Background for Optimal Transport

Michael Wilson

Abstract

This document is a brief introduction to the theoretical foundations of optimal transport and Wasserstein Distances.

1 Introduction

The Wasserstein distance is a distance between probability measures. Optimal Transport between measures involves finding a push-forward measure (associated with a continuous map) that minimizes the Wasserstein distance. For this work, we will restrict ourselves to considering the squared Wasserstein 2-Distance between probability measures μ_0 and μ_1;

$$W_2^2(\mu_0, \mu_1) = \inf_{\pi \in \Pi(\mu_0, \mu_1)} E||X - Y|| = \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\mathcal{X} \times \mathcal{Y}} ||x - y|| d\gamma(x, y) \tag{1}$$

If γ is absolutely continuous with respect to the Lebesgue measure, then we say γ has a density $f(x, y) = \frac{\delta \gamma}{\delta \lambda}$, and can write this expectation as a Riemann integral;

$$= \inf_{\gamma \in \Pi(\mu_0, \mu_1)} \int_{\mathcal{X}} \int_{\mathcal{Y}} ||x - y|| f(x, y) dx dy \tag{2}$$