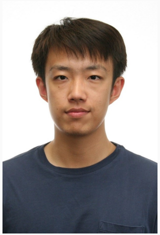
Partially Linear Additive Gaussian Graphical Models

Presented by: Sinong Geng
Princeton University

June 13th, 2019 @ ICML 2019

Authors



Sinong Geng
Princeton University

Minhao Yan
Cornell University

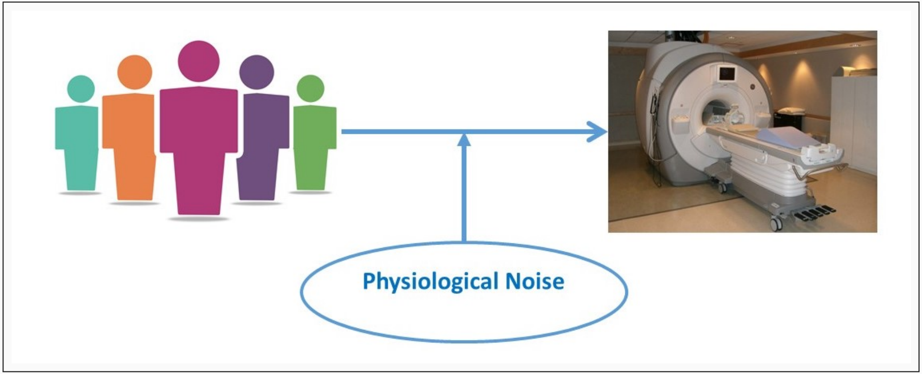
Mladen Kolar
The University of Chicago

Sanmi Koyejo
University of Illinois

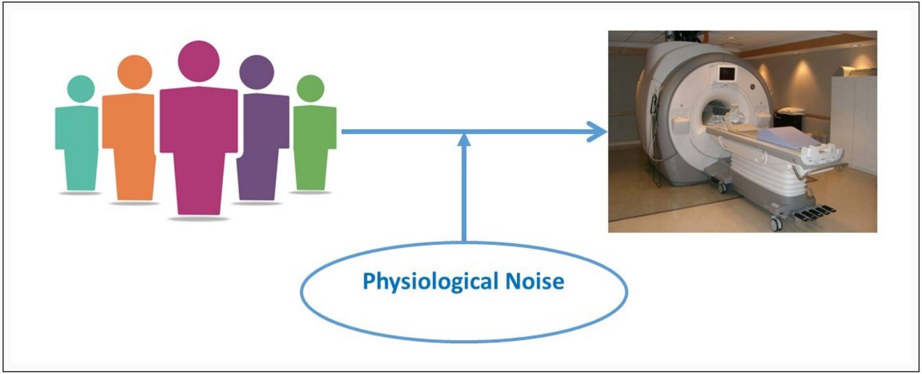
Poster: Partially Linear Additive Gaussian Graphical Models

Thu Jun 13th 06:15 – 09:00 PM @ Pacific Ballroom

Brain Functional Connectivity Analysis

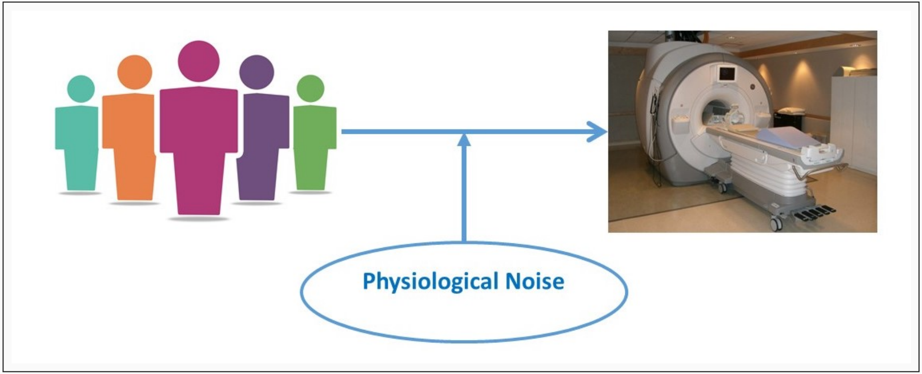


Brain Functional Connectivity Analysis



Estimation is distorted by physiological noise [Van Dijk et al., 2012, Goto et al., 2016].

Brain Functional Connectivity Analysis



Estimation is distorted by physiological noise [Van Dijk et al., 2012, Goto et al., 2016].

The noise sources are observable e.g. motion, breathing ...

Model Formulation: Goals

→ A general formulation of the effects caused by the noise.

→ Stronger theoretical guarantees compared tp methods with hidden variables.

Model Formulation

- → $\mathbf{Z}$ denotes the observed fMRI data, and random variable G, the physiological noise.
- → $\mathbf{Z} \mid G = g$ follows a Gaussian graphical model [Yang et al., 2015] with a parameter matrix, denoted by $\boldsymbol{\Omega}(g)$:

$$\mathrm{P}(\mathbf{Z} = \mathbf{z}; \Omega(g) \mid G = g) \propto \exp\left\{ \sum_{j=1}^{p} \Omega_{jj}(g) z_j \sum_{j=1}^{p} \sum_{j'>j}^{p} \Omega_{jj'}(g) z_j z_{j'} - \frac{1}{2} \sum_{j}^{p} z_j^2 \right\}.$$

- → Parameter matrices are additive:

$$\boldsymbol{\Omega}(g) := \boldsymbol{\Omega}_0 + \mathbf{R}(g).$$

Model Formulation: $\boldsymbol{\Omega}(\boldsymbol{g})$

Goals:

- Identifiable parameters
- A general formulation

Model Formulation: $\boldsymbol{\Omega}(g)$

Goals:

- Identifiable parameters
- A general formulation

Assumptions:

- $\mathbf{R}(g) = 0$ for any g satisfying $|g| \leq g^*$.
- $\mathbf{R}(g)$, and $\Omega(g)$ are smooth enough to be recovered by kernel methods.

Model Formulation: $\boldsymbol{\Omega}(g)$

Goals:

- Identifiable parameters
- A general formulation

Assumptions:

- $\mathbf{R}(g) = 0$ for any g satisfying $|g| \leq g^*$.
- $\mathbf{R}(g)$, and $\boldsymbol{\Omega}(g)$ are smooth enough to be recovered by kernel methods.

Existing assumptions:

- $\mathbf{R}(g) = 0$ [Van Dijk et al., 2012, Power et al., 2014].
- $\mathbb{E}(\mathbf{R}(g)) = 0$ [Lee and Liu, 2015, Geng et al., 2018].

Parameter Estimation

Log Pseudo Likelihood:

- We summarize the varying effects as $M_{ij} := \mathbf{x}_{ij}^\top \Omega_{i\cdot j}$, where $\mathbf{x}_{ij}^\top$ denotes the i^{th} row vector of $\mathbf{x}_j$.
- $$\begin{aligned} &\ell_{PL}\left(\{\mathbf{z}_i, g_i\}_{i\in[n]}; \mathbf{R}(\cdot), \Omega_0\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{p}\left\{z_{ij}\left(\mathbf{x}_{ij}^\top \Omega_{0\cdot j} + M_{ij}\right) - \frac{1}{2}z_{ij}^2 \right. \\ &\quad \left. - \frac{1}{2}\left(\mathbf{x}_{ij}^\top \Omega_{0\cdot j} + M_{ij}\right)^2\right\}. \end{aligned}$$

Parameter Estimation

- Pseudo-Profile Likelihood [Fan et al., 2005]
- Suppose that Assumptions are satisfied. Then, for any $\epsilon > 0$, with probability of at least $1-\epsilon$, there exists $C_4 > 0$, so that $\hat{\Omega}_0$ shares the same structure with the underlying true parameter Ω_0^*, if for some constant $C_5 > 0$,

$$C_5\sqrt{\frac{\log p}{n}} \geq \lambda \geq \frac{4}{\alpha} C_4 \sqrt{\frac{\log p}{n}},$$
$$r := 4C_2\lambda \leq \left\|\Omega_{0S}^*\right\|_\infty,$$

and $n \geq \left(64C_5C_2^2C_3/\alpha\right)^2 \log p$.

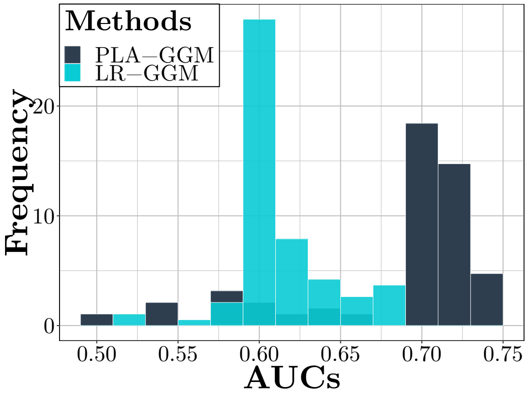
Parameter Estimation

Sparsistency: The underlying structure can be recovered with a high probability.

$\sqrt{n}$ Convergence: The smallest scale of the non-zero component that the PPL method can distinguish from zero converges to zero at a rate of $\sqrt{n}$.

Overall Performance

- → LR-GGM
- → fMRI dataset with control subjects and those with Schizophrenia.
- → Diagnosis using the recovered structure by two different methods.



Thank you!

References I

J. Fan, T. Huang, et al. Profile likelihood inferences on semiparametric varying-coefficient partially linear models. *Bernoulli*, 11(6):1031–1057, 2005.

S. Geng, M. Kolar, and O. Koyejo. Joint nonparametric precision matrix estimation with confounding. *arXiv preprint arXiv:1810.07147*, 2018.

M. Goto, O. Abe, T. Miyati, H. Yamasue, T. Gomi, and T. Takeda. Head motion and correction methods in resting-state functional mri. *Magnetic Resonance in Medical Sciences*, 15(2):178–186, 2016.

References II

W. Lee and Y. Liu. Joint estimation of multiple precision matrices with common structures. *The Journal of Machine Learning Research*, 16(1):1035–1062, 2015.

J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *Neuroimage*, 84:320–341, 2014.

K. R. Van Dijk, M. R. Sabuncu, and R. L. Buckner. The influence of head motion on intrinsic functional connectivity mri. *Neuroimage*, 59(1):431–438, 2012.

References III

E. Yang, P. Ravikumar, G. I. Allen, and Z. Liu. Graphical models via univariate exponential family distributions. *Journal of Machine Learning Research*, 16(1):3813–3847, 2015.