Generative modeling project : Neural Optimal Transport

Paul Barbier[1] **and Bastien Le Chenadec**[1]

[1]École des Ponts ParisTech, Master MVA

Contribution Statement

1 Introduction

Optimal Transport (OT) is a mathematical framework that aims to find the most efficient way to transport a distribution of mass to another. This framework has been used extensively in the context of generative models, for instance as a loss function in the training of Generative Adversarial Networks (GANs) or by learning a mapping between two distributions. In this project, we aim to study the paper "Neural Optimal Transport" (Korotin, 2023) [1] which introduces an algorithm to train a neural network to learn the optimal transport between two distributions.

2 Background on Optimal Transport

Let μ and ν be two probability distributions on $\mathcal{X}$ and $\mathcal{Y}$ respectively (typically $\mathcal{X}, \mathcal{Y} = \mathbb{R}^n, \mathbb{R}^m$). To give a meaning to "efficiently" transporting mass, we need to define a cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ that quantifies the cost of transporting a unit of mass in $\mathcal{X}$ to one in $\mathcal{Y}$. The (Monge) optimal transport problem consists in finding a **transport map** $T^* : \mathcal{X} \to \mathcal{Y}$ such that :

$$T^* \in \underset{T\#\mu=\nu}{\text{Argmin}} \int_{\mathcal{X}} c(x, T(x)) d\mu(x) \tag{1}$$

where $T\#\mu$ is the pushforward distribution of μ by T, defined by $(T\#\mu)(A) = \mu(T^{-1}(A))$ for any measurable set $A \subset \mathcal{Y}$. This formulation calls for a deterministic mapping from $\mathcal{X}$ to $\mathcal{Y}$, which is not always desirable or feasible under general assumptions. Kantorovich introduced a more general OT problem that aims at finding a **transport plan** $\pi^* \in \Pi(\mu, \nu)$ in the set of joint distributions on $\mathcal{X} \times \mathcal{Y}$ with marginals μ and ν such that :

$$\pi^* \in \underset{\pi\in\Pi(\mu,\nu)}{\text{Argmin}} \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) d\pi(x, y) \tag{2}$$

In general the solution to the Kantorovich problem is stochastic, but in some cases it may be deterministic in which case it is also a solution to the Monge problem. Following this idea of stochasticity in the solution, weak OT was introduced as a relaxation of the Kantorovich problem, where the cost function is of the form $C : \mathcal{X} \times \mathcal{P}(\mathcal{Y}) \to \mathbb{R}$. In this case the weak OT problem writes :

$$\pi^* \in \underset{\pi\in\Pi(\mu,\nu)}{\text{Argmin}} \int_{\mathcal{X}} C(x, \pi(\cdot|x)) d\pi(x) \tag{3}$$

where $\pi(\cdot|x)$ is the conditional distribution of π given x and $d\pi(x)$ is the marginal distribution of π on $\mathcal{X}$.

Building on this framework, (Korotin, 2023) [1] introduce **stochastic maps** $T : \mathcal{X} \times \mathcal{Z} \to \mathcal{Y}$ where $\mathcal{Z}$ is a latent space corresponding to the randomness in the transport. They show that the weak optimal transport problem can be reformulated and solved by a SGAD algorithm. This approach is particularly interesting in the context of generative modeling, as it allows to learn a stochastic mapping between two distributions.

3 Conclusion

References

[1] Alexander Korotin, Daniil Selikhanovych, and Evgeny Burnaev. Neural optimal transport. *arXiv (Cornell University)*, 1 2022. doi: 10.48550/arxiv.2201.12220. URL https://arxiv.org/abs/2201.12220.

APPENDIX

A FIGURES