

TEORÍA DE TRANSPORTE ÓPTIMO

PRÁCTICA 12: BARICENTROS DE WASSERSTEIN

Los ejercicios son los de la sección de ejercicios del Capítulo 12 de las notas del curso, con la misma numeración; las referencias a teoremas, proposiciones y secciones remiten a las notas (<https://github.com/jfbonder/transporte-optimo>).

Cuaderno: *Baricentros de Wasserstein* (abrir en Colab).

Ejercicio 12.1 (Traslaciones y escala). Sean $(\mu_i, \lambda_i)_{i=1}^N$ con $\mu_i \in \mathcal{P}_2(\mathbb{R}^d)$.

- Probar que si $T(x) = Ax + b$ es afín con A lineal e invertible y μ tiene media m_μ , entonces $T_\# \mu$ tiene media $Am_\mu + b$.
- Probar que si $\bar{\mu}$ es un baricentro de (μ_i, λ_i) y $v_i \in \mathbb{R}^d$, entonces $\tau_{\bar{v}} \# \bar{\mu}$ con $\bar{v} = \sum \lambda_i v_i$ es un baricentro de $(\tau_{v_i} \# \mu_i, \lambda_i)$, y que $\mathcal{F}$ cambia en $\sum_i \lambda_i |v_i - \bar{v}|^2$. *Sugerencia:* Lema 10.1.4.
- Deducir que la media de todo baricentro es $\sum \lambda_i m_{\mu_i}$ y que para calcular baricentros se puede suponer que todas las μ_i están centradas. Probar también que $D_{\lambda} \# \bar{\mu}$ es baricentro de $(D_{\lambda} \# \mu_i, \lambda_i)$ para la dilatación D_{λ} .

Ejercicio 12.2 (Dos medidas: el valor mínimo). Sean $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ y $t \in (0, 1)$. Probar que

$$\min_{\nu} (1-t)W_2^2(\mu_1, \nu) + tW_2^2(\nu, \mu_2) = t(1-t)W_2^2(\mu_1, \mu_2),$$

y que el mínimo se alcanza en el punto μ_t de la geodésica (Proposición 12.1.3). *Sugerencia:* para la desigualdad $\geq$, escribir $W_2(\mu_1, \mu_2) \leq W_2(\mu_1, \nu) + W_2(\nu, \mu_2)$ y minimizar $(1-t)a^2 + tb^2$ sujeto a $a + b = W_2(\mu_1, \mu_2)$.

Ejercicio 12.3 (Existencia para el problema multimarginal). Completar la demostración del Teorema 12.5.1: probar que $\Pi(\mu_1, \dots, \mu_N)$ es no vacío, ajustado y débilmente cerrado en $\mathcal{P}((\mathbb{R}^d)^N)$, y que el funcional $\gamma \mapsto \int c d\gamma$ es semicontinuo inferiormente, de modo que el mínimo se alcanza (repetir el argumento del Teorema 2.3.9).

Ejercicio 12.4 (Pegado de N planes). Sean $\mu, \mu_1, \dots, \mu_N \in \mathcal{P}(\mathbb{R}^d)$ y $\pi_i \in \Pi(\mu, \mu_i)$. Desintegrando $\pi_i = \pi_i^y \otimes \mu$ (Teorema A.2.1), se define $\tilde{\gamma} = (\bigotimes_{i=1}^N \pi_i^y) \otimes \mu$, es decir

$$\int \Phi d\tilde{\gamma} = \int_{\mathbb{R}^d} \left(\int_{(\mathbb{R}^d)^N} \Phi(y, x_1, \dots, x_N) d\pi_1^y(x_1) \cdots d\pi_N^y(x_N) \right) d\mu(y).$$

Probar que $\tilde{\gamma}$ es una medida de probabilidad bien definida (medibilidad en y del integrando) y que $(\mathbf{p}_0, \mathbf{p}_i) \# \tilde{\gamma} = \pi_i$ para todo i , siguiendo la demostración del Lema A.3.1. Deducir que $\gamma = (\mathbf{p}_1, \dots, \mathbf{p}_N) \# \tilde{\gamma} \in \Pi(\mu_1, \dots, \mu_N)$.

Ejercicio 12.5 (Detalles de la Proposición 12.6.1). Sea $G = \sum_i \lambda_i F_{\mu_i}^{[-1]}$ con $\mu_i \in \mathcal{P}_2(\mathbb{R})$.

- Probar que cada $F_{\mu_i}^{[-1]}$, y por lo tanto G , es no decreciente y continua por la izquierda en $(0, 1)$ (Ejercicio 8.2).
- Sea $\bar{\mu} = G_\# \mathcal{L}^1|_{(0,1)}$. Probar que $F_{\bar{\mu}}^{[-1]} = G$ y que $\bar{\mu} \in \mathcal{P}_2(\mathbb{R})$.
- Deducir que el baricentro de N gaussianas unidimensionales $N(m_i, \sigma_i^2)$ es la gaussiana $N(\sum \lambda_i m_i, (\sum \lambda_i \sigma_i^2)^2)$, y que el baricentro de N medidas uniformes en intervalos es uniforme en un intervalo.

Ejercicio 12.6 (Gaussianas con covarianzas que conmutan). Sean $\mu_i = N(0, \Sigma_i)$ con Σ_i definidas positivas que conmutan dos a dos (por ejemplo, todas diagonales).

- Probar que $\bar{\Sigma} = (\sum_i \lambda_i \Sigma_i^{1/2})^2$ resuelve la ecuación de punto fijo (12.6.1).
- Verificar directamente el criterio de la Proposición 12.4.1: los mapas $T_i(x) = \Sigma_i^{1/2} \bar{\Sigma}^{-1/2} x$ son gradientes de funciones convexas, transportan $N(0, \bar{\Sigma})$ en μ_i y satisfacen $\sum \lambda_i T_i = \text{id}$.

(c) Comparar con el Ejercicio 12.5(c): el caso unidimensional es el caso $d = 1$ de este ejercicio.

Ejercicio 12.7 (El criterio multimarginal). Sea $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ en las hipótesis de la Proposición 12.4.1: $T_i = \nabla u_i$ con $(T_i)_\# \mu = \mu_i$ y $\sum_i \lambda_i T_i = \text{id}$. Sea $\gamma = (T_1, \dots, T_N)_\# \mu \in \mathcal{P}((\mathbb{R}^d)^N)$.

- (a) Probar que $\gamma \in \Pi(\mu_1, \dots, \mu_N)$ y que $A_\# \gamma = \mu$, donde $A(x_1, \dots, x_N) = \sum \lambda_i x_i$.
- (b) Probar que $\int c d\gamma = \mathcal{F}(\mu)$ para el costo c de (12.5.1), y deducir del Teorema 12.5.1 que γ es un plan multimarginal óptimo.
- (c) Comparar con el Ejercicio 7.4: allí el plan óptimo estaba concentrado en un hiperplano $\{\sum x_i = v\}$; aquí, en el gráfico de $x \mapsto (T_1(x), \dots, T_N(x))$.

Ejercicio 12.8 (Cuantización y baricentros de Diracs). (a) Probar directamente (sin usar el Teorema 12.3.2) que el baricentro de masas de Dirac δ_{x_i} es único y es $\delta_{\bar{x}}$, $\bar{x} = \sum \lambda_i x_i$.

- (b) Sea $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ y $\nu = \sum_{k=1}^N b_k \delta_{y_k}$. Probar que $W_2^2(\mu, \nu) \geq \sum_k b_k \text{Var}(\mu|_k)$ donde $\mu|_k$ es la restricción normalizada de μ a la celda que el plan óptimo envía a y_k , con igualdad si y sólo si cada y_k es el baricentro de su celda. Deducir la condición de optimalidad del Ejercicio 10.9(c).

Ejercicio 12.9 ($W_2^2(\mu, \cdot)$ no es convexo a lo largo de geodésicas). Sean $\mu = \frac{1}{2}(\delta_{(-1,0)} + \delta_{(1,0)})$, $\nu_0 = \frac{1}{2}(\delta_{(-1,1)} + \delta_{(1,-1)})$ y $\nu_1 = \frac{1}{2}(\delta_{(1,1)} + \delta_{(-1,-1)})$ en $\mathbb{R}^2$.

- (a) Probar que $\nu_t = \frac{1}{2}(\delta_{(-1+2t,1)} + \delta_{(1-2t,-1)})$ es una geodésica de W_2 entre ν_0 y ν_1 (¿es la única?).
- (b) Calcular $W_2^2(\mu, \nu_t) = 1 + 4 \min\{t^2, (1-t)^2\}$ y concluir que $t \mapsto W_2^2(\mu, \nu_t)$ no es convexa: vale 1 en los extremos y 2 en $t = \frac{1}{2}$.
- (c) Comparar con el Lema 12.3.1 (convexidad a lo largo de interpolaciones lineales) y con el Ejercicio 10.11. Este ejemplo muestra que la unicidad del baricentro no puede probarse por un argumento de convexidad geodésica.

Ejercicios computacionales. Los ejercicios siguientes se resuelven con el cuaderno *Baricentros de Wasserstein* (véase el Apéndice E); los enunciados figuran también al final del cuaderno. Las referencias a secciones son a las del cuaderno.

Ejercicio 12.10 (Rigidez en dimensión uno). Para N medidas empíricas en $\mathbb{R}$ con n átomos cada una, mostrar que el plan multimarginal óptimo del Teorema 12.5.1 es el comonótono (empareja los k -ésimos datos ordenados de cada medida) y que $B_\# \gamma^*$ es exactamente el promedio de cuantiles de la Proposición 12.6.1. Verificarlo resolviendo el programa lineal multimarginal para n chico.

Ejercicio 12.11 (Traslaciones). Probar que si se traslada cada μ_i por un vector v_i , el baricentro se traslada por $\bar{v} = \sum_i \lambda_i v_i$ y el valor mínimo de $\mathcal{F}$ cambia en $\sum_i \lambda_i |v_i - \bar{v}|^2$ (cf. Ejercicio 12.1). Deducir que para calcular baricentros basta centrar todas las medidas, y verificarlo numéricamente.

Ejercicio 12.12 (Convexidad del funcional). Elegir dos medidas ν_0, ν_1 (por ejemplo, dos de las gaussianas de la Sección 3 del cuaderno) y graficar $s \mapsto \mathcal{F}((1-s)\nu_0 + s\nu_1)$ y $s \mapsto \mathcal{F}(\nu_s)$, donde ν_s es la geodésica de W_2 entre ellas. La primera es convexa por el Lema 12.3.1. ¿Lo es la segunda? Comparar con el Ejercicio 12.9.

Ejercicio 12.13 (Baricentros de formas). Discretizar dos o tres figuras planas (un disco, un cuadrado, un triángulo) como medidas uniformes sobre una grilla y calcular su baricentro resolviendo el programa lineal del Teorema 12.5.1 (`ot.lp.barycenter`). Observar cómo crece el tamaño del problema con la grilla; en el Capítulo 13 se ve la versión entrópica, que escala mucho mejor.