

Le transport optimal pour le machine learning

Encadrants :

Gabriel Peyré (ENS) et Marco Cuturi (ENSAE)
(gabriel.peyre@ens.fr, marco.cuturi@ensae.fr)

Sujet :

Le transport optimal (TO) est un champ de recherche très actif à l'intersection des probabilités, des équations aux dérivées partielles et de l'optimisation. Le TO propose une manière canonique de définir, à partir d'une distance donnée sur un espace d'observations, une nouvelle distance entre les mesures décrivant des probabilités sur ces observations. Cette distance de TO a de nombreux intérêts. Elle permet en particulier de manipuler et d'interpoler d'une façon géométrique des distributions de probabilité, ce qui ouvre des perspectives très riches en imagerie et en machine learning. On pourra consulter le livre [2] pour un tour d'horizon du transport optimal, des algorithmes de calculs, et des applications.

L'émergence de nouveaux algorithmes de calculs ouvrent la porte à l'application du transport optimal pour la résolution de problème en grande dimension pour le machine learning. Cependant, les distance de TO ne sont pas estimables efficacement en grande dimension, car le nombre d'échantillons nécessaires croît exponentiellement vite avec la dimension, comme détaillé dans [1]. Le but de ce projet est de comprendre cette analyse théorique et de faire des tests numériques pour la confirmer. On pourra ensuite contraster ces résultats pour le transport optimal avec ceux, plus simples, obtenus pour des distances Euclidiennes [3].

Plusieurs sujets de stages reliés à ce projet sont possibles, en France (Nice, Paris, Bordeaux, Rennes, Nancy), en Europe (Cambridge, Pise, Munich, Berlin, Bonn), en Amérique du Nord (Montreal, Alberta, MIT, NYU) ou au Japon (Tokyo).

Références

- [1] J. Weed, F. Bach. Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. Technical Report, Arxiv-1707.00087, 2017
- [2] G. Peyré, M. Cuturi, Computational Optimal Transport <https://optimaltransport.github.io/>
- [3] Sriperumbudur, B., Fukumizu, K., Gretton, A., Schoelkopf, B., et Lankriet, G., On the empirical estimation of integral probability metrics, Electronic Journal of Statistics, 6 , pp.1550-1599, 2012.