Optimal Transport

Preliminaries and Applications

Kirby

Department of Mathematics
University of Utah

Feburary 2022

Optimal Transportation

History

Originally a question of optimally moving ammunition from factory to battlefield by Gaspard Monge in 1781 during Napoleonic France.

Optimal Transportation

History

Originally a question of optimally moving ammunition from factory to battlefield by Gaspard Monge in 1781 during Napoleonic France.

Formulation

Let X and Y be measure spaces with measures μ, ν respectively. In order to transport $x \in X$ to $y \in Y$, we let $\mu(X) = \nu(Y)$. The optimal transport problem is minimizing the total effort (through nonnegative $c : X \times Y \to \mathbb{R}$) over μ and ν.

Monge Formulation

Transport Map

We call a function $T : X \to Y$ a *transport map* if $T_{\#}\mu = \nu$ (that is, $T_{\#}\mu$ is the pushforward of ν). Furthermore, we say T transports μ to ν.

Monge Formulation

Transport Map

We call a function $T : X \to Y$ a *transport map* if $T_{\#}\mu = \nu$ (that is, $T_{\#}\mu$ is the pushforward of ν). Furthermore, we say T transports μ to ν.

Monge Formulation

Let

$$\mathcal{T}(\mu,\nu) = \{T : X \to Y | T_{\#}\mu = \nu, \ T \text{ measurable}\} \tag{1}$$

Then we say that $\mathbb{M}(\mu,\nu)$ is the minimizer

$$\mathbb{M}(\mu,\nu) = \inf_{T \in \mathcal{T}(\mu,\nu)} \int_X c(x, T(x)) d\mu(x). \tag{2}$$

Monge Formulation Restrictions

Splitting Mass

Because we have $T : X \to Y$ with $T_{\#}\mu = \nu$, T is bijective, so it cannot *split mass*. This means for $y_1, y_2 \in Y$, a point $x \in X$ cannot have $T(x) = \{y_1, y_2\}$.

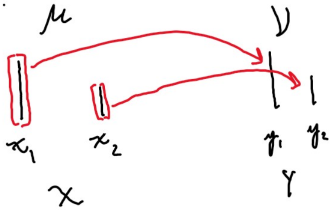

Figure: If we let $\mu = \frac{2}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2}$ and $\nu = \frac{2}{3}\delta_{y_1} + \frac{1}{3}\delta_{y_2}$ the only valid transport map is in red.

Kantorovich Formulation

Transport Plan

We call a measure $\pi \in \mathcal{P}(X \times Y)$ a *transport plan* if for all measurable $A \subset X$ and $B \subset Y$,

$$\pi(A \times Y) = \mu(A) \quad \pi(X \times B) = \nu(B). \tag{3}$$

Kantorovich Formulation

Transport Plan

We call a measure $\pi \in \mathcal{P}(X \times Y)$ a *transport plan* if for all measurable $A \subset X$ and $B \subset Y$,

$$\pi(A \times Y) = \mu(A) \quad \pi(X \times B) = \nu(B). \tag{3}$$

Kantorovich Formulation

Let

$$\Pi(\mu, \nu) = \{\pi | \pi \in \mathcal{P}(X \times Y) \text{ constrained by (3)}\} \tag{4}$$

Then we say that $\mathbb{K}(\mu, \nu)$ is the minimizer

$$\mathbb{K}(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} c(x, y) d\pi(x, y). \tag{5}$$

Kantorovich Formulation as a Relaxation

We no longer have to be restricted to not splitting mass!

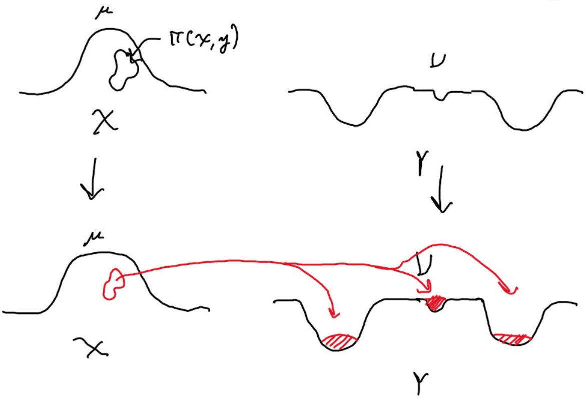

Figure: Splitting mass is no problem as now $\pi(x, y)$ can be defined as necessary to provide an optimal transport scheme.

Shipper's Problem

It costs $c(x_1, y_1)$ dollars for a factory to transport coal from mine x_1 to factory y_1. I tell the mine owner that I can charge them $\phi(x_1)$ dollars to pick up at mine x_1 and charge $\psi(y_1)$ to deliver to factory y_1, then, in order for the factory owner to agree to my terms,

$$\phi(x) + \psi(y) \leq c(x, y)$$

for every location x and destination y. I can make this sum as close to $c(x, y)$ as I want, so I can maximize my profit while minimizing the factory owner's effort.

Deriving Duality

For a measure π, we see that we have $I[\pi, \phi, \psi]$ where

$$\sup_{\phi \in C(X), \psi \in C(Y)} \left\{ \int_X \phi(x) d\mu + \int_Y \psi(y) d\nu - \int_{X \times Y} (\phi(x) + \psi(y) d\pi) \right\}$$

is 0 when $\pi \in \Pi(\mu, \nu)$ and $+\infty$ otherwise. Now, writing this as a Kantorovich Problem,

$$\mathbb{K}(\mu, \nu) = \inf_{\pi \geq 0} \left\{ \int_{X \times Y} c(x, y) d\pi(x, y) + \sup_{\phi \in C(X), \psi \in C(Y)} I[\pi, \phi, \psi] \right\} \quad (6)$$

We can informally$\star$ exchange the sup and inf to yield

$$\sup_{\phi, \psi} \left\{ \int_X \phi(x) d\mu + \int_Y \psi(y) d\nu + \inf_{\pi \in \Pi} \int_{X \times Y} c(x, y) - (\phi(x) + \psi(y) d\pi) \right\} \quad (7)$$

Kantorovich Duality

Following from (7), we see that

$$\inf_{\pi\in\Pi}\int_{X\times Y} c(x,y)-(\phi(x)+\psi(y)d\pi)) = \begin{cases} 0 & \text{if } \phi(x)+\psi(y)\le c(x,y) \\ +\infty & \text{otherwise.} \end{cases}$$

This gives us the dual problem

Dual Problem

For $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ and a nonnegative cost function $c : X \times Y \to \mathbb{R}$, the Kantorovich dual problem is

$$\sup_{\phi\in C(X),\psi\in C(Y)} \left\{ \int_X \phi(x)d\mu(x) + \int_Y \psi(Y)d\nu(y) \right\} \tag{8}$$

subject to $\phi(x)+\psi(y)\le c(x,y)$.

Discussing Informal Exchange

There's more to the proof of Kantorovich Duality in general, but they all rely on a rigorous minimax principle.

In Villani, this is done using Fenchel-Rockafeller Duality and Legendre Transforms in section 1.1.

Some more important assumptions is that X, Y are taken to be Polish spaces and c is taken to be lower-semicontinuous in order to guarantee existence of the Kantorovich problem.

This is definitely something I can talk about more in dms, so if you have specific questions, ping me in a proper channel or message me at kirby#4923.

Wasserstein Distance

Let's start by considering Kantorovich-Monge solutions to the transport problem with cost functions $c(x, y) = d(x, y)^p$ where d is the distance on X and Y. Let $\Omega \subset X$ and set $x_0 \in X$.

$$\mathcal{P}_p(X) = \left\{ \mu \in \mathcal{P}(X) : \int_\Omega d(x, x_0)^p < +\infty \right\}$$

is the admissable class of measures μ, even if Ω is unbounded.

Lemma (Kantorovich-Monge forms a metric.)

Let $W_p(\mu, \nu) = \left[\inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y)^p d\pi(x, y) \right]^{1/p}$ defined for $\mu, \nu \in \mathcal{P}_P(X)$ be the Wasserstein Distance. This forms a metric on X.

Wasserstein Topology

Before showing that the Wasserstein Metric has a corresponding space, we see that if $x_n \to x$ in X then $W_p(\delta_{x_n}, \delta_x) = d(x_n, x)^{\inf(1,p)} \to 0$. In general, this looks like

Theorem (Wasserstein distances metrize weak convergence)

Let $p \in (0, \infty)$, let $(\mu_k)_{k=1}^{\infty}$ be a sequence of measures in $\mathcal{P}_p(X)$ and $\mu \in \mathcal{P}(X)$, then the following are equivalent[a]:

1. $W_p(\mu_k, \mu) \to 0$
2. *$\mu_k \to \mu$ in the weak sense, meaning for $h \in C_b(X)$,*

$$\lim_{k \to \infty} \int h d\mu_k = \int h d\mu.$$

[a] The theorem in Villani is much stronger and gives more equivalences, but due to time, I just want to make this statement.

Furthermore, we usually look at $p = 2$, and say $\mathbb{W}_2$ is the space $\mathcal{P}_2(X)$ endowed with the metric W_2.

Continuity Equation

Continuity Equation.

The linear transport equation