

**Seminario Conjunto: Aprendizaje de Máquinas y Modelamiento Estocástico.
Mini Workshop: Computational Optimal Transport**

Lugar: Sala (B04, piso -1, Beauchef 851)

Fecha: Lunes 26, noviembre, 2018.

Hora: 1500 - 1700hrs

Presentadora: Dr. Elsa Cazelles

Título: Statistical properties of barycenters in the Wasserstein space

Hora: 1500hrs

Abstract:

In this work, we discuss the analysis of data in the form of probability measures on $\mathbb{R}^d$. The aim is to provide a better understanding of the usual statistical tools on this space endowed with the Wasserstein distance. The first order statistical analysis is a natural notion to consider, consisting of the study of the Fréchet mean (or barycenter). In particular, we focus on the case of discrete data (or observations) sampled from absolutely continuous probability measures (a.c.) with respect to the Lebesgue measure. We thus introduce an estimator of the barycenter of random measures, penalized by a convex function, making it possible to enforce its a.c. Another estimator is regularized by adding entropy when computing the Wasserstein distance. We are particularly interested in controlling the variance of these estimators. Thanks to these results, the principle of Goldenshluger and Lepski allows us to obtain an automatic calibration of the regularization parameters. We then apply this work to the registration of multivariate densities, especially for flow cytometry data. We also propose a test statistic that can compare two multivariate distributions, efficiently in terms of computational time. Finally, we perform a second-order statistical analysis to extract the global geometric tendency of a dataset, also called the main modes of variation. For that purpose, we propose algorithms allowing to carry out a geodesic principal components analysis in the space of Wasserstein.

Presentador: Gonzalo Ríos

Título: Bayesian learning with Wasserstein barycenters

Hora: 1600hrs

Abstract:

In this work, we introduce a novel paradigm for Bayesian learning based on optimal transport theory. Namely, we propose to use the Wasserstein barycenter of the posterior law on models, as an alternative to the maximum a posteriori estimator and Bayesian model average. We exhibit conditions granting the existence and consistency of this estimator, discuss some of its basic and specific properties, and provide insight into the theoretical advantages of the proposed estimator. Finally, we show methods and practical implementations for calculating this estimator, relying on standard sampling methods, empirical approximations and present novel stochastic gradient-based methods. This work is in collaboration with Julio Backhoff, Joaquin Fontbona and Felipe Tobar.