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

INTERNSHIP PROPOSAL
2017 - 2018

Romain COUILLET & Pierre COMON
romain.couillet@centralesupelec.fr
pierre.comon@gipsa-lab.grenoble-inp.fr

A Random Matrix Study of Random Tensors

Project context and application

Random matrix theory has recently provided a set of original tools to understand, improve and propose new algorithms for a wide variety of statistical machine learning methods for (sometimes not so) large dimensional datasets, far beyond sample covariance matrix-based models; this is the case for instance of kernel random matrices [1,2], random feature maps and simple random neural network structures [3], deep nets [4], robust statistics [5,6], SVM [7], community detection on graphs [8,9], semi-supervised learning [10], etc.

Practical applications however sometimes call for more advanced structures than matrices and particularly tensors [11-14], as in video processing, medical imaging, hypergraphs, multi-modal data processing, etc. The spectral analysis of tensors per se is quite delicate and of limited practical interest as it involves NP-hard questions (also the very definition of the spectrum of tensors is still a subject of definition controversies), but multiple applications (in detection, estimation, clustering) bypass these issues by developing polynomial-time methods involving matrix representations, called "unfoldings", of the tensors. These unfoldings naturally lead to structured random matrices, quite unlike those conventionally met in random matrix theory (even when the tensor entries are iid). The subject of the internship is to explore those new matrix structures in the context of detection, estimation or classification questions for tensors. Few works exist so far in the literature that handle those questions in depth. Notable exceptions are [15-17] that consider the question of source detection for different randomness constraints in the tensors, with [17] addressing the question of structured models of matrix unfoldings.

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 Pierre Comon (professor at GIPSA-lab, expert in tensor theory). 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 (open positions are to be fulfilled).

Related domains

Random matrix theory, large dimensional statistics, data processing, machine learning, tensors.

Main steps

- Getting to grasp with the random matrix analysis methods for structured random matrices (likely through concentration of measure theory or moment-based techniques).
- Analysis of simple random matrix structures of unfolded random tensors.
- Theoretical analysis of statistical inference methods on tensors (detection, estimation, classification). These inference methods will drive the models studied in the previous items.
- Practical validation in Python/Matlab. Tests and comparison 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

Pierre COMON
http://www.gipsa-lab.grenoble-inp.fr/~pierre.comon/
pierre.comon@gipsa-lab.grenoble-inp.fr

References

[1] El Karoui, Noureddine. "The spectrum of kernel random matrices." *Annals of Stats* 38, no. 1 (2010): 1-50.
[2] Couillet, Romain, and Florent Benaych-Georges. "Kernel spectral clustering of large dimensional data." *Electronic Journal of Statistics* 10, no. 1 (2016): 1393-1454.
[3] Louart, Cosme, Zhenyu Liao, and Romain Couillet. "A Random Matrix Approach to Neural Networks." *arXiv preprint arXiv:1702.05419* (2017).
[4] Choromanska, Anna, Mikael Henaff, Michael Mathieu, Gérard Ben Arous, and Yann LeCun. "The loss surfaces of multilayer networks." In *Artificial Intelligence and Statistics*, pp. 192-204. 2015.
[5] 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.
[6] 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.
[7] Liao, Zhenyu, and Romain Couillet. "A Large Dimensional Analysis of Least Squares Support Vector Machines." *arXiv preprint arXiv:1701.02967* (2017).
[8] 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.
[9] Ali, Hafiz Tiomoko, and Romain Couillet. "Spectral community detection in heterogeneous large networks." *arXiv preprint arXiv:1611.01096* (2016).
[10] Mai, Xiaoyi, and Romain Couillet. "A random matrix analysis and improvement of semi-supervised learning for large dimensional data."
[11] P.Comon, "Tensors: a brief introduction," IEEE Signal Processing Magazine, 31, May 2014.
[12] W. Hackbusch. Tensor Spaces and Numerical Tensor Calculus. Springer, Berlin, Heidelberg, 2012.
[13]J. M. Landsberg. Tensors: Geometry and Applications, volume 128 of Graduate Studies in Mathematics. AMS publ., 2012.
[14]L. Qi and Z. Luo, Tensor Analysis, SIAM publ., 2017
[15] Richard, Emile, and Andrea Montanari. "A statistical model for tensor PCA." In *Advances in Neural Information Processing Systems*, pp. 2897-2905. 2014.
[16] Tomioka, Ryota, and Taiji Suzuki. "Spectral norm of random tensors." *arXiv:1407.1870* (2014).
[17] Bayer, Rémy, and Philippe Loubatan. "Large deviation analysis of the CPD detection problem based on random tensor theory." In *Signal Processing Conf. (EUSIPCO), 2017 25th European*, pp. 658-662. IEEE, 2017.