Barycenters

$$\bar{x} \in \operatorname{argmin}_{x \in X} \sum_{i=1}^{K} \frac{\lambda_i}{2} d(x, x_i)^2$$

Can be viewed as generalizing the idea of linear combinations to a metric space (X, d) that doesn't have a linear structure

$$\lambda_i \in \mathbb{R} \qquad \sum_{i=1}^{K} \lambda_i = 1$$

Classic Wasserstein Barycenter Problem

Agueh Carlier (2011)

Given $\overset{\text{Probability}}{\vee}$ measures

$\mu_1, \dots, \mu_K \in \mathbb{P}_2(\mathbb{R}^d)$

coefficients $\lambda_1, \dots, \lambda_K$

$\lambda_i \in [0,1]$, $\sum_{i=1}^{K} \lambda_i$

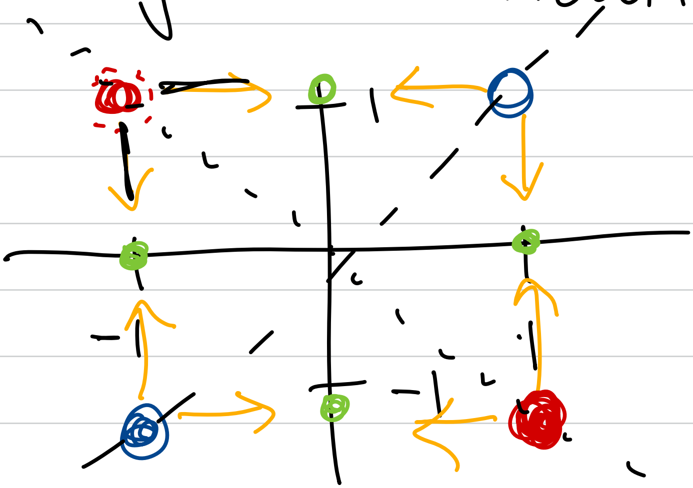
$$\bar{\nu} \in \underset{\nu \in \mathbb{P}_2(\mathbb{R}^d)}{\operatorname{argmin}} \sum_{i=1}^{K} \frac{\lambda_i}{2} W_2^2(\mu_i, \nu)$$

$$t \longmapsto \sum_{i=1}^{K} \frac{\lambda_i}{2} W_2^2(\mu_i, t\nu_1 + (1-t)\nu_0)$$

is convex in t

This almost guarantees uniqueness

Nonuniqueness only happens in degenerate situations



$$\mu_1 = \frac{1}{2}\left(\delta_{(-1,1)} + \delta_{(1,-1)}\right)$$

$$\mu_2 = \frac{1}{2}\left(\delta_{(1,1)} + \delta_{(-1,-1)}\right)$$

If any of the μ_i are a.c. then there is a unique Barycenter.

Optimality Conditions / Characterization of Barycenters

$$\bar{\nu} \in \underset{\nu}{\operatorname{argmin}} \underbrace{\sum_{i=1}^{K} \frac{\lambda_i}{2} W_2^2(\mu_i, \nu)}_{B(\nu)}$$

$$\frac{B(t\nu + (1-t)\bar{\nu}) - B(\bar{\nu})}{t} \geq 0$$

$$\lim_{t \to 0^+} \frac{B(t\nu + (1-t)\bar{\nu}) - B(\bar{\nu})}{t}$$

Let f_i be an optimal potential for the transport of $\bar{\nu}$ to μ_i

$$W_2^2(\mu_i, \bar{\nu}) = \int_{\mathbb{R}^d} f_i \, d\nu + f_i^c \, d\mu_i$$

$$W_2^2(\mu_i, t\nu + (1-t)\bar{\nu}) \geq \int_{\mathbb{R}^d} f_i \, d(t\nu + (1-t)\bar{\nu}) + \int_{\mathbb{R}^d} f_i^c \, d\mu_i$$

$$\lim_{t \to 0^+} \frac{B(t\nu + (1-t)\bar{\nu}) - B(\bar{\nu})}{t} \geq \sum_{i=1}^{K} \lambda_i \int_{\mathbb{R}^d} f_i \left[\frac{d(t\nu + (1-t)\bar{\nu}) - d\bar{\nu}}{t} \right]$$

$$= \sum_{i=1}^{K} \lambda_i \int_{\mathbb{R}^d} f_i \, d(\omega - \bar{\omega})$$

Other direction is more annoying but under mild assumptions on the μ_i (e.g. μ_i a.c.) We actually have

$$\lim_{t \to 0^+} \frac{B(t\omega + (1-t)\bar{\omega}) - B(\bar{\omega})}{t}$$

$$= \sum_i \lambda_i \int_{\mathbb{R}^d} f_i \, d(\omega - \bar{\omega})$$

Since $\bar{\omega}$ is optimal

$$\int_{\mathbb{R}^d} \sum_{i=1}^{K} \lambda_i f_i(y)\, d(\nu(y) - \bar{\nu}(y))$$

≥ 0 for all $\nu \in \mathbb{P}_2(\mathbb{R}^d)$

$$\Rightarrow \operatorname{spt}(\bar{\nu}) \subset \operatorname*{argmin}_{y \in \mathbb{R}^d} \sum_{i=1}^{K} \lambda_i f_i(y)$$

If we replace f_i by $f_i + c_i$

We can assume that

$$\inf_{y \in \mathbb{R}^d} \sum_{i=1}^{K} \lambda_i f_i(y) = 0$$

<u>Thm</u> A point $\bar{\nu} \in \mathbb{P}_2(\mathbb{R}^d)$ is a barycenter of $\mu_1, \ldots, \mu_K$

w/ weights $\lambda_1, \dots, \lambda_K \in [0,1]$

if and only if there exist optimal Kantorovich potentials $f_1, \dots, f_K$ where f_i transports $\bar{\nu}$ to μ_i s.t.

$$\sum_{i=1}^{K} \lambda_i f_i(y) = 0 \text{ on } \operatorname{spt}(\bar{\nu})$$

and $$\sum_{i=1}^{K} \lambda_i f_i(y) \geq 0$$

for all $y \in \mathbb{R}^d$

The condition $\sum_{i=1}^{K} \lambda_i f_i(y) = 0$ on $\operatorname{spt}(\bar{\nu})$ means transport from $\bar{\nu}$ to μ_i is regular (Transport map is Lipschitz)

map from $\bar{\nu}$ to μ_i is $\nabla \varphi_i$ where $\varphi_i(y) = \frac{1}{2}|y|^2 - f_i(y)$

$$\sum_{i=1}^{K} \lambda_i \varphi_i(y) = \frac{1}{2}|y|^2 \text{ on } \operatorname{spt}(\bar{\nu})$$

$$\sum_{i=1}^{K} \lambda_i \varphi_i(y) \leq \frac{1}{2}|y|^2 \text{ else}$$

$$0 < D^2\left(\sum_{i=1}^{K} \lambda_i \varphi_i(y)\right) < I$$

$$0 < D^2 \varphi_i < \frac{1}{\lambda_i} I$$

Signed Barycenter Problem

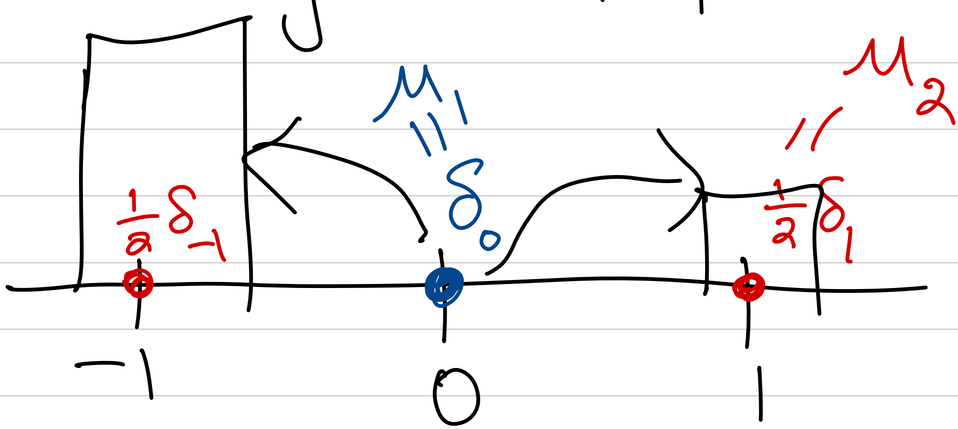
$$\bar{\nu} \in \operatorname*{argmin}_{\nu \in \mathbb{P}_2(\mathbb{R}^d)} \underbrace{\sum_{i=1}^{K} \lambda_i W_2^2(\mu_i, \nu)}_{B(\nu)}$$

$$\lambda_i \in \mathbb{R}, \quad \sum_{i=1}^{K} \lambda_i = 1$$

w/ signed weights

$$t \longmapsto B(t\nu_1 + (1-t)\nu_0)$$

is in general not convex



Assume $\lambda_1 > 0$, $\lambda_2 < 0$

$$t \longmapsto \underbrace{(1-t)\delta_{-1} + t\delta_{1}}_{\mu_t}$$

$B(\mu_t)$ is maximized at $t = 1/2$

and minimized at

$t = 0, 1$

$t \mapsto B(\omega_t)$ is $\underline{\underline{\text{not}}}$ convex.

$\underline{\text{Thm}}$ (Tornabene et al. '24)

If $\lambda_1 > 0$ and $\lambda_2, \ldots, \lambda_K < 0$ then

$$t \longmapsto B\left(\left(tT_1 + (1-t) T_0 \right)_{\#} \mu_1 \right)$$

is 1-strongly convex where T_1, T_0 are arbitrary optimal maps

$$W_2^2(\mu_i, \nu) \leq LOT_e(\mu_i, \nu)$$

$$W_2^2(\mu_i, \nu) \leq \int_{\mathbb{R}^d} |X_i(z) - Y(z)|^2 \, de(z)$$

where $X_{i\#}e = \mu_i$, $Y_\# e = \nu$ are the OT maps

$$W_2^2(\mu_i, \nu) = \inf_{X_{i\#}e = \mu_i} \int_{\mathbb{R}^d} |X_i(z) - Y(z)|^2 \, de(z)$$

$$-W_2^2(\mu_i, \nu) =$$

$$\sup_{X_{i\#}\ell=\mu_i} -\frac{1}{2}\int |X_i - Y|^2 d\ell$$

$$-W_2^2(\mu_i, Y_{\#}\ell) =$$

$$\sup_{X_{i\#}\ell=\mu_i} -\frac{1}{2}\int |X_i - Y|^2 d\ell$$

$$= -\frac{1}{2}\int |Y|^2 d\ell +$$

$$\sup_{X_{i\#}\ell=\mu_i} \int \left[(X_i, Y) - \frac{1}{2}|X_i|^2\right] d\ell$$

$$= -\frac{1}{2}\int |Y(z)|^2 \, d\rho(z) - \frac{1}{2}\int |x|^2 \, d\mu_i(x) + \underbrace{\sup_{X_i \# \rho = \mu_i} \int \left(X_i(z), Y(z) \right) d\rho(z)}_{\text{convex in } Y}$$

$$-W_2^2(\mu_i, Y_\# \rho) = -\frac{1}{2}\int |Y|^2 \, d\rho + \overset{\text{convex}}{F_i(Y)}$$

$$W_2^2(\mu_1, Y_\# \mu_1) =$$

$$\frac{1}{2}\int |Y(z)-z|^2 d\mu_1(z)$$

$$B(Y_\# \mu_1) =$$

$$\frac{\lambda_1}{2}\int |Y(z)-z|^2 d\mu_1(z)$$

$$-\sum_{i=2}^{K} \frac{|\lambda_i|}{2}\int |Y(z)|^2 d\mu_1(z)$$

$$+ \sum_{i=2}^{k} |\lambda_i| \, F_i(Y)$$

$$\begin{tikzcd}
\mu \arrow[rr, "T"] \arrow[dr, shift left, "X^{-1}"] & & \nu \arrow[ll, bend right, "T^{-1}"'] \\
& e \arrow[ul, shift left, "X"] \arrow[ur, "Y"'] &
\end{tikzcd}$$

$$Y \circ X^{-1} \qquad e = X^{-1}_{\#}\mu$$

$$\inf_{S_{\#} e = \mu} \int |S(z) \; Y(z)|^2 \, de$$

$$\inf_{S_\# e = \mu} \int \left| S(Y^{-1}) - \mathrm{id} \right|^2 d\omega$$

choose $S = T^{-1} \circ Y$

$$\int \left| T^{-1} \circ Y \circ Y^{-1} - \mathrm{id} \right|^3 d\omega$$

$$= \int \left| T^{-1} - \mathrm{id} \right|^2 d\omega$$

$$\int |X - Y|^2 \, de$$

$$\int |T - \mathrm{id}|^2$$

$$\int |\underbrace{X \circ Y^{-1}} - \mathrm{id}|^2$$

$$\nabla \varphi = X \circ Y^{-1}$$

$$\underset{\sim}{\nabla\varphi} \circ Y = \underset{\sim}{X}$$

If f_i is the OT potential from ν to μ_i

f - - - -