Convolutional Sparse Coding VS Non-negative Matrix Factorization

Supervision:

- Matthieu Kowalski, University Paris-Saclay, L2S matthieu.kowalski@universite-paris-saclay.fr
- Charles Soussen, CentraleSupelec, L2S charles.soussen@centralesupelec.fr
- Thomas Moreau, Parietal, Inria thomas.moreau@inria.fr

Context Recent breakthroughs in machine learning have relied on supervised paradigms with large annotated databases. For example, impressive classification performances have been obtained on the ImageNet database in computer vision. ImageNet contains millions of images belonging to hundreds of categories such as plant species, animals, and others. However, it is not always possible to develop enough labeled data to build an informative representation of the samples. For medical applications, **even if labels are available, they are often weak**. For example, a doctor diagnosed cancer on a scanner, but the label does not precisely say its localization in the image. An analogy with object detection would be to localize bikes in pictures when the training labels only tell if a bike is present or not anywhere in the image. In this weak to no supervision context, statistical insights can be learned by studying the data distribution. **Unsupervised learning** is a set of methods to infer common structures from unlabeled data. These structures can, in turn, be used to compare the data and highlight commonalities across them.

Applications where data are signals are numerous, especially in healthcare. For specific medical records, the scale of the collected databases makes the annotations by doctors too costly or not even possible. Even if some annotations exist for the considered task, they may be hard to assess. For instance, when predicting a patient's consciousness in a coma from electroencephalographic recordings, there is no consensus on the ground truth. Unsupervised learning can, without any annotations, highlight statistical effects in the structure of the signal, which can be used, for example, to stratify the subjects in coherent groups. However, still today, **unsupervised learning methods for signals remain a computational and a statistical challenge**.

Methods A particular type of unsupervised technique for signals is the so-called **inverse problem**. These problems occur naturally when studying the source of an observed signal, knowing the physics of the observation of this signal. The resolution of inverse problems is a major challenge in many fields, from astrophysics to bio-imaging. This is typically the case for source localization in MEG signals, for example. These problems are usually ill-posed, and we rarely have a ground truth to evaluate them. The most classic resolution approach is to formulate them as optimization problems integrating a prior on the solutions' structure and to perform minimization by iterative methods. Computational efficiency and the quality of the solutions depend strongly on the prior choice.

Based on the short time-frequency transform of the signal, the **non-negative matrix factorization (NMF)** is a classical way to encode such prior by learning frequency patterns in the signal. Starting from a non-negative matrix, the (power) spectrogram, NMF seeks a low-rank decomposition leading to a dictionary of spectrum patterns and a matrix of time activation of these patterns. Each of these components only encodes additive effects. Such methods have been applied with success for audio source separation [4] to learn specific percussive dictionaries. To use such an approach for inverse problems, the Low-rank Time-frequency synthesis (LRTFS) has been proposed in [1]. The **convolutional sparse coding (CSC)** model is another established framework to encode priors for signals. It learns **shift-invariant patterns** to sparsely reconstruct a signal, which corresponds to recurrent structures present in the data. While recent advances have improved the computational tractability of these methods, it is not yet widely used to solve inverse problems.

The idea of this internship is to compare the two approaches from an unsupervised dictionary/pattern learning point of view. Furthermore, mathematical links between the CSC and the LRTFS will be explored using the convolutional interpretation of the short-time Fourier transform.

Environment The internship will take place between the Group of Inverse Problem (GPI) at the Laboratoire des Signaux et Systèmes (L2S) and the Inria Saclay, in the Parietal team. The GPI is used to study ill inverse problems as well as from theoretical aspects (Bayesian models, convex/non-convex optimization), algorithms implementation and acceleration (especially on GPU), and a wide range of applications (tomography, source separation/localization, astrophysics...). Parietal is a large team focused on mathematical methods for statistical modeling of brain function using neuroimaging data (fMRI, MEG, EEG). Particular topics of interest include machine learning techniques, numerical and parallel optimization, applications to human cognitive neuroscience, and scientific software development.

This internship will also take place in the context of the ANR BMWs (Bayesian Meets Wavelets), which aims to develop new frameworks for joint processing of EEG/MEG data within a common space-time formulation. This ANR is joint with the Institut de Mathematics de Marseille and the Systems Neuroscience Institute.

The internship may lead to a Ph.D. position.

Requirements

- Strong mathematical background. Knowledge in numerical optimization is a plus.
- Good programming skills in Python. Knowledge of a deep-learning library is a plus.

References

[1] Févotte, C. and Kowalski, M. (2014) Low-rank time-frequency synthesis. Advances in Neural Information Processing Systems, 27, 3563-3571.

[2] Dupré La Tour, T., Moreau, T., Jas, M. and Gramfort, A. (2018). Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals. Advances in Neural Information Processing Systems (NIPS).

[3] Jas, M., Dupré La Tour, T., Şimşekli, U. and Gramfort, A. (2017). Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding. Advances in Neural Information Processing Systems (NIPS), pages 1099–1108.

[4] Laroche C., Kowalski M., Papadopoulos H. and Richard G. (2018), Hybrid Projective Nonnegative Matrix Factorization with Drum Dictionaries for Harmonic/Percussive Source Separation, IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol. 26, N. 9, pp. 1499-1511