GEOMETRIC METHODS FOR STATISTICAL LEARNING AND HIGH-DIMENSIONAL DATA

Mauro Maggioni (*Duke University)*

12:00 h., miércoles, 17 de septiembre de 2014,
Aula 520 (Depto. Matemáticas UAM)

Abstract

We discuss a family of ideas, algorithms, and results for analyzing various new and classical problems in the analysis of high-dimensional data sets. Many of these approaches are inspired by ideas in harmonic analysis. In the first part of the talk we discuss Laplacians and diffusion processes on graphs associated with high-dimensional point clouds, and their applications to the problem of parametrizing point clouds. In the second part of the talk we discuss novel multiscale geometric decompositions and approximations of point clouds, and exploit such decompositions to perform a variety of tasks in signal processing and statistical learning. In particular, we discuss the problem of dictionary learning, where one is interested in constructing, given a training set of signals (e.g. images), a set of vectors (dictionary) such that the signals admit a sparse representation in terms of the dictionary vectors. Finally, we discuss extensions of these constructions that enable one to learning estimators for probability measures generating the data, and for learning certain types of stochastic dynamical system in high-dimensions.