

Proposition de Thèse

Analyse en Composantes Principales Tensorielle - Tensor PCA

Un outil essentiel en analyse de données et intelligence artificielle est l'analyse en composantes principales (PCA), qui permet de retrouver un signal vectoriel caché dans un bruit matriciel aléatoire de grande taille. Plus récemment différents auteurs se sont intéressés à généraliser cette analyse en rang plus élevé, un problème appelé *analyse en composantes principales tensorielle* (en anglais *Tensor PCA*). Dans sa version la plus simple (spiked tensor model), il s'agit à nouveau de retrouver un signal vectoriel (c'est à dire un tenseur de rang un) caché dans un bruit aléatoire tensoriel de rang 3 ou plus. Le paysage aléatoire correspondant à ce problème a été récemment exploré et des asymptotes sur le nombre de points critiques ont été établies [2]. Elles reposent sur des techniques de matrices aléatoires.

Indépendamment de cela des techniques analytiques nouvelles pour étudier les aspects plus spécifiques des tenseurs aléatoires ont été introduites depuis une dizaine d'années par R. Gurau, V. Rivasseau et leurs collaborateurs [3] dans un contexte très différent, celui de la géométrie aléatoire discrète et de la quantification de la gravité. Les outils correspondants, spécifiquement tensoriels, tels les graphes meloniques et l'énumération combinatoire de triangulations, n'ont pas encore été appliqués en analyse en composantes principales tensorielle. Ce sera précisément l'objectif de la thèse.

La thèse sera effectuée sous la direction de V. Rivasseau, *en cotutelle entre le laboratoire de physique théorique de Paris-Sud (LPT, équipe de physique mathématique de Benedetti-Rivasseau) et le Laboratoire de Vision et d'Ingénierie des Contenus du CEA (LVIC, équipe de Tamaazousti-Lahoche)*. Elle fait partie d'un plan de coopération à plus long terme de ces laboratoires sur l'interface physique théorique et intelligence artificielle.

Les conséquences potentielles vont en effet de problèmes très concrets comme l'analyse d'images et de textes pratiquée au LVIC jusqu'aux aspects plus abstraits étudiés au LPT (quantification de la gravité, holographie, matière condensée...). Ce sujet de thèse convient donc à un étudiant très polyvalent avec une très bonne formation mathématique qui désire effectuer un travail pionnier dans ce sujet, et qui est prêt à s'investir non seulement dans les aspects théoriques du problème mais aussi dans ses applications.

References

- [1] A. Montanari, E. Richard, *A statistical model for tensor PCA*, arXiv:1411.1076.
- [2] G. Ben Arous, Song Mei, A. Montanari, M. Nica, *The landscape of the spiked tensor model*, arXiv:1711.05424.
- [3] R. Gurau, *Random Tensors*, Oxford University Press (2016).