Stochastic algorithms for optimal transport

Antoine Godichon-Baggioni and Olivier Wintenberger,
antoine.godichon_baggioni@upmc.fr, olivier.wintenberger@upmc.fr
Laboratoire de Probabilités, Statistique et Modélisation
Sorbonne-Université, 75005 Paris, France

Regularized optimal transport is more and more studied in machine learning as a competitive tool to compare probability measures. We focus here on the well-known Kantorovitch formulation of the transport optimal problem. More precisely, considering probability measures μ, ν on $\mathcal{X}$, $\mathcal{Y}$ and a cost function $c : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}^d$, the aim is to minimize the function

$$W_\epsilon\left(\mu,\nu\right) := \min_{\pi\in\Pi(\mu,\nu)} \int_{\mathcal{X}\otimes\mathcal{Y}} c(x,y)d\pi(x,y) + \epsilon KL(\pi|\mu\times\nu),$$

where $\pi(\mu,\nu)$ is the set of probability measure on $\mathcal{X}\times\mathcal{Y}$ whose marginals are μ and ν. Furthermore, $\epsilon \geq 0$ and $KL(.|.)$ is the Kullback-Leibler divergence defined by

$$KL(\pi|\mu\times\nu) = \int_{\mathcal{X}\times\mathcal{Y}} \left(\log\left(\frac{d\pi}{d(\mu\otimes\nu)}(x,y)\right) - 1\right) d\pi(x,y).$$

Remark that the term $\epsilon KL(\pi|\mu\times\nu)$ can be seen as a regularization term. The aim of the internship is to adapt usual stochastic optimization methods (stochastic gradient methods for instance) to approximate the solution. In this aim, several lines can be devised. The first phase might be a probabilistic one, studying the course notes http://www.math.columbia.edu/~mnutz/docs/EOT_lecture_notes.pdf to familiarize with optimal transport domain. The second one might be an online learning one: one can study the approximation of the optimal transport by Stochastic Gauss-Newton algorithm, based on the recent work of (Bercu et al., 2021) in the particular case where $\mathcal{Y}$ is a discrete space.

References

Bercu, B., Bigot, J., Gadat, S., and Siviero, E. (2021). A stochastic gauss-newton algorithm for regularized semi-discrete optimal transport. *arXiv preprint arXiv:2107.05291*.