

TEORÍA DE TRANSPORTE ÓPTIMO

PRÁCTICA 10: DISTANCIAS DE WASSERSTEIN

Los ejercicios son los de la sección de ejercicios del Capítulo 10 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: *Distancias de Wasserstein* (abrir en Colab).

Ejercicio 10.1 (Momentos). Decidir a qué espacios $\mathcal{P}_p(\mathbb{R})$ pertenecen las medidas $\delta_2, \frac{1}{4}\delta_{-2} + \frac{3}{4}\delta_2, \mathcal{L}^1|_{[0,1]}, e^{-x}\mathcal{L}^1|_{[0,\infty)}, \frac{1}{\pi(1+x^2)}\mathcal{L}^1$ y $\sum_{n \geq 1} \frac{6}{\pi^2 n^2} \delta_n$, y calcular $W_p(\mu, \delta_0)$ cuando sea finito. Probar que $\mathcal{P}_p(\mathbb{R}^d) \subsetneq \mathcal{P}_q(\mathbb{R}^d)$ para $q < p$.

Ejercicio 10.2 (Escala, traslaciones y medias). Sean $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d), p \geq 1$.

- Probar que $W_p(D_{\lambda\#}\mu, D_{\lambda\#}\nu) = \lambda W_p(\mu, \nu)$ para la dilatación $D_{\lambda}(x) = \lambda x, \lambda > 0$, y que W_p es invariante por traslaciones simultáneas y por isometrías de $\mathbb{R}^d$.
- Probar que $W_p(\mu, \nu) \geq |m_{\mu} - m_{\nu}|$, donde m denota la media. *Sugerencia:* Jensen.
- Deducir que si $\nu = \tau_{v\#}\mu$ es una traslación de μ , entonces $W_p(\mu, \nu) = |v|$ para todo $p \geq 1$, y la traslación es un mapa óptimo. Comparar con el Lema 10.1.4: para $p = 2$ hay una descomposición exacta, ¿vale algo análogo para $p \neq 2$?

Ejercicio 10.3 (Masa que se escapa). Sea $\mu_n = (1 - \frac{1}{n})\delta_0 + \frac{1}{n}\delta_{a_n}$ con $a_n \rightarrow +\infty$.

- Probar que $\mu_n \rightarrow \delta_0$ para toda elección de a_n , y que $W_p(\mu_n, \delta_0)^p = \frac{1}{n} a_n^p$.
- Elegir a_n de modo que $W_q(\mu_n, \delta_0) \rightarrow 0$ pero $W_p(\mu_n, \delta_0) \rightarrow +\infty$, para $1 \leq q < p$ dados. Interpretar en términos del Teorema 10.3.1: ¿qué momento deja de converger?
- Probar que $\mathcal{P}_p(\mathbb{R}^d)$ con la topología de W_p no es localmente compacto: la bola $\{W_p(\mu, \delta_0) \leq 1\}$ no es compacta.

Ejercicio 10.4 (Convergencia contra funciones Lipschitz). Sean $\mu_n, \mu \in \mathcal{P}_1(\mathbb{R}^d)$ tales que $\int \varphi d\mu_n \rightarrow \int \varphi d\mu$ para toda φ 1-Lipschitz. Probar que $\mu_n \rightarrow \mu$ (convergencia contra C_b). *Sugerencia:* aproximar $g \in C_b$ uniformemente sobre compactos por funciones Lipschitz acotadas (Ejercicio 2.5), y usar que la familia $\{\mu_n\}$ es ajustada: la función $\varphi(x) = |x|$ es 1-Lipschitz, así que $\int |x| d\mu_n$ está acotado. ¿Es cierta la recíproca? (Comparar con el Ejercicio 10.3.)

Ejercicio 10.5 (La fórmula del área vía Kantorovich–Rubinstein). Sean $\mu, \nu \in \mathcal{P}_1(\mathbb{R})$.

- Probar que para toda φ 1-Lipschitz con soporte compacto, $\int \varphi d(\mu - \nu) = - \int_{\mathbb{R}} \varphi'(x) (F_{\mu}(x) - F_{\nu}(x)) dx$ (integrar por partes, o usar Fubini), y deducir $W_1(\mu, \nu) \leq \int_{\mathbb{R}} |F_{\mu} - F_{\nu}| dx$ a partir del Teorema 10.6.2.
- Probar la igualdad $W_1(\mu, \nu) = \int_{\mathbb{R}} |F_{\mu} - F_{\nu}| dx$ tomando φ con $\varphi' = -\text{sgn}(F_{\mu} - F_{\nu})$, primero para medidas con soporte compacto y luego en general por aproximación. Comparar con la Proposición 8.4.5.
- Probar que para $p > 1$, $W_p^p(\mu, \nu) \neq \int_{\mathbb{R}} |F_{\mu} - F_{\nu}|^p dx$ en general (el área entre las pseudoinversas y el área entre las distribuciones coinciden, pero las integrales de las potencias no). Dar un ejemplo.

Ejercicio 10.6 (Transporte en el rectángulo). Sea $\Omega = [0, 3] \times [0, 2]$ y sean $\mu_0 = \rho_0 \mathcal{L}^2, \mu_1 = \rho_1 \mathcal{L}^2$ con

$$\rho_0 = \begin{cases} \frac{2}{9}, & x \leq 1 \text{ o } (1 < x < 2, y \leq 1), \\ \frac{1}{9}, & x \geq 2 \text{ o } (1 < x < 2, y > 1), \end{cases} \quad \rho_1 = \begin{cases} \frac{2}{9}, & x \geq 2 \text{ o } (1 < x < 2, y \leq 1), \\ \frac{1}{9}, & x \leq 1 \text{ o } (1 < x < 2, y > 1). \end{cases}$$

- Representar gráficamente ρ_0 y ρ_1 y verificar que son densidades de probabilidad.
- Probar que $u(x, y) = -x$ es 1-Lipschitz y que $\int u d(\mu_0 - \mu_1) = \frac{4}{9}$. Deducir $W_1(\mu_0, \mu_1) \geq \frac{4}{9}$.

- (c) Para cada y fijo, comparar las funciones de distribución de $\rho_0(\cdot, y)$ y $\rho_1(\cdot, y)$ y probar, usando el Ejercicio 8.3, que existe un plan que mueve la masa sólo hacia la derecha a lo largo de rectas horizontales. Probar que ese plan satura $u(x) - u(x') = |x - x'|$ en su soporte y concluir que es óptimo y que $W_1(\mu_0, \mu_1) = \frac{4}{9}$.
- (d) ¿Es único el plan óptimo? ¿Es único el potencial de Kantorovich?

Ejercicio 10.7 (Del disco a una densidad radial, con costo $|x - y|$). Sean μ y ν las medidas del Ejercicio 9.3 y $c(x, y) = |x - y|$.

- (a) Probar que $u(x) = -|x|$ es 1-Lipschitz y que $u(x) - u(y) = |x - y|$ si y sólo si $y = tx$ con $t \geq 1$ (mismo rayo, más lejos del origen).
- (b) Deducir que el mapa radial T del Ejercicio 9.3 es también óptimo para $c(x, y) = |x - y|$, y que $W_1(\mu, \nu) = \int (|T(x)| - |x|) d\mu = \int |y| d\nu - \int |x| d\mu$.
- (c) Probar que, a diferencia del caso cuadrático, el plan óptimo no es único: cualquier plan que mueva la masa a lo largo de rayos alejándola del origen es óptimo.

Ejercicio 10.8 (La distancia SNCF). Sea X un conjunto y $P: X \rightarrow (0, \infty)$ acotada. Se define $d(x, y) = 0$ si $x = y$ y $d(x, y) = P(x) + P(y)$ si $x \neq y$.

- (a) Probar que d es una distancia en X (todo viaje pasa por un nodo central: así funciona una red ferroviaria centralizada en París, de donde el nombre, por la *Société Nationale des Chemins de fer Français*).
- (b) Probar que para $\mu, \nu \in \mathcal{P}(X)$,

$$\min_{\Pi(\mu, \nu)} \int d d\pi = \int_X P d|\mu - \nu|.$$

Sugerencia: la masa común $\mu \wedge \nu$ conviene dejarla quieta; el resto paga P al salir y P al llegar. Para la cota inferior usar el Ejercicio 6.3 con $\varphi = P$ sobre un conjunto donde se concentra $(\mu - \nu)^+$ y $\varphi = -P$ en su complemento; verificar que φ es 1-Lipschitz para d .

Ejercicio 10.9 (Distancia a un conjunto). Sea $\Omega \subset \mathbb{R}^d$ compacto, $\Sigma \subset \Omega$ cerrado y $\mu = f\mathcal{L}^d \in \mathcal{P}(\Omega)$.

- (a) Probar que $\int_\Omega d(x, \Sigma) d\mu(x) = \min\{W_1(\mu, \nu) : \nu \in \mathcal{P}(\Sigma)\}$, y que el mínimo se alcanza en $\nu = (p_\Sigma)_\# \mu$ para cualquier selección medible p_Σ de la proyección sobre Σ .
- (b) Probar el análogo con $d(x, \Sigma)^p$ y W_p^p , $p > 1$.
- (c) Si Σ es finito, $\#\Sigma = N$, se obtiene el problema de *cuantización óptima*

$$\min\{W_p^p(\mu, \nu) : \#\text{sop } \nu \leq N\} = \min_{\#\Sigma \leq N} \int_\Omega d(x, \Sigma)^p d\mu.$$

Para $p = 2$, probar que en un $\Sigma = \{x_1, \dots, x_N\}$ óptimo cada x_i es el baricentro de su celda de Voronoi $V_i = \{x : |x - x_i| \leq |x - x_j| \forall j\}$ respecto de μ .

Ejercicio 10.10 (La distancia W_∞). Para μ, ν con soporte compacto se define

$$W_\infty(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \pi\text{-ess sup } |x - y|.$$

- (a) Probar que el ínfimo se alcanza y que W_∞ es una distancia.
- (b) Probar que $W_p(\mu, \nu) \nearrow W_\infty(\mu, \nu)$ cuando $p \rightarrow \infty$.
- (c) Probar que si $p_n \rightarrow \infty$ y γ_n es óptimo para $|x - y|^{p_n}$, entonces, salvo subsucesión, $\gamma_n \rightarrow \gamma$ con γ óptimo para W_∞ . *Sugerencia:* semicontinuidad inferior de $\pi \mapsto \pi\text{-ess sup } |x - y|$ respecto de la convergencia débil.
- (d) Probar que $d_H(\text{sop } \mu, \text{sop } \nu) \leq W_\infty(\mu, \nu)$, donde d_H es la distancia de Hausdorff, y dar un ejemplo con desigualdad estricta.

Ejercicio 10.11 (La interpolación lineal no es geodésica). Sean $\mu = \delta_0$, $\nu = \delta_1$ en $\mathbb{R}$ y $\nu_t = (1 - t)\mu + t\nu$.

- (a) Calcular $W_2(\nu_s, \nu_t) = \sqrt{|t - s|}$ y deducir que la curva $t \mapsto \nu_t$ tiene longitud infinita en $(\mathcal{P}_2, W_2)$, mientras que la geodésica $\mu_t = \delta_t$ tiene longitud 1.

- (b) Para $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ cualesquiera, probar que $W_2^2(\nu_s, \nu_t) \leq |t-s| W_2^2(\mu, \nu)$: la interpolación lineal es siempre $\frac{1}{2}$ -Hölder. Probar que si μ y ν son discretas con soportes disjuntos, la longitud es infinita. (★) Mostrar en cambio que si $\mu = f\mathcal{L}^d$ y $\nu = g\mathcal{L}^d$ con f, g suaves y positivas en un dominio acotado y regular, la longitud es finita: existe un campo ξ con $\nabla \cdot (f\xi) = f - g$ y $\int |\xi|^2 f < \infty$, y entonces $W_2(\nu_s, \nu_t) \leq C|t-s|$ (comparar con la fórmula de Benamou–Brenier, [1, Sec. 5.4]).
- (c) Probar que $t \mapsto W_2^2(\nu_t, \mu)$ es convexa (es la restricción del funcional del Lema 12.3.1 a un segmento) y comparar con el Ejercicio 12.9.

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

Ejercicio 10.12 (Escala). Probar numéricamente, y luego a mano, que $W_p(\lambda_{\#}\mu, \lambda_{\#}\nu) = \lambda W_p(\mu, \nu)$ para la dilatación $x \mapsto \lambda x$, $\lambda > 0$, y que W_p es invariante por traslaciones simultáneas de ambas medidas (Ejercicio 10.2).

Ejercicio 10.13 (Traslaciones). Para $\nu = (\tau_v)_{\#}\mu$, con $\tau_v(x) = x + v$, se tiene $W_p(\mu, \nu) \leq |v|$. Probar que para $p = 2$ hay siempre igualdad y verificarlo numéricamente con medidas discretas. (*Sugerencia:* probar primero la descomposición $W_2^2(\mu, \nu) = |m_{\mu} - m_{\nu}|^2 + W_2^2(\tilde{\mu}, \tilde{\nu})$, donde m_{μ} denota la media de μ y $\tilde{\mu}, \tilde{\nu}$ las medidas centradas.)

Ejercicio 10.14 (La interpolación lineal tiene longitud infinita). Para $\mu = \delta_0$ y $\nu = \delta_1$ en $\mathbb{R}$, mostrar que $W_2((1-s)\mu + s\nu, (1-t)\mu + t\nu) = \sqrt{|t-s|}$ y deducir que la curva $t \mapsto (1-t)\mu + t\nu$ tiene longitud infinita en $(\mathcal{P}_2(\mathbb{R}), W_2)$. Comparar con la Sección 5.1 del cuaderno y con el Ejercicio 10.11.

Ejercicio 10.15 (Un tercer exponente). Extender la construcción de la Sección 4 del cuaderno para exhibir, para cada par $1 \leq q < p$, una sucesión μ_n con $W_q(\mu_n, \delta_0) \rightarrow 0$ pero $W_p(\mu_n, \delta_0) \rightarrow \infty$ (cf. Ejercicio 10.3).

Ejercicio 10.16 (Kantorovich–Rubinstein con el solver exacto). El solver exacto devuelve también potenciales duales (`ot.emd(a, b, M, log=True)`). Comparar los potenciales (u, v) así obtenidos con la función 1-Lipschitz φ de la Sección 6 del cuaderno: no tienen por qué coincidir (¿por qué?), pero el valor dual $\sum_i a_i u_i + \sum_j b_j v_j$ debe ser $W_1(\mu, \nu)$. ¿Es $v = -u$ sobre los puntos comunes a ambos soportes? ¿Por qué no necesariamente?

REFERENCIAS

- [1] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*, volume 87 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Cham, 2015.