

Séminaire de Probabilités et Statistique

Mardi 05 Janvier à 14h00

Laboratoire Dieudonné
Salle de conférence - LJAD

Adeline Fermanian

Sorbonne Université - LPSM

Learning with signatures

Sequential or temporal data arise in many fields of research, such as quantitative finance, medicine or computer vision. We will be concerned with a novel approach for sequential learning, called the signature method, and rooted in rough path theory. Its basic principle is to represent multi-dimensional paths by a graded feature set of their iterated integrals, called the signature. After a general overview of signatures in machine learning, we will focus on two specific problems. On the one hand, in order to combine signatures with machine learning algorithms, it is necessary to truncate these infinite series. Therefore, we define an estimator of the truncation order and provide theoretical guarantees in a linear functional regression setting. On the other hand, the signature method presents a great deal of flexibility, which can make precise applications challenging. We therefore derive a canonical collection of choices that provide a domain-agnostic starting point. We derive these choices as a result of an extensive empirical study on 26 datasets, and go on to show competitive performance against current benchmarks for multivariate time series classification.