

Propriétés d'Optimisation de la Distance de Sliced Wasserstein Discrète

Eloi TANGUY, MAP5, Université Paris-Cité - Paris

La distance de Sliced Wasserstein (SW) est devenue une alternative courante à la distance de Wasserstein pour la comparaison des mesures de probabilité. Les applications répandues incluent le traitement d'images, l'adaptation de domaine et les modèles génératifs, où il est typique d'optimiser certains paramètres afin de minimiser SW, qui dans la pratique sert de fonction de perte entre des mesures de probabilité discrètes. Ces problèmes d'optimisation présentent tous le même sous-problème, qui consiste à minimiser la distance SW entre deux mesures discrètes uniformes avec le même nombre de points en fonction du support (c'est-à-dire une matrice de données) de l'une des mesures.

Après une présentation du Transport Optimal en dimension 1 et quelques détours par un problème de reconstruction de mesures discrètes à partir de projections, nous nous attaquerons aux propriétés de l'énergie définie par la distance de SW entre une mesure discrète dont nous varions le support, et une mesure discrète de référence. Nous étudions la régularité et les propriétés d'optimisation de cette énergie, ainsi que son approximation de Monte-Carlo (estimation de l'espérance de SW en utilisant des échantillons sur les projections), ainsi que les propriétés statistiques asymptotiques et non asymptotiques de l'approximation de Monte-Carlo. Enfin, nous montrons que dans un certain sens, les méthodes de descente de gradient stochastique minimisant ces énergies convergent vers des points critiques (généralisés), avec une extension à l'entraînement des réseaux neuronaux génératifs.