DIMENSION REDUCTION FOR OPTIMAL TRANSPORT AND NORMALIZING FLOWS

MARTIN LOTZ

1. MSc Project description

Suppose we are given a space $\mathcal{X}$ and a probability measure μ on $\mathcal{X}$ (typically with a density ρ). Three of the most common problems are:

(1) generate samples from μ;
(2) evaluate the density ρ at a point $\boldsymbol{x}$;
(3) evaluate an integral $\mathbb{E}_\mu[f] = \int_{\mathcal{X}} f(\boldsymbol{x})\rho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

These problems are all crucial in the context of Markov Chain Monte Carlo (MCMC) methods. When the space $\mathcal{X}$ can be described using algebraic equations (for example, when considering configuration spaces of molecules), then one can speed up sampling and integrating on $\mathcal{X}$ considerably by intersecting with random subspaces, using ideas from integral geometry [1]. The project will explore this idea in to case studies.

A popular approach is to model the distribution as the push-forward measure of a simpler distribution via an invertible map $T\colon \mathcal{Z} \to \mathcal{X}$ [3]. These ideas are closely related to Generative Adversarial Networks (GANs). The map T is assumed to be a composition of simpler maps, often implemented in the form of a deep neural network. However, the invertibility assumption seems more restrictive than necessary, and one obtains smoother constructions by allowing T to be merely surjective. While one can artificially augment the data [2], a more principled approach is to replace the Jacobian of the transformation by an integral of a normal Jacobian, using the co-area formula of integral geometry, and then to suitably approximate this integral. This approach will be tested extensively using data sets for image interpolation and generation.

A second application is in inverse optimal transport. Given two probability distributions μ and ν on $\mathcal{X}$ and $\mathcal{Y}$, an optimal transport plan with respect to a cost function $c(x, y)$ is a way to move particles from $\mathcal{X}$ to $\mathcal{Y}$ in a way that preserves measure and minimizes the overall cost. In some situations (migration flows), one observes a transport plan and is interested in recovering the cost function [4]. Solving the resulting optimization problem using MCMC is very expensive, and the potential for speeding up the computation using random intersections will be explored. This problem can also be cast in the framework of normalizing flows, where the map is parametrized by the (unknown, but learnable) cost function.

2. Relation to PhD project

The proposed project is linked to a PhD project that aims to develop efficient Mutual Information (MI) estimation algorithms. Mutual Information, or the related concepts of Entropy and Kullback-Leibler divergence, measures the distance between probability distributions and is used as objective function for the optimization problems arising in the MSc project. Besides giving the student the opportunity to learn a set of essential skills needed for the PhD project, the MSc project will also provide useful case studies and applications. The associated PhD project is supported by BT Wireless Research.

References

[1] Paul Breiding and Orlando Marigliano. Random points on an algebraic manifold, 2018.
[2] Chin-Wei Huang, Laurent Dinh, and Aaron Courville. Augmented normalizing flows: Bridging the gap between generative flows and latent variable models, 2020.
[3] George Papamakarios, Eric Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan. Normalizing flows for probabilistic modeling and inference. *arXiv preprint arXiv:1912.02762*, 2019.
[4] Andrew M Stuart and Marie-Therese Wolfram. Inverse optimal transport. *SIAM Journal on Applied Mathematics*, 80(1):599–619, 2020.