

Thematic school: Optimization & algorithms for high-dimensional machine learning and inference

Monday 7 October 2024 - Friday 11 October 2024

Institut de Mathématiques de Toulouse (IMT)

Scientific Programme

Valentina Ros - High-dimensional random landscapes.

Abstract: I will discuss how to characterize properties of random Gaussian landscapes in high-dimension, with a particular focus on the distribution of stationary points. Using simple denoising problems as an example, I will consider first the case of quadratic Gaussian landscapes on spheres, where the landscape problem can be mapped into a problem of random matrix theory. I will then review results on highly non-convex landscapes (known in the physics literature as “p-spin models” with $p > 2$), which can be obtained with counting formalisms such as Kac-Rice. If time permits, I will discuss how to characterize the local landscape geometry, i.e., the connectivity between stationary points at fixed overlap with each others, using tools of large deviation theory.

Prerequisite: I will try to be self-consistent; some previous knowledge of random matrix theory can be useful (eigenvalue density, outliers, BBP transition).

Cédric Févotte - Majorization-minimization for non-negative matrix factorization

Abstract: Data is often available in matrix form, in which columns are samples, and processing of such data often entails finding an approximate factorization of the matrix into two factors. The first factor (the “dictionary”) yields recurring patterns characteristic of the data. The second factor (“the activation matrix”) describes in which proportions each data sample is made of these patterns. Nonnegative matrix factorisation (NMF) is a popular unsupervised learning technique for analysing data with nonnegative values, with applications in many areas such as in text information retrieval, recommender systems, audio signal processing, and hyperspectral imaging. The talk will provide a tutorial on NMF for data processing, with a focus on majorization-minimization algorithms for NMF with the beta-divergence, a continuous family of loss functions that takes the quadratic loss, KL divergence and Itakura-Saito divergence as special cases. The tutorial will also address regularised variants of NMF (sparsity, smoothness) and will present some applications in imaging (remote sensing, medical imaging).

Webpage

<https://www.irit.fr/~Cedric.Fevotte/>

Cynthia Rush - An Introduction to Approximate Message Passing Algorithms

Abstract: Approximate Message Passing (AMP) refers to a class of iterative algorithms that have been successfully applied to a number of high-dimensional statistical estimation problems like linear regression, generalized linear models, and low-rank matrix estimation, and a variety of engineering and computer science applications such as imaging, communications, and deep learning. AMP algorithms have two features that make them particularly attractive: they can easily be tailored to take advantage of prior information on the structure of the signal, such as sparsity, and under suitable assumptions on a design matrix, AMP theory provides precise asymptotic guarantees for statistical procedures in the high-dimensional regime. In this tutorial, I will present the main ideas of AMP from a statistical perspective to illustrate the power and flexibility of the AMP framework. We will discuss the algorithms’ use for characterizing the exact performance of a large class of statistical estimators in the high-dimensional proportional asymptotic regime and we will discuss the algorithms’ applications in engineering fields like wireless communication and signal processing.

Edouard Pauwels - An introduction to optimization for deep learning

Abstract: The mini-course presents a self-contained view of the basics of optimization for deep network training. This does not require prior knowledge of artificial neural networks or optimization. We describe how training reduces to optimization and what are the main algorithmic building blocks in this context: algorithmic differentiation, sub-sampling and gradient algorithms. The presentation aims at providing an overview of the specific optimization developments for neural network training algorithms and a few theoretical results.

Prerequisite: basic familiarity with derivative calculus and related analysis principles, discrete random variable, linear algebra.

François Malgouyres - Properties of the Landscape in Neural Network Optimization

Abstract: After reviewing fundamental aspects of non-convex optimization, I will discuss three types of findings that describe the properties of the landscape of the objective function during the optimization of neural network weights. The first set of results offers a detailed depiction of the landscape of deep linear neural networks, shedding light on the implicit regularization that occurs in this context. The second set of findings establishes favorable properties of the landscape for wide neural networks. The final set of results provides a geometric description of neural networks, illustrating the influence of the geometry on implicit regularization.