Barycenter in Wasserstein space existence and consistency

Thibaut Le Gouic, Jean-Michel Loubes

Barycenter in Wasserstein spaces: existence and consistency

Thibaut Le Gouic in collaboration with Jean-Michel Loubes

École centrale de Marseille
tle-gouic@centrale-marseille.fr,
http://tle-gouic.perso.centrale-marseille.fr

Abstract. We study barycenters in the Wasserstein space $\mathcal{P}_p(E)$ of a locally compact geodesic space (E, d). In this framework, we define the barycenter of a measure $\mathbb{P}$ on $\mathcal{P}_p(E)$ as its Fréchet mean. The paper establishes its existence and states consistency with respect to $\mathbb{P}$. We thus extends previous results on $\mathbb{R}^d$, with conditions on $\mathbb{P}$ or on the sequence converging to $\mathbb{P}$ for consistency.

Keywords: barycenter, Wasserstein space, geodesic spaces

1 Introduction

The Fréchet mean of a Borel probability measure μ, defined on a metric space (E, d), as the minimizer of

$$x \mapsto \mathbb{E}d^2(x, X), X \sim \mu$$

provides a natural extension of the barycenter as it coincides on $\mathbb{R}^d$ with the barycenter $\sum_{i=1}^n \lambda_i x_i$ of the points $(x_i)_{1\leq i\leq n}$, with weights $(\lambda_i)_{1\leq i\leq n}$ if

$$\mu = \sum_{i=1}^n \lambda_i x_i.$$

Its existence is a straightforward consequence of the local compactness of a geodesic space (E, d) when assumed. But it is not obvious in more general cases.

Mimicking the Fréchet mean, for any $p \geq 1$, we define a p-barycenter (or simply a barycenter) of a measure μ, as any minimizer of $x \mapsto \mathbb{E}d^p(x, X)$, where $X \sim \mu$.

The Wasserstein space $\mathcal{P}_p(E)$ of a locally compact geodesic space (E, d) is the set of all Borel probability measure on (E, d) such that $\mathbb{E}d^p(x, X) < \infty$, for some $x \in E$, endowed with the p-Wasserstein metric defined between two measures μ, ν as

$$W_p^p(\mu, \nu) = \inf_{\pi \in \Gamma(\mu,\nu)} \int d^p(x, y) d\pi(x, y), \quad (1)$$

where $\Gamma(\mu,\nu)$ is the set of measures on $E \times E$ with marginals μ and ν. Since the Wasserstein space of a locally compact space geodesic space is geodesic but not locally compact (unless (E,d) is compact), the existence of a barycenter is not as straightforward.

This paper presents its existence and study consistency properties. Several works has already been achieved in this field. An important one is the demonstration of existence and uniqueness of the barycenter of measures $\mathbb{P}$ on $\mathcal{P}_2(\mathbb{R}^d)$, with $d \in \mathbb{N}^*$, and $\mathbb{P}$ finitely supported on Dirac masses:

$$\mathbb{P} = \sum_{i=1}^{n} \lambda_i \delta_{\mu_i},$$

such that $\mu_i \in \mathcal{P}_p(\mathbb{R}^d)$, for $1 \leq i \leq n$, with one μ_i *vanishing on small sets.* In this case (when the measure $\mathbb{P}$ is finitely supported on Dirac masses), the barycenter of $\mathbb{P}$ is also the minimizer of

$$\nu \mapsto \sum i = 1^n \lambda_i W_p^p(\nu, \mu_i),$$

which is how the problem is more classically posed.

This *vanishing* property is said to be satisfied for a probability measure if it gives probability 0 to sets of Hausdorff dimension less than $d-1$. Any measure absolutely with respect to the Lebesgue measure *vanishes on small sets.* This work of [AC] has been extended to compact Riemannian manifolds, with the condition to *vanish on small sets* being replaced by absolute continuity with respect to the volume measure by [KP]. Since the Wasserstein space of a compact space is also compact, the existence in this setting can be easily obtained, but their work provides, among other results, an interesting extension, to our concern, to the work of [AC], by showing a dual problem called the *multidimensional* problem, for any $\mathbb{P}$ of the form

$$\sum_{i=1}^{n} \lambda_i \delta_{\mu_i}.$$

The same dual problem has been used in a previous work to show existence of barycenter whenever there exists a measurable (not necessarily unique) barycenter application on (E^n, d^n) that associate the barycenter of $\sum_{i=1}^n \lambda_i \delta_{x_i}$ to every n-uplets $(x_1, ..., x_n)$. It is a first step toward the proof of existence of barycenter for any $\mathbb{P}$.

Two statistical problems arise from the notion of barycenter. The first one can be stated as follows. Given $(\mu_i)_{i \geq 1}$, and $(\lambda_j^J)_{1 \leq j \leq J}$, it would be useful that the (or any) barycenter of $\mathbb{P}_J = \sum_{j=1}^J \lambda_j^J \delta_{\mu_j}$ converges to a barycenter of a limit measure of the sequence $(\mathbb{P}_J)_{J \geq 1}$. [BK] studied this problem in the case where $(\mu_j)_{j \geq 1}$ have compact support, are absolutely continuity with respect to the Lebesgue measure and are indexed on a compact set Θ of $\mathbb{R}^d$. They state more precisely that given a probability measure on Θ, one can induce a probability measure $\mathbb{P}$ on $\mathcal{P}_p(\mathbb{R}^d)$, and if the $(\mu_j)_{j \geq 1}$ are chosen randomly under $\mathbb{P}^{\otimes \infty}$, the (unique) barycenter of $\frac{1}{J}\sum_{j=1}\delta_{\mu_j}$ converges to the barycenter of $\mathbb{P}$, $\mathbb{P}$-almost surely.

In a previous work [BLGL], the authors produced a similar result under the assumptions that the $(\mu_j)_{j\geq 1}$ are *admissible deformations*, which is a similar condition.

The second statistical problem rising from this framework is the following. Given $(\lambda_i)_{i\geq 1}$ and $(\mu_j^n)_{1\leq j\leq J}$ converging to some $(\mu_j)_{1\leq j\leq J}$, a question of our interest is whether the barycenter of $\sum_{j=1}^{J}\lambda_j\delta_{\mu_j^n}$ converges to a barycenter of the limit $(\mu_j)_{1\leq j\leq J}$. The problem has been answered positively in [BLGL], up to a subsequence, since the barycenter is not unique. Although these two problems are presented differently, they can be formulated into one problem. Does the (or any) barycenter of $\mathbb{P}_n$ converges to the barycenter of $\mathbb{P}$ when $\mathbb{P}_n$ converges to $\mathbb{P}$?

This paper presents a positive result of [LGL], that implies, in particular, existence of the barycenter for any Borel probability measure $\mathbb{P}\in\mathcal{P}_p(\mathcal{P}_p(E))$.

2 Existence of barycenter

Let (E,d) be a geodesic locally compact space. For any $p\geq 1$, define by $\mathcal{P}_p(E)$ the Wasserstein space of E, by the space of all Borel probability measures such that for all $x\in E$, $\mathbb{E}d^p(x,X)<\infty$, endowed with the Wasserstein metric defined in (1). Denote thus by $\mathcal{P}_p(\mathcal{P}_p(E))$ the p-Wasserstein space of Borel measures on $\mathcal{P}_p(E)$ endowed with the Wasserstein metric.

For a measure $\mathbb{P}\in\mathcal{P}_p(\mathcal{P}_p(E))$, we define a *barycenter* as a minimizer of the function

$$\mu\mapsto\mathbb{E}(W_p^p(\tilde{\mu},\mu)),$$

where $\tilde{\mu}\sim\mathbb{P}$. Remark that $\mathbb{E}(W_p^p(\tilde{\mu},\mu))=W_p^p(\mathbb{P},\delta_\mu)$ where the two notations W_p refer to the Wasserstein distance but on different spaces respectively $\mathcal{P}_p(E)$ and $\mathcal{P}_p(\mathcal{P}_p(E))$.

Then [LGL] proves the following result.

Theorem 1. *Let $\mathbb{P}\in\mathcal{P}_p(\mathcal{P}_p(E))$ be a measure on $\mathcal{P}_p(E)$. Then there exists a barycenter of $\mathbb{P}$.*

This result is a consequence of the existence of barycenter for $\mathbb{P}$ finitely supported, showed in [BLGL] or [LG], and the consistency result of [LGL] presented above.

3 Consistency of the barycenter

Since the barycenter is not necessarily unique for a given $\mathbb{P}$, the continuity of the barycenter with respect to $\mathbb{P}$ does not make sense. However, it is interesting to know for a sequence of measure $(\mathbb{P}_n)_{n\geq 1}\subset\mathcal{P}_p(E)$ converging in $\mathcal{P}_p(\mathcal{P}_p(E))$ to $\mathbb{P}$, a sequence of their barycenter converges to a barycenter of $\mathbb{P}$. [LGL] provides a positive answer.

Theorem 2. *Let $(\mathbb{P}_n)_{n\geq 1} \subset \mathcal{P}_p(\mathcal{P}_p(E))$ be a sequence of measures on $\mathcal{P}_p(E)$, and set μ_n a barycenter of $\mathbb{P}_n$, for all $n \in \mathbb{N}$. Suppose that $W_p(\mathbb{P}, \mathbb{P}_n) \to 0$. Then, the sequence $(\mu_n)_{n\geq 1}$ is compact in $\mathcal{P}_p(E)$ and any limit is a barycenter of $\mathbb{P}$.*

Proof (Main ideas). The proof is in three steps.

The first step is to show that the sequence $(\mu_n)_{n\geq 1}$ is tight. It is indeed a consequence of the fact that balls on (E, d) are compact together with applying a Markov inequality to these balls.

The second step uses Skorokhod representation theorem and lower semicontinuity of $\nu \mapsto W_p(\mu, \nu)$ for any μ, to show that any weak limit of the sequence $(\mu_n)_{n\geq 1}$ is a barycenter of $\mathbb{P}$.

The final step shows that the convergence of the $(\mu_n)_{n\geq 1}$ holds actually in $\mathcal{P}_p(E)$.

Applying this result to a constant sequence provides the following corollary.

Corollary 1. *The set of all barycenters of a given measure $\mathbb{P} \in \mathcal{P}_p(\mathcal{P}_p(E))$ is compact.*

An interesting and immediate corollary follows from the assumption that $\mathbb{P}$ has a unique barycenter.

Corollary 2. *Suppose $\mathbb{P} \in \mathcal{P}_p(\mathcal{P}_p(E))$ has a unique barycenter. Then for any sequence $(\mathbb{P}_n)_{n\geq 1} \subset \mathcal{P}_p(\mathcal{P}_p(E))$ converging to $\mathbb{P}$, any sequence $(\mu_n)_{n\geq 1}$ of their barycenters converges to the barycenter of $\mathbb{P}$.*

On $E = \mathbb{R}^d$ and for $p = 2$, there exists a simple condition that ensures that the barycenter is unique.

Proposition 1. *Let $\mathbb{P} \in \mathcal{P}_2(\mathcal{P}_2(\mathbb{R}^d))$ such that there exists a set $A \in \mathcal{P}_2(\mathbb{R}^d)$ of measures such that for all $\mu \in A$,*

$$B \in \mathcal{B}(\mathbb{R}^d), \dim(B) \leq d - 1 \implies \mu(B) = 0, \tag{2}$$

and $\mathbb{P}(A) > 0$, then, $\mathbb{P}$ admits a unique barycenter.

Proof. It is a consequence of the fact that if ν satisfies (2), then $\mu \mapsto W_2(\mu, \nu)$ is strictly convex and this so is $\mu \mapsto \mathbb{E}W_2^2(\mu, \tilde{\mu})$.

4 Statistical applications

Previous results imply that the two statistical problems mentioned in the introduction have positive answers. Define

$$\mathbb{P}_J = \sum_{i=1}^{J} \lambda_i^J \delta_{\mu_j}$$

with measure $\mu_j \in \mathcal{P}_p(E)$ and weights λ_j so that $\mathbb{P}_J$ converges to some measure $\mathbb{P}$, then Theorem 2 states that the barycenter (or any barycenter if not unique) of $\mathbb{P}_J$ converges to the barycenter of $\mathbb{P}$ (provided $\mathbb{P}$ has a unique barycenter).

Also, given

$$\mathbb{P}_n = \sum_{j=1}^{J} \lambda_j \delta_{\mu_j^n}$$

with positive weights λ_j and measures $(\mu_j^n)_{1 \leq j \leq J, n \geq 1} \subset \mathcal{P}_p(E)^J$ converging to some limit measures $(\mu_j)_{1 \leq j \leq J} \in \mathcal{P}_p(E)^J$, then, Theorem 2 states that the barycenter (or any if not unique) converges to the barycenter of $\sum_{j=1}^{J} \lambda_j \delta_{\mu_j^n}$ (if unique). These applications are further developed in [LGL].

Acknowledgments

The author would like to thank Anonymous Referee #2 for the detailed review.

References

[AC] Agueh, Carlier, *Barycenters in the Wasserstein space*, IAM Journal on Mathematical Analysis (2011)

[KP] Kim, Pass, *Wasserstein Barycenters over Riemannian manifolds*, (2014)

[BK] Bigot, Klein, *Consistent estimation of a population barycenter in the Wasserstein space*, (2012)

[BLGL] Boissard, Le Gouic, Loubes, *Distribution's template estimate with Wasserstein metrics*, Bernoulli Journal (2015)

[LGL] Le Gouic, Loubes, *Barycenter in Wasserstein spaces: existence and consistency*, (2015)

[LG] Le Gouic, *Localisation de masse et espaces de Wasserstein*, Thesis manuscript (2013)