1(a,c) shows different clusters of connections for each group. These results are consistent with the fact the prefrontal cortex has more control of other regions and behavior in the control subjects. In the cocaine addicted subjects, the cerebellum has been shown compensatory responses to the reduced prefrontal cortex control.

The directionality of the interactions between brain regions is shown in Figure 1(b,d). While GGMs are unable to show the directionality of the interactions, SEMs provide us such information.

Note that both models (GGMs and SEMs) share the same structure. This suggests that assuming Gaussianity of fMRI datasets is a reasonable assumption, since for Gaussian datasets this is the expected behavior. Eventhough more research has to be performed in order to conclude on this topic. Finally, the GGM structure is a way to validate the SEM structure, since the GGM structure has been shown to have good generalization performance in a cross-validation setting [7].

4 Conclusions

Our initial experiments show that by adding sparseness and local constancy priors to both Gaussian graphical models and structural equation models allow for finding functional interactions in the whole brain. Both models are suitable for fully exploratory research without using a priori set of brain regions.

In our experiments, the results have been neuropsychologically validated. Further experiments on larger datasets are necessary. We plan to test the ability of both methods to perform classification between drug-addicted and controls subjects.

References

[1] M. Hampson, B. Peterson, P. Skudlarski, J. Gatenby, and J. Gore. Detection of functional connectivity using temporal correlations in MR images. *Human Brain Mapping*, 2002.

[2] C. Buchel and K. Friston. Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex*, 1997.

[3] J. Kim, W. Zhu, L. Chang, P. Bentle, and T. Ernst. Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*, 2007.

[4] A. Roebroeck, E. Formisano, and R. Goebel. Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage*, 2005.

[5] J. Burge and T. Lane. Learning class-discriminative dynamic Bayesian networks. *ICML*, 2005.

[6] V. Smith, J. Yu, T. Smulders, A. Hartemink, and E. Jarvis. Computational inference of neural information flow networks. *PLoS Computational Biology*, 2006.

[7] J. Honorio, L. Ortiz, D. Samaras, N. Paragios, and R. Goldstein. Sparse and locally constant Gaussian graphical models. *NIPS*, 2009.

[8] R. Goldstein, D. Tomasi, N. Alia-Klein, L. Zhang, F. Telang, and N. Volkow. The effect of practice on a sustained attention task in cocaine abusers. *NeuroImage*, 2007.