Projection of a Wasserstein barycenter on a fixed discrete support

1 Setting

Let $\mathcal{P}_2(\mathbb{R})$ be the set of probability measures on $\mathbb{R}$ with finite second moment. For $\mu \in \mathcal{P}_2(\mathbb{R})$ let F_μ be its distribution function and

$$Q_\mu(t) = \inf\{x \in \mathbb{R} : F_\mu(x) \geq t\}, \qquad t \in (0,1),$$

its (left-continuous) quantile function. It is well known that, on the real line, the L_2 Wasserstein distance is the L_2 distance between quantile functions:

$$W_2^2(\mu,\nu) = \int_0^1 \big(Q_\mu(t) - Q_\nu(t)\big)^2\, dt. \tag{1}$$

Moreover, a function $Q : (0,1) \to \mathbb{R}$ is the quantile function of some $\nu \in \mathcal{P}_2(\mathbb{R})$ if and only if it is non-decreasing, left-continuous and square integrable; in that case ν is the image of the Lebesgue measure λ on $(0,1)$ through Q, $\nu = Q_\#\lambda$.

Let $\mu_1, \ldots, \mu_n \in \mathcal{P}_2(\mathbb{R})$ and $w_1, \ldots, w_n \geq 0$ with $\sum_i w_i = 1$. The (Wasserstein) barycenter of the μ_i is

$$\bar\mu = \operatorname*{arg\,min}_{\nu \in \mathcal{P}_2(\mathbb{R})} \sum_{i=1}^n w_i\, W_2^2(\mu_i,\nu), \qquad Q_{\bar\mu} = \sum_{i=1}^n w_i\, Q_{\mu_i}.$$

Even when all the μ_i are supported on a common finite set (for instance the categories $\{1,\ldots,7\}$ of a Likert scale), $\bar\mu$ is in general supported on a larger set of non-integer values.

Let $S = \{s_1 < s_2 < \cdots < s_K\} \subset \mathbb{R}$ be a finite set (the *domain* of the variable) and

$$\mathcal{P}(S) = \{\nu \in \mathcal{P}_2(\mathbb{R}) :\ \nu(S) = 1\}$$

the set of distributions supported on S. We look for the distribution of $\mathcal{P}(S)$ that best represents the barycenter.

Definition 1 (Nearest-point map)**.** Let $m_k = (s_k + s_{k+1})/2$, $k = 1, \ldots, K-1$, be the midpoints of S, and set $m_0 = -\infty$, $m_K = +\infty$. The map $T_S : \mathbb{R} \to S$ is

$$T_S(x) = s_k \quad \text{if } x \in (m_{k-1}, m_k], \qquad k = 1, \ldots, K.$$

$T_S(x)$ is a point of S at minimum distance from x: $|x - T_S(x)| = d(x,S) = \min_{s\in S} |x - s|$. The minimizer is unique except at the midpoints $x = m_k$, which are equidistant from s_k and s_{k+1} (the definition assigns them to s_k). Moreover T_S is non-decreasing and left-continuous.

The *simplified barycenter* is $\tilde\mu = (T_S)_\#\bar\mu$, i.e.

$$\tilde\mu(\{s_k\}) = \bar\mu\big((m_{k-1}, m_k]\big), \qquad k = 1, \ldots, K :$$

each value s_k receives the mass of the barycenter lying between the two adjacent midpoints. This is the rule implemented in `DDAWASS::simplifyD()` and `DDAWASS::WD.vec.mean.simplified()`.

2 Results

Lemma 2 (Quantile function of the projection)**.** *For every $\mu \in \mathcal{P}_2(\mathbb{R})$, $Q_{(T_S)_\#\mu} = T_S \circ Q_\mu$.*

Proof. $T_S \circ Q_\mu$ is non-decreasing and left-continuous, being the composition of two non-decreasing left-continuous functions, and it is bounded because T_S takes values in the finite set S. Hence it is the quantile function of the measure $(T_S \circ Q_\mu)_\#\lambda = (T_S)_\#(Q_\mu)_\#\lambda = (T_S)_\#\mu$. □

Theorem 3 (Best approximation on the domain)**.** *For every $\mu \in \mathcal{P}_2(\mathbb{R})$ and every $\nu \in \mathcal{P}(S)$,*

$$W_2^2\big(\mu, (T_S)_\#\mu\big) = \int_{\mathbb{R}} d(x,S)^2 \, d\mu(x) \;\le\; W_2^2(\mu,\nu).$$

Hence $(T_S)_\#\mu$ is a Wasserstein projection of μ on $\mathcal{P}(S)$. The projection is unique if μ gives no mass to the midpoints $\{m_1, \dots, m_{K-1}\}$; otherwise the minimizers are exactly the distributions obtained by splitting the mass of each midpoint m_k in any proportion between s_k and s_{k+1}.

Proof. Let $\nu \in \mathcal{P}(S)$. Its quantile function Q_ν takes values in S (because $\nu(S) = 1$), so for every $t \in (0,1)$

$$\big(Q_\mu(t) - Q_\nu(t)\big)^2 \;\ge\; d\big(Q_\mu(t), S\big)^2 = \big(Q_\mu(t) - T_S(Q_\mu(t))\big)^2.$$

Integrating over $(0,1)$ and using (1) and Lemma 2,

$$W_2^2(\mu,\nu) \;\ge\; \int_0^1 \big(Q_\mu(t) - T_S(Q_\mu(t))\big)^2 dt = W_2^2\big(\mu, (T_S)_\#\mu\big) = \int_0^1 d\big(Q_\mu(t), S\big)^2 dt = \int_{\mathbb{R}} d(x,S)^2 \, d\mu(x),$$

the last equality being the change of variables $x = Q_\mu(t)$ ($\mu = (Q_\mu)_\#\lambda$).

Uniqueness. Equality holds iff $Q_\nu(t)$ is a nearest point of $Q_\mu(t)$ for almost every t. If $Q_\mu(t) \notin \{m_k\}$ the nearest point is unique, so $Q_\nu(t) = T_S(Q_\mu(t))$. The set $\{t : Q_\mu(t) = m_k\}$ is an interval of length $\mu(\{m_k\})$; on it Q_ν may take either value s_k or s_{k+1}, and, Q_ν being non-decreasing, it equals s_k on an initial sub-interval and s_{k+1} on the rest: the mass $\mu(\{m_k\})$ is split between s_k and s_{k+1}, and every split gives the same value of the integral. □

Remark 4. The proof uses only the pointwise inequality $|x - y| \ge d(x,S)$ for $y \in S$, so Theorem 3 holds for every Wasserstein distance W_p, $p \ge 1$: $(T_S)_\#\mu$ minimizes $W_p(\mu,\nu)$ over $\nu \in \mathcal{P}(S)$, with $W_p^p(\mu, (T_S)_\#\mu) = \int d(x,S)^p \, d\mu$.

The next result shows that projecting the barycenter is the same as computing the barycenter directly in the constrained set $\mathcal{P}(S)$. It relies on the Hilbert (flat) structure of $(\mathcal{P}_2(\mathbb{R}), W_2)$ and is specific to $p = 2$.

Theorem 5 (Constrained barycenter)**.** *For every $\nu \in \mathcal{P}_2(\mathbb{R})$,*

$$\sum_{i=1}^n w_i W_2^2(\mu_i, \nu) = \sum_{i=1}^n w_i W_2^2(\mu_i, \bar\mu) + W_2^2(\bar\mu, \nu). \tag{2}$$

Consequently

$$\tilde\mu = (T_S)_\#\bar\mu \;\in\; \operatorname*{arg\,min}_{\nu \in \mathcal{P}(S)} \sum_{i=1}^n w_i \, W_2^2(\mu_i, \nu),$$

i.e. the simplified barycenter is a barycenter of $\mu_1, \dots, \mu_n$ constrained to the domain S, and the set of constrained barycenters coincides with the set of Wasserstein projections of $\bar\mu$ on $\mathcal{P}(S)$ (described in Theorem 3).

Proof. Write $Q_i = Q_{\mu_i}$, $\bar{Q} = Q_{\bar{\mu}} = \sum_i w_i Q_i$ and $Q = Q_\nu$, and let $\langle f, g\rangle = \int_0^1 fg\,dt$. By (1),

$$\sum_i w_i \|Q_i - Q\|^2 = \sum_i w_i \|Q_i - \bar{Q}\|^2 + \|\bar{Q} - Q\|^2 + 2\Big\langle \sum_i w_i (Q_i - \bar{Q}),\, \bar{Q} - Q \Big\rangle,$$

and the last term vanishes because $\sum_i w_i (Q_i - \bar{Q}) = \bar{Q} - \bar{Q} = 0$. This proves (2). The first term on the right-hand side of (2) does not depend on ν, hence minimizing $\sum_i w_i W_2^2(\mu_i, \nu)$ over $\nu \in \mathcal{P}(S)$ is equivalent to minimizing $W_2^2(\bar{\mu}, \nu)$ over $\nu \in \mathcal{P}(S)$, whose solutions are given by Theorem 3. □

Remark 6. The set $\mathcal{P}(S)$ is not convex in the quantile representation (a convex combination of S-valued quantile functions is not S-valued), but the proof does not need convexity: identity (2) is purely algebraic. Equation (2) also gives the loss due to the constraint:

$$\min_{\nu \in \mathcal{P}(S)} \sum_i w_i W_2^2(\mu_i, \nu) - \min_{\nu \in \mathcal{P}_2(\mathbb{R})} \sum_i w_i W_2^2(\mu_i, \nu) = W_2^2(\bar{\mu}, \tilde{\mu}) = \int d(x, S)^2 \, d\bar{\mu}(x).$$

Proposition 7 (Size of the approximation)**.** *Let $h = \max_k (s_{k+1} - s_k)$ and suppose $\bar{\mu}([s_1, s_K]) = 1$ (this holds when all the μ_i are supported on S, since then $s_1 \le Q_i \le s_K$). Then*

$$W_2(\bar{\mu}, \tilde{\mu}) \le h/2, \qquad \big|\mathrm{E}_{\tilde{\mu}}[X] - \mathrm{E}_{\bar{\mu}}[X]\big| \le W_1(\bar{\mu}, \tilde{\mu}) \le W_2(\bar{\mu}, \tilde{\mu}) \le h/2.$$

For an integer scale ($h = 1$) the simplified barycenter is within $1/2$ from the exact one, and its mean differs from the exact mean by at most $1/2$.

Proof. For $x \in [s_1, s_K]$, $d(x, S) \le h/2$, hence by Theorem 3 $W_2^2(\bar{\mu}, \tilde{\mu}) = \int d(x, S)^2 d\bar{\mu} \le h^2/4$. For the means, $|\mathrm{E}_{\bar{\mu}} X - \mathrm{E}_{\bar{\mu}} X| = \big|\int_0^1 (T_S(\bar{Q}) - \bar{Q})\,dt\big| \le \int_0^1 |T_S(\bar{Q}) - \bar{Q}|\,dt = W_1(\bar{\mu}, \tilde{\mu})$, and $W_1 \le W_2$ by Jensen's inequality. □

Remark 8 (What the simplified barycenter does not preserve)**.** The projection is optimal with respect to the Wasserstein distance, but it does not preserve the mean (nor the variance) of the exact barycenter; the error is bounded by Proposition 7. If one wants a distribution on S with the same mean as $\bar{\mu}$, the problem $\min\{W_2(\bar{\mu}, \nu) : \nu \in \mathcal{P}(S),\ \mathrm{E}_\nu X = \mathrm{E}_{\bar{\mu}} X\}$ is a different (constrained) problem, whose solution is in general not given by the nearest-point rule. Other intuitive rules, such as splitting the mass of each point $x \in (s_k, s_{k+1})$ between s_k and s_{k+1} proportionally to the distances (which preserves the mean), are not optimal for W_2.

3 Algorithm

Given the exact barycenter $\bar{\mu} = \sum_l \pi_l \delta_{x_l}$ (computed from the averaged quantile functions) and the domain S:

1. compute the midpoints $m_k = (s_k + s_{k+1})/2$, $k = 1, \dots, K-1$;
2. for $k = 1, \dots, K$, set $\tilde{\pi}_k = \sum_{l:\ m_{k-1} < x_l \le m_k} \pi_l$ (with $m_0 = -\infty$, $m_K = +\infty$);
3. return $\tilde{\mu} = \sum_k \tilde{\pi}_k \delta_{s_k}$.

The cost is $O(L \log K)$, L being the number of support points of $\bar{\mu}$. By Theorems 3 and 5, $\tilde{\mu}$ is both the distribution on S closest to the exact barycenter and a barycenter of the data constrained to S; the convention of assigning the midpoints to the lower value selects one of the (equivalent) solutions when $\bar{\mu}$ has atoms exactly at midpoints.