Distanza di Wasserstein

Consideriamo le misure di Borel su $\mathbb{R}^n$ unitarie e con secondo momento finito:

1) $$P_2(\mathbb{R}^n) = \left\{ \mu ;\ \mu \text{ misura di Borel su } \mathbb{R}^n \text{ con } \mu(\mathbb{R}^n) = 1 \text{ e } \int_{\mathbb{R}^n} |x|^2 d\mu < \infty \right\}.$$

La distanza di Wasserstein $W_2 : P_2(\mathbb{R}^n) \times P_2(\mathbb{R}^n) \to [0,\infty)$ è definita come

2) $$W_2(\mu,\nu) = \min \left\{ \left(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 d\pi \right)^{1/2} ;\ \pi \in \Pi(\mu,\nu) \right\}.$$

Elenchiamo senza dimostrazione alcuni fatti di base:

1. $(P_2(\mathbb{R}^n), W_2)$ è uno spazio metrico. Inoltre la mappa $\mathbb{R}^n \ni x \mapsto \delta_x \in P_2(\mathbb{R}^n)$, delta di Dirac in x, è una isometria.

2. Questo spazio metrico è completo e separabile.

3. Se $K \subset \mathbb{R}^n$ è un compatto,

3) $$\mathcal{P}(K) = \{ \mu \in \mathcal{P}(\mathbb{R}^n) ;\ \operatorname{spt}(\mu) \subset K \} \subset P_2(\mathbb{R}^n)$$

è compatto.

4. La convergenza $\mu_K \xrightarrow[K\to\infty]{W_2} \mu$, $\mu_K, \mu \in P_2(\mathbb{R}^n)$ equivale ad A) + B):

A) $\mu_K \to \mu$ debolmente (dualità con funzioni continue e limitate)

B) $$\lim_{K\to\infty} \int_{\mathbb{R}^n} |x|^2 d\mu_K = \int_{\mathbb{R}^n} |x|^2 d\mu.$$

5. Siano $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^n)$ tali che $\mu_0 << \mathcal{L}^n$.

Sia $T = \nabla\varphi$ la soluzione del problema di Monge per μ_0, μ_1.

Per $t \in [0,1]$ prendiamo la combinazione lineare - convessa

4) $$T_t = (1-t)\, Id + tT$$

e definiamo la curva $\gamma : [0,1] \to \mathcal{P}_2(\mathbb{R}^n)$

5) $$\gamma(t) = (T_t)_\# \mu_0 , \qquad t \in [0,1]$$

Chiaramente $\gamma(0) = \mu_0$ e $\gamma(1) = \mu_1$.

<u>Teorema</u> La curva γ è l'unica geodetica (a velocità costante su $[0,1]$) che collega μ_0 a μ_1.

Nel seguito useremo diffusamente la curva di misure $(T_t)_\# \mu$. Ma non useremo la sua proprietà geodetica.

Displacement convexity

Esistono tre principali tipi di funzionali di tipo "energie" definiti sullo spazio di misure.

(1) Energia potenziale. Assegnata una funzione potenziale $V: \mathbb{R}^n \xrightarrow{\text{Borel}} \mathbb{R} \cup \{+\infty\}$ (limitata dal basso) si definisce il funzionale $\mathcal{V}: P(\mathbb{R}^n) \to \mathbb{R} \cup \{+\infty\}$

6) $$\mathcal{V}(\mu) = \int_{\mathbb{R}^n} V(x)\, d\mu.$$

(2) Energia di interazione. Assegnata una funzione di interazione $W: \mathbb{R}^n \xrightarrow{\text{Borel}} \mathbb{R} \cup \{+\infty\}$ (limitata dal basso) si definisce il funzionale $\mathcal{W}: P(\mathbb{R}^n) \times P(\mathbb{R}^n) \to \mathbb{R} \cup \{\infty\}$

7) $$\mathcal{W}(\mu, \nu) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} W(x-y)\, d\mu(x)\, d\nu(y)$$

significativa anche nel caso $\mu = \nu$.

(3) Energia interna. Assegnata una funzione $U: \mathbb{R}^+ \xrightarrow{\text{Borel}} \mathbb{R} \cup \{\infty\}$ con proprietà opportune (ad es $U \geq 0$) si definisce il funzionale $\mathcal{U}$

8) $$\mathcal{U}(\mu) = \begin{cases} \int_{\mathbb{R}} U(\rho)\, dx & \text{se } \mu = \rho\, \mathcal{L}^n \\ +\infty & \text{altrimenti} \end{cases}$$

Ci interessa l'energia interna e vogliamo capire quando verifica la seguente proprietà di convessità.

<u>Def</u> Siano $\mu, \nu \in P_2(\mathbb{R}^n)$ tali che $\mu, \nu \ll \mathcal{L}^n$. Sia $\gamma(t) = ((1-t)\,Id + t\nabla\varphi)_{\#} \in P_2(\mathbb{R}^n)$ la geodetica da μ a ν. Diciamo che l'energia interna $\mathcal{U}$ è <u>displacement convessa</u> se la funzione $t \mapsto \mathcal{U}(\gamma(t))$ è convessa.

<u>Def</u> Diciamo che la funzione interna $U : \mathbb{R}^+ \to \mathbb{R}^+$ verifica la proprietà MC_n, $n \in \mathbb{N}$, se $s \mapsto s^n U(s^{-n})$, $s>0$, è noncrescente e convessa.

MC sta per McCann. Possiamo aggiungere $U(0) = 0$. Vediamo degli esempi:

(1) Entropia: $U(\rho) = \rho \log \rho$, $\rho > 0$. È negativa per $0 < \rho < 1$. Abbiamo

9) $$s^n U(s^{-n}) = s^n s^{-n} \log s^{-n} = -n \log s \,, \quad s>0$$

che è decrescente e convessa.

(2) Entropia di Rényi: $U(\rho) = -\rho^\alpha$ con $\alpha \in \mathbb{R}$ da discutere. Abbiamo

10) $$s^n U(s^{-n}) = s^n(-s^{-n\alpha}) = -s^{n(1-\alpha)} = \phi(s).$$

Per $\alpha \leq 1$ ϕ è decrescente. Inoltre

$$(11)\qquad \begin{aligned} \phi'(s) &= -n(1-\alpha)\, s^{n(1-\alpha)-1} \\ \phi''(s) &= \underbrace{-n(1-\alpha)}_{\geq 0 \text{ per } \alpha \leq 1} [n(1-\alpha)-1]\, s^{n(1-\alpha)-2} \end{aligned}$$

e dunque $\phi'' \geq 0$ per $n(1-\alpha)-1 \leq 0 \iff \alpha \geq 1-\frac{1}{n}$.

Dunque l'Entropia di Rényi verifica MC_n per

$$(12)\qquad 1-\frac{1}{n} \leq \alpha \leq 1.$$

(3) Potenza positiva; $U(\rho) = \rho^{\beta}$. In questo caso

$$13)\qquad s^n U(s^{-n}) = s^n s^{-n\beta} = s^{n(1-\beta)} = \phi(s)$$

che è decrescente per $\beta \geq 1$. Inoltre

$$14)\qquad \begin{aligned} \phi'(s) &= n(1-\beta)\, s^{n(1-\beta)-1} \\ \phi''(s) &= \underbrace{n(1-\beta)[n(1-\beta)-1]}_{\geq 0 \quad \beta \geq 1} s^{n(1-\beta)-2} \end{aligned}$$

e per $\beta \geq 1$ è anche convessa.

Siano $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^n)$ con $\mu_0 \ll \mathcal{L}^n$ e sia $T := \nabla\varphi : \mathbb{R}^n \to \mathbb{R}^n$ la soluzione del problema di Monge con $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ convessa. Siano poi $D = \mathrm{dom}(\varphi) = \{x \in \mathbb{R}^n : \varphi(x) \in \mathbb{R}\}$,

15) $$D_0 = \left\{ x \in D : \begin{array}{l} H\varphi(x) \text{ esiste nel senso} \\ \text{q.o. di Alexandrov} \end{array} \right\},$$

$$\Sigma = \{x \in D_0 : \det H\varphi(x) = 0\}.$$

Posto $T = \nabla\varphi$, per $t \in [0,1)$ formiamo $T_t : \mathbb{R}^n \to \mathbb{R}^n$

16) $$T_t(x) = (1-t)x + tT(x) = \nabla\varphi_t(x)$$

dove $\varphi_t = (1-t)\dfrac{|x|^2}{2} + t\varphi$ è convessa.

<u>Lemma 1</u> Per ogni $t \in [0,1)$, T_t è iniettiva su D_0, la misura $\mu_t = (T_t)_\# \mu_0$ è concentrata su $T_t(D_0)$ ed inoltre, se $\mu_0 = \rho\,\mathcal{L}^n$ allora $\mu_t = \rho_t\,\mathcal{L}^n$ dove

17) $$\rho_t = \frac{\rho}{\det \nabla T_t} \circ T_t^{-1}, \qquad \text{q.o. su } T_t(D_0).$$

Dim. In primo luogo

18) $$\begin{aligned} \langle T_t(x) - T_t(y), x-y \rangle &= \langle (1-t)(x-y) + t(T(x) - T(y)), x-y \rangle \\ &= (1-t)|x-y|^2 + t\underbrace{\langle T(x) - T(y), x-y\rangle}_{\geq 0 \text{ perché } T = \nabla\varphi \text{ è monotono}} \geq (1-t)|x-y|^2 \end{aligned}$$

Si deduce che

19) $$|T_t(x) - T_t(y)| \geq (1-t)|x-y|,$$

e dunque, per $t<1$ la mappa $T_t : D_0 \to T_t(D_0)$ è iniettiva e inoltre $T_t^{-1} : T_t(D_0) \to D_0$ è Lipschitz con $Lip(T_t^{-1}) = \frac{1}{1-t}$.

Sia ora $A \subset T_t(D_0)$ di Borel con $\mathcal{L}^n(A) = 0$. Allora

20) $$\mu_t(A) = (T_t)_\# \mu_0(A) = \mu_0(T_t^{-1}(A)) = 0$$

perché $\mathcal{L}^n(T_t^{-1}(A)) = 0$ essendo T_t^{-1} Lipschitz, ed essendo $\mu_0 << \mathcal{L}^n$. Questo prova che $\mu_t << \mathcal{L}^n$, ovvero $\mu_t = \rho_t \mathcal{L}^n$.

Proviamo la formula 17). Per $A \subset T_t(D_0)$ di Borel:

21) $$\begin{aligned}\int_A \rho_t(x)\,dx &= \mu_t(A) = (T_t)_\# \mu_0(A) \\ &= \mu_0(T_t^{-1}(A)) = \int_{T_t^{-1}(A)} \rho_0(y)\,dy \\ &= \int_A \rho_0(T_t^{-1}(x)) \det(\nabla T_t^{-1}(x))\,dx \\ &= \int_A \frac{\rho_0}{\det(\nabla T_t)} \circ T_t^{-1}(x)\,dx .\end{aligned}$$

Sopra abbiamo usato la formula del cambio di variabile $y = T_t^{-1}(x)$, valida per T_t^{-1} Lipschitz. Abbiamo anche usato il fatto che $\det \nabla T_t \neq 0$.

$\square$

Quando $t=1$ la situazione è più delicata.
Enunciamo senza dimostrazione il seguente lemma.

Lemma 2 Siano $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^n)$ con $\mu_0 \ll \mathcal{L}^n$ e $\mu_1 \ll \mathcal{L}^n$.
Sia $T = \nabla\varphi \in \mathcal{T}(\mu_0,\mu_1)$ la soluzione del problema di Monge.
Siano

22) $$\begin{aligned} D &= \{x \in \mathbb{R}^n : \nabla\varphi(x) \text{ esiste}\} \\ D_0 &= \{x \in \mathbb{R}^n : H\varphi(x) \text{ esiste nel senso di Alexandrov}\} \\ \Sigma &= \{x \in D_0 : \det H\varphi(x) = 0\}. \end{aligned}$$

Allora:

i) $T : D_0 \setminus \Sigma \to \mathbb{R}^n$ è 1-1.

ii) $\mathcal{L}^n(T(\Sigma)) = 0$

iii) Se $\mu_0 = \rho_0 \mathcal{L}^n$ e $\mu_1 = \rho_1 \mathcal{L}^n$ allora vale la formula di Monge-Ampère

23) $$\rho_1(\nabla\varphi)\det H\varphi = \rho_0 \quad \text{q.o. su } \{\varphi \in \mathbb{R}\}.$$

Ci serve anche il seguente risultato sulla concavità della funzione determinante. Sia $S_n^+(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : A = A^T \text{ ed } A \geq 0\}$ l'insieme delle matrici semidefinite positive. È un insieme convesso.

Lemma La funzione $S_n^+(\mathbb{R}) \ni A \to \det(A)^{1/n} \in \mathbb{R}$ è concava:

24) $$\det(A+B)^{\frac{1}{n}} \geq \det(A)^{\frac{1}{n}} + \det(B)^{\frac{1}{n}}.$$

Dim. Senza perdere di generalità si può supporre $A>0$. Basta considerare $A_\varepsilon = A + \varepsilon I$ con $\varepsilon > 0$ e poi fare $\varepsilon \downarrow 0$. Detta $C = \sqrt{A} > 0$ si fattorizza

25) $$A + B = C(I + \underbrace{C^{-1} B C^{-1}}_{\substack{\| \\ D \in S_n^+(\mathbb{R})}})C.$$

Per le proprietà del determinante basta provare il lemma per $A = I$ matrice identità. La tesi è:

26) $$\det(I+B)^{1/n} \geq 1 + \det(B)^{1/n}.$$

Diagonalizzando B, si può supporre B diagonale con autovalori $\lambda_1, \dots, \lambda_n \geq 0$. La (26) diventa

27) $$\prod_{i=1}^{n} (1+\lambda_i)^{1/n} \geq 1 + \prod_{i=1}^{n} \lambda_i^{1/n}.$$

Dalla disuguaglianza Media-geometrica $\leq$ Media-aritmetica:

28) $$\left(\prod_{i=1}^{n} \frac{1}{1+\lambda_i}\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^{n} \frac{1}{1+\lambda_i},$$

29) $$\left(\prod_{i=1}^{n} \frac{\lambda_i}{1+\lambda_i}\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^{n} \frac{\lambda_i}{1+\lambda_i},$$

e sommando membro a membro

30) $$\left(\prod_{i=1}^{n} \frac{1}{1+\lambda_i}\right)^{1/n} + \left(\prod_{i=1}^{n} \frac{\lambda_i}{1+\lambda_i}\right)^{1/n} \leq 1,$$

da cui si ottiene la (27).

□

<u>Teorema</u> (Convessità dell'Energia interna). Consideriamo un funzionale $\mathcal{U} : P_2(\mathbb{R}^n) \to \mathbb{R} \cup \{\infty\}$

(31) $$\mathcal{U}(\mu) = \begin{cases} \int_{\mathbb{R}^n} U(\rho)\, dx & \text{se } \mu = \rho\, \mathcal{L}^n \\ +\infty & \text{altrimenti} \end{cases}$$

dove $U : [0,\infty) \to \mathbb{R}$ verifica

i) $U(0) = 0$

ii) $s \mapsto s^n U(s^{-n})$ è convessa e decrescente (non crescente)

Siano $\mu_0, \mu_1 \in P_2(\mathbb{R}^n)$ con $\mu_0, \mu_1 << \mathcal{L}^n$ e sia $T \in \mathcal{T}(\mu_0, \mu_1)$ la soluzione del problema di Monge.

Allora $\mathcal{U}$ è convessa lungo la curva $t \mapsto \mu_t = \big((1-t)Id + tT\big)_\# \mu_0$, per $t \in [0,1]$.

Dim. Dal Lemma 1 sappiamo che $\mu_t = \rho_t \mathcal{L}^n$ con ρ_t che verifica (17). Dunque

(32) $$\mathcal{U}(\mu_t) = \int_{T_t(D_0)} U\left(\frac{\rho_0}{\det \nabla T_t} \circ T_t^{-1}\right) dy \qquad t \in [0,1),$$

dove D_0 è come in (15). Con il cambio di variabile $T_t^{-1}(y) = x$, che è Lipschitz (vedi dim. Lemma 1) si trova

(33) $$\mathcal{U}(\mu_t) = \int_{D_0} U\left(\frac{\rho_0}{\det \nabla T_t}\right) \det \nabla T_t \, dx, \quad t \in [0,1).$$

Per il Lemma (2) l'insieme $\Sigma = \{\det \nabla T = 0\}$ ha misura nulla. (Qui $t=1$). In (33) si può sostituire D_0 con $D_0 \setminus \Sigma$.

Quando $t=1$, la formula (33) si deduce dal Lemma 2. Omettiamo questi dettagli.

Ora osserviamo la funzione

$$[0,1] \ni t \mapsto f(t) = U\left(\frac{\rho_0}{\det \nabla T_t}\right) \det \nabla T_t \in \mathbb{R}.$$

$$= \rho_0 \, s^n \, U(s^{-n})$$

con $s = \left(\frac{\det \nabla T_t}{\rho_0}\right)^{1/n} = s(t)$. Ora $t \mapsto s(t)$ è concava per il lemma sui determinanti, mentre $s \mapsto \underbrace{s^n U(s^{-n})}_{= g(s)}$ è decrescente e convessa.

Dunque

$$g(\underbrace{s((1-t)\alpha + t\beta)}_{\geq (1-t)s(\alpha) + t s(\beta)}) \overset{g\downarrow}{\leq} g\left((1-t)s(\alpha) + t s(\beta)\right) \overset{g \text{ convessa}}{\leq}$$

$$\leq (1-t)\, g(s(\alpha)) + t\, g(s(\beta)).$$

□

Disuguaglianza di Brunn - Minkowski

Teorema Siano $A, B \subset \mathbb{R}^n$ due insiemi compatti. Per ogni $t \in [0,1]$ vale la disuguaglianza

$$\mathcal{L}^n((1-t)A + tB)^{1/n} \geq (1-t)\,\mathcal{L}^n(A)^{\frac{1}{n}} + t\,\mathcal{L}^n(B)^{\frac{1}{n}}.$$

Dim. Osserviamo preliminarmente che $(1-t)A + tB = \{(1-t)x + ty : x \in A \text{ e } y \in B\}$ è un insieme compatto e dunque di Borel.

Consideriamo le misure

$$\mu_0 = \frac{1}{\mathcal{L}^n(A)} \mathcal{L}^n \llcorner A \quad \text{e} \quad \mu_1 = \frac{1}{\mathcal{L}^n(B)} \mathcal{L}^n \llcorner B$$

chiaramente $\mu_0, \mu_1 \in P_2(\mathbb{R}^n)$ e $\mu_0, \mu_1 \ll \mathcal{L}^n$ con densità $\rho_0 = 1/\mathcal{L}^n(A)$ e $\rho_1 = 1/\mathcal{L}^n(B)$.

Sia $T \in \mathcal{T}(\mu_0, \mu_1)$ la soluzione del problema di Monge con costo quadratico. Sia poi

$$\mu_t = \left((1-t)\,\mathrm{Id} + tT\right)_{\#} \mu_0$$

la "geodetica" da μ_0 a μ_1. Osserviamo che se $K \subset \mathbb{R}^n \setminus \{(1-t)A + tB\}$ allora

$$\begin{aligned}\mu_t(K) &= \mu_0(\{x \in \mathbb{R}^n : (1-t)\underset{\substack{\cap\\ A}}{x} + t\underset{\substack{\cap\\ B}}{\underbrace{T(x)}} \in K\}) \\ &= \mu_0(\emptyset) = 0,\end{aligned}$$

ovvero μ_t è concentrata su $(1-t)A + tB$.

Consideriamo l'entropia di Rényi

$$U(s) = -s^{1-\frac{1}{n}}, \quad s \geq 0$$

che verifica $U(0)=0$, con $s \mapsto s^n U(s^{-n})$ decrescente e convessa (proprietà MC_n). Dunque il funzionale

$$\mathcal{U}(\mu) = -\int_{\mathbb{R}^n} \rho^{1-\frac{1}{n}} dx, \quad \mu = \rho \mathcal{L}^n$$

è convesso lungo μ_t: $\quad (|A| = \mathcal{L}^n(A))$

$$\mathcal{U}(\mu_t) \leq -(1-t)\int_A |A|^{\frac{1}{n}-1} dx - t\int_B |B|^{\frac{1}{n}-1} dx =$$

$$= -(1-t)|A|^{\frac{1}{n}} - t|B|^{\frac{1}{n}}.$$

Sappiamo che $\mu_t \ll \mathcal{L}^n$. Sia $\mu_t = \rho_t \mathcal{L}^n$ con $\rho_t = 0$ fuori da $(1-t)A + tB := C$. Dunque

$$-\mathcal{U}(\mu_t) = |C| \int_C \rho_t^{1-\frac{1}{n}} \frac{dx}{|C|}$$

$$\overset{\text{Hölder}}{\leq} |C| \left(\int_C \rho_t \frac{dx}{|C|} \right)^{1-\frac{1}{n}} = |C|^{\frac{1}{n}}$$

perché $\int_C \rho_t \, dx = 1$. Riordinando le disuguaglianze si ottiene la tesi.

$\square$

Sia $E \subset \mathbb{R}^n$ un insieme compatto (limitato). Un modo per misurare l'area della frontiera ∂E è il contenuto di Minkowski:

$$\mathcal{M}(E) = \liminf_{r \downarrow 0} \frac{\mathcal{L}^n(E + rB) - \mathcal{L}^n(E)}{r}$$

dove $B = \{x \in \mathbb{R}^n : |x| < 1\}$ è la palla unitaria.

Se ∂E è regolare, ad esempio localmente il grafico di una funzione Lipschitz, allora $\mathcal{M}(E) = H^{n-1}(\partial E)$ è la nozione classica di area di ipersuperficie.

<u>Teorema</u> (Disuguaglianza isoperimetrica) Sia $E \subset \mathbb{R}^n$ un insieme compatto con $\mathcal{L}^n(E) = \mathcal{L}^n(B)$. Allora

$$\mathcal{M}(E) \geq \mathcal{M}(B).$$

Dim. Per la disuguaglianza di Brunn-Minkowski:

$$\mathcal{L}^n(E + rB)^{1/n} \geq \mathcal{L}^n(E)^{1/n} + \mathcal{L}^n(rB)^{1/n} = \mathcal{L}^n(E)^{\frac{1}{n}} + r\,\mathcal{L}^n(B)^{\frac{1}{n}}$$

Quindi

$$\mathcal{M}(E) \geq \liminf_{r \downarrow 0} \frac{1}{r}\left\{ \left(\mathcal{L}^n(E)^{\frac{1}{n}} + r\,\mathcal{L}^n(B)^{\frac{1}{n}}\right)^n - \mathcal{L}^n(E)\right\}$$

$$= \frac{d}{dr}\Big|_{r=0} \left(\mathcal{L}^n(E)^{\frac{1}{n}} + r\,\mathcal{L}^n(B)^{\frac{1}{n}}\right)^n = n\,\mathcal{L}^n(E)^{\frac{n-1}{n}}\,\mathcal{L}^n(B)^{\frac{1}{n}}.$$

$$= n\,\mathcal{L}^n(B).$$

D'altra parte

$$\mathcal{M}(B) = \lim_{r \downarrow 0} \frac{\mathcal{L}^n(B + rB) - \mathcal{L}^n(B)}{r}$$

$$= \lim_{r \downarrow 0} \frac{\mathcal{L}^n((1+r)B) - \mathcal{L}^n(B)}{r}$$

$$= \lim_{r \downarrow 0} \frac{(1+r)^n - 1}{r} \mathcal{L}^n(B) = n \mathcal{L}^n(B).$$

$\square$