Extensions du transport optimal: divergence de Sinkhorn et Gromov-Wasserstein "unbalanced"

Thibault SÉJOURNÉ, DMA, ENS, PSL - Paris

Gabriel PEYRÉ, DMA, ENS, PSL - Paris
François-Xavier VIALARD, LIGM, UPEM - Marne La Vallée

Le transport optimal, théorie permettant de définir des distances entre distributions, est un outil de choix dans les domaines de l'apprentissage machine et de l'estimation statistique car il prend en compte la géométrie de l'espace sous-jacent. Ces distances souffrent cependant de trois limitations pouvant être problématiques : (i) elles sont coûteuses à calculer, (ii) se limitent à la comparaison de probabilités et (iii) comparent des mesures définies sur le même espace. Ces contraintes peuvent être gênantes pour passer à l'échelle dans les calculs, pour être insensible aux "outliers" géométriques (dûs à des données bruitées), ou comparer des graphes (tels que des molécules aux structures différentes). Pour pallier ces limitations ont été proposés la régularisation entropique [1], le transport non-équilibré [2] et la distance de Gromov-Wasserstein [3].
Dans cette présentation j'introduirai d'abord la formulation non-équilibrée du transport, ainsi que sa variante entropique. Je détaillerai une variante de l'algorithme de Sinkhorn permettant de calculer le dual du problème grâce à une modification mineure de l'algorithme dans sa version équilibrée, avec une convergence linéaire. Je montrerai que le gain en temps de calcul dû à l'ajout d'une régularisation entropique se fait au détriment des propriétés mathématiques, mais qu'il est possible de corriger ce problème en définissant la divergence de Sinkhorn, une fonctionnelle convexe, positive, définie et métrisant la convergence en loi (travail joint avec Jean Feydy et Alain Trouvé) [4,5].
Dans un second temps, je présenterai la distance de Gromov-Wasserstein qui est un problème d'optimisation quadratique non convexe comparant des espaces munis d'une métrique et d'une mesure positive. Je définirai deux généralisations non-équilibrée de cette distance, l'une étant une borne supérieure pour l'autre. Je montrerai que la première définit une distance entre espaces métriques mesurés, et que pour la seconde il est possible de la calculer grâce à une régularisation entropique comme une suite de problème de transport non-équilibrés [6].

Références :

[1] Cuturi, M. (2013). Sinkhorn distances : Lightspeed computation of optimal transport. Advances in neural information processing systems, 26, 2292-2300.
[2] Liero, M., Mielke, A., & Savaré, G. (2018). Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. Inventiones mathematicae, 211(3), 969-1117.
[3] Mémoli, F. (2011). Gromov–Wasserstein distances and the metric approach to object matching. Foundations of computational mathematics, 11(4), 417-487.
[4] Feydy, J., Séjourné, T., Vialard, F. X., Amari, S. I., Trouvé, A., & Peyré, G. (2019, April). Interpolating between optimal transport and MMD using Sinkhorn divergences. In The 22nd International Conference on Artificial Intelligence and Statistics (pp. 2681-2690). PMLR.
[5] Séjourné, T., Feydy, J., Vialard, F. X., Trouvé, A., & Peyré, G. (2019). Sinkhorn divergences for unbalanced optimal transport. arXiv preprint arXiv :1910.12958.
[6] Séjourné, T., Vialard, F. X., & Peyré, G. (2020). The Unbalanced Gromov Wasserstein Distance : Conic Formulation and Relaxation. arXiv preprint arXiv :2009.04266.

Contact : thibault.sejourne@ens.fr