APMTH 221 FINAL PROJECT PROPOSAL

JIAFENG (KEVIN) CHEN AND FRANCISCO RIVERA

1. Collaboration

Francisco and Kevin will work together to research the relevant optimization techniques for efficiently computing Wasserstein distances and solving optimal transport problems. Then, they will split up techniques, and through implementations of them, empirically (and when possible, theoretically) assess their performance. Time permitting, they will then attempt to create a new algorithm that performs competitively with state of the art benchmarks.

2. Problem Statement

Given two probability distributions p_X and p_Y, the p^{th} *Wasserstein distance* is the following optimization problem:

$$W_p(p_X, p_Y) = \left(\inf_{(X,Y)} \left\{ \mathbb{E}_{(X,Y)} \left[d(X,Y)^p \right] : X \sim p_X, Y \sim p_Y \right\} \right)^{1/p},$$

where the minimum is taken over all joint distributions with marginals p_X, p_Y. The Wasserstein distance is a probabilistic view of a class of problems called *optimal transport problems* [2]. If we think of a probability distribution as a collection of "mass" at different points in M (with density of the mass proportional to the probability measure), then the Wasserstein distance captures the cost of transforming one distribution into the other in units of mass times distance. This analogy is particularly interpretable if we consider discrete probability measures, which are collections of point masses.

Solving the optimization problem underpinning this metric is in general hard and remains an unsolved problem. We will aim to test the effectiveness of different optimization techniques to this problem, with particular focus on discrete distributions.

3. Motivation

Numerous important applications exist for the Wasserstein metric. For instance, if we consider images as a discrete probabilistic distribution over $\mathbb{R}^2$,[1] the Wasserstein distance turns out to mimick human perceptions of similarity: Images that are similar to humans are close in Wasserstein distances. Other distances over distributions, such as the well-known Kullback-Leibler divergence, do not have such properties. Thus, Wasserstein distances are extremely important in computer vision, machine learning, and statistics.

4. Data

As this is a largely theoretical and methodological project, we shall rely on numerical simulation and analysis. We will generate test distributions using simulation and do not need to concern ourselves with finding real data. When necessary, we will use standard distributions found in the literature for consistency.

5. Deliverables

We will implement optimization techniques to calculate Wasserstein distances and attempt a novel algorithm that performs competitively. The deliverable will be a survey of recent literature and a report of our contributions.

Date: February 21, 2018.

[1]For example, we may consider black-and-white images, where the probability mass is proportional to the pixel intensity.

6. Next Steps

First, we will review the current literature and familiarize ourselves with cutting-edge techniques in solving optimal transport problems.

We have the following tentative ideas which we may or may not individually implement, but from which we will source our work,

- We will use linear programs to calculate Wasserstein distances for discrete metric spaces.[2] We will explore average complexity for randomly generated test cases.
- Training a neural network to hot-start an optimization algorithm for the discrete case, following recent literature on the creative use of deep learning techniques for difficult optimization problems as in [1].
- Exploring other relevant optimization techniques such as simulated annealing and genetic algorithms.

References

[1] Paul Dütting, Zhe Feng, Harikrishna Narasimhan, and David C Parkes. Optimal auctions through deep learning. *arXiv preprint arXiv:1706.03459*, 2017.

[2] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport. Technical report, 2017.

[2] This is the Kantorovich problem, see [2]