Network Modeling and Functional Data Methods for Brain Functional Connectivity Studies

Kehui Chen

Department of Statistics,
University of Pittsburgh

Nonparametric Statistics Workshop,

Ann Arbor, Oct 06, 2016

Brain connectivity studies

"Over the past 20 years, neuroimaging has become a predominant technique in systems neuroscience. One might envisage that over the next 20 years the neuroimaging of distributed processing and connectivity will play a major role in disclosing the brains functional architecture and operational principles." – Friston 2011, Brain connectivity

Brain connectivity studies

- Human connectome project (HCP): publicaly available data sets, processing pipelines and code.
- Structure connectivity, functional connectivity, effective connectivity.

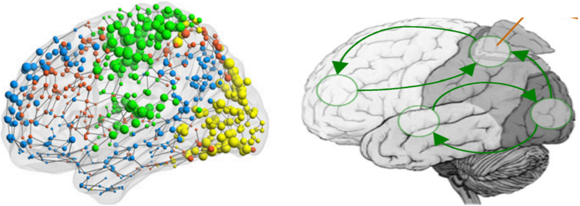


Figure: Crossley 2013, Friston 2011

- Functional connectivity is some type of statistical dependence among measured time series.

Functional connectivity measure

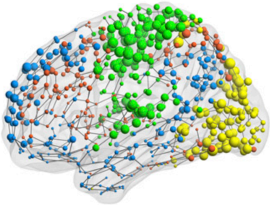
- Various methods have been proposed to characterize the functional connectivity measured with different modalities (EEG/MEG, fMRI etc).
- Correlation, partial correaltion, mutual information, coherence, phase locking values, among others.
- Functional data approach for correlation (He et al. 2016).
- Wang et al. 2014 (Frontiers in neuroscience): 42 methods, no-single measure is optimal for all types of data.
- Bastos et al. 2016 (Frontiers in systems Neuroscience): more on neuronal oscillartory.

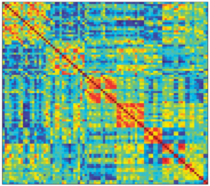
Brain segregation and integration

- Long history of functional segregation and the localization of function.
- Accurate parcellation provides
  - efficient comparisons across subjects and studies.
  - a foundation for illuminating the functional and structural organization of the brain
  - a means to reduce data complexity while improving statistical sensitivity and power for many neuroimaging studies.
- A single function may involve many specialized areas and their functional integration among them.
- Areas differ from their neighbors in microstructural architecture, functional specialization, connectivity with other areas.
- (Glasser 2016, Nature) used multi-source and multi-methods to detect 180 areas.

Community detection & brain parcellation

- $G = (V, E)$, represented by an adjacency matrix $A = (a_{ij})_{m \times m}$
- Node: depends on resolution
- Edge: depends on the measure and threshold
- Most modularity based approaches: more inter-community links, less intra-community links
- Stochastic block model: the connectivity patten is similar within communities.





Special features in brain connectivity problem

- The nodes are naturally embedded in a three-dimensional brain space
- Connectivity between adjacent nodes is sometimes over-represented due to technical reasons.
- Account for the spurious connectivity in adjacent nodes by removing the effect of spatial location so as to recover functionally distinct brain regions (“communities”).
- Especially relavant when doing subdivision parcellation (small areas).
- Feature adjusted stochastic block model (FASBM)

The stochastic block model

- Data: adjacency matrix $A \in \{0,1\}^{n\times n}$, where A_{ij} indicates the presence/absence of an edge between node i and node j.
- $A_{ii} = 0$, $A_{ij} = A_{ji}$, $\forall\ i,j$.
- Each node i belongs to a community with label $r_i \in \{1,\ldots,K\}$.
- Given $R = (r_1,\ldots,r_n)$ and B,

$$A_{ij} \sim \text{Bernoulli}(B_{r_i,r_j}), \ \ \text{independently}.$$

- $B \in [0,1]^{K\times K}$, symmetric, is the community-wise connectivity.
- Nodes in the same community have similar connectivity patterns.

Feature adjusted stochastic block model

The network Y is generated by

$$Y_{ij} = \begin{cases} \text{exponential family with mean } \mu_{ij} & \text{if} \quad i < j \\ 0 & \text{if} \quad i = j \\ Y_{ji} & \text{if} \quad i > j \end{cases}$$

FASBM is formulated as

$$E(Y_{ij}) = \mu_{ij} = g^{-1}(\theta_{r_i r_j} + f(\beta^T z_{ij})), \quad \text{with } \|\beta\| = 1.$$

Feature adjusted stochastic block model

$$E(Y_{ij}) = \mu_{ij} = g^{-1}(\theta_{r_i r_j} + f(\beta^T z_{ij})), \quad \text{with } \|\beta\| = 1.$$

- SBM is a special case of FASBM
- Model relational data with single parameter exponential family; allow for a scaling parameter.
- Nonparametric estimation of the node feature effect.
- Can take multiple node features.

Model Fitting

- Maximum likelihood estimation in terms of (g, θ, β, f) : but no close form solution.
- Alternates between two stages of maximization:
  1. First with respect to the parameters in the block model component, g and θ,
  2. and then with respect to the parameters in the single-index model component, f and β.
- We adapt the labeling switch algorithms (Bickel& Chen 2009) for the SBM to stage 1 and the estimation procedures for fitting single-index models (Carrol et al 1997) to stage 2.
- Paper and matlab package available upon request.

Simulation setting

- Generate network with m nodes and K community.
- $Y_{ij} \sim \text{Bernoulli}(g^{-1}(\theta_{r_i r_j} + f(d_{ij})))$.
- $\theta = \text{logit}\begin{pmatrix} 0.5 & 0.2 \\ 0.2 & 0.2 \end{pmatrix}$ and $\theta = \text{logit}\begin{pmatrix} 0.5 & 0.2 & 0.2 \\ 0.2 & 0.3 & 0.2 \\ 0.2 & 0.2 & 0.1 \end{pmatrix}$.
- Compute d_{ij} as distance between node i and node j
- Let $f = a\sin(-8d_{ij})$, with a taking different values, $0, 1.4$ or 1.8

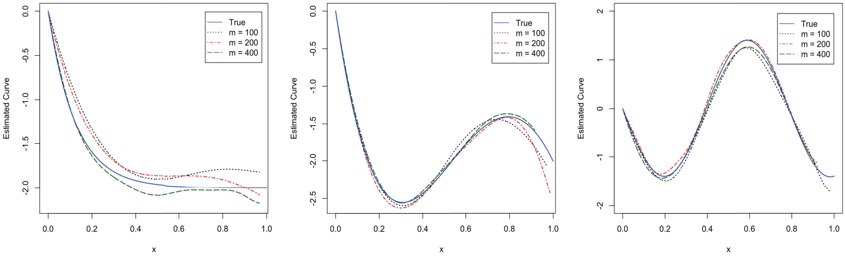
Simulation result for $K = 2$

		$m = 100$			$m = 200$			$m = 400$		
		FASBML	SBML	SPEC	FASBML	SBML	SPEC	FASBML	SBML	SPEC
$a = 0$	$\overline{Mis_p}$	**0.012**	**0.012**	0.041	**0.0004**	**0.0004**	0.006	**0**	**0**	0.0003
	NMI	**0.924**	**0.924**	0.783	**0.997**	**0.997**	0.955	**1**	**1**	0.998
$a = 1.4$	$\overline{Mis_p}$	0.157	0.443	**0.128**	**0.012**	0.469	0.079	**0.0001**	0.481	0.045
	NMI	**0.592**	0.038	0.470	**0.962**	0.005	0.625	**0.999**	0.002	0.75
$a = 1.8$	$\overline{Mis_p}$	**0.174**	0.461	0.182	**0.036**	0.469	0.132	**0.005**	0.482	0.105
	NMI	**0.524**	0.007	0.349	**0.908**	0.004	0.464	**0.989**	0.002	0.549

Simulation result for $K = 3$

		$m=100$			$m=200$			$m=400$		
		FASBML	SBML	SPEC	FASBML	SBML	SPEC	FASBML	SBML	SPEC
$a=0$	$\overline{Mis_p}$	**0.262**	0.265	0.298	**0.073**	0.075	0.185	**0.011**	**0.011**	0.074
	NMI	**0.546**	0.544	0.404	**0.825**	0.824	0.545	**0.954**	0.953	0.753
$a=1.4$	$\overline{Mis_p}$	**0.380**	0.535	0.407	**0.167**	0.524	0.352	**0.038**	0.534	0.335
	NMI	**0.332**	0.099	0.272	**0.682**	0.117	0.331	**0.910**	0.117	0.351
$a=1.8$	$\overline{Mis_p}$	**0.421**	0.566	0.450	**0.197**	0.573	0.434	**0.020**	0.592	0.436
	NMI	**0.260**	0.053	0.209	**0.625**	0.050	0.244	**0.919**	0.038	0.270

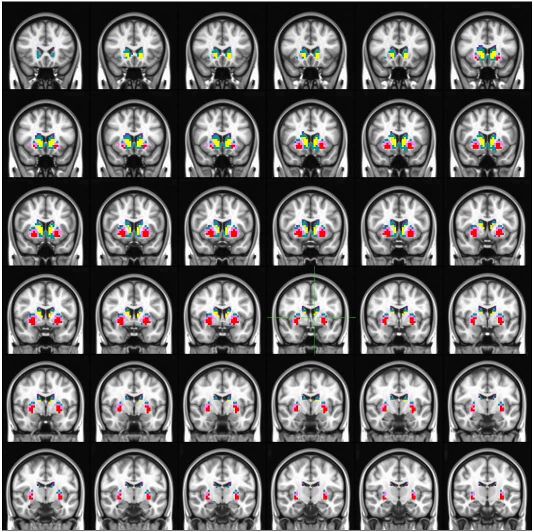
Fitted nonparametric functions

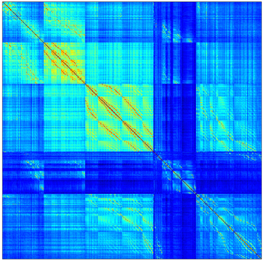


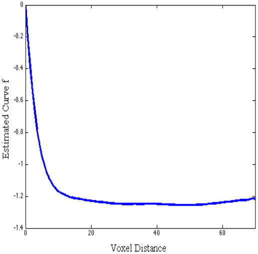
(a) $f(x) = 2\exp(-8x) - 2$
(b) $f(x) = 10x^4 - 42x^3 + 50x^2 - 20x$
(c) $f(x) = 1.4sin(-8x)$

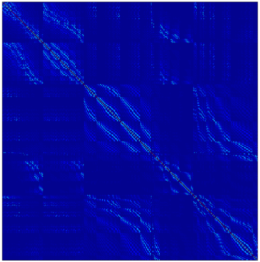
Data application

- Data collected at the University of Pittsburgh Medical Center
- Rs-fMRI data were preprocessed using AFNI and FSL, normalized to MNI152 template.
- We conduct our analyses on $m = 448$ gray matter voxels in the basal ganglia mask.
- The basal ganglia subserve a wide range of functions, including motor, cognitive, motivational, and emotional processes.
- We tried to parcellatesubdivisions in basal ganglia solely based on the functional connectivity matrix.

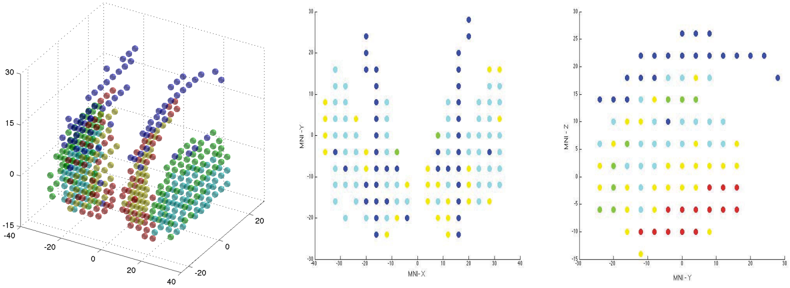








$$E(Y_{ij}) = \theta_{g_i g_j} + f(z_{ij}).$$



Community 1 (yellow) corresponds to caudate body, Community 3 (green) is putamen, Community 5 (cyan) is pallidum and Community 2 (red), 4 (blue) could be caudate head, but also spread out to places outside of caudate.

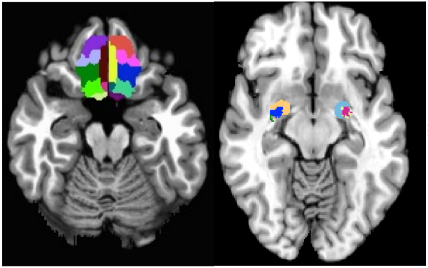
Functional connectivity and cognitive study

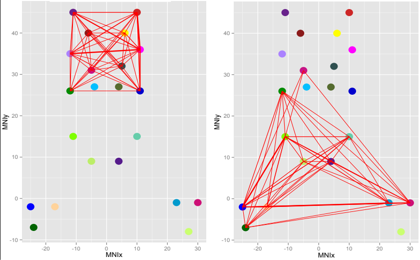
- How the complex functional connectivity patterns are related to the capacity for information processing.
- How does functional connectivity (network organization) develop with age, and interact with covaraites such as behavioral measures, cognitive function and mental distorter status?
- Topological measures: small worldness (global clustering coefficient); High-degree hub nodes; Community structure/modules; Connectivity threshold functions (Petersen et.al 2016), other measures?
- Group analysis of brain functional connectivity can borrow ideas and tools from functional data analysis.

Adaptation of localized FPCA (ongoing work)

- The nodes in brain network has natural spatial embedding and features spatial smoothness.
- Extract modes of variation in brain networks that are localized to regions of interest.
- Localized functional principal component analysis (Chen & Lei 2015).

Preliminary results on rs-fMRI connecitivty





THANK YOU!

Reference

- Friston, K. J. “Functional and effective connectivity: a review.” Brain connectivity 1.1 (2011): 13-36.
- Crossley, N. A., et al. “Cognitive relevance of the community structure of the human brain functional coactivation network.” Proceedings of the National Academy of Sciences 110.28 (2013): 11583-11588.
- He, J., et al. “Beyond Pearson Correlation for quantification of BOLD fMRI functional connectivity: a functional data approach.”
- Wang, H. E., et al. “A systematic framework for functional connectivity measures.” Frontiers in neuroscience 8 (2014): 405.
- Bastos, A. M., and Schoffelen, JM. “A tutorial review of functional connectivity analysis methods and their interpretational pitfalls.” Frontiers in systems neuroscience (2016).
- Glasser, M. F., et al. “A Multi-modal parcellation of human cerebral cortex.” Nature (2016), 536, 171-181.

Reference

- Bickel, P.J., and Chen. A. "A nonparametric view of network models and NewmanGirvan and other modularities." Proceedings of the National Academy of Sciences 106.50 (2009): 21068-21073.
- Carroll, R. J., et al. "Generalized partially linear single-index models." Journal of the American Statistical Association 92.438 (1997): 477-489.
- Petersen, A., et al. "Connectivity threshold functions for network representation and visualization, with applications to brain networks", Brain Connectivity (2016).
- Chen, K., and Lei, J. "Network cross-validation for determining the number of communities in network data." Journal of the American Statistical Association (2016), to appear.
- Chen, K., and Lei, J. "Localized Functional Principal Component Analysis." Journal of the American Statistical Association 110.511 (2015): 1266-1275.

Network cross-validation for selecting K

- Adapt the NCV method developed in Chen & Lei (2016).
- Block-wise node-pair splitting
- Very flexible, can be used for stochastic block model, degree-corrected block model and extendable to FASBM.
- R code and Matlab code are avaiable.