

An Introduction to Topological Data Analysis

Frédéric Chazal* Pascal Massart[†] Bertrand Michel[‡]

Abstract

Topological Data Analysis (TDA) is a recent and fast growing field at the crossing of mathematics, computer science and statistics. It is mainly motivated by the idea that topology and geometry provide a powerful approaches to infer, analyze and exploit robust qualitative and quantitative information about the structure of data represented, in general, as point clouds or samples in Euclidean or more general metric spaces. With the emergence and development of persistent homology theory, computational topology and geometry have brought new efficient mathematical and computational tools to infer, analyze and exploit the topological and geometric structure of complex data. The first part of the lecture is an introduction to TDA and persistent homology. It will give a presentation of the mathematical foundations of the field with a particular focus on stability properties of persistent homology that play a fundamental role in the understanding of the statistical behavior topological descriptors of data. The second part of the lecture will focus on the statistical aspects of TDA, from a mathematical and practical point of view. The last part of the lecture will be about a statistical analysis of the Mapper algorithm, a popular algorithm in the field of TDA

*INRIA Paris Saclay

[†]Université Paris Sud

[‡]École Centrale Nantes