Embedding problems in Wasserstein spaces

Benoît Kloeckner

Université de Grenoble I, Institut Fourier

September 2010

Contents

Contents

Transport plans

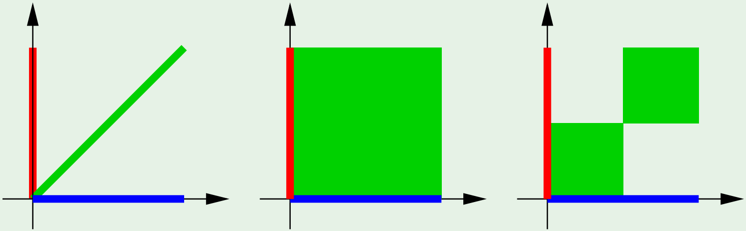
X is a polish (= complete separable) metric space,
$c : X \times X \to [0, +\infty)$ a "cost" function.

Definition

Given μ, ν probability measures, $\Gamma(\mu, \nu)$ is the set of transport plans, that is measures Π on $X \times X$ that projects on μ and ν:

$$\Pi(A \times X) = \mu(A) \quad \textit{and} \quad \Pi(X \times B) = \nu(B)$$

Three examples



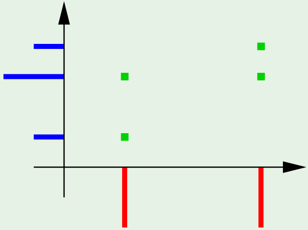
Optimal transport problem

Given μ and ν, find a transport plan $\Pi \in \Gamma(\mu, \nu)$ that is *optimal*, *i.e.* minimizes total cost:

$$\int_{X \times X} c(x,y)\,\Pi(dx\,dy)$$

On the line, with quadratic cost

If $X = \mathbb{R}$ and $c(x,y) = |x-y|^2$, the only optimal transport plan is the non-decreasing rearrangement.



Wasserstein space

Optimal transport yields a metric space of measures: the *Wasserstein space* $\mathscr{W}_2(X)$ is the set of probability measures μ having finite second moment:

$$\int_X d(x_0, x)^2 \, \mu(dx) < +\infty \quad \text{for some } x_0 \in X$$

endowed with the metric

$$W(\mu, \nu) := \left(\inf_{\Pi \in \Gamma(\mu,\nu)} \int_{X \times X} d(x, y)^2 \, \Pi(dx \, dy) \right)^{1/2}$$

Geometric questions

We aim to study the geometry of $\mathscr{W}_2(X)$, assuming prior knowledge on X.

First embedding question

Which metric spaces embed in $\mathscr{W}_2(X)$, either isometrically, bi-Lipschitz, or quasi-isometrically?
The *rank* is the maximal dimension of a Euclidean space isometric embedding.

Geometric questions

We aim to study the geometry of $\mathscr{W}_2(X)$, assuming prior knowledge on X.

Second embedding question

Into which spaces does $\mathscr{W}_2(X)$ embed, either isometrically, bi-Lipschitz, or quasi-isometrically?

Isometries question

What is the isometry group of $\mathscr{W}_2(X)$? Is it (canonically) isomorphic to the isometry groups of X?

Miscellany results

Proposition

The space $\mathscr{W}_2(\mathbb{R})$ isometrically embeds into $L^2([0,1])$.

Simply use the inverse distribution function.

Miscellany results

Given a map $\phi : X \to X$, one defines a map $\phi_\# : \mathscr{W}_2(X) \to \mathscr{W}_2(X)$ by push-forward:

$$\phi_\# \mu(A) = \mu(\phi^{-1}(A)).$$

Moreover, if ϕ is an isometry then $\phi_\#$ also is, so that we have:

$$\# : \operatorname{Isom}(X) \hookrightarrow \operatorname{Isom}(\mathscr{W}_2(X)).$$

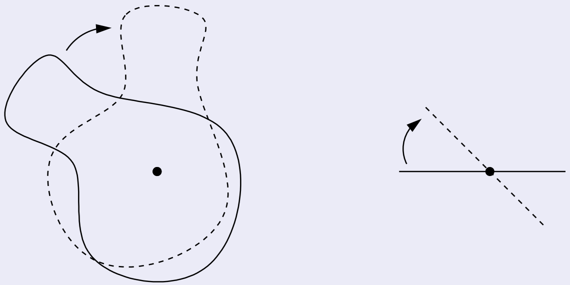
Theorem (joint with Jérôme Bertrand)

If X is a tree or a negatively curved, simply connected manifold, then $\#$ is onto.

Miscellany results

Theorem

- *The isometry group of* $\mathscr{W}_2(\mathbb{R}^n)$ *is larger than that of* $\mathbb{R}^n$.
- *If* $n > 1$, *given any* μ *and any isometry* Φ *of the Wasserstein space, there is an isometry* ϕ *of* $\mathbb{R}^n$ *such that* $\Phi(\mu) = \phi_\#\mu$.



- *This does not hold if* $n = 1$: $\mathscr{W}_2(\mathbb{R})$ *has "exotic" isometries.*

Positive embedding results

Proposition

If X contains a complete geodesic, then $\mathscr{W}_2(X)$ contains

- *bi-Lipschitz embedding of Euclidean spaces of arbitrary dimension,*
- *isometric embeddings of Euclidean balls of arbitrary dimension and diameter.*

Positive embedding results

Theorem

For all n, there is a bi-Lipschitz embedding of X^n into $\mathscr{W}_2(X)$.

From this we can deduce a dynamical statement. Recall that topological entropy is a measure of exponential spreading speed of a map's orbits.

Corollary

Assume X is compact. If ϕ is a map acting on X with positive entropy, then $\phi_\#$ acts on $\mathscr{W}_2(X)$ with infinite entropy.

A non-embedding result

Theorem (joint with Jérôme Bertrand)

If X is a visible Hadamard space, then $\mathscr{W}_2(X)$ has rank 1.

Hadamard spaces are roughly speaking the simply connected and non-positively curved spaces.
The *visibility* condition means that any two boundary points are connected by a geodesic.

Examples of spaces satisfying the hypotheses

- trees,
- simply connected manifolds with sectional curvature bounded above by $\kappa < 0$,
- in particular (real, complex, etc.) hyperbolic spaces,
- the I_{pq} buildings, ...

Contents

First idea

To embed X^n into $\mathscr{W}_2(X)$, one can try the map

$$F(x_1, x_2, \dots, x_n) = \frac{1}{n}\delta_{x_1} + \frac{1}{n}\delta_{x_2} + \dots \frac{1}{n}\delta_{x_n}.$$

However, F is invariant by permutation so this is not an embedding at all.

Embedding half-cones

First twist in the idea:

$$\begin{array}{rrcl} G: & \mathbb{R}^n_< & \to & \mathscr{W}_2(\mathbb{R}) \\ & (x_1, x_2, \dots, x_n) & \mapsto & \frac{1}{n}\delta_{nx_1} + \frac{1}{n}\delta_{nx_2} + \dots \frac{1}{n}\delta_{nx_n} \end{array}$$

where

$$\mathbb{R}^n_< := \{(x_1 < x_2 < \dots < x_n)\}$$

is an isometric embedding of a Euclidean open half-cone.
Composing with a complete geodesic of X and restricting, we get bi-Lipschitz embeddings of $\mathbb{R}^{n-1}$ and isometric embeddings of arbitrarily large balls of $\mathbb{R}^n$.

Embedding powers

Second twist in the idea: change the weights.

Theorem

$$\begin{array}{rrcl} H: & X^n & \to & \mathscr{W}_2(X) \\ & x=(x_1,\dots,x_n) & \mapsto & \frac{\alpha}{2}\delta_{x_1}+\frac{\alpha}{4}\delta_{x_2}+\cdots+\frac{\alpha}{2^n}\delta_{x_n} \end{array}$$

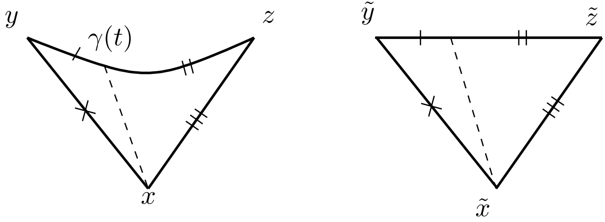
(where $\alpha = 1/(1-2^{-n})$ is a normalizing constant), is a bi-Lipschitz embedding with explicit and absolute constants.

The upper bound is trivial. The proof of the lower bound consist in using combinatorial arguments to prove that in any optimal transport plan between $H(x)$ and $H(y)$, a quantity of mass at least 2^{-n} has to travel from x_i to y_i, where i is chosen so as to maximize the distance.

Contents

Curvature

- Geodesics are assumed to be globally minimizing, constant speed:
$$d(\gamma_s, \gamma_t) = v|s-t|$$
- If every pair of points is linked by a geodesic segment, one says that X is *geodesic*.
- The CAT(0) inequality means any triangles in X is thinner than the Euclidean triangle with the same side lengths.



Hadamard spaces

Definition

The polish metric space X is a Hadamard space if it is:

- *locally compact,*
- *geodesic,*
- *globally* $\mathrm{CAT}(0)$*, and therefore simply connected.*

A hadamard space is always contractible, uniquely geodesic, and its distance is convex along pairs of geodesics.
Every measure μ in $\mathscr{W}_2(X)$ has a unique barycenter, that is a point x minimizing $W(\mu, \delta_x)$.

Examples

Some Hadamard spaces

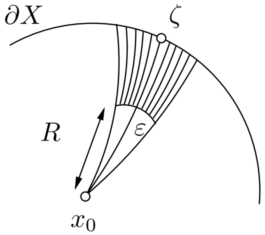
- Euclidean spaces $\mathbb{R}^n$,
- hyperbolic spaces $\mathbb{R}\mathrm{H}^n$, $\mathbb{C}\mathrm{H}^n$, $\mathbb{H}\mathrm{H}^n$, $\mathbb{O}\mathrm{H}^2$,
- symmetric spaces of non-compact type, like $\mathrm{SL}(n;\mathbb{R})/\mathrm{SO}(n)$,
- simply-connected manifolds with non-positive sectional curvature,
- trees, I_{pq} buildings,
- products of Hadamard spaces,
- gluings of two Hadamard spaces along convex, isometric subsets.

Geodesic boundary

- A *ray* is a geodesic parametrized by $[0, +\infty)$. Two rays are *asymptotic* if they stay at bounded distance. The asymptote class of a ray γ is denoted by $\gamma_{+\infty}$.
- The *geodesic boundary* of X is

$$\partial X := \mathscr{R}_1(X)/\sim \quad = \{\text{unit rays}\}/\text{asymptote relation}$$

and $\bar{X} := X \cup \partial X$ is compact once endowed with the *cone topology*.



Asymptotic distance

- Given two rays γ, β, the number

$$d_\infty(\gamma, \beta) = \lim_{t \to \infty} \frac{d(\gamma_t, \beta_t)}{t}$$

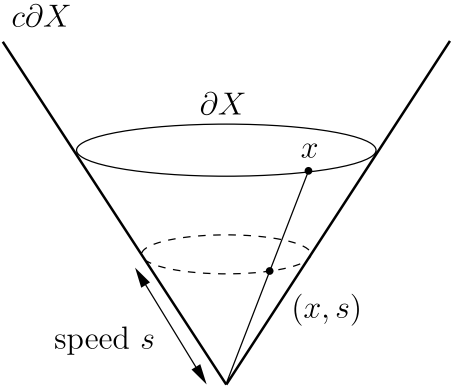
  exists and defines a distance on ∂X.
- Recall that X is *visible* if for all $\zeta, \xi \in \partial X$, there is some complete geodesic γ such that $\gamma_{-\infty} = \zeta$ and $\gamma_{+\infty} = \xi$. This is equivalent to $d_\infty \equiv 2$ on ∂X.

Cone over the boundary

The asymptote relation $\sim$ and the asymptotic distance d_∞ extends readily to non-unit rays.

The cone over the boundary is then the topological space

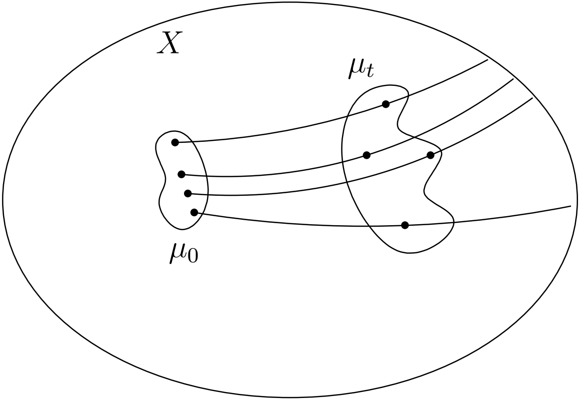
$$c\partial X = \frac{X \times [0, +\infty)}{(x,0) = (y,0)} \simeq \mathscr{R}(X)/\sim$$



Contents

Displacement interpolation

A geodesic ray (μ_t) in $\mathscr{W}_2(X)$ always has a *displacement interpolation*: there is a probability measure μ on $\mathscr{R}(X)$ such that $\mu_t = e_{t\#}\mu$ where e_t is evaluation at time t.



Boundary of the Wassertein space

There are two natural definitions of $\partial\mathscr{W}_2(X)$.

- Copy the case of X and take $\mathscr{R}_1(\mathscr{W}_2(X))/\sim$.
- Given a ray (μ_t), use displacement interpolation to define its endpoint as

$$\mu_\infty := e_{\infty\#}\mu \in \mathscr{P}_1(c\partial X)$$

  where $\mathscr{P}_1(c\partial X)$ is the set of probability measures having unitary square mean second coordinate.

Asymptotic distance between geodesics

The two previous definitions coincide.

Theorem (Asymptotic formula)

Let (μ_t), (σ_t) be two rays of $\mathscr{W}_2(X)$. Then $(\mu_t) \sim (\sigma_t)$ if and only if $\mu_\infty = \sigma_\infty$ and

$$\lim_{t \to +\infty} \frac{W(\mu_t, \sigma_t)}{t} = W_\infty(\mu_\infty, \sigma_\infty)$$

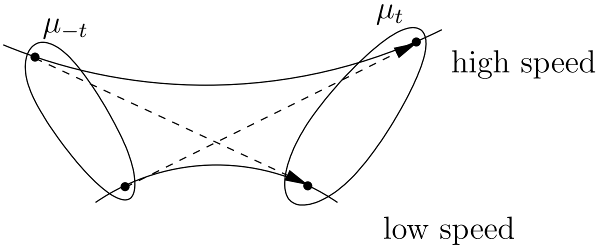
where W_∞ is the quadratic Wasserstein distance associated to the metric d_∞ on $c\partial X$.

Proof of the Rank Theorem

Assume X is Hadamard and visible, and prove that $\mathbb{R}^2$ does not embed isometrically in $\mathscr{W}_2(X)$.

Step 1

A complete geodesic $(\mu_t)_{t\in\mathbb{R}}$ has its asymptotic measure μ_∞ concentrated in $\mathscr{P}(\partial X) \subset \mathscr{P}_1(c\partial X)$.



Proof of the Rank Theorem

Assume X is Hadamard and visible, and prove that $\mathbb{R}^2$ does not embed isometrically in $\mathscr{W}_2(X)$.

Step 2

The restriction of W_∞ to $\mathscr{P}(\partial X)$ has no rectifiable curves.

Since X is visible, given $\mu_\infty, \sigma_\infty \in \mathscr{P}(\partial X)$ one has

$$W_\infty(\mu_\infty, \sigma_\infty) = \|\mu_\infty - \sigma_\infty\|_{TV}^{1/2}$$

and a square-rooted distance has no rectifiable curve.

Proof of the Rank Theorem

Assume X is Hadamard and visible, and prove that $\mathbb{R}^2$ does not embed isometrically in $\mathscr{W}_2(X)$.

Step 3

The boundary of the euclidean plane, endowed with d_∞, is a rectifiable curve.

This boundary is the unit circle with the chordal metric.

Many more spaces are ruled out by the proof, like open cones in $\mathbb{R}^n$ and small perturbations of Euclidean plane.