Atelier
« Méthodes de calcul pour le transport optimal
18—22 juillet 2016

Workshop
"Computational Optimal Transportation"
July 18—22, 2016

# Stochastic optimization for large-scale optimal transport

## Aude Genevay[*]

aude.genevay@gmail.com

Optimal transport (OT) defines a powerful framework to compare probability distributions in a geometrically faithful way. However, the practical impact of OT is still limited because of its computational burden. We propose a new class of stochastic optimization algorithms to cope with large-scale problems routinely encountered in machine learning applications. These methods are able to manipulate arbitrary distributions (either discrete or continuous) by simply requiring to be able to draw samples from them, which is the typical setup in high-dimensional learning problems. This alleviates the need to discretize these densities, while giving access to provably convergent methods that output the correct distance without discretization error. These algorithms rely on two main ideas : (a) the dual OT problem can be re-cast as the maximization of an expectation; (b) entropic regularization of the primal OT problem results in a smooth dual optimization optimization which can be addressed with algorithms that have a provably faster convergence. We instantiate these ideas in three different setups: (i) when comparing a discrete distribution to another, we show that incremental stochastic optimization schemes can beat Sinkhorn's algorithm, the current state-of-the-art finite dimensional OT solver; (ii) when comparing a discrete distribution to a continuous density, a semi-discrete reformulation of the dual program is amenable to averaged stochastic gradient descent, leading to better performance than approximately solving the problem by discretization ; (iii) when dealing with two continuous densities, we propose a stochastic gradient descent over a reproducing kernel Hilbert space (RKHS). This is currently the only known method to solve this problem, apart from computing OT on finite samples. We backup these claims on a set of discrete, semi-discrete and continuous benchmark problems.

*This is joint work with G. Peyré, M. Cuturi, and F. Bach.*

[*]MIDO, Université Paris-Dauphine, Place de Lattre de Tassigny, 75016 Paris, FRANCE