University of Grenoble-Alpes, GIPSA-lab
University of Paris-Saclay, L2S @ CentraleSupélec

INTERNSHIP PROPOSAL
2017 - 2018

Romain COUILLET & Sophie ACHARD
romain.couillet@centralesupelec.fr
sophie.achard@gipsa-lab.grenoble-inp.fr

Random Matrices for Machine Learning on Brain Graphs

Project context and application

Tremendous progress in understanding the functioning of the brain activities have recently been provided by adequate signal processing of fMRI (functional Magetic-Reasonance Imaging) time series (BOLD) acquisitions. These time series are mapped to regions on the brain, thereby creating a (possibly dynamic) elementary graph of the brain network with connections inferred from the BOLD data [1]. The BOLD data being rather few in number and intrinsically noisy, graph reconstructions relate to some extent to correlation matrix estimation under the so-called random matrix theory (RMT) regime of equally large and number data [2,3]. Improved correlation estimates may therefore be obtained using advanced RMT techniques.

Once a brain graph model is set, further analyses, such as activity pattern region clustering or discriminative graph classification and regression, can be performed (for instance for neurodegenerative disease progress assessment) [4-5], for which recent works within the spectrum of random matrix theory also provide highly performing methods [6-13].

The objective of the internship is to propose a random-matrix improved framework for "machine learning for brain graphs". The challenging program will consist precisely in (i) identifying the "large dimensional" limitations of conventional methods in fMRI data processing for brain graph analysis, (ii) proposing adequate random-matrix analysis models and study their characteristics, (iii) proposing and analyzing the performance of improved methods for classification and regression on brain graphs, and (iv) testing these methods on actual fMRI acquisitions. As such, the internship has a dual theoretical and applied flavor, which calls for a profound interest of the applicant in applied mathematics, notably in biostatistics applications.

The internship will take place in Grenoble, within the GIPSA-Lab of the University of Grenoble-Alpes with Romain Couillet (professor affiliated both at GIPSA-lab, Grenoble and CentraleSupélec, Gif, expert in random matrix theory) and Sophie Achard (senior CNRS researcher at GIPSA-lab, expert in functional brain signal processing). A possibility to hold the internship at CentraleSupélec can be envisaged but is strongly not advised.

The internship may conduct to a PhD thesis position as of September 2018.

Related domains

Random matrix theory, large dimensional statistics, data processing, machine learning, brain graphs, fMRI.

Main steps

- Getting to grasp with the random matrix analysis methods for matrices related to graphs, as well as with elementary fMRI brain graph processing.
- Theoretical analysis of statistical inference (machine learning) methods for simple random matrix-based correlation or covariance matrices, adapted to brain graph models (in particular Gaussian Markov random field models).
- Proposal of improved methods.
- Practical validation in Python/Matlab. Tests and comparisons versus real datasets.

Requirements

Basic knowledge on random matrix theory (having followed a class on the topic is strongly advised) and asymptotic probability/statistics; basic knowledge on statistical inference and machine learning; ideally good coding skills in either Python or Matlab.

Contact

Romain COUILLET
http://romaincouillet.hebfree.org/
romain.couillet@centralesupelec.fr

Sophie ACHARD
http://www.gipsa-lab.grenoble-inp.fr/~sophie.achard/
sophie.achard@gipsa-lab.grenoble-inp.fr

References

[1] Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186-198.
[2] Couillet, R., & Debbah, M. (2011). *Random matrix methods for wireless communications*. Cambridge University Press.
[3] Tao, T. (2012). *Topics in random matrix theory* (Vol. 132, p. 282). Providence, RI: American Mathematical Society.
[4] F. D. V. Fallani, J. Richiardi, M. Chavez and S. Achard, Graph analysis of functional brain networks : practical issues in translational neuroscience. Philosophical Transactions of the Royal Society B : Biological Sciences 369.1653 , 20130521. 2014.
[5] J. Richiardi, S. Achard, B. Horst, and D. Van De Ville. Machine learning with brain graphs. IEEE Signal Processing Magazine, 30(3): 58-70, 2013.
[6] Alfred O. Hero, Bala Rajaratnam. Large Scale Correlation Screening. arxiv:https://arxiv.org/abs/1102.1204
[7] Couillet, Romain, and Florent Benaych-Georges. "Kernel spectral clustering of large dimensional data." *Electronic Journal of Statistics* 10, no. 1 (2016): 1393-1454.
[8] El Karoui, Noureddine, Derek Bean, Peter J. Bickel, Chinghway Lim, and Bin Yu. "On robust regression with high-dimensional predictors." *National Academy of Sciences* 110, no. 36 (2013): 14557-14562.
[9] Couillet, Romain, Frédéric Pascal, and Jack W. Silverstein. "The random matrix regime of Maronna's M-estimator with elliptically distributed samples." *Journal of Multivariate Analysis* 139 (2015): 56-78.
[10] Liao, Zhenyu, and Romain Couillet. "A Large Dimensional Analysis of Least Squares Support Vector Machines." *arXiv preprint arXiv:1701.02967* (2017).
[11] Krzakala, Florent, Cristopher Moore, Elchanan Mossel, Joe Neeman, Allan Sly, Lenka Zdeborová, and Pan Zhang. "Spectral redemption in clustering sparse networks." *Proceedings of the National Academy of Sciences* 110, no. 52 (2013): 20935-20940.
[12] Ali, Hafiz Tiomoko, and Romain Couillet. "Spectral community detection in heterogeneous large networks." *arXiv preprint arXiv:1611.01096* (2016).
[13] Mai, Xiaoyi, and Romain Couillet. "A random matrix analysis and improvement of semi-supervised learning for large dimensional data."