Sparse Brain Network using Penalized Linear Regression

Hyekyoung Lee [a,f*], Dong Soo Lee [a,c,d,e], Hyejin Kang [a,d],
Boong-Nyun Kim [b] and Moo K. Chung [f,g,h]

[a]Department of Nuclear Medicine,
[b]Department of Neuropsychiatry, Seoul National University College of Medicine,
[c]Interdisciplinary Program in Cognitive Science,
[d]Institute of Radiation Medicine, Medical Research Center,
[e]Department of Molecular Medicine and Biopharmaceutical Sciences,
[f]Department of Brain and Cognitive Sciences, Seoul National University, Seoul, Korea;
[g]Department of Biostatistics and Medical Informatics, [h]Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, USA.

ABSTRACT

Sparse partial correlation is a useful connectivity measure for brain networks, especially, when it is hard to compute the exact partial correlation due to the small-n large-p situation. In this paper, we consider a sparse linear regression model with a l_1-norm penalty for estimating sparse brain connectivity based on the partial correlation. For the numerical experiments, we construct the sparse brain networks of 97 regions of interest (ROIs) extracted from FDG-PET data for the autism spectrum disorder (ASD) children and the pediatric control (PedCon) subjects. To validate the results, we check their reproducibilities by leave-one-out cross validation and compare the clustered structures derived from the brain networks of ASD and PedCon.

Keywords: Brain Connectivity, Compressed Sensing, Partial Correlation, LASSO

1. INTRODUCTION

The majority of connectivity analyses in brain imaging has been on thresholding correlation in detecting focal regions of correlated voxels. The main limitation of connectivity analyses based on correlation is that they fail to explicitly factor out the confounding effect of other regions. To remedy this limitation, partial correlation has been naturally introduced. Unfortunately, this type of problem usually belongs to the small-n large-p setting so it is not feasible to estimate the exact partial correlation. So far the majority of literature have used the penalized likelihood method in imposing the sparseness on the partial correlation estimation.[1] In this paper, we introduce a different approach based on the penalized linear regression for estimating sparse partial correlation.[2] The penalized linear regression with l_1-norm, which is also known as least absolute shrinkage and selection operator (LASSO), is usually formulated as the convex optimization to find the sparsest solution of the under-determined linear regression problem.[3]

The proposed model is applied to the 97 regions of interest (ROIs) extracted from FDG-PET data for 26 autistim spectrum disorder (ASD) children and 11 pediatric control (PedCon) subjects. It is generally known that ASD has the global patterns of underconnectivity and the local patterns of overconnectivity in the key brain regions. The differences between ASD and PedCon are mostly found in connectivities between lobes, especially, connection between secondary association cortices such as frontal and parietal regions.[4,5] Dense internal and sparse external linkages are properties of a cluster (also called a community or module). Thus, after estimating the partial correlation matrices by the penalized linear regression, we seek the modular structures of ASD and PedCon brain network and observe their differences based on the lobe structures.

*Send correspondence to Hyekyoung Lee. 28 Yongun-dong, Jongno-gu, Seoul, 110-744, Korea. E-mail: leehk@postech.ac.kr

2. DATA ACQUISITION AND PREPROCESSING

There are twenty six children with ASD (24 boys, mean age: 6.0 ± 1.8 years) and eleven children with PedCon (8 boys, mean age: 9.73 ± 2.55 years). All PET scans were obtained from ECAT EXACT 47 (Siemens-CTI, Knoxville, USA) PET scanner with an intrinsic resolution of 5.2 mm FWHM. PET images were preprocessed using Statistical Parametric Mapping (SPM) package. After spatial normalization to the standard template space, mean FDG uptake within 97 ROIs were extracted. The values of FDG uptake were globally normalized to the individual's total gray matter mean count.

3. SPARSE BRAIN NETWORK ESTIMATION

We formulate the sparse brain connectivity based on partial correlation in the penalized linear regression framework.

3.1 The Penelized Linear Regression for Estimating the Partial Correlation

Suppose that $\{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_p\}$ is the n-dimensional data vector measured at the p selected ROIs on the FDG-PET images of n subjects. We assume $\boldsymbol{f}_i$ are centered and normalized. It is known that the partial correlation is estimated by the linear regression estimation as follows[2] :

$$\boldsymbol{f}_i = \sum_{j \neq i} \beta_{ij} \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \text{ for } i = 1, ..., p, \tag{1}$$

where $\boldsymbol{\epsilon}_i$ is uncorrelated with all variables except $\boldsymbol{f}_i$ and β_{ij} is the measure of relationship between data vectors $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$ given all other data vectors. When $\text{var}(\boldsymbol{\epsilon}_i) = (1/\pi_{ii})$ and $\text{cov}(\boldsymbol{\epsilon}_i, \boldsymbol{\epsilon}_j) = \pi_{ij}/(\pi_{ii}\pi_{jj})$, the partial correlation θ_{ij} is given by $\theta_{ij} = -\pi_{ij}/\sqrt{\pi_{ii}\pi_{jj}} = \beta_{ij}\sqrt{(\pi_{ii}/\pi_{jj})}$.

Now we change the linear model in Eq.(1) to a matrix form $\boldsymbol{x} = \boldsymbol{Ab}$ as shown in Fig.1. Note that the elements of $\boldsymbol{X}$ follows Gaussian distribution with mean 0 and variance 1 because its column vectors are centered and normalized. When $n \ll p$, estimating the partial correlation in the linear model $\boldsymbol{x} = \boldsymbol{Ab}$ fall under a high-dimension-small-sample-size situation. Thus, we should consider shrinkage methods to regularize the model parameters and one solution is to add the sparseness penalty such as l_1 norm to the model parameters like

$$\hat{\boldsymbol{b}} = \arg\min_{b} \parallel \boldsymbol{x} - \boldsymbol{Ab} \parallel_2^2 + \lambda \parallel \boldsymbol{b} \parallel_1 \tag{2}$$

where $\parallel \cdot \parallel_1$ is a l_1 norm (sum of absolute values of elements) and λ is the regularization parameter. It is known as LASSO. To solve the problem, we exploited the coordinate descent learning and the active-set algorithm.[2]

3.2 Validating the Estimated Network

To validate the constructed network, we checked the reproducibility of the network obtained from different datasets using the leave-one-out cross-validation (see Fig.1). When λ increases, zero elements (the sparseness) increase, but irreproducible elements decrease (its number is near zero). Thus, we can tell that our penalized linear regression method for the partial correlation estimation finds the consistent brain connectivity.

4. SIGNIFICANCE TEST FOR FINDING GROUP DIFFERENCES

After calculating the threshold which maximizes the number of clusters, we partitioned the ROIs (nodes) of graph by the agglomerative hierarchical clustering (see Fig.2(c),(d)). Fig.3 supports that the autism brain network has local overconnectivity and long-range underconnectivity compared to the normal control, because, although the number of edges and clusters of ASD and PedCon were not different, the number of edges connected between lobes and the sum of anatomical distances of edges were significantly larger than one of ASD , while the number of edges connected within a lobe of PedCon was significantly smaller than one of ASD.

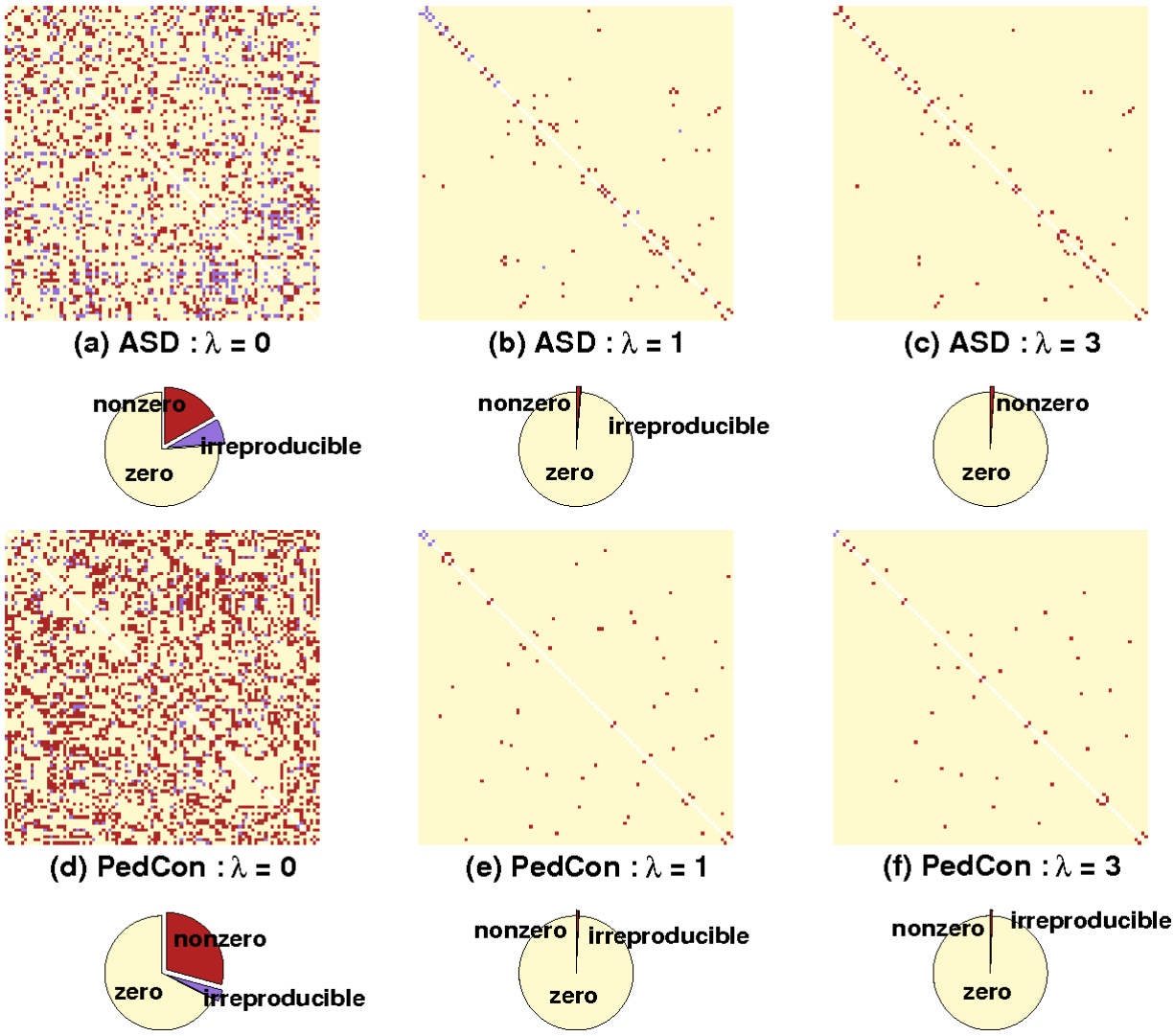

Figure 1. Partial correlation matrices of (a)-(c) ASD and (d)-(f) PedCon varying $\lambda = 0, 1, 3$ from left to right. If the standard deviation of the partial correlation elements obtained by cross validation is more than 0.1, we call it irreproducible elements and represent by the purple dots. We divided the reproducible partial correlations into zero and nonzero elements, which are represented by the yellow and red dots. The ratio of 3 kinds of elements is drawn in the pie chart.

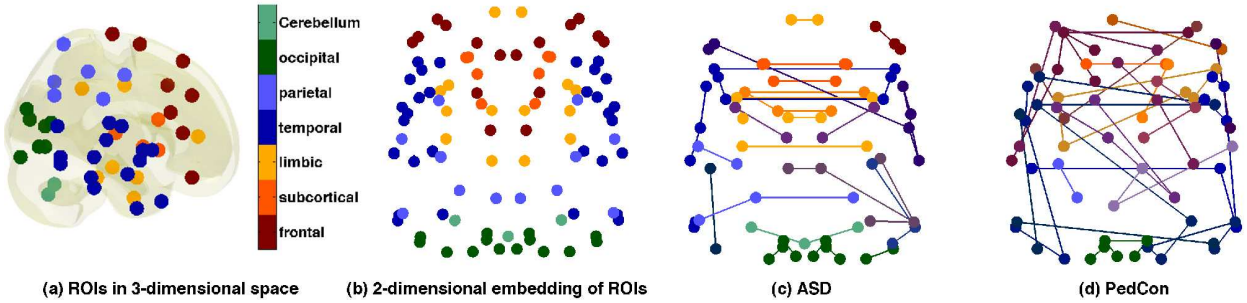

Figure 2. Visualization of ROIs in (a) 3-dimensional and (b) 2-dimensional space. Each lobe is represented by different color as shown in colorbar (a). Clustered brain networks for (c) ASD and (d) PedCon. In these figures, the color represents the cluster. The color of nodes in (c) is more similar to one in (b).

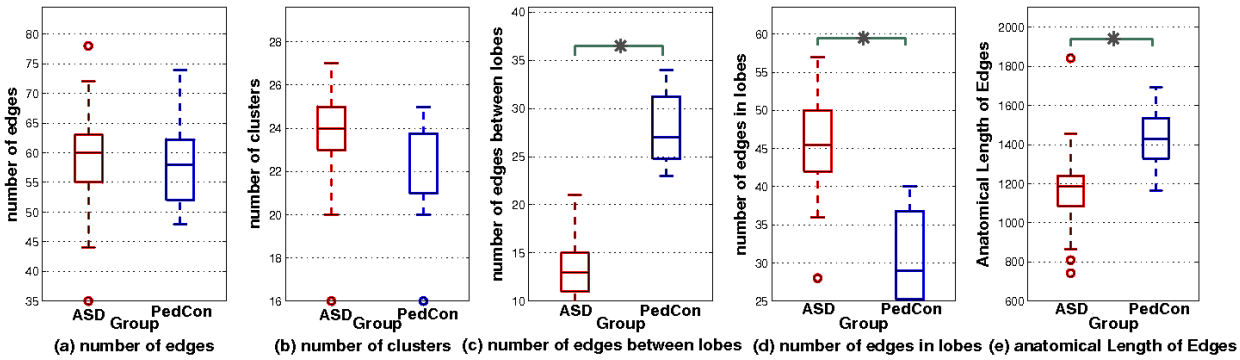

Figure 3. Significance tests on the number of (a) edges ($p = 0.5489$), (b) clusters ($p = 0.0406$), (c) edges connected between lobes ($p < 0.001$) and (d) edges connected in lobe ($p < 0.001$) and (e) sum of anatomical length of edges ($p < 0.001$). The red box in the left is for autism and the blue box in the right is for PedCon. The asterisk (*) represents the significant difference obtained by the Wilcoxon rank sum test with the significance level $p < 0.01$.

ACKNOWLEDGMENTS

This work was supported by WCU-grant from the government of Korea to the Department of Brain and Cognitive Sciences, Seoul National University and the National Research Foundation of Korea Grant funded by the Korean Government[NRF-2009-351- D00026].

REFERENCES

[1] Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., Chen, K., and Reiman, E., "Learning brain connectivity of alzheimers disease by sparse inverse covariance estimation," *NeuroImage* **50**, 935–949 (2010).

[2] Peng, J., Wang, P., Zhou, N., and Zhu, J., "Partial correlation estimation by joint sparse regression models," *Journal of the American Statistical Association* **104**, 735–746 (2009).

[3] Chen, S., Donoho, D., and Saunders, M., "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing* **20**, 33–61 (1999).

[4] Courchesne, E., Pierce, K., Schumann, C., Redcay, E., Buckwalter, J., Kennedy, D., and Morgan, J., "Mapping early brain development in autism," *Neuron* **56**, 399–413 (2007).

[5] Just, M., Cherkassky, V., Keller, T., and Minshew, N., "Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity," *Brain* **127**, 1811–1821 (2004).