How to compute WOT?

Guillaume Carlier, Hugo Malamut, Maxime Sylvestre
Université Paris-Dauphine PSL (CEREMADE)

May 28, 2024

Optimal transport

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. Let $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ l.s.c., lower bounded. The optimal transport problem is the following

$$\inf_{\pi \in \Pi(\mu,\nu)} \int c(x,y) d\pi(x,y) \tag{OT}$$

where $\Pi(\mu, \nu)$ is the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ, ν.
Usually the cost $c(x,y) = \|x - y\|^2$ is used.

Entropic optimal transport

To compute the optimal transport problem the entropy regularized problem is often used to compute optimal transport [Cut13]

$$\inf_{\pi \in \Pi(\mu,\nu)} \int c(x,y) d\pi(x) + \epsilon H(\pi \mid \mu \otimes \nu) \tag{EOT$_\epsilon$}$$

where $H(\pi \mid \mu \otimes \nu)$ is the relative entropy of π with respect to $\mu \otimes \nu$. The Sinkhorn algorithm computes the solution to the dual problem.

$$\sup_{\phi,\psi} \int \phi d\mu + \int \psi d\nu - \epsilon \int \exp\left(\frac{\phi(x) + \psi(y) - c(x,y)}{\epsilon}\right) d\mu \otimes \nu(x,y)$$

which by duality is equivalent to the entropic optimal transport problem. We moreover have the following relationship:

$$\frac{d\pi^*}{d\mu \otimes \nu}(x,y) = \exp\left(\frac{\phi(x) + \psi(y) - c(x,y)}{\epsilon}\right)$$

Sinkhorn's Algorithm

Sinkhorn's algorithm is essentially an alternatig minimization algorithm on each variables ϕ, ψ.
Indeed fixing ψ_k and maximizing in ϕ the first order conditions grant that

$$\phi_{k+1}(x) = \epsilon \ln \left(\int \exp \left(\frac{\psi_k(y) - c(x,y)}{\epsilon} \right) d\nu(y) \right)$$

similarily for ψ we have

$$\psi_{k+1}(y) = \epsilon \ln \left(\int \exp \left(\frac{\phi_{k+1}(x) - c(x,y)}{\epsilon} \right) d\mu(x) \right)$$

Towards weak optimal transport

Both problems require to minimize a convex function in π over $\Pi(\mu, \nu)$. And in fact they have a common formulation.

If we disintegrate π with respect to the projection onto the first variable we can write $\pi = \mu \otimes \pi_x$ and thus

$$\int c d\pi = \int \left(\int c(x, y) d\pi_x(y) \right) d\mu(x), \quad H(\pi \mid \mu \otimes \nu) = \int H(\pi_x \mid \nu) d\mu(x)$$

where the second equality holds by additivity of the entropy.

Note that now both functions are of the form

$$\int c(x, \pi_x) d\mu(x)$$

where for every x the function $c(x, .)$ is convex.

Weak Optimal Transport

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. Let $c : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ jointly l.s.c., lower bounded and convex in the second variable. We study the following problem called weak optimal transport [GRST15]

$$\inf_{\pi \in \Pi(\mu,\nu)} \int c(x, \pi_x) d\mu(x) \tag{WOT}$$

Existence of solutions by [VBP19] and a review of applications by [BVP20].

Examples

Entropy regularized optimal transport :

$$c(x,\pi_x) = \int c(x,y)d\pi_x + \epsilon H(\pi_x \mid \nu)$$

Vector Quantile Regression [CCG15] :

$$c(x,\pi_x) = \begin{cases} \int \|x - p_E(y)\|^2 d\pi_x & \text{if } \int p_{E^\perp}(y)d\pi_x = 0 \\ +\infty & \text{else} \end{cases}$$

Monopolist problem [DDT17]:

$$c(x,\pi_x) = \inf_{q \le \pi_x} \|x - \int y dq\|$$

Examples

Strassen Theorem [GJ20]:

$$c(x, \pi_x) = \|x - \int y d\pi_x\|^2$$

Martingale optimal transport [BHLP13]:

$$c(x, \pi_x) = c^0(x, \pi_x) + \theta \left(x - \int y \pi_x \right)$$

Bass martingale [BVST23] :

$$c(x, \pi_x) = \begin{cases} W_2^2(\pi_x, \mathcal{N}(x, 1)) & \text{if } \int y \pi_x = 0 \\ +\infty & \text{else} \end{cases}$$

Framework

We restrict ourselves to **compactly supported measures** μ, ν and study for $\epsilon \geq 0$ the following problem

$$\inf_{\substack{\pi \in \Pi(\mu,\nu) \\ \int f(y) d\pi_x = 0}} \int c(x, \pi_x) d\mu(x) + \epsilon H(\pi \mid \mu \otimes \nu) \tag{WOT$_\epsilon$}$$

where c has a **continuity modulus** with respect to W_∞ which is uniform in x, f is vector valued, continuous. We further assume that $\int f \otimes f d\nu$ is **invertible** (constraint qualification).

Dual formulation

As in the optimal transport case the problem WOT_ϵ admits a dual formulation. Under the hypothesis made before there is **no duality gap** and the following dual problem **has solutions** in L^∞.

$$\max_{\substack{\phi,\lambda\in L^\infty(\mu)),\psi\in L^\infty(\nu)\\ \beta\in L^\infty(\mu\otimes\nu)}} \int \phi d\mu + \int \psi d\nu - \sigma_{\epsilon,\mu\otimes\nu}(\phi+\psi+\lambda f + c_x^*(\beta_x) - \beta)$$

where

$$\sigma_{\epsilon,\mu\otimes\nu}(a) = \begin{cases} \epsilon \int \exp\left(\frac{a(x,y)}{\epsilon}\right) d\mu\otimes\nu(x,y) & \text{if } \epsilon > 0 \\ 1_{a\leq 0} & \text{else} \end{cases}$$

Convergence as $\epsilon \to 0$

The gamma convergence for EOT$_\epsilon$ [CDPS17] introduced a block approximation. Here to derive rates of convergence we introduce a slice approximation.
Let $(A_i)_i$ a partition of $\mathbb{R}^d$ the slice approximation π^A of π is defined as $\pi^A = \mu \otimes \pi_x^A$ where

$$\frac{d\pi_x^A}{d\nu} = \sum_i \frac{\pi_x(A_i)}{\nu(A_i)}$$

If the diameter of A_i is less than δ we have $H(\pi^A \mid \mu \otimes \nu) = O(-\ln(\delta))$.

Thus since f is continuous and λ is bounded, as soon as c is **uniformly hölder** we get

$$WOT_\epsilon(\mu, \nu) - WOT_0(\mu, \nu) = O(\epsilon \ln(\epsilon)).$$

SISTAlgorithm

Introduced in [CDGS20] the algorithm to compute WOT_ϵ is an extension of the Sinkhorn algorithm.
If $F(\phi, \psi, \lambda, \beta)$ is the dual function then the algorithm is the following:

$$\phi_{k+1} \in \arg\min_{\phi} F(\phi, \psi_k, \lambda_k, \beta_k)$$

$$\psi_{k+1} \in \arg\min_{\psi} F(\psi_{k+1}, \psi, \lambda_k, \beta_k)$$

$$(\lambda_{k+1}, \beta_{k+1}) = (\lambda_k, \beta_k) - \nabla_{\lambda,\beta} F(\phi_{k+1}, \psi_{k+1}, \lambda_k, \beta_k)$$

This algorithm was also successfully used in [GNT23] for geodesics extrapolation in Wasserstein space.

Numerical examples - Projection onto convex order

In [GRST15] they studied a WOT problem with the following cost $c(x, p) = \|x - \int y dp(y)\|^2$.
They proved that if the solution is π^* then $T_\sharp \mu$ is the solution to

$$\inf_{\tilde{\mu} \leq_{cv} \nu} W_2^2(\mu, \tilde{\mu})$$

where $T(x) = \int y d\pi_x$ and $\{. \leq_{cv} \nu\}$ is the set of measures $\tilde{\mu}$ such that there is a Martingale coupling $\tilde{\mu}$ and ν.

Numerical examples - Brenier Strassen interpolation

Let $\mu = \mathcal{N}(0,1)$ and $\nu = \mathcal{N}(-1,2) + \mathcal{N}(1,2)$. For $t \in [0,1]$ using the cost $c(x,p) = t \int \|x-y\|^2 dp + (1-t)\|x - \int y dp\|^2$ we interpolate between a diffusion from μ to ν and the optimal transport.

Thank you for your attention !

- Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner.
  Model-independent bounds for option prices: A mass transport approach, 2013.
- Julio Daniel Backhoff-Veraguas and Gudmund Pammer.
  Applications of weak transport theory, 2020.
- Julio Backhoff-Veraguas, Walter Schachermayer, and Bertram Tschiderer.
  The bass functional of martingale transport, 2023.
- Guillaume Carlier, Victor Chernozhukov, and Alfred Galichon.
  Vector quantile regression: An optimal transport approach, 2015.
- Guillaume Carlier, Arnaud Dupuy, Alfred Galichon, and Yifei Sun.
  Sista: learning optimal transport costs under sparsity constraints, 2020.

References ii

- Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer.
  Convergence of entropic schemes for optimal transport and gradient flows, 2017.
- Marco Cuturi.
  Sinkhorn distances: Lightspeed computation of optimal transportation distances, 2013.
- Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos.
  Strong duality for a multiple-good monopolist.
  Econometrica, 85(3):735–767, 2017.

References iii

- Nathael Gozlan and Nicolas Juillet.
  On a mixture of brenier and strassen theorems.
  Proceedings of the London Mathematical Society, 120(3):434–463, March 2020.
- Thomas Gallouët, Andrea Natale, and Gabriele Todeschi.
  From geodesic extrapolation to a variational bdf2 scheme for wasserstein gradient flows, 2023.
- Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, and Prasad Tetali.
  Kantorovich duality for general transport costs and applications, 2015.
- Julio Backhoff Veraguas, Mathias Beiglboeck, and Gudmund Pammer.
  Existence, duality, and cyclical monotonicity for weak transport costs, 2019.