Università di Pisa

Dipartimento di Matematica
Corso di Laurea in Matematica

Introduzione al Trasporto Ottimo

Seminario di Calcolo delle Variazioni

Marco Inversi

15 febbraio 2020

Indice

Notazione

1. Sia $(\mathbb{X}; \mathcal{T})$ uno spazio topologico. Denoteremo con $\mathcal{B}(\mathbb{X})$ la σ-algebra dei boreliani in $\mathbb{X}$ indotta dalla topologia $\mathcal{T}$. Adotteremo la seguente notazione:
   - $\mathscr{M}(\mathbb{X})$ è l'insieme delle misure boreliane su $\mathbb{X}$ finite con segno;
   - $\mathscr{M}_+(\mathbb{X})$ è l'insieme delle misure boreliane su $\mathbb{X}$ finite e non negative;
   - $\mathscr{P}(\mathbb{X})$ è l'insieme delle misure di probabilità boreliane su $\mathbb{X}$.
2. Data $\mu \in \mathscr{M}(\mathbb{X})$, adotteremo la seguente notazione:
   - $|\mu|$ è la variazione totale di μ (ricordiamo che $|\mu| \in \mathscr{M}_+(\mathbb{X})$);
   - μ^+ e μ^- sono rispettivamente la parte positiva e la parte negativa della misura μ (ricordiamo che $\mu^+, \mu^- \in \mathscr{M}_+(\mathbb{X})$ e che $\mu = \mu^+ - \mu^-$);
   - il supporto di μ è l'insieme

     $$\operatorname{supp}(\mu) \coloneqq \{x \in \mathbb{X} \mid \ |\mu|(U) > 0 \ \forall U \text{ intorno aperto di } x\}.$$

     Ricordiamo che se $(\mathbb{X}; \mathcal{T})$ è metrizzabile il supporto di μ è chiuso in $\mathbb{X}$.
3. Dati μ in $\mathscr{M}(\mathbb{X})$ e $E \in \mathcal{B}(\mathbb{X})$, diremo che μ è concentrata in E se vale $|\mu|(E^c) = 0$ o, equivalentemente, se per ogni boreliano $E' \subseteq E^c$ vale che $\mu(E') = 0$. Ricordiamo che una misura è concentrata in ogni insieme boreliano contenente il suo supporto.
4. Dati $\mu \in \mathscr{M}(\mathbb{X})$ e $E \in \mathcal{B}(\mathbb{X})$, si pone $\mu \llcorner E$ la misura in $\mathscr{M}(\mathbb{X})$ definita da

   $$\mu \llcorner E(B) \coloneqq \mu(E \cap B).$$

Capitolo 1

Il problema del trasporto ottimo

1.1 Preliminari

Iniziamo definendo il contesto in cui ambienteremo il problema e la nozione di push-forward di misure; mostreremo una semplice formula di cambio di variabili negli integrali, di cui faremo largo uso nel seguito.

1.1.1 Generalità sugli spazi metrici

Definizione 1.1.1 (Spazio polacco)**.**
Uno spazio metrico $(\mathbb{X}; d)$ completo e separabile è detto spazio metrico polacco.

Mostriamo il seguente risultato di regolarità interna delle misure boreliane negli spazi polacchi.

Lemma 1.1.2 (Ulam)**.**
Siano $(\mathbb{X}; d)$ uno spazio polacco, $\mu \in \mathscr{M}_+(\mathbb{X})$ ed $\varepsilon > 0$. Esiste un compatto $K \subseteq \mathbb{X}$ tale che $\mu(\mathbb{X} \setminus K) < \varepsilon$.

Dimostrazione. Sia $D = \{x_i \mid i \in \mathbb{N}\}$ un denso numerabile in $\mathbb{X}$. Per ogni $k \in \mathbb{N}^+$ vale che

$$\mathbb{X} = \bigcup_{i \in \mathbb{N}} \overline{B}(x_i; 2^{-k}).$$

Essendo μ una misura finita, per ogni $k \in \mathbb{N}$ esiste $n(k) \in \mathbb{N}$ tale che

$$\mu \left(\mathbb{X} \setminus \bigcup_{i \leq n(k)} \overline{B}(x_i; 2^{-k}) \right) < \frac{\varepsilon}{2^{k+1}}.$$

Introduciamo l'insieme

$$K \coloneqq \bigcap_{k \in \mathbb{N}} \left(\bigcup_{i \leq n(k)} \overline{B}(x_i; 2^{-k}) \right).$$

Notiamo che K è totalmente limitato; inoltre K è chiuso, perchè è intersezione di una famiglia numerabile di chiusi (ciascun elemento di tale famiglia è unione finita di un palle chiuse); essendo $(\mathbb{X}; d)$ completo, vale che K è completo. Concludiamo che K è compatto. Infine, notiamo che

$$\mu(\mathbb{X} \setminus K) \leq \sum_{k \in \mathbb{N}} \mu \left(\mathbb{X} \setminus \bigcup_{i \leq n(k)} \overline{B}(x_i; 2^{-k}) \right) < \varepsilon.$$

□

Presentiamo anche il seguente lemma di regolarizzazione delle funzioni semicontinue inferiormente per inf-convoluzione. Nel seguito faremo larghissimo uso in seguito.

Lemma 1.1.3. *Siano $(\mathbb{X}; d_X)$ uno spazio metrico e $c : \mathbb{X} \to [0, +\infty]$ una funzione semicontinua inferiormente. Esiste una successione $(c_k)_k$ tale che per ogni $k \in \mathbb{N}$ valgono le seguenti proprietà:*

- $c_k : \mathbb{X} \to [0, k]$;
- c_k *è* k*-lipschitziana;*
- $0 \leq c_k(x) \leq c_{k+1}(x) \leq c(x)$ *per ogni* $x \in \mathbb{X}$;
- $\sup_{k \in \mathbb{N}} c_k(x) = c(x)$ *per ogni* $x \in \mathbb{X}$.

Dimostrazione. Fissiamo $k \in \mathbb{N}^+$ e per ogni $x \in \mathbb{X}$ definiamo

$$c_k(x) \coloneqq \inf_{x' \in \mathbb{X}} \{c(x') \wedge k + k d_X(x; x')\}.$$

Mostriamo che la successione $(c_k)_k$ ha le proprietà richieste. Fissiamo $k \in \mathbb{N}$; è immediato verificare che $c_k(x) \in [0, k]$ per ogni $x \in \mathbb{X}$. Per ogni $x' \in \mathbb{X}$ poniamo

$$g_{k,x'}(x) \coloneqq c(x') \wedge k + k d_X(x; x')$$

e notiamo che $g_{k,x'} : \mathbb{X} \to [0, +\infty)$ è una funzione k-lipschitziana; per definizione, vale che

$$c_k(x) = \inf_{x' \in \mathbb{X}} g_{k,x'}(x);$$

allora c_k è k-lipschitziana (è estremo inferiore puntuale di funzioni con questa proprietà). Fissiamo $x \in \mathbb{X}$: essendo $g_{k,x'}(x) \leq g_{k+1,x'}(x) \leq c(x)$ per ogni $x' \in \mathbb{X}$, la proprietà passa all'estremo inferiore su $x' \in \mathbb{X}$. Infine, fissiamo $x \in \mathbb{X}$; osserviamo che la disuguaglianza

$$\sup_{k \in \mathbb{N}} c_k(x) \leq c(x)$$

è ovvia. Per provare la disuguaglianza opposta, possiamo supporre che l'estremo superiore sia finito (altrimenti è ovvio). Per ogni $k \in \mathbb{N}$ esiste $x'_k \in \mathbb{X}$ tale che

$$c(x'_k) \wedge k + k d_X(x; x'_k) \leq c_k(x) + 2^{-k}.$$

Essendo $\sup_{k \in \mathbb{N}} c_k(x) < +\infty$, deve valere che $d_X(x; x'_k) \to 0$. Inoltre, abbiamo anche che

$$c(x'_k) \wedge k \leq c_k(x) + 2^{-k}.$$

Dalla semicontinuità inferiore di c deduciamo che

$$c(x) \leq \liminf_{k \to +\infty} c(x'_k) = \liminf_{k \to +\infty} c(x'_k) \wedge k \leq \lim_{k \to +\infty} c_k(x) + 2^{-k} = \sup_{k \in \mathbb{N}} c_k(x).$$

□

Osservazione 1.1.4. Vale la pena notare che il risultato presentato nel lemma 1.1.3 è praticamente ottimale, perchè in generale l'estremo superiore puntuale di funzioni continue è una funzione semicontinua inferiormente.

1.1.2 Push forward di misure

Introduciamo la seguente definizione in un contesto del tutto generale, anche se nel seguito lavoreremo sempre in spazi metrici.

Definizione 1.1.5 (Push forward)**.**
Siano dati due spazi topologici $\mathbb{X}, \mathbb{Y}$ e una funzione boreliana $f : \mathbb{X} \to \mathbb{Y}$. Definiamo l'operatore di push-forward $f_\# : \mathscr{M}(\mathbb{X}) \to \mathscr{M}(\mathbb{Y})$ tale che per ogni misura boreliana $\mu \in \mathscr{M}(\mathbb{X})$, $f_\#\mu$ è la misura boreliana in $\mathscr{M}(\mathbb{Y})$ definita da

$$f_\#\mu(B) = \mu(f^{-1}(B)).$$

Osservazione 1.1.6. L'operatore di push-forward (vedi 1.1.5) è ben definito ed è lineare (rispetto alla misura); è immediato osservare che $f_\#(\mathscr{M}_+(\mathbb{X})) \subseteq \mathscr{M}_+(\mathbb{Y})$ e $f_\#(\mathscr{P}(\mathbb{X})) \subseteq \mathscr{P}(\mathbb{Y})$. Inoltre, se $\mathbb{Z}$ è uno spazio topologico e $g : \mathbb{Y} \to \mathbb{Z}$ è una funzione boreliana, vale che $(g \circ f)_\# = g_\# \circ f_\#$.

Presentiamo i seguenti risultati per misure boreliane non negative, perchè ne faremo uso in questo contesto. Precisiamo che formule analoghe alle seguenti valgono anche per misure boreliane con segno.

Proposizione 1.1.7 (Formula di cambio di variabili)**.**
Siano dati due spazi topologici $\mathbb{X}, \mathbb{Y}$, una funzione boreliana $f : \mathbb{X} \to \mathbb{Y}$, una misura $\mu \in \mathscr{M}_+(\mathbb{X})$ e una funzione boreliana $\varphi : \mathbb{Y} \to [-\infty, +\infty]$. Allora vale che φ è $f_\#\mu$-integrabile se e solo se $\varphi \circ f$ è μ-integrabile. In tal caso, vale la formula di cambio di variabili

$$\int_{\mathbb{X}} \varphi \circ f \, d\mu = \int_{\mathbb{Y}} \varphi \, df_\#\mu. \tag{1.1}$$

Dimostrazione. Per la definizione di push forward (vedi 1.1.5), la formula (1.1) vale se φ è l'indicatrice di un boreliano in $\mathbb{Y}$. Per linearità, la formula (1.1) vale se φ è una funzione semplice non negativa su $\mathbb{Y}$. Sia $\varphi : \mathbb{Y} \to [0, +\infty]$ una funzione boreliana. Esiste una successione $(\varphi_n)_n$ di funzioni semplici tali che

- $0 \leq \varphi_n(y) \leq \varphi_{n+1}(y) \leq \varphi(y)$ per ogni $y \in \mathbb{Y}$ per ogni $n \in \mathbb{N}$;
- per ogni $y \in \mathbb{Y}$ vale che $\varphi(y) = \sup_n \varphi_n(y)$.

Notiamo che $(\varphi_n \circ f)_n$ è una successione di funzioni semplici non negative che tende puntualmente crescendo a $\varphi \circ f$. Sappiamo che per ogni $n \in \mathbb{N}$ vale che

$$\int_{\mathbb{X}} \varphi_n \circ f \, d\mu = \int_{\mathbb{Y}} \varphi_n \, df_\#\mu.$$

Passando al limite con il teorema di Beppo Levi, troviamo che la formula (1.1) vale per φ. Sia $\varphi : \mathbb{Y} \to [-\infty, +\infty]$ una funzione boreliana. Consideriamo la decomposizione in parte positiva e negativa $\varphi = \varphi^+ - \varphi^-$. Per definizione φ è $f_\#\mu$-integrabile se e solo se

$$\min\left\{\int_{\mathbb{Y}} \varphi^+ \, df_\#\mu; \int_{\mathbb{Y}} \varphi^- \, df_\#\mu\right\} < +\infty;$$

essendo φ^+, φ^- non negative, è equivalente richiedere che

$$\min\left\{\int_{\mathbb{X}} \varphi^+ \circ f \, d\mu; \int_{\mathbb{X}} \varphi^- \circ f \, d\mu\right\} < +\infty,$$

che è la definizione di integrabilità di $\varphi \circ f$ rispetto alla misura μ. In tal caso, si ha

$$\begin{aligned}\int_{\mathbb{Y}} f \, df_{\#}\mu &= \int_{\mathbb{Y}} \varphi^{+} \, df_{\#}\mu - \int_{\mathbb{Y}} \varphi^{-} \, df_{\#}\mu \\ &= \int_{\mathbb{X}} \varphi^{+} \circ f \, d\mu - \int_{\mathbb{X}} \varphi^{-} \circ f \, d\mu \\ &= \int_{\mathbb{X}} \varphi \circ f \, d\mu,\end{aligned}$$

intendendo le somme nell'algebra di $\overline{\mathbb{R}}$. □

La seguente proposizione fornisce un criterio utile per capire se due misure sono una il push forward dell'altra.

Proposizione 1.1.8. *Siano $\mathbb{X}, \mathbb{Y}$ spazi metrici, $\mu \in \mathscr{M}_{+}(\mathbb{X})$, $\nu \in \mathscr{M}_{+}(\mathbb{Y})$ e $T : \mathbb{X} \to \mathbb{Y}$ una funzione boreliana. Si ha che $\nu = T_{\#}\mu$ se e solo se*

$$\int_{\mathbb{Y}} \varphi \, d\nu = \int_{\mathbb{X}} \varphi \circ T \, d\mu \quad \forall \varphi \in C_b(\mathbb{Y}). \tag{1.2}$$

Dimostrazione. La proposizione 1.1.7 implica che vale la condizione (1.2).

Viceversa, supponiamo che valga la condizione (1.2). Notiamo che è sufficiente mostrare che per ogni aperto $A \subseteq \mathbb{Y}$ vale che $\nu(A) = T_{\#}\mu(A)$; infatti, se due misure boreliane coincidono sugli aperti, allora sono uguali. Dato un aperto $A \subseteq \mathbb{Y}$, notiamo che la funzione $\mathbb{1}_A$ è semicontinua inferiormente; per il lemma 1.1.3 esiste una successione $(\varphi)_n \subseteq C_b(\mathbb{Y})$ tale che

- $0 \le \varphi_n(y) \le \varphi_{n+1}(y) \le \mathbb{1}_A(y)$ per ogni $y \in \mathbb{Y}$ per ogni $n \in \mathbb{N}$,
- $\sup_n \varphi_n(y) = \mathbb{1}_A(y)$ per ogni $y \in \mathbb{Y}$.

Notiamo che $(\varphi_n \circ T)_n$ è una successione di funzioni che converge puntualmente crescendo a $\mathbb{1}_{T^{-1}(A)}$. Per ogni $n \in \mathbb{N}$ vale

$$\int_{\mathbb{Y}} \varphi_n(y) \, d\nu = \int_{\mathbb{X}} \varphi_n \circ T \, d\mu;$$

passando al limite con il teorema di Beppo Levi, troviamo che

$$\int_{\mathbb{Y}} \mathbb{1}_A(y) \, d\nu = \int_{\mathbb{X}} \mathbb{1}_{T^{-1}(A)}(x) \, d\mu,$$

ovvero $\nu(A) = \mu(T^{-1}(A))$. □

1.2 Il problema del trasporto ottimo di massa

1.2.1 Formulazione di Monge

Sia data una certa quantità di massa distribuita in un certo modo; vogliamo spostarla per distribuirla in un altro modo. Trasportare un'unità di massa ha un costo (è ragionevole immaginare che sia proporzionale alla distanza); ci chiediamo quale sia il modo più economico (in termini di costo) per spostare la massa dalla vecchia alla nuova configurazione. La formulazione originale del problema di trasporto ottimo di Monge (1781) nasce con queste motivazioni. Possiamo descrivere una distribuzione di massa come una misura; allora, siano dati due spazi topologici $\mathbb{X}, \mathbb{Y}$, due misure di probabilità $\mu \in \mathscr{P}(\mathbb{X})$, $\nu \in \mathscr{P}(\mathbb{Y})$ e una funzione costo $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ ($c(x; y)$ indica il costo che ha spostare un'unità di massa dal punto $x \in \mathbb{X}$ al punto $y \in \mathbb{Y}$). Data una mappa boreliana $T : \mathbb{X} \to \mathbb{Y}$, diciamo che è una mappa di trasporto ammissibile se $T_{\#}\mu = \nu$. In altri termini, abbiamo una distribuzione di massa μ in $\mathbb{X}$; spostiamo ciascun punto seguendo la mappa T (cioè il punto x viene mandato in $T(x)$); la distribuzione di massa che otteniamo in $\mathbb{Y}$ è data proprio da $T_{\#}\mu$: chiediamo che la mappa T sistemi in punti disposti secondo μ in $\mathbb{X}$ in modo che rispettino la distribuzione ν in $\mathbb{Y}$. Per ogni mappa boreliana $T : \mathbb{X} \to \mathbb{Y}$ consideriamo il funzionale

$$\mathcal{C}(T) := \int_{\mathbb{X}} c(x; T(x)) \; d\mu;$$

$\mathcal{C}(T)$ rappresenta il costo complessivo per spostare la massa distribuita in $\mathbb{X}$ secondo μ in $\mathbb{Y}$ distribuendola secondo $T_{\#}\mu$. La formulazione di Monge del problema di trasporto ottimo consiste nello studio del problema

$$\inf \left\{ \int_{\mathbb{X}} c(x; T(x)) \; d\mu \;\middle|\; T : \mathbb{X} \to \mathbb{Y} \text{ è boreliana, } T_{\#}\mu = \nu \right\}.$$

Questa formulazione presenta molti problemi: innanzitutto, il funzionale da minimizzare è completamente non lineare e asimmetrico rispetto a $\mathbb{X}$ e $\mathbb{Y}$. Inoltre il problema non è ben posto, nel senso che l'insieme delle mappe di trasporto ammissibili potrebbe essere vuoto: basti pensare al caso in cui $\mu = \delta_{x_0}$ e $\nu = \frac{1}{2}(\delta_{y_0} + \delta_{y_1})$. In questo caso, è verificato il vincolo della conservazione della massa, ma una mappa di trasporto non può disporre in due punti distinti la massa concentrata in un unico punto. Per di più, è impossibile trattare questo problema con il metodo diretto.

Per affrontare il problema, bisogna procedere in maniera più generale, dando un'interpretazione diversa (e meno intuitiva) alla situazione che cerchiamo di descrivere. Si formula un problema in versione rilassata e soltanto alla fine si scopre la regolarità della soluzione.

1.2.2 Formulazione di Kantorovich

Siano dati $\mathbb{X}, \mathbb{Y}$ spazi topologici. Denotiamo con $p_X : \mathbb{X} \times \mathbb{Y} \to \mathbb{X}$ e $p_Y : \mathbb{X} \times \mathbb{Y} \to \mathbb{Y}$ le proiezioni coordinate. Siano date $\mu \in \mathscr{P}(\mathbb{X})$ e $\nu \in \mathscr{P}(\mathbb{Y})$.

Definizione 1.2.1 (Piani di trasporto)**.**
Si dice che $\pi \in \mathscr{P}(\mathbb{X} \times \mathbb{Y})$ è un piano di trasporto se $(p_X)_{\#}\pi = \mu$ e $(p_Y)_{\#}\pi = \mu$, cioè se π è una probabilità boreliana su $\mathbb{X} \times \mathbb{Y}$ che ha μ e ν come marginali. Per l'insieme dei piani di trasporto utilizzeremo la notazione $\Gamma(\mu; \nu)$.

Osservazione 1.2.2. Notiamo che una probabilità boreliana π su $\mathbb{X} \times \mathbb{Y}$ è un piano di trasporto se e solo se per ogni $A \in \mathcal{B}(\mathbb{X})$ vale che $\mu(A) = \pi(A \times \mathbb{Y})$ e per ogni $B \in \mathcal{B}(\mathbb{Y})$ vale che $\nu(B) = \pi(\mathbb{X} \times B)$.

Ricordiamo che una probabilità boreliana sul $\mathbb{X} \times \mathbb{Y}$ è univocamente determinata dal valore che assume sui rettangoli boreliani, cioè sugli insiemi della forma $A \times B$, con $A \in \mathcal{B}(\mathbb{X}), B \in \mathcal{B}(\mathbb{Y})$ (infatti tali insiemi generano la σ-algebra $\mathcal{B}(\mathbb{X} \times \mathbb{Y})$). Dato un piano di trasporto π e due insiemi $A \in \mathcal{B}(\mathbb{X}), B \in \mathcal{B}(\mathbb{Y})$, interpretiamo la quantità $\pi(A \times B)$ come la massa presente in A spostata in B.

Possiamo presentare la formulazione di Kantorovich del problema di trasporto ottimo.

Definizione 1.2.3 (Funzionale costo)**.**
Sia data una mappa $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ boreliana; diremo che c è una funzione costo. Definiamo il funzionale costo $\mathcal{C} : \Gamma(\mu; \nu) \to [0, +\infty]$ tale che

$$\mathcal{C}(\pi) := \int_{\mathbb{X}\times\mathbb{Y}} c(x; y)\ d\pi.$$

Poniamo

$$I(\mu; \nu) := \inf\{\mathcal{C}(\pi) \mid \pi \in \Gamma(\mu; \nu)\}.$$

Osservazione 1.2.4. Innanzitutto notiamo che $I(\mu; \nu)$ è ben definito perchè $\Gamma(\mu; \nu)$ è non vuoto (infatti la probabilità prodotto $\mu \times \nu$ è un piano di trasporto). Precisiamo anche che la formulazione di Kantorovich è più generale di quella di Monge, nel senso che è possibile associare ad una mappa di trasporto ammissibile $T : \mathbb{X} \to \mathbb{Y}$ il piano di trasporto

$$\pi := (Id \times T)_{\#}\mu.$$

La verifica che π sia effettivamente un piano di trasporto è immediata; inoltre, per la formula di cambio di variabili (vedi 1.1.7) vale che

$$\int_{\mathbb{X}\times\mathbb{Y}} c(x; y)\ d\pi = \int_{\mathbb{X}} c(x; T(x))\ d\mu,$$

quindi il funzionale costo definito nel senso di Monge coincide con quello definito nel senso di Kantorovich.

$I(\mu; \nu)$ **è un minimo? Se sì, è raggiunto in un unico punto? É possibile dire qualcosa sull'eventuale punto di minimo? Tale punto di minimo è rappresentato da una mappa di trasporto?**

Nel seguito daremo una risposta a queste domande in un contesto ragionevolmente generale.

1.3 Esistenza dei piani ottimali

Vogliamo mostrare che il problema di Kantorovich ammette soluzione. Utilizzeremo la strategia del metodo diretto: dobbiamo introdurre una nozione di convergenza di misure che renda semicontinuo inferiormente il funzionale costo e compatto il suo spazio di definizione.

1.3.1 Convergenza debole di misure

Definizione 1.3.1 (Convergenza debole)**.**
Siano $\mathbb{Z}$ uno spazio topologico, $(\mu_n)_n$ una successione di misure in $\mathscr{M}(\mathbb{Z})$ e μ_∞ una misura in $\mathscr{M}(\mathbb{Z})$. Si dice che $(\mu_n)_n$ converge debolmente a μ_∞ se per ogni funzione test $\varphi \in C_b(\mathbb{Z})$ vale che

$$\lim_{n\to+\infty} \int_{\mathbb{Z}} f\ d\mu_n = \int_{\mathbb{Z}} f\ d\mu_\infty.$$

Scriveremo $\mu_n \rightharpoonup \mu_\infty$.

La definizione 1.3.1 è del tutto generale: mostreremo che quella appena introdotta è la nozione "giusta" di convergenza di misure per studiare l'esistenza del minimo del problema di Kantorovich.

Innanzitutto mostriamo che gli insiemi $\mathscr{M}_+(\mathbb{X})$ e $\mathscr{P}(\mathbb{X})$ sono chiusi per convergenza debole di misure. Questo risultato non è sorprendente, se interpretiamo la convergenza debole di misure come convergenza puntuale dei funzionali definiti dall'integrazione di funzioni continue e limitate rispetto alle misure considerate.

Lemma 1.3.2. *Sia $(\mathbb{Z}; d_Z)$ uno spazio metrico. Gli spazi $\mathscr{M}_+(\mathbb{Z})$ e $\mathscr{P}(\mathbb{Z})$ sono chiusi rispetto alla convergenza debole di misure (vedi 1.3.1).*

Dimostrazione. Siano $(\mu_n)_n$ una successione in $\mathscr{M}_+(\mathbb{Z})$ e $\mu_\infty \in \mathscr{M}(\mathbb{Z})$ misure tali che $\mu_n \rightharpoonup \mu_\infty$. Dobbiamo provare che μ_∞ è non negativa. Sia $A \subseteq \mathbb{Z}$ un aperto. Essendo la funzione $\mathbb{1}_A$ semicontinua inferiormente, per il lemma 1.1.3 esiste una successione $(\varphi)_k \subseteq C_b(\mathbb{Y})$ tale che

- $0 \leq \varphi_k(y) \leq \varphi_{k+1}(y) \leq \mathbb{1}_A(y)$ per ogni $y \in \mathbb{Y}$ per ogni $k \in \mathbb{N}$,
- $\sup_k \varphi_k(y) = \mathbb{1}_A(y)$ per ogni $y \in \mathbb{Y}$.

Per ogni $k \in \mathbb{N}$ vale che

$$\liminf_{n\to+\infty} \mu_n(A) = \liminf_{n\to+\infty} \int_{\mathbb{Z}} \mathbb{1}_A \, d\mu_n \geq \liminf_{n\to+\infty} \int_{\mathbb{Z}} \varphi_k \, d\mu_n = \int_{\mathbb{Z}} \varphi_k \, d\mu_\infty.$$

Sia $\mu_\infty = \mu_\infty^+ - \mu_\infty^-$ la decomposizione di μ_∞ in parte positiva e negativa. Dobbiamo mostrare che la parte negativa è nulla. Ricordiamo che per ogni $k \in \mathbb{N}$ vale che

$$\int_{\mathbb{Z}} \varphi_k \, d\mu_\infty = \int_{\mathbb{Z}} \varphi_k \, d\mu_\infty^+ - \int_{\mathbb{Z}} \varphi_k \, d\mu_\infty^-;$$

poichè la successione $(\varphi_k)_k$ assume valori in $[0,1]$ e le misure $\mu_\infty^+, \mu_\infty^-$ sono finite, per il teorema di convergenza dominata, vale che

$$\begin{aligned}
\liminf_{n\to+\infty} \mu_n(A) &\geq \lim_{k\to+\infty} \int_{\mathbb{Z}} \varphi_k \, d\mu_\infty \\
&= \lim_{k\to+\infty} \left[\int_{\mathbb{Z}} \varphi_k \, d\mu_\infty^+ - \int_{\mathbb{Z}} \varphi_k \, d\mu_\infty^- \right] \\
&= \int_{\mathbb{Z}} \mathbb{1}_A \, d\mu_\infty^+ - \int_{\mathbb{Z}} \mathbb{1}_A \, d\mu_\infty^- \\
&= \mu_\infty(A).
\end{aligned}$$

Dalla convergenza debole di misure, segue anche che

$$\lim_{n\to+\infty} \mu_n(\mathbb{Z}) = \lim_{n\to+\infty} \int_{\mathbb{Z}} 1 \, d\mu_n = \int_{\mathbb{Z}} 1 \, d\mu_\infty = \mu_\infty(\mathbb{Z}).$$

Per ogni chiuso $C \subseteq \mathbb{Z}$ vale che

$$\limsup_{n\to+\infty} \mu_n(C) = \lim_{n\to+\infty} \mu_n(\mathbb{Z}) - \liminf_{n\to+\infty} \mu_n(C^c) \leq \mu_\infty(\mathbb{Z}) - \mu_\infty(C^c) = \mu_\infty(C).$$

In particolare, abbiamo che μ_∞ è non negativa su tutti gli insiemi chiusi. Ricordiamo che esistono due insiemi boreliani disgiunti A, B tali che $A \cup B = \mathbb{Z}$, μ_∞^+ è supportata in A e μ_∞^- è supportata in B. Essendo $(\mathbb{Z}; d_Z)$ uno spazio metrico (in particolare è numerabilmente chiuso, cioè ogni aperto si scrive come unione numerabile di chiusi) e μ_∞^- una misura boreliana non negativa, vale che

$$\mu_\infty^-(B) = \sup\{\mu_\infty^-(C) \mid C \subseteq B,\ C \text{ chiuso}\}.$$

Siccome μ_∞^+ è concentrata in B^c, si ha che

$$\begin{aligned} 0 &\geq -\mu_\infty^-(B) \\ &= -\sup\{\mu_\infty^-(C) \mid C \subseteq B,\ C \text{ chiuso}\} \\ &= -\sup\{-\mu_\infty(C) \mid C \subseteq B,\ C \text{ chiuso}\} \\ &= \inf\{\mu_\infty(C) \mid C \subseteq B,\ C \text{ chiuso}\} \\ &\geq 0, \end{aligned}$$

essendo μ_∞ positiva su tutti i chiusi. Allora vale che $\mu_\infty^-(B) = 0$, da cui segue che $\mu_\infty^-(\mathbb{X}) = 0$. Concludiamo che μ_∞ è una misura non negativa.

Se $(\mu_n)_n$ è una successione di misure in $\mathscr{P}(\mathbb{X})$ che converge debolmente a μ_∞, allora μ_∞ è non negativa; inoltre abbiamo anche provato che

$$1 = \lim_{n\to+\infty} \mu_n(\mathbb{X}) = \mu_\infty(\mathbb{X}).$$

□

1.3.2 Semicontinuità inferiore

Siano dati due spazi metrici $\mathbb{X}, \mathbb{Y}$, due misure di probabilità $\mu \in \mathscr{P}(\mathbb{X}), \nu \in \mathscr{P}(\mathbb{Y})$ e una funzione costo $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$. Nel seguito, doteremo $\mathbb{X} \times \mathbb{Y}$ della distanza $d_{X\times Y} : (\mathbb{X} \times \mathbb{Y}) \times (\mathbb{X} \times \mathbb{Y}) \to [0, +\infty)$ tale che

$$d_{X\times Y}((x; y); (x'; y')) := d_X(x; x') + d_Y(y; y').$$

Lemma 1.3.3. *L'insieme $\Gamma(\mu; \nu)$ definito in 1.2.1 è chiuso in $\mathscr{M}(\mathbb{X} \times \mathbb{Y})$ rispetto alla convergenza debole di misure.*

Dimostrazione. Sia $(\pi_n)_n$ una successione in $\Gamma(\mu; \nu)$ che converge debolmente ad una misura π_∞ in $\mathscr{M}(\mathbb{X} \times \mathbb{Y})$. Abbiamo provato nel lemma 1.3.2 che π_∞ è una misura di probabilità. Dobbiamo mostrare che $(p_X)_\#\pi = \mu$ e $(p_Y)_\#\pi = \nu$. Notiamo che possiamo interpretare ogni funzione test $\varphi \in C_b(\mathbb{X})$ come una funzione $\tilde{\varphi} \in C_b(\mathbb{X} \times \mathbb{Y})$ indipendente dalla seconda variabile, ovvero $\tilde{\varphi} = \varphi \circ p_X$. In tal caso, vale che

$$\begin{aligned} \int_{\mathbb{X}\times\mathbb{Y}} \varphi \circ p_X \; d\pi_\infty &= \int_{\mathbb{X}\times\mathbb{Y}} \tilde{\varphi} \; d\pi_\infty \\ &= \lim_{n\to+\infty} \int_{\mathbb{X}\times\mathbb{Y}} \tilde{\varphi} \; d\pi_n \\ &= \lim_{n\to+\infty} \nu(\mathbb{Y}) \int_{\mathbb{X}} \varphi(x) \; d\mu \\ &= \int_{\mathbb{X}} \varphi \; d\mu. \end{aligned}$$

Utilizzando la proposizione 1.1.8, deduciamo che $\mu = (p_X)_\#\pi_\infty$. In maniera del tutto analoga, possiamo mostrare che l'altra marginale di π_∞ è ν. □

Possiamo finalmente enunciare e dimostrare il teorema di semicontinuità inferiore del funzionale costo sotto ipotesi molto generali.

Teorema 1.3.4 (Semicontinuità inferiore)**.**
Supponiamo che la funzione costo c sia semicontinua inferiormente. Sia $\mathcal{C} : \Gamma(\mu;\nu) \to [0,+\infty]$ il funzionale costo definito in 1.2.3. $\mathcal{C}$ è semicontinuo inferiormente rispetto alla convergenza debole di misure.

Dimostrazione. Sia $(\pi_n)_n$ una successione di misure in $\Gamma(\mu;\nu)$ che converge debolmente ad una misura π_∞. Dobbiamo mostrare che

$$\mathcal{C}(\pi_\infty) \leq \liminf_{n\to+\infty} \mathcal{C}(\pi_n).$$

Abbiamo provato nel lemma 1.3.3 che $\pi_\infty \in \Gamma(\mu;\nu)$. Per il lemma 1.1.3 esiste una successione $(c_k)_k$ di funzioni lipschitziane e limitate tali che per ogni $(x;y) \in \mathbb{X}\times\mathbb{Y}$ vale che

- $0 \leq c_k(x;y) \leq c_{k+1}(x;y) \leq c(x;y)$ per ogni $k \in \mathbb{N}$,
- $\sup_{k\in\mathbb{N}} c_k(x;y) = c(x;y)$.

Per ogni $k \in \mathbb{N}$, dalla convergenza debole delle misure (c_k una funzione test ammissibile) segue che

$$\liminf_{n\to+\infty} \mathcal{C}(\pi_n) = \geq \liminf_{n\to+\infty} \int_{\mathbb{X}\times\mathbb{Y}} c_k \; d\pi_n = \int_{\mathbb{X}\times\mathbb{Y}} c_k \; d\pi_\infty.$$

Passando all'estremo superiore in k con il teorema di Beppo Levi, troviamo che

$$\liminf_{n\to+\infty} \mathcal{C}(\pi_n) \geq \sup_{k\in\mathbb{N}} \int_{\mathbb{X}\times\mathbb{Y}} c_k \; d\pi_\infty = \int_{\mathbb{X}\times\mathbb{Y}} c \; d\pi_\infty = \mathcal{C}(\pi_\infty).$$

□

1.3.3 Compattezza

Abbiamo ottenuto un risultato di semicontinuità inferiore del funzionale costo assumendo soltanto che $\mathbb{X}, \mathbb{Y}$ siano spazi metrici; per ottenere un teorema di compattezza, invece, dobbiamo supporre che siano spazi polacchi (vedi 1.1.1).

Dobbiamo provare un risultato di compattezza nello spazio delle misure non negative. Iniziamo analizzando il caso in cui le misure siano definite su uno spazio compatto: in questo caso preliminare, gli ingredienti fondamentali sono il teorema di Banach-Alaoglu e il teorema di rappresentazione di Riesz.

Lemma 1.3.5. *Sia $(\mathbb{Z}; d_Z)$ uno spazio metrico compatto e $(\mu_n)_n$ una successione in $\mathscr{M}(\mathbb{Z})$ tale che*

$$\sup_{n\in\mathbb{N}} |\mu_n|(\mathbb{Z}) < +\infty.$$

Esiste una sottosuccessione (non rinominata) e una misura $\mu_\infty \in \mathscr{M}(\mathbb{Z})$ tale che $\mu_n \rightharpoonup \mu_\infty$ nel senso della definizione 1.3.1.

Dimostrazione. Se $(\mathbb{Z}; d_Z)$ è uno spazio metrico compatto, per il teorema di rappresentazione di Riesz, lo spazio $\mathscr{M}(\mathbb{Z})$ dotato della norma $\|\mu\| := |\mu|(\mathbb{Z})$ è canonicamente isometrico al duale di $C(\mathbb{Z})$. Per la precisione, la mappa $J : \mathscr{M}(\mathbb{Z}) \to C(\mathbb{Z})'$ tale che per ogni $\mu \in \mathscr{M}(\mathbb{Z})$ per ogni $\varphi \in C(\mathbb{Z})$ vale che

$$J_\mu(\varphi) := \int_{\mathbb{Z}} \varphi \, d\mu$$

è un'isometria lineare bigettiva. Notiamo che, in questo caso, la nozione di convergenza debole in $\mathscr{M}(\mathbb{Z})$ definita in 1.3.1 corrisponde alla convergenza debole * in $C(\mathbb{Z})'$. La successione $(\mu_n)_n$ è limitata in $\mathscr{M}(\mathbb{Z})$, allora la successione $(J_{\mu_n})_n$ è limitata in $C(\mathbb{Z})'$; per il teorema di Banach-Alaoglu esistono una sottosuccessione (non rinominata) ed un funzionale $L \in C(\mathbb{Z})'$ tale che $J_{\mu_n} \overset{*}{\rightharpoonup} L$ in $C(\mathbb{Z})'$ (precisiamo che $\mathbb{Z}$ è totalmente limitato, quindi è separabile; segue che $C(\mathbb{Z})$ è separabile e allora si può applicare il teorema di Banach-Alaoglu in versione sequenziale). Essendo J surgettiva, esiste μ_∞ tale che $L = J_{\mu_\infty}$. Per definizione di convergenza debole *, per ogni $\varphi \in C(\mathbb{Z})$ vale che

$$\lim_{n\to+\infty} \int_{\mathbb{Z}} \varphi \, d\mu_n = \int_{\mathbb{Z}} \varphi \, d\mu_\infty.$$

In altri termini, $\mu_n \rightharpoonup \mu_\infty$ nel senso della definizione 1.3.1. □

Se lo spazio di ambientazione non è compatto, abbiamo bisogno di un risultato più potente. Possiamo dare la seguente definizione in un contesto del tutto generale.

Definizione 1.3.6 (Equi-tightness)**.**
Siano $(\mathbb{Z}; \mathcal{T})$ uno spazio topologico e $\mathscr{F} \subseteq \mathscr{M}_+(\mathbb{Z})$ una famiglia di misure. Si dice che $\mathscr{F}$ è equi-tight ("equi-tesa") se per ogni $\varepsilon > 0$ esiste un compatto K contenuto in $\mathbb{Z}$ tale che per ogni $\mu \in \mathscr{F}$ vale che $\mu(\mathbb{Z} \setminus K) < \varepsilon$.

Teorema 1.3.7 (Prokhorov)**.**
Siano $(\mathbb{Z}; d_Z)$ uno spazio polacco e $\mathscr{F}$ una famiglia in $\mathscr{P}(\mathbb{Z})$. Sono fatti equivalenti:

- *$\mathscr{F}$ è equi-tight (vedi 1.3.6);*
- *$\mathscr{F}$ è relativamente compatta rispetto alla convergenza debole.*

Dimostrazione. Mostriamo soltanto che la prima condizione implica la seconda (nel seguito utilizzeremo solo questo risultato). Sia $(\mu_n)_n$ una successione contenuta in $\mathscr{F}$; dobbiamo mostrare che esistono una sottosuccessione (non rinominata) e $\mu_\infty \in \mathscr{P}(\mathbb{Z})$ tale che $\mu_n \rightharpoonup \mu_\infty$ (basta mostrare che $\mu_\infty \in \mathscr{M}(\mathbb{Z})$ e utilizzare il lemma 1.3.2). Essendo la famiglia $\mathscr{F}$ equi-tight, esiste una successione $(K_k)_k$ non decrescente di sottospazi compatti, tali che, detto

$$\omega_k := \sup_{\mu \in \mathscr{F}} \mu(\mathbb{Z} \setminus K_k),$$

vale che $\omega_k \to 0$. Per ogni $n, k \in \mathbb{N}$ denotiamo con $\nu_n^k := \mu_n \llcorner K_k$. Fissato $k \in \mathbb{N}$, possiamo interpretare $(\nu_n^k)_n$ come una successione limitata di misure in $\mathscr{M}_+(K_k)$. Per il lemma 1.3.5, esistono una sottosuccessione $(\nu_{n_h}^k)_h$ ed una misura ν^k tali che $\nu_{n_h}^k \rightharpoonup \nu^k$ debolmente in $\mathscr{M}(K_k)$. Per il lemma 1.3.2, $\nu^k \in \mathscr{M}_+(K_k)$. Inoltre, vale che

$$\nu^k(K_k) = \lim_{h\to+\infty} \nu_{n_h}^k(K_k) = \lim_{h\to+\infty} \mu_{n_h}(K_k) = 1 - \lim_{h\to+\infty} \mu_{n_h}(\mathbb{Z} \setminus K_k) \geq 1 - \omega_k.$$

Dalla stessa stima, segue anche che

$$\nu^k(K_k) \leq 1.$$

Possiamo interpretare ν^k come una misura in $\mathscr{M}_+(\mathbb{Z})$ concentrata in K_k. Scrivendo esplicitamente gli integrali, è ovvio che $\nu^k_{n_h} \rightharpoonup \nu^k$ debolmente in $\mathscr{M}(\mathbb{Z})$.

Utilizzando un procedimento diagonale, possiamo trovare una sottosuccessione $(\mu_{n_p})_p$ ed una famiglia di misure $(\nu^k)_k \subseteq \mathscr{M}_+(\mathbb{Z})$ tale che per ogni $k \in \mathbb{N}$ valgono le seguenti proprietà:

- ν^k è concentrata in K_k;
- $\mu_{n_p} \llcorner K_k \rightharpoonup \nu^k$ debolmente in $\mathscr{M}(\mathbb{Z})$ se $p \to +\infty$;
- $1 - \omega_k \leq \nu^k(K_k) \leq 1$.

Fissiamo $k \in \mathbb{N}$; per ogni $p \in \mathbb{N}$ notiamo che $\mu_{n_p} \llcorner (K_{k+1} \setminus K_k)$ è una misura in $\mathscr{M}_+(\mathbb{Z})$ e coincide con $\mu_{n_p} \llcorner K_{k+1} - \mu_{n_p} \llcorner K_k$. Per differenza, deduciamo che $\mu_{n_p} \llcorner (K_{k+1} \setminus K_k) \rightharpoonup \nu^{k+1} - \nu^k$ in $\mathscr{M}(\mathbb{Z})$; essendo $\mathscr{M}_+(\mathbb{Z})$ chiuso rispetto alla convergenza debole, deduciamo che $\nu^{k+1} - \nu^k$ è una misura non negativa. In altri termini, abbiamo provato che, per ogni $B \in \mathcal{B}(\mathbb{Z})$, la successione $(\nu^k(B))_k$ è crescente (magari non strettamente). Definiamo $\nu : \mathcal{B}(\mathbb{Z}) \to [0, +\infty]$ tale che

$$\nu(B) := \sup_{k \in \mathbb{N}} \nu^k(B).$$

Innanzitutto, notiamo che ν ha valori in $[0, 1]$. Vogliamo provare che è una misura σ-additiva: è banale osservare che ν è finitamente additiva sui disgiunti ed è monotona; inoltre, è facile verificare che ν è σ-subadditiva, cioè che se $(A_n)_n$ è una successione qualsiasi di boreliani vale che

$$\nu\left(\bigcup_n A_n\right) \leq \sum_n \nu(A_n).$$

Da queste proprietà segue facilmente che ν è una misura σ-additiva.

Per concludere, mostriamo che $\mu_{n_p} \rightharpoonup \nu$ debolmente in $\mathscr{M}(\mathbb{Z})$. Sia $\varphi \in \mathcal{C}_b(\mathbb{Z})$ una funzione test; dobbiamo provare che

$$\limsup_{p \to +\infty} \left| \int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d\nu \right| = 0.$$

Fissiamo $p, k \in \mathbb{N}$. Vale che

$$\begin{aligned} \left| \int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d\nu \right| \leq & \left| \int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d(\mu_{n_p} \llcorner K_k) \right| \\ & + \left| \int_{\mathbb{Z}} \varphi \, d(\mu_{n_p} \llcorner K_k) - \int_{\mathbb{Z}} \varphi \, d\nu^k \right| \\ & + \left| \int_{\mathbb{Z}} \varphi \, d\nu^k - \int_{\mathbb{Z}} \varphi \, d\nu \right|; \end{aligned}$$

Stimiamo separatamente ciascun addendo. Per il primo termine, vale che

$$\begin{aligned}\left|\int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d(\mu_{n_p} \llcorner K_k)\right| &= \left|\int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{K_k} \varphi \, d\mu_{n_p}\right| \\ &\leq \int_{\mathbb{Z}\setminus K_k} |\varphi| \, \, d\mu_{n_p} \\ &\leq \omega_k \sup_{\mathbb{Z}} |\varphi| \, .\end{aligned}$$

Per il terzo termine, ricordando che $\nu - \nu^k \in \mathscr{M}_+(\mathbb{Z})$, vale che

$$\begin{aligned}\left|\int_{\mathbb{Z}} \varphi \, d\nu - \int_{\mathbb{Z}} \varphi \, d\nu^k\right| &= \left|\int_{\mathbb{Z}} \varphi \, d(\nu - \nu^k)\right| \\ &\leq \int_{\mathbb{Z}} |\varphi| \, \, d(\nu - \nu^k) \\ &\leq \sup_{\mathbb{Z}} |\varphi| \, (\nu(\mathbb{Z}) \setminus \nu^k(\mathbb{Z})) \\ &\leq \omega_k \sup_{\mathbb{Z}} |\varphi| \, .\end{aligned}$$

Per il secondo addendo, fissato $k \in \mathbb{N}$, dalla convergenza debole in $\mathscr{M}_+(\mathbb{Z})$, deduciamo che

$$\limsup_{p \to +\infty} \left|\int_{\mathbb{Z}} \varphi \, d(\mu_{n_p} \llcorner K_k) - \int_{\mathbb{Z}} \varphi \, d\nu^k\right| = 0.$$

Allora, per ogni $k \in \mathbb{N}$ vale che

$$\limsup_{p \to +\infty} \left|\int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d\nu\right| \leq 2\omega_k \sup_{\mathbb{Z}} |\varphi| \, .$$

Passando al limite in k e ricordando che $\omega_k \to 0$, troviamo che

$$\limsup_{p \to +\infty} \left|\int_{\mathbb{Z}} \varphi \, d\mu_{n_p} - \int_{\mathbb{Z}} \varphi \, d\nu\right| \leq 0.$$

□

Osservazione 1.3.8. Precisiamo che nella dimostrazione data del teorema di Prokhorov (vedi 1.3.7) non abbiamo mai utilizzato la separabilità e la completezza dello spazio: infatti, queste condizioni non sono necessarie per mostrare l'implicazione che ci interessa; quindi, questa è valida sotto la sola ipotesi che $(\mathbb{Z}; d_Z)$ sia uno spazio metrico. Per essere precisi, abbiamo soltanto utilizzato l'esistenza di una famiglia crescente di insiemi compatti $(K_k)_k$ che invade $\mathbb{Z}$ in misura uniformemente in $\mathscr{F}$, cioè tale che

$$\limsup_{k \to +\infty} \left(\sup_{\mu \in \mathscr{F}} \mu(\mathbb{Z} \setminus K_k)\right) = 0;$$

l'esistenza di tale successione è data dall'assunzione che $\mathscr{F}$ si equi-tight (vedi 1.3.6) e non dalle proprietà topologiche dello spazio.

La compattezza e la completezza, invece, sono essenziali per mostrare l'altra implicazione (che non abbiamo esaminato perchè non ne faremo uso nel seguito); in questo caso, infatti, abbiamo bisogno di scrivere $\mathbb{Z}$ come unione crescente di compatti $(K_k)_k$ dal punto di vista insiemistico, cioè

$$\mathbb{Z} = \bigcup_k K_k.$$

L'esistenza di un'esaustione (numerabile) in compatti segue dalle proprietà topologiche dello spazio.

Corollario 1.3.9. *Siano* $(\mathbb{X}; d_X), (\mathbb{Y}, d_Y)$ *spazi polacchi e* $\mu \in \mathscr{P}(\mathbb{X}), \nu \in \mathscr{P}(\mathbb{Y})$. *L'insieme* $\Gamma(\mu;\nu)$ *definito in 1.2.1 è sequenzialmente compatto in* $\mathscr{M}(\mathbb{X} \times \mathbb{Y})$ *rispetto alla convergenza debole.*

Dimostrazione. Abbiamo già mostrato che $\Gamma(\mu;\nu)$ è chiuso in $\mathscr{M}(\mathbb{X} \times \mathbb{Y})$ rispetto alla convergenza debole (vedi 1.3.3); dobbiamo soltanto provare la relativa compattezza. Per il teorema di Prokhorov (vedi 1.3.7), è sufficiente mostrare che $\Gamma(\mu;\nu)$ è una famiglia equi-tight. Fissiamo $\varepsilon > 0$; per il lemma di Ulam (vedi 1.1.2: in questo passaggio è essenziale lavorare in spazi polacchi), esistono dei compatti $K_X \subseteq \mathbb{X}$ e $K_Y \subseteq \mathbb{Y}$ tali che

$$\mu(\mathbb{X} \setminus K_X) < \frac{\varepsilon}{2}, \quad \nu(\mathbb{Y} \setminus K_Y) < \frac{\varepsilon}{2}.$$

Notiamo che $K_X \times K_Y$ è compatto in $\mathbb{X} \times \mathbb{Y}$. Data $\pi \in \Gamma(\mu;\nu)$, notiamo che

$$\begin{aligned}\pi((\mathbb{X} \times \mathbb{Y}) \setminus (K_X \times K_Y)) &\leq \pi((\mathbb{X} \setminus K_X) \times \mathbb{Y}) + \pi(\mathbb{X} \times (\mathbb{Y} \setminus K_Y)) \\ &= \mu(\mathbb{X} \setminus K_X) + \nu(\mathbb{Y} \setminus K_Y) < \varepsilon.\end{aligned}$$

□

1.3.4 Soluzione del problema di Kantorovich

I risultati di compattezza e semicontinuità inferiore mostrati permettono di dimostrare facilmente l'esistenza della soluzione del problema di Kantorovich.

Siano $(\mathbb{X}; d_X), (\mathbb{Y}, d_Y)$ spazi polacchi, $\mu \in \mathscr{P}(\mathbb{X}), \nu \in \mathscr{P}(\mathbb{Y})$ e $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ una funzione semicontinua inferiormente.

Teorema 1.3.10 (Esistenza del minimo)**.**
Il funzionale $\mathcal{C} : \Gamma(\mu;\nu) \to [0, +\infty]$ *definito in 1.2.3 ammette minimo.*

Dimostrazione. Lo spazio $\Gamma(\mu;\nu)$ è debolmente compatto (vedi 1.3.9) e il funzionale $\mathcal{C}$ è debolmente semicontinuo inferiormente in $\Gamma(\mu;\nu)$ (vedi 1.3.4). Allora $\mathcal{C}$ ammette minimo in $\Gamma(\mu;\nu)$ per il teorema di Weierstrass. □

Definizione 1.3.11 (Piani ottimali)**.**
Denotiamo con

$$\Gamma_0(\mu;\nu) := \arg\min\{\mathcal{C}(\pi) \mid \pi \in \Gamma(\mu;\nu)\};$$

diremo che l'insieme $\Gamma_0(\mu;\nu)$ (non vuoto per il teorema 1.3.10) è formato dai piani di trasporto ottimali.

Capitolo 2

Il problema duale

2.1 Formulazione duale

Siano $(\mathbb{X}; d_X), (\mathbb{Y}; d_Y)$ spazi polacchi, $\mu \in \mathscr{P}(\mathbb{X}), \nu \in \mathscr{P}(\mathbb{Y})$ e $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ una funzione costo semicontinua inferiormente.

Definizione 2.1.1 (Insieme di contatto)**.**
Denotiamo con

$$I_c := \{(\Phi; \Psi) \in \mathrm{Lip}_b(\mathbb{X}) \times \mathrm{Lip}_b(\mathbb{Y}) \mid \Phi(x) + \Psi(y) \leq c(x; y)\ \forall (x; y) \in \mathbb{X} \times \mathbb{Y}\}.$$

Date $(\Phi; \Psi) \in I_c$, definiamo

$$\Gamma_{\Phi;\Psi} := \{(x; y) \in \mathbb{X} \times \mathbb{Y} \mid \Phi(x) + \Psi(y) = c(x; y)\}$$

l'insieme di contatto tra $(\Phi; \Psi)$ e c.

Definizione 2.1.2. Introduciamo il funzionale $\mathcal{J} : I_c \to [0, +\infty]$ tale che

$$\mathcal{J}(\Phi; \Psi) := \int_{\mathbb{X}} \Phi \; d\mu + \int_{\mathbb{Y}} \Psi \; d\nu.$$

Osservazione 2.1.3. Notiamo che per ogni $\pi \in \Gamma(\mu; \nu)$ per ogni $(\Phi; \Psi) \in I_c$ vale che

$$\mathcal{J}(\Phi; \Psi) = \int_{\mathbb{X}} \Phi \; d\mu + \int_{\mathbb{Y}} \Psi \; d\nu = \int_{\mathbb{X} \times \mathbb{Y}} [\Phi(x) + \Psi(y)] \; d\pi \leq \int_{\mathbb{X} \times \mathbb{Y}} c \; d\pi = \mathcal{C}(\pi).$$

In particolare, vale che

$$\sup_{(\Phi;\Psi) \in I_c} \mathcal{J}(\Phi; \Psi) \leq \inf_{\pi \in \Gamma(\mu;\nu)} \mathcal{C}(\pi).$$

Mostreremo che, sotto certe ipotesi (blande) sulla funzione c, vale l'uguaglianza. Questo risultato è noto come teorema di Kantorovich-Rubinstein.

2.2 c-dualità

La soluzione del problema duale è elementare, ma non immediata; inoltre è ricca di nuove idee delle quali proveremo a dare un'interpretazione qualitativa.

D'ora in avanti assumeremo sempre che la funzione costo c sia a valori reali.

2.2.1 Definizioni

Dobbiamo introdurre alcune definizioni, che si riveleranno essere quelle "giuste" per affrontare il problema duale.

Definizione 2.2.1 (c-ciclica monotonia)**.**
Un insieme $\Gamma \subseteq \mathbb{X} \times \mathbb{Y}$ è detto c-ciclicamente monotono se per ogni $N \in \mathbb{N}^+$ per ogni permutazione σ di $\{1; \ldots; N\}$ per ogni $(x_1; y_1), \ldots, (x_N; y_N) \in \mathbb{X} \times \mathbb{Y}$ vale che

$$\sum_{i=1}^{N} c(x_i; y_{\sigma(i)}) \geq \sum_{i=1}^{N} c(x_i; y_i).$$

Se la quantità $c(x; y)$ indica il costo dello spostamento di un'unità di massa dal punto $x \in \mathbb{X}$ al punto $y \in \mathbb{Y}$, in un insieme c-ciclicamente monotono Γ ogni coppia di punti $(x; y)$ è stata disposta nella maniera più economica possibile tra tutti i punti di Γ. In un certo senso, la nozione di c-ciclica monotonia di un insieme quantifica il fatto che una certa quantità di massa è stata riarrangiata nella maniera ottimale.

Osservazione 2.2.2. Data $(\Phi; \Psi) \in I_c$, l'insieme di contatto Γ è c-ciclicamente monotono. Infatti, presi $N \in \mathbb{N}^+$, una permutazione σ di $\{1; \ldots; N\}$ e dei punti $(x_1; y_1), \ldots, (x_N; y_N) \in \mathbb{X} \times \mathbb{Y}$ vale che

$$\begin{aligned}\sum_{i=1}^{N} c(x_i; y_{\sigma(i)}) &\geq \sum_{i=1}^{N} \Phi(x_i) + \Psi(y_{\sigma(i)}) \\ &= \sum_{i=1}^{N} \Phi(x_i) + \Psi(y_i) \\ &= \sum_{i=1}^{N} c(x_i; y_i).\end{aligned}$$

Presentiamo una generalizzazione di alcune nozioni classiche di analisi convessa.

Definizione 2.2.3 (c-affinità)**.**
Fissati $x \in \mathbb{X}, \alpha \in \mathbb{R}$, diremo che la funzione $c(x; \cdot) + \alpha : \mathbb{Y} \to \mathbb{R}$ è c-affine nella seconda variabile; fissati $y \in \mathbb{Y}, \beta \in \mathbb{R}$ diremo che la funzione $c(\cdot; y) + \beta : \mathbb{X} \to \mathbb{R}$ e c-affine nella prima variabile.

Definizione 2.2.4 (c-coniugata)**.**
Data una funzione $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ definiamo la funzione $\Phi^c : \mathbb{Y} \to \overline{\mathbb{R}}$ tale che

$$\Phi^c(y) := \inf_{x \in \mathbb{X}} \{c(x; y) - \Phi(x)\}.$$

Diremo che Φ^c è la c-coniugata di Φ. Analogamente, se $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ si definisce $\Psi^c : \mathbb{X} \to \overline{\mathbb{R}}$ la c-coniugata di Ψ come

$$\Psi_c(x) := \inf_{y \in \mathbb{Y}} \{c(x; y) - \Psi(y)\}.$$

Osservazione 2.2.5. Nel contesto della definizione 2.2.4, se Φ non è identicamente $-\infty$, allora Φ^c non assume mai il valore $+\infty$, cioè $\Phi^c : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$. Se esiste $y_0 \in \mathbb{Y}$ e $\alpha \in \mathbb{R}$ tali che $\Phi(x) \geq c(x; y) + \alpha$ per ogni $x \in \mathbb{X}$, allora vale che

$$\Phi^c(y_0) \geq \inf_{x \in \mathbb{X}} \{c(x; y_0) - \Phi(x)\} \geq -\alpha > -\infty.$$

Ovviamente, valgono risultati analoghi per Ψ.

Se Φ non è identicamente $-\infty$, dalla definizione 2.2.4 segue che Φ^c assume valori in $\mathbb{R} \cup \{-\infty\}$ e che è la più grande funzione $\eta : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ compatibile con il vincolo

$$\Phi(x) + \eta(y) \leq c(x; y) \ \forall (x; y) \in \mathbb{X} \times \mathbb{Y}.$$

Quindi, in un certo senso, fissata $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ la funzione Φ^c ottimizza la disuguaglianza del problema duale. Analogamente, se Φ^c non è identicamente $-\infty$, allora Φ^{cc} assume valori in $\mathbb{R} \cup \{-\infty\}$ ed è la più grande funzione $\gamma : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ compatibile con il vincolo

$$\gamma(x) + \Phi^c(y) \leq c(x; y) \ \forall (x; y) \in \mathbb{X} \times \mathbb{Y}.$$

Quindi, abbiamo ancora ottimizzato la disuguaglianza del problema duale. Allora, è lecito chiedersi cosa succede iterando questa costruzione (ammesso che ciò sia possibile). I risultati seguenti mostreranno che il processo è stazionario già alla terza iterazione, cioè $\Phi^c = \Phi^{ccc}$.

Definizione 2.2.6 (c-concavità)**.**
Una funzione $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ è detta c-concava se è estremo inferiore puntuale di una famiglia di funzioni c-affini nella prima variabile (vedi 2.2.3). Analogamente, una funzione $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ è detta c-concava se è estremo inferiore puntuale di una famiglia di funzioni c-affini nella seconda variabile.

Lemma 2.2.7. *Sia $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ una funzione che non vale identicamente $-\infty$. Allora Φ è c-concava se e solo se è la c-coniugata di una funzione $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$.*

Dimostrazione. Supponiamo che Φ sia c-concava; allora esistono due famiglie $\{y_i\}_{i \in I} \subseteq \mathbb{Y}, \{\alpha_i\}_{i \in I} \subseteq \mathbb{R}$ tali che per ogni $x \in \mathbb{X}$ vale

$$\Phi(x) = \inf_{i \in I} \{c(x; y_i) + \alpha_i\}.$$

Per ogni $y \in \mathbb{Y}$ denotiamo

$$I_y := \{i \in I \mid y_i = y\};$$

definiamo $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ tale che

$$\Psi(y) := \begin{cases} -\inf\limits_{i \in I_y} \alpha_i & \text{se } I_y \neq \emptyset, \\ -\infty & \text{se } I_y = \emptyset. \end{cases}$$

Notiamo che Ψ è ben definita, cioè non assume mai il valore $+\infty$. Sia $y \in \mathbb{Y}$ tale che $I_y \neq \emptyset$; essendo Φ non identicamente $-\infty$, esiste $x \in \mathbb{X}$ tale che

$$\Phi(x) = \inf_{i \in I} \{c(x; y_i) + \alpha_i\} \geq K > -\infty.$$

Deduciamo che

$$K \leq \inf_{i \in I_y} \{c(x; y_i) + \alpha_i\} = c(x; y) + \inf_{i \in I_y} \alpha_i;$$

segue che

$$-\inf_{i \in I_y} \alpha_i \leq c(x; y) - K < +\infty.$$

Verifichiamo che $\Phi = \Psi^c$. Notiamo che $\{I_y\}_{y\in\mathbb{Y}}$ forma una partizione di I; allora, per ogni $x \in \mathbb{X}$ abbiamo che

$$\begin{aligned}\Psi^c(x) &= \inf_{y\in\mathbb{Y}}\{c(x;y) - \Psi(y)\} \\ &= \inf_{y\in\mathbb{Y},\ I_y\neq\emptyset}\left\{c(x;y) + \inf_{i\in I_y}\{\alpha_i\}\right\} \\ &= \inf_{i\in I}\{c(x;y_i) + \alpha_i\} \\ &= \Phi(x).\end{aligned}$$

Viceversa, supponiamo che esista $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ tale che $\Phi = \Psi^c$. Per definizione, si ha

$$\Phi(x) = \inf_{y\in\mathbb{Y}}\{c(x;y) - \Psi(y)\}.$$

Innanzitutto, osserviamo che Ψ non è identicamente $-\infty$, perchè Φ non assume mai il valore $+\infty$. Notiamo che vale

$$\Phi(x) = \inf_{y\in\mathbb{Y},\ \Psi(y)\neq-\infty}\{c(x;y) - \Psi(y)\};$$

questo basta a concludere, perchè abbiamo scritto Φ come estremo inferiore puntuale di funzioni c-affini nella prima variabile. □

Teorema 2.2.8. *Sia $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ una funzione che non vale identicamente $-\infty$. Allora Φ^{cc} è ben definita e vale $\Phi^{cc} \geq \Phi$. Inoltre l'uguaglianza vale l'uguale per ogni $x \in \mathbb{X}$ se e solo se Φ è c-concava.*

Dimostrazione. Essendo Φ non identicamente $-\infty$, notiamo che Φ^c non assume mai il valore $+\infty$; quindi possiamo ben definire $\Phi^{cc} : \mathbb{X} \to \overline{\mathbb{R}}$ (vedi 2.2.4). Notiamo che per ogni $y \in \mathbb{Y}$ vale che

$$\begin{aligned}c(x;y) - \Phi^c(y) &= c(x;y) - \inf_{x'\in\mathbb{X}}\{c(x';y) - \Phi(x')\} \\ &\geq c(x;y) - (c(x;y) - \Phi(x)) \\ &= \Phi(x).\end{aligned}$$

Passando all'estremo inferiore su $y \in \mathbb{Y}$, troviamo che $\Phi^{cc}(x) \geq \Phi(x)$.

Supponiamo che valga l'uguaglianza $\Phi^{cc} = \Phi$; allora Φ è la c-coniugata di Φ^c e quindi è c-concava per il lemma 2.2.7. Viceversa, supponiamo che Φ sia c-concava; per il lemma 2.2.7 è equivalente assumere che esista una funzione $\Psi : \mathbb{Y} \to \mathbb{R} \cup \{-\infty\}$ tale che $\Phi = \Psi^c$. Allora abbiamo che $\Phi^c = \Psi^{cc} \geq \Psi$. Dalla definizione di c-coniugata, segue banalmente che $\Phi^{cc} \leq \Psi^c = \Phi$. Essendo l'altra disuguaglianza sempre verificata, deduciamo che $\Phi = \Phi^{cc}$. □

Definizione 2.2.9 (c-sopradifferenziale)**.**
Dati $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$, definiamo il c-sopradifferenziale di Φ in x come l'insieme

$$\partial^c\Phi(x) := \{y \in \mathbb{Y} \mid \Phi(x') \leq \Phi(x) + c(x';y) - c(x;y)\ \forall x' \in \mathbb{X}\}.$$

Diremo che la coppia $(x;y) \in \operatorname{Graph}(\partial^c\Phi)$ se $y \in \partial^c\Phi(x)$.

Lemma 2.2.10. *Sia $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ una funzione non identicamente $-\infty$. Per ogni $x \in \mathbb{X}$ tale che $\partial^c\Phi(x) \neq \emptyset$, vale che $\Phi(x) > -\infty$.*

Dimostrazione. Supponiamo che esista $x \in \mathbb{X}$ tale che $\Phi(x) = -\infty$ ed esista $\partial^c\Phi(x) \neq \emptyset$; allora esiste $y \in \mathbb{Y}$ tale che per ogni $x' \in \mathbb{X}$ vale che

$$\Phi(x') \leq \Phi(x) + c(x'; y) - c(x; y) = -\infty.$$

Questo è assurdo, perchè stiamo assumendo che Φ non sia identicamente $-\infty$. □

Lemma 2.2.11. *Sia data una funzione $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ non identicamente $-\infty$. Sia $x \in \mathbb{X}$ tale che $\partial^c\Phi(x) \neq \emptyset$. Si ha che $y \in \partial^c\Phi(x)$ se e solo se $\Phi(x) + \Phi^c(y) = c(x; y)$.*

Dimostrazione. Innanzitutto osserviamo che, per il lemma 2.2.10, $\Phi(x) > -\infty$. Ricordiamo anche che Φ^c assume valori in $\mathbb{R} \cup \{-\infty\}$; in particolare, la somma $\Phi(x) + \Phi^c(y)$ è ben definita nell'algebra di $\overline{\mathbb{R}}$. Per definizione, si ha che $y \in \partial^c\Phi(x)$ se e solo se per ogni $x' \in \mathbb{X}$ vale che

$$\Phi(x') \leq \Phi(x) + c(x'; y) - c(x; y);$$

questo è equivalente a richiedere che

$$c(x; y) - \Phi(x) \leq \inf_{x' \in \mathbb{X}} \{c(x; y) - \Phi(x')\} = \Phi^c(y),$$

ovvero che

$$c(x; y) \leq \Phi(x) + \Phi^c(y).$$

Per concludere, notiamo che la disuguaglianza opposta è sempre verificata. □

2.2.2 Supporto di un piano ottimale

Lemma 2.2.12. *Siano $\mathbb{X}_1, \ldots, \mathbb{X}_k$ spazi metrici e $\gamma_1 \in \mathscr{P}(\mathbb{X}_1), \ldots, \gamma_k \in \mathscr{P}(\mathbb{X}_k)$. Esiste uno spazio di probabilità $(\Omega; \mathcal{A}; \mathbb{P})$ e delle mappe misurabili $T_i : \Omega \to \mathbb{X}_i$ tali che $(T_i)_\#\mathbb{P} = \gamma_i$ per ogni $i \in \{1; \ldots; n\}$.*

Dimostrazione. La costruzione è elementare: basta prendere

$$\Omega := \prod_{i=1}^{k} \mathbb{X}_i, \ \mathcal{A} := \bigotimes_{i=1}^{k} \mathcal{B}(\mathbb{X}_i), \ \mathbb{P} := \bigotimes_{i=1}^{k} \gamma_i,$$

cioè lo spazio prodotto, dotato della σ-algebra prodotto e della probabilità prodotto. La mappa $T_i : \Omega \to \mathbb{X}_i$ data dalla proiezione sull'i-esima coordinata soddisfa le richieste. □

Il seguente teorema è fondamentale perchè lega l'ottimalità di un piano di trasporto con la c-ciclica monotonia del suo supporto, che è una proprietà "geometrica". Se interpretiamo la nozione di c-ciclica monotonia di un insieme come ottimalità della disposizione di tutti i punti di quell'insieme (rispetto al costo dato dallo spostamento), non è sorprendente che un piano di trasporto ottimale sia supportato in un insieme che gode di questa proprietà. La dimostrazione è molto esplicativa e si basa sull'idea che, se il supporto di un piano di trasporto non è c-ciclicamente monotono, cioè la massa non è stata spostata nella maniera più economica possibile, allora è possibile muoverla diversamente per ottenere la stessa disposizione ad un costo globalmente inferiore.

Teorema 2.2.13. *Supponiamo che $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty)$ sia continua. Sia $\pi \in \Gamma_0(\mu; \nu)$ (π è un piano di trasporto ottimale) tale che*

$$\int_{\mathbb{X} \times \mathbb{Y}} c \, d\pi < +\infty.$$

Allora il supporto di π è c-ciclicamente monotono.

Dimostrazione. Supponiamo per assurdo che esistano $N \in \mathbb{N}^+$, una permutazione σ di $\{1; \dots; N\}$ e dei punti $(x_1; y_1), \dots (x_N; y_N)$ in $\mathrm{supp}(\pi)$ tali che

$$\sum_{i=1}^{N} c(x_i; y_{\sigma(i)}) < \sum_{i=1}^{N} c(x_i; y_i).$$

Dalla continuità di c segue che per ogni $i \in \{1; \dots; N\}$ esistono un intorno U_i di x_i in $\mathbb{X}$ e un intorno V_i di y_i in $\mathbb{Y}$ tale che per ogni $x'_i \in U_i$ per ogni $y'_i \in V_i$ vale che

$$\sum_{i=1}^{N} c(x'_i; y'_{\sigma(i)}) < \sum_{i=1}^{N} c(x'_i; y'_i).$$

Vogliamo costruire una misura $\theta \in \mathscr{M}(\mathbb{X} \times \mathbb{Y})$ con le seguenti proprietà:

- $(p_X)_\#\theta = 0$ e $(p_Y)_\#\theta = 0$;
- detta θ^- la parte negativa di θ, si ha $\theta^- \leq \pi$;
- c è integrabile rispetto a θ e vale che
$$\int_{\mathbb{X}\times\mathbb{Y}} c \, d\theta < 0.$$

Se costruiamo una misura θ con queste proprietà, è immediato trovare l'assurdo: infatti, dalle prime due proprietà segue immediatamente che $\theta + \pi \in \Gamma(\mu; \nu)$; dalla terza proprietà, segue che

$$\int_{\mathbb{X}\times\mathbb{Y}} c \, d(\theta + \pi) = \int_{\mathbb{X}\times\mathbb{Y}} c \, d\theta + \int_{\mathbb{X}\times\mathbb{Y}} c \, d\pi < \int_{\mathbb{X}\times\mathbb{Y}} c \, d\pi.$$

Questo è assurdo perchè π è un piano di trasporto ottimale in $\Gamma(\mu; \nu)$.

Per costruire tale misura, definiamo

$$z_i := \pi(U_i \times V_i).$$

Siccome $(x_i; y_i) \in \mathrm{supp}(\pi)$, vale per definizione che $z_i > 0$. Poniamo

$$\pi_i := \frac{\pi \llcorner (U_i \times V_i)}{z_i}$$

e notiamo che $\pi_i \in \mathscr{P}(\mathbb{X} \times \mathbb{Y})$. Per il lemma 2.2.12, esiste uno spazio di probabilità $(\Omega; \mathcal{A}; \mathbb{P})$ e delle mappe misurabili $T_i : \Omega \to U_i \times V_i$ tali che $\pi_i = (T_i)_\#\mathbb{P}$. Componendo T_i con le proiezioni coordinate, si trova immediatamente che esistono delle mappe $X_i : \Omega \to U_i, Y_i : \Omega \to Y_i$ misurabili tali che $T_i = X_i \times Y_i$. Notiamo anche che $X_i \times Y_{\sigma(i)} : \Omega \to U_i \times V_i$ è misurabile. Possiamo definire

$$\theta := \sum_{i=1}^{N} (X_i \times Y_{\sigma(i)})_\#\mathbb{P} - \sum_{i=1}^{N} (X_i \times Y_i)_\#\mathbb{P}.$$

É ovvio che $\theta \in \mathscr{M}(\mathbb{X} \times \mathbb{Y})$; mostriamo che soddisfa le tre proprietà richieste.

- Mostriamo la prima proprietà. Vale che

$$
\begin{aligned}
(p_X)_\#\theta &= \sum_{i=1}^N [p_X \circ (X_i \times Y_{\sigma(i)})]_\#\mathbb{P} - \sum_{i=1}^N [p_X \circ (X_i \times Y_i)]_\#\mathbb{P} \\
&= \sum_{i=1}^N (X_i)_\#\mathbb{P} - \sum_{i=1}^N (X_i)_\#\mathbb{P} = 0;
\end{aligned}
$$

$$
\begin{aligned}
(p_Y)_\#\theta &= \sum_{i=1}^N [p_Y \circ (X_i \times Y_{\sigma(i)})]_\#\mathbb{P} - \sum_{i=1}^N [p_Y \circ (X_i \times Y_i)]_\#\mathbb{P} \\
&= \sum_{i=1}^N (Y_{\sigma(i)})_\#\mathbb{P} - \sum_{i=1}^N (Y_i)_\#\mathbb{P} = 0.
\end{aligned}
$$

- Mostriamo la terza proprietà. Innanzitutto verifichiamo che c è integrabile rispetto a θ. Osserviamo che, detta $\theta = \theta^+ - \theta^-$ l'unica decomposizione di θ in parte positiva e parte negativa, si ha (per unicità) che

$$
\theta^+ = \sum_{i=1}^N (X_i \times Y_{\sigma(i)})_\#\mathbb{P}, \;\; \theta^- = \sum_{i=1}^N (X_i \times Y_i)_\#\mathbb{P}.
$$

Essendo c continua, a meno di restringere gli intorni, possiamo supporre che sia limitata in $U_i \times V_i$ e in $U_i \times V_{\sigma(i)}$. Allora vale che

$$
\int_{\mathbb{X}\times\mathbb{Y}} c \; d\theta^+ = \sum_{i=1}^N \int_{U_i \times V_{\sigma(i)}} c \; d(X_i \times Y_{\sigma(i)})_\#\mathbb{P} < +\infty,
$$

perchè $(X_i \times Y_{\sigma(i)})_\#\mathbb{P}$ è una misura di probabilità su $U_i \times V_{\sigma(i)}$. Analogamente, possiamo verificare che

$$
\int_{\mathbb{X}\times\mathbb{Y}} c \; d\theta^- < +\infty.
$$

Utilizzando la formula di cambio di variabili (vedi 1.1.7), troviamo che

$$
\begin{aligned}
\int_{\mathbb{X}\times\mathbb{Y}} c \; d\theta &= \int_{\mathbb{X}\times\mathbb{Y}} c \; d\theta^+ - \int_{\mathbb{X}\times\mathbb{Y}} c \; d\theta^- \\
&= \sum_{i=1}^N \int_\Omega c(X_i; Y_{\sigma(i)}) \; d\mathbb{P} - \sum_{i=1}^N \int_\Omega c(X_i; Y_i) \; d\mathbb{P} \\
&= \int_\Omega \left[\sum_{i=1}^N c(X_i; Y_{\sigma(i)}) - \sum_{i=1}^N c(X_i; Y_i) \right] \; d\mathbb{P} < 0 :
\end{aligned}
$$

infatti, per la costruzione effettuata, per ogni $\omega \in \Omega$ l'integranda è strettamente negativa e $\mathbb{P}$ è una misura di probabilità su Ω.

- Esaminiamo la seconda condizione. Notiamo che

$$
\theta^- = \sum_{i=1}^N (X_i \times Y_i)_\#\mathbb{P} = \sum_{i=1}^N \pi_i \le \frac{N}{\min_i z_i}\pi,
$$

da cui segue che

$$\frac{\min_i z_i}{N}\theta^- \leq \pi.$$

Se poniamo

$$t := \frac{\min_i z_i}{N},$$

osserviamo che $t > 0$ e le due proprietà già verificate (cioè la prima e la terza elencate in precedenza) sono ancora soddisfatte dalla misura $t\theta$. Allora è sufficiente sostituire la misura θ con $t\theta$, che verifica anche la seconda condizione.

□

Il teorema seguente mostra una proprietà particolarmente significativa degli insiemi c-ciclicamente monotoni, che acquisisce ulteriore importanza alla luce del teorema 2.2.13.

Teorema 2.2.14 (Rockafeller generalizzato).
Siano $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty)$, $\Gamma \subseteq \mathbb{X} \times \mathbb{Y}$ *un insieme c-ciclicamente monotono e* $(x_0; y_0) \in \Gamma$. *Esiste una funzione* $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ *con le seguenti proprietà:*

- $\Phi(x_0) = 0$, *quindi* Φ *non è identicamente* $-\infty$*;*
- Φ *è c-concava;*
- $\Gamma \subseteq Graph\,(\partial^c\Phi)$.

In particolare, Φ *è finita in tutti i punti di* $p_X(\Gamma)$.

Dimostrazione. Procediamo in maniera euristica. Supponiamo che $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ sia una funzione che soddisfa le proprietà richieste. Per definizione (vedi 2.2.9), per ogni $(x^*; y^*) \in \Gamma$ per ogni $x \in \mathbb{X}$ vale che

$$\Phi(x) \leq c(x; y^*) - c(x^*; y^*) + \Phi(x^*). \tag{2.1}$$

Prendiamo $N \in \mathbb{N}^+$ e $(x_1; y_1), \ldots, (x_N; y_N) \in \Gamma$; iterando la disuguaglianza (2.1), troviamo che

$$\begin{aligned}\Phi(x) &\leq c(x; y_N) - c(x_N; y_N) + \Phi(x_N)\\ &\leq c(x; y_N) - c(x_N; y_N) + c(x_N; y_{N-1}) - c(x_{N-1}; y_{N-1}) + \Phi(x_{N-1})\\ &\leq c(x; y_N) - c(x_N; y_N) + c(x_N; y_{N-1}) - c(x_{N-1}; y_{N-1}) + \Phi(x_{N-1})\\ &\qquad + \cdots + c(x_1; y_0) - c(x_0; y_0)\\ &= c(x; y_N) + \sum_{i=1}^{N} c(x_i; y_{i-1}) - \sum_{i=0}^{N} c(x_i; y_i).\end{aligned}$$

Questa relazione suggerisce una formula per definire Φ in modo da soddisfare le richieste del teorema. Definiamo $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ tale che

$$\Phi(x) := \inf\left\{c(x; y_N) + \sum_{i=1}^{N} c(x_i; y_{i-1}) - \sum_{i=0}^{N} c(x_i; y_i)\right\}, \tag{2.2}$$

dove l'estremo inferiore è fatto su tutte le possibili le scelte di $N \in \mathbb{N}^+$ e di $(x_1; y_1), \ldots,$ $(x_N; y_N) \in \Gamma$. Siccome c assume valori reali, è ovvio che Φ ha valori in $\mathbb{R} \cup \{-\infty\}$ ed è

c-concava (vedi 2.2.6). Mostriamo che $\Gamma \subseteq \mathrm{Graph}\,(\partial^c\Phi)$. Fissiamo $(x^*; y^*) \in \Gamma$ e $x \in \mathbb{X}$; dobbiamo verificare che vale la disuguaglianza (2.1). Dalla definizione di Φ, fissando $(x_N; y_N) \in \Gamma$, troviamo che

$$\Phi(x) \leq c(x; y_N) - c(x_N; y_N) + \Phi(x_N),$$

che è la disuguaglianza (2.1), chiamando $(x_N; y_N) = (x^*; y^*)$.

Infine verifichiamo che $\Phi(x_0) = 0$. Dalla definizione di Φ e dal fatto che Γ è c-ciclicamente monotono deduciamo che $\Phi(x_0) \geq 0$ (precisiamo che utilizziamo la disuguaglianza della definizione 2.2.1 soltanto per permutazioni σ di $\{0; \ldots; N\}$ tali che $\sigma(i) = i + 1$ per ogni $i = 0, \ldots, N-1$ e $\sigma(N) = 0$). La disuguaglianza $\Phi(x_0) \leq 0$ si ottiene scegliendo $N = 1$ e $(x_1; y_1) = (x_0; y_0)$ nella definizione di Φ.

In conclusione, notiamo che per ogni $(x; y) \in \Gamma$ vale che $y \in \partial^c\Phi(x)$, quindi $\Phi(x) > -\infty$ per il lemma 2.2.10. □

2.2.3 Soluzione del problema duale

Siamo pronti per affrontare il problema duale.

Definizione 2.2.15 (Formula di dualità)**.**
Sia $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ una funzione costo. Diremo che vale la formula di dualità per c se si ha che

$$\sup_{(\Phi;\Psi)\in I_c} \mathcal{J}(\Phi; \Psi) = \inf_{\pi\in\Gamma(\mu;\nu)} \mathcal{C}(\pi).$$

Abbiamo già osservato che per ogni funzione costo c la disuguaglianza

$$\sup_{(\Phi;\Psi)\in I_c} \mathcal{J}(\Phi; \Psi) \leq \inf_{\pi\in\Gamma(\mu;\nu)} \mathcal{C}(\pi)$$

è ovvia. Vogliamo individuare una classe di funzioni costo abbastanza ampia per cui sia valida la disuguaglianza opposta, implicando la formula di dualità.

Corollario 2.2.16. *Sia $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty)$ una funzione costo lipschitziana e limitata. Allora vale la formula di dualità e l'estremo superiore nella formula 2.2.15 è un massimo, raggiunto da una coppia di funzioni $(\Phi; \Phi^c)$ lipschitziane e limitate.*

Dimostrazione. Abbiamo provato nel teorema 1.3.10 l'esistenza di un piano di trasporto ottimale π. Poniamo

$$\Gamma := \mathrm{supp}(\pi);$$

abbiamo mostrato nel teorema 2.2.13 che Γ è c-ciclicamente monotono; per il teorema 2.2.14, fissato un punto a caso $(x_0; y_0) \in \Gamma$, esiste una funzione $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ c-concava, tale che $\Phi(x_0) = 0$ e $\Gamma \subseteq \mathrm{Graph}\,(\partial^c\Phi)$. Da quest'ultima proprietà segue che per ogni $x \in \mathbb{X}$ vale che

$$\Phi(x) \leq c(x; y_0) - c(x_0; y_0) + \Phi(x_0) \leq c(x; y_0);$$

essendo c limitata, deduciamo che Φ è limitata dall'alto. Per definizione di c-coniugata (vedi 2.2.4), per ogni $y \in \mathbb{Y}$ vale che

$$\Phi^c(y) \leq c(x_0; y) - \Phi(x_0) = c(x; y_0);$$

essendo c limitata, deduciamo che Φ^c è limitata dall'alto. Applicando il risultato provato in 2.2.8, per ogni $x \in \mathbb{X}$ vale che

$$\Phi(x) = \Phi^{cc}(x) = \inf_{y\in\mathbb{Y}}\{c(x;y) - \Phi^c(y)\} \geq -\sup_{y\in\mathbb{Y}}\Phi^c(y).$$

Essendo Φ^c limitata dall'alto, deduciamo che Φ è limitata dal basso. In maniera del tutto analoga, si mostra che Φ^c è limitata dal basso. Dalla definizione di Φ (vedi 2.2), notiamo che Φ è estremo inferiore puntuale di una famiglia di funzioni equi-lipschitziane uniformemente limitate dal basso (abbiamo verificato che Φ è limitata dal basso); allora Φ è lipschitziana. Dalla definizione di Φ^c (vedi 2.2.4), deduciamo che Φ^c è estremo inferiore puntuale di una famiglia di funzioni equi-lipschitziane uniformemente limitate dal basso (abbiamo verificato che Φ^c è limitata dal basso); allora Φ^c è lipschitziana.

In conclusione, ricordando il risultato del lemma 2.2.11, per ogni $(x;y) \in \mathrm{Graph}\,(\partial^c\Phi)$ vale che

$$\Phi(x) + \Phi^c(y) = c(x;y). \tag{2.3}$$

Siccome $\mathrm{supp}(\pi) = \Gamma \subseteq \mathrm{Graph}\,(\partial^c\Phi)$, l'equazione 2.3 vale quasi ovunque in $\mathbb{X}\times\mathbb{Y}$ rispetto alla misura π. Segue che

$$\mathcal{C}(\pi) = \int_{\mathbb{X}\times\mathbb{Y}} c\ d\pi = \int_{\mathbb{X}} \Phi\ d\mu + \int_{\mathbb{Y}} \Phi^c\ d\nu = \mathcal{J}(\Phi;\Phi^c).$$

Allora vale la formula di dualità è l'estremo superiore è raggiunto dalla coppia $(\Phi;\Phi^c)$.
□

Teorema 2.2.17 (Kantorovich-Rubinstein)**.**
Sia $c : \mathbb{X}\times\mathbb{Y} \to [0,+\infty]$ una funzione costo semicontinua inferiormente. Vale la formula di dualità (vedi 2.2.15). Inoltre, ogni piano di trasporto ottimale π tale che

$$\int_{\mathbb{X}\times\mathbb{Y}} c\ d\pi < +\infty$$

è concentrato su un insieme c-ciclicamente monotono e σ-compatto (cioè che si scrive come unione numerabile di compatti).

Dimostrazione. **Step 1**: Mostriamo che vale la formula di dualità. Abbiamo provato nel teorema 1.3.10 l'esistenza di un piano ottimale di trasporto π. Abbiamo mostrato che esiste una famiglia $(c_k)_k$ di funzioni non negative, lipschitziane e limitate che tende crescendo a c in maniera puntuale (vedi 1.1.3). Per ogni $k \in \mathbb{N}$ denotiamo con $\mathcal{C}_k$ il funzionale costo definito in 1.2.3 e con $\mathcal{J}_k$ il funzionale duale definito in 2.1.2 relativi al costo c_k. Per ogni $k \in \mathbb{N}$ vale che

$$\min_{\Gamma(\mu;\nu)} \mathcal{C} \geq \sup_{I_c} \mathcal{J} \geq \sup_{I_{c_k}} \mathcal{J}_k = \max_{I_{c_k}} \mathcal{J}_k = \min_{\Gamma(\mu;\nu)} \mathcal{C}_k,$$

avendo utilizzato nelle ultime due disuguaglianze il risultato provato in 2.2.16. Poniamo

$$\gamma := \min_{\Gamma(\mu;\nu)} \mathcal{C},\ \gamma_k := \min_{\Gamma(\mu;\nu)} \mathcal{C}_k\ \forall k \in \mathbb{N};$$

è banale osservare che la successione $(\gamma_k)_k$ è non decrescente. Allora, per provare la formula di dualità è sufficiente mostrare che $\gamma_k \to \gamma$. Possiamo scegliere una successione $(\pi_k)_k \subseteq \Gamma(\mu;\nu)$ tale che π_k è un piano di trasporto ottimale per il costo $\mathcal{C}_k$. Avendo provato che $\Gamma(\mu;\nu)$ è compatto rispetto alla convergenza debole di misure (vedi 1.3.9), esiste un piano di trasporto π tale che, a meno di passare a sottosuccessioni (non rinominate), $\pi_k \rightharpoonup \pi$. Fissiamo $p \in \mathbb{N}$; ricordando che $c_p \in \mathrm{Lip}_b(\mathbb{X} \times \mathbb{Y})$, per la convergenza debole di misure abbiamo che

$$\begin{aligned}\lim_{k\to+\infty} \gamma_k = \lim_{k\to+\infty} \mathcal{C}_k(\pi_k) &= \lim_{k\to+\infty} \int_{\mathbb{X}\times\mathbb{Y}} c_k \ d\pi_k \\ &\geq \limsup_{k\to+\infty} \int_{\mathbb{X}\times\mathbb{Y}} c_p \ d\pi_k = \int_{\mathbb{X}\times\mathbb{Y}} c_p \ d\pi.\end{aligned}$$

Passando all'estremo superiore in p con il teorema di Beppo Levi, otteniamo che

$$\lim_{k\to+\infty} \gamma_k \geq \sup_{p\in\mathbb{N}} \int_{\mathbb{X}\times\mathbb{Y}} c_p \ d\pi = \int_{\mathbb{X}\times\mathbb{Y}} c \ d\pi \geq \gamma.$$

Essendo l'altra disuguaglianza ovvia, concludiamo che vale la formula di dualità. Per la precisione, abbiamo provato che π è un piano di trasporto ottimale per la funzione costo c.

Step 2: Sia π un piano ottimale tale che

$$\int_{\mathbb{X}\times\mathbb{Y}} c \ d\pi < +\infty.$$

Mostriamo che vale la condizione di ottimalità. Abbiamo già osservato che esiste una successione $(\Phi_k;\Psi_k)_k \subseteq \mathrm{Lip}_b(\mathbb{X}) \times \mathrm{Lip}_b(\mathbb{Y})$ tale che per ogni $k \in \mathbb{N}$ vale che

$$\int_{\mathbb{X}} \Phi_k \ d\mu + \int_{\mathbb{Y}} \Psi_k \ d\nu = \min_{\Gamma(\mu;\nu)} \mathcal{C}_k$$

e inoltre

$$\Phi_k(x) + \Psi_k(y) \leq c_k(x;y) \leq c(x;y) \quad \forall (x;y) \in \mathbb{X} \times \mathbb{Y}.$$

Avendo anche provato che

$$\lim_{k\to+\infty} \left\{ \int_{\mathbb{X}} \Phi_k \ d\mu + \int_{\mathbb{Y}} \Psi_k \ d\nu \right\} = \lim_{k\to+\infty} \left\{ \int_{\mathbb{X}\times\mathbb{Y}} [\Phi_k + \Psi_k] \ d\pi \right\} = \int_{\mathbb{X}\times\mathbb{Y}} c \ d\pi$$

ed essendo c una funzione in $L^1(\mathbb{X} \times \mathbb{Y})$ rispetto alla misura π, deduciamo che $(c - \Phi_k - \Psi_k)_k$ converge a 0 in $L^1(\mathbb{X} \times \mathbb{Y})$ rispetto alla misura π. Allora esistono una sottosuccessione (non rinominata) ed un insieme $\Gamma \subseteq \mathbb{X} \times \mathbb{Y}$ tale che $\pi(\Gamma^c) = 0$ in cui $c(x;y) - \Phi_k(x) - \Psi_k(y) \to 0$. Per il lemma di Ulam (vedi 1.1.2) possiamo anche assumere che Γ sia σ-compatto (a meno di eliminare un altro insieme di misura π nulla). Rimane da provare che Γ è c-ciclicamente monotono. Fissiamo $\mathbb{N} \in \mathbb{N}^+$, σ una permutazione di $\{1;\dots;N\}$ e $(x_1;y_1),\dots(x_N;y_N)$ punti in Γ. Per ogni $k \in \mathbb{N}$ vale che

$$\sum_{i=1}^{N} c(x_i;y_{\sigma(i)}) \geq \sum_{i=1}^{N} \Phi_k(x_i) + \Psi_k(y_{\sigma(i)}) = \sum_{i=1}^{N} \Phi_k(x_i) + \Psi_k(y_i).$$

Essendo nell'insieme di convergenza puntuale, se prendiamo il limite per $k \to +\infty$, troviamo

$$\sum_{i=1}^{N} c(x_i;y_{\sigma(i)}) \geq \sum_{i=1}^{N} c(x_i;y_i).$$

$\square$

2.2.4 Caratterizzazione dei piani ottimali

Abbiamo enunciato condizioni necessarie sulla "forma" del supporto dei piani di trasporto ottimali. Per completezza, enunciamo anche il seguente risultato, che afferma che tali condizioni sono sostanzialmente sufficienti.

Teorema 2.2.18. *Sia $c : \mathbb{X} \times \mathbb{Y} \to [0, \infty]$ una funzione costo semicontinua inferiormente e supponiamo che esistano $\alpha \in L^1(\mathbb{X}; \mu)$ e $\beta \in L^1(\mathbb{Y}; \nu)$ tali che per ogni $(x; y) \in \mathbb{X} \times \mathbb{Y}$ vale che $c(x; y) \leq \alpha(x) + \beta(y)$.*

- *Un piano di trasporto π è ottimale per la funzione costo c se e solo se è concentrato su un insieme c-ciclicamente monotono e σ-compatto.*
- *Esiste una funzione $\Phi : \mathbb{X} \to \mathbb{R} \cup \{-\infty\}$ c-concava, tale che $\Phi \in L^1(\mathbb{X}; \mu)$, $\Phi^c \in L^1(\mathbb{Y}; \nu)$ e inoltre*

$$\min_{\Gamma(\mu;\nu)} \mathcal{C} = \int_{\mathbb{X}} \Phi \, d\mu + \int_{\mathbb{Y}} \Phi^c \, d\nu.$$

Definizione 2.2.19. Dato uno spazio metrico $(\mathbb{X}; d)$ e $p \in [1 + \infty)$ definiamo l'insieme

$$\mathscr{P}_p(\mathbb{X}) := \left\{ \mu \in \mathscr{P}(\mathbb{X}) \;\middle|\; \int_{\mathbb{X}} d(x; x_0)^p \, d\mu < +\infty \text{ per qualche } x_0 \in \mathbb{X} \right\}.$$

Osservazione 2.2.20. Notiamo che nella definizione 2.2.19, per la disuguaglianza triangolare, se $\mu \in \mathscr{P}_p(\mathbb{X})$, allora per ogni $x' \in \mathbb{X}$ vale che

$$\int_{\mathbb{X}} d(x; x')^p \, d\mu < +\infty.$$

Osservazione 2.2.21. Il teorema 2.2.18 si applica nel caso particolare in cui $\mathbb{X} = \mathbb{Y} = \mathbb{R}^d$, $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ e $c(x; y) = |x - y|^p$. Infatti è sufficiente osservare che

$$|x - y|^p \leq 2^{p-1} \left(|x|^p + |y|^p\right)$$

e stiamo assumendo che

$$\int_{\mathbb{R}^d} |x|^p \, d\mu + \int_{\mathbb{R}^d} |y|^p \, d\nu < +\infty.$$

Per riassumere, abbiamo mostrato che, data una funzione costo $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ semicontinua inferiormente, la c-ciclica monotonia e la σ-compattezza del supporto di un piano di trasporto ottimale π sono necessarie se $c \in L^1(\mathbb{X} \times \mathbb{Y}; \pi)$ e diventano sufficienti se esistono due funzioni $\alpha \in L^1(\mathbb{X}; \mu), \beta \in L^1(\mathbb{Y}; \nu)$ tali che $c(x; y) \leq \alpha(x) + \beta(y)$ per ogni $x \in \mathbb{X}, y \in \mathbb{Y}$. É stato provato che queste condizioni sono ancora sufficienti se assume valori finiti, oppure se $c : \mathbb{X} \times \mathbb{Y} \to [0, +\infty]$ è continua. L'esempio seguente (dovuto ad Aldo Pratelli) mostra che qualche ipotesi di continuità/finitezza sulla funzione costo deve essere fatta per dedurre l'ottimalità di un piano di trasporto dalla c-ciclica monotonia del suo supporto.

Esempio 2.2.22. Siano $\mathbb{X} = \mathbb{Y} = [0, 1]$ e μ, ν la misura di Lebesgue in $[0, 1]$; fissiamo $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Definiamo la funzione costo $c : [0, 1]^2 \to [0, +\infty]$ tale che

$$c(x; y) := \begin{cases} 1 & \text{se } y = x, \\ 2 & \text{se } y = x + \alpha \mod (1), \\ +\infty & \text{altrimenti.} \end{cases}$$

Definiamo le mappe $T_1, T_2 : \mathbb{X} \to \mathbb{Y}$ tali che

$$T_1(x) := x, \quad T_2(x) := x + \alpha \mod (1).$$

Se definiamo le misure

$$\pi_1 := (Id \times T_1)_\#(\mathscr{L} \llcorner [0, 1]), \quad \pi_1 := (Id \times T_2)_\#(\mathscr{L} \llcorner [0, 1]),$$

è immediato osservare che π_1, π_2 sono piani di trasporto (le marginali sono tutte misure di Lebesgue in $[0, 1]$) e sono concentrati rispettivamente sui grafici di T_1 e T_2. É immediato osservare che

$$\mathcal{C}(\pi_1) = 1 < 2 = \mathcal{C}(\pi_2)$$

e che π_1 è un piano di trasporto ottimale per il costo c. Vogliamo mostrare che $\Gamma := \text{Graph}\,(T_2)$ è c-ciclicamente monotono. Se così non fosse, esisterebbero un minimo intero $N \geq 2$, una permutazione σ di $\{1; \ldots ; N\}$ e dei punti $(x_1; y_1), \ldots (x_N; y_N) \in \Gamma$ tali che

$$\sum_{i=1}^{N} c(x_i; y_i) > \sum_{i=1}^{N} c(x_i; y_{\sigma(i)}) \tag{2.4}$$

Notiamo che possiamo assumere che σ sia uno shift degli indici, cioè $\sigma(i) = i-1$ per ogni $i = 2, \ldots, N$ e $\sigma(1) = N$: è possibile ridursi a questo caso, osservando che per almeno uno dei cicli disgiunti in cui si decompone σ deve valere la disuguaglianza stretta (2.4); dalla minimalità di N si deduce che σ è un N-ciclo e, a meno di rinominare gli elementi, possiamo ricondurci al caso in cui sia uno shift degli indici. Dalla minimalità di N segue che i punti scelti in Γ sono tutti distinti. Dalla formula di T_2 segue che $y_i = x_i + \alpha \mod (1)$ per ogni $i \in \{1; \ldots ; N\}$. Ricordiamo la formula che definisce c: essendo il termine sinistro nella disequazione (2.4) finito (vale $2N$), per ogni $i \in \{1; \ldots ; N\}$ deve valere $y_{\sigma(i)} = x_i$ oppure $y_{\sigma(i)} = x_i + \alpha \mod (1)$. Se esiste i_0 tale che $y_{\sigma(i)} = x_i + \alpha \mod (1)$, deduciamo che $x_{\sigma(i)} + \alpha = x_i + \alpha \mod (1)$, ovvero $x_i = x_{\sigma(i)}$; allora $y_i = y_{\sigma(i)}$ e questo è assurdo per la minimalità di N. Allora vale $y_{\sigma(i)} = x_i$ per ogni $i \in \{1; \ldots ; N\}$, ovvero $x_i = x_{i-1} + \alpha \mod (1)$ per ogni $i \in \{2; \ldots ; N\}$ e $x_1 = x_N + \alpha \mod (1)$. Concatenando queste uguaglianze, troviamo che

$$x_1 = x_1 + N\alpha \mod (1),$$

che è assurdo perchè α è irrazionale.

Capitolo 3

Teorema di Brenier

Abbiamo studiato l'esistenza della soluzione del problema di trasporto ottimo formulato da Kantorovich. Vogliamo mostrare che, almeno in un caso molto particolare (ma anche significativo), riusciamo a ottenere la soluzione del problema di Monge e a scrivere il piano di trasporto ottimale in maniera elegante. Nel seguito assumeremo che $\mathbb{X} = \mathbb{Y} = \mathbb{R}^n$ e inoltre

$$c(x; y) = \frac{1}{2} |x - y|^2 .$$

3.1 Sull'esistenza della mappa di trasporto

Definizione 3.1.1 (Grafico misurabile)**.**
Un sottoinsieme $\Gamma \subseteq \mathbb{X} \times \mathbb{Y}$ si dice grafico misurabile se è boreliano e per ogni $x \in \mathbb{X}$ esiste un unico $y \in \mathbb{Y}$ tale che $(x; y) \in \Gamma$.

Lemma 3.1.2. *Se $\Gamma \subseteq \mathbb{X} \times \mathbb{Y}$ è un grafico misurabile (vedi 3.1.1) e π è un piano di trasporto concentrato su Γ, allora esiste una mappa di trasporto T tale che $\pi = (Id \times T)_{\#}\mu$.*

Dimostrazione. Siccome π è concentrata su Γ, esiste un boreliano $\Gamma_1 \subseteq \Gamma$ tale che $\pi(\mathbb{X} \times \mathbb{Y} \setminus \Gamma_1) = 0$. Per la regolarità interna di π (vedi 1.1.2), esiste una successione non decrescente $(K_k)_k$ di compatti in Γ_1 tale che

$$\pi\left(\mathbb{X} \times \mathbb{Y} \setminus \bigcup_k K_k\right) = 0.$$

Per ogni $k \in \mathbb{N}$ definiamo $C_k \coloneqq p_X(K_k)$ e sia $T_k : C_k \to \mathbb{Y}$ tale che $(x; T_k(x)) \in K_k$ per ogni $x \in C_k$. Notiamo che K_k è il grafico della funzione T_k; siccome K_k è compatto, T_k è continua. Inoltre, è banale osservare che $(T_k)_{|C_l} = T_l$ per ogni $l \leq k$. Notiamo anche che

$$\mu(\mathbb{X} \setminus C_k) = \pi(p_X^{-1}(\mathbb{X} \setminus C_k)) \leq \pi(\mathbb{X} \times \mathbb{Y} \setminus K_k).$$

Precisiamo che la successione $(C_k)_k$ è crescente; allora l'insieme $\mathbb{X} \setminus \bigcup_k C_k$ è μ-trascurabile. Scegliamo $y_0 \in \mathbb{Y}$ a caso; sia $T : \mathbb{X} \to \mathbb{Y}$ tale che

$$T(x) \coloneqq \begin{cases} T_k(x) & x \in \bigcup_l C_l, \\ y_0 & x \notin \bigcup_l C_l. \end{cases}$$

La mappa T è ben definita ed è boreliana. Per concludere, è sufficiente provare che $(Id \times T)_{\#}\mu = \pi$; basta verificare che per ogni $k \in \mathbb{N}$ vale $(Id \times T)_{\#}(\mu \llcorner C_k) = \pi \llcorner K_k$ e poi prendere il limite per $k \to +\infty$. Fissato $k \in \mathbb{N}$, possiamo verificare che l'integrale di una funzione boreliana φ non negativa rispetto ad entrambe le misure coincide e concludere con la proposizione 1.1.8. Per la formula di cambio di variabili (vedi 1.1.7) vale che

$$\int_{\mathbb{X}\times\mathbb{Y}} \varphi \; d\left[(Id \times T)_{\#}(\mu \llcorner C_k)\right] = \int_{C_k} \varphi(x; T(x)) \; d\mu.$$

Inoltre, abbiamo che

$$\begin{aligned}\int_{\mathbb{X}\times\mathbb{Y}} \varphi \; d[\pi \llcorner K_k] &= \int_{\mathbb{X}\times\mathbb{Y}} \varphi(x;y)\mathbb{1}_{K_k}(x;y) \; d\pi \\ &= \int_{\mathbb{X}\times\mathbb{Y}} \varphi(x;T(x))\mathbb{1}_{C_k}(x) \; d\pi \\ &= \int_{\mathbb{X}} \varphi(x;T(x))\mathbb{1}_{C_k}(x) \; d\mu,\end{aligned}$$

dove abbiamo usato il fatto che $\varphi(x;y)\mathbb{1}_{K_k}(x;y) = \varphi(x;T(x))\mathbb{1}_{C_k}(x)$. □

Precisiamo che la definizione 3.1.1 è del tutto generale e il lemma 3.1.2 vale anche se $\mathbb{X}$ e $\mathbb{Y}$ sono spazi polacchi, con una dimostrazione analoga.

Se la funzione costo è il quadrato della distanza euclidea, la nozione di c-concavità è legata a quella di concavità tradizionale.

Osservazione 3.1.3. Sia $\Phi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ una funzione c-concava; allora Φ è semicontinua superiormente e funzione $\Phi - \frac{1}{2}|\cdot|^2 : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ è concava. Infatti, se vale

$$\Phi(x) = \inf_{i\in I}\{c(x;y_i) + \alpha_i\},$$

segue immediatamente che Φ è semicontinua superiormente (è estremo puntuale inferiore di funzioni continue, perchè consideriamo una funzione costo continua). Inoltre, vale che

$$\Phi(x) - \frac{1}{2}|x|^2 = \inf_{i\in I}\left\{\frac{1}{2}|y_i|^2 + \alpha_i - < x, y_i >\right\};$$

quindi $\Phi - \frac{1}{2}|\cdot|^2$ è concava e semicontinua superiormente perchè è estremo puntuale inferiore di funzioni affini.

Alla luce di questo risultato, è essenziale ricordare il seguente teorema che applicheremo allo studio della differenziabilità delle funzioni convesse.

Teorema 3.1.4 (Rademacher).
Siano $\Omega \subseteq \mathbb{R}^d$ un aperto e $T : \Omega \to \mathbb{R}$ una funzione lipschitziana. Allora T è differenziabile quasi ovunque in Ω rispetto alla misura di Lebesgue.

Sia data una funzione $T : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convessa; denotiamo con

$$\Omega := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$$

il dominio di finitezza di f. La convessità di f garantisce che Ω è convesso. Inoltre f è localmente limitata nella parte interna di Ω; dalla monotonia dei rapporti incrementali segue che per ogni palla $B(x;R)$ relativamente compatta in Ω per ogni $r \in (0,R)$ vale che

$$\operatorname{Lip}(f;B(x;r)) \leq \frac{1}{R-r}\left(\sup_{B(x;r)} f - \inf_{B(x;r)} f\right) < +\infty,$$

da cui segue che f è localmente lipschitziana nella parte interna del suo dominio di finitezza. Da questa discussione e dal teorema di Rademacher (vedi 3.1.4) si deduce il risultato seguente.

Corollario 3.1.5. *Ogni funzione $T : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convessa è differenziabile quasi ovunque (rispetto alla misura di Lebesgue) nella parte interna del suo dominio di finitezza.*

Possiamo finalmente enunciare e dimostrare il teorema di Brenier e il criterio di ottimalità di Knott-Smith.

Teorema 3.1.6 (Brenier, Knott-Smith)**.**
Siano $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^n)$ tali che μ è assolutamente continua rispetto alla misura di Lebesgue.

- *Ogni piano di trasporto ottimale π (ne esiste almeno uno) è indotto da una mappa di trasporto T, cioè vale*
$$\pi = (Id \times T)_{\#}\mu.$$
Inoltre vale che $T = \nabla\psi$, dove $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ è una funzione semicontinua inferiormente, convessa, finita μ-quasi ovunque e differenziabile per μ-quasi ogni $x \in \mathbb{R}^n$. In particolare, T è una mappa di trasporto ottimale, cioè risolve il problema di Monge.

- *Viceversa, se $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ è una funzione semicontinua inferiormente, convessa, finita μ-quasi ovunque, differenziabile per μ-quasi ogni $x \in \mathbb{R}^n$ e tale che $|\nabla\psi| \in L^2(\mathbb{R}^n;\mu)$, allora la misura*
$$\pi := (Id \times \nabla\psi)_{\#}\mu$$
è un piano di trasporto ottimale da μ a $\nu := (\nabla\psi)_{\#}\mu \in \mathscr{P}_2(\mathbb{R}^n)$.

Dimostrazione. **1)** Abbiamo mostrato nel teorema 1.3.10 che esiste un piano di trasporto ottimale $\pi \in \Gamma(\mu;\nu)$. Notiamo che

$$\begin{aligned}
\mathcal{C}(\pi) &\leq \mathcal{C}(\mu \otimes \nu)\\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{2}|x-y|^2 \; d\mu \otimes \nu\\
&\leq \int_{\mathbb{R}^n \times \mathbb{R}^n} (|x|^2 + |y|^2) \; d\mu \otimes \nu\\
&= \int_{\mathbb{R}^n} |x|^2 \; d\mu + \int_{\mathbb{R}^n} |y|^2 \; d\nu < +\infty,
\end{aligned}$$

essendo $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^n)$. Applicando il teorema 2.2.13 troviamo che $\Gamma := \operatorname{supp}(\pi)$ è c-ciclicamente monotono; per il teorema di Rockafeller generalizzato (vedi 2.2.14), esiste una funzione $\Phi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ non identicamente $-\infty$, c-concava e tale che

$\Gamma \subseteq \mathrm{Graph}\,(\partial^c\Phi) \subseteq \mathbb{R}^n_x \times \mathbb{R}^n_y$. Per il lemma 2.2.10, la funzione Φ è finita su $p_{\mathbb{R}^n_x}(\Gamma)$; notiamo che

$$\mu(p_{\mathbb{R}^n_x}(\Gamma)) = \pi(p^{-1}_{\mathbb{R}^n_x}(p_{\mathbb{R}^n_x}(\Gamma))) \geq \pi(\Gamma) = 1,$$

essendo π concentrata in Γ. Abbiamo osservato in 3.1.3 che la funzione $\Phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ tale che

$$\Psi(x) \coloneqq \frac{1}{2}\,|x|^2 - \Phi(x)$$

è convessa e semicontinua inferiormente. Detto $D(\Psi)$ il suo dominio di finitezza, segue che $p_{\mathbb{R}^n_x}(\Gamma) \subseteq D(\Phi)$ e quindi $D(\Phi)$ è un insieme di misura μ piena. Inoltre $D(\Psi)$ è un insieme convesso; detta Ω la parte interna di $D(\Psi)$, vale che $\mu(\Omega) = 1$ (infatti il bordo di un insieme convesso ha misura di Lebesgue nulla, quindi ha misura μ nulla per assoluta continuità). Per quanto enunciato in 3.1.5, segue che Ψ è differenziabile quasi ovunque in Ω; è immediato osservare che la stessa proprietà vale per Φ.

Sia $x \in \Omega$ un punto in cui $\nabla\Phi(x)$ esiste; vogliamo mostrare che esiste un unico punto $y \in \mathbb{R}^n$ tale che $(x; y) \in \mathrm{Graph}\,(\partial^c\Phi)$. Questa condizione equivale a dire che esiste un unico $y \in \mathbb{R}^n$ tale che la quantità $c(x'; y) - \Phi(x')$ ha un minimo in $x' = x$. Differenziando questa relazione rispetto ad x' e valutando in x (ciò è possibile per la scelta di x), deduciamo che $(x - y) - \nabla\Phi(x) = 0$, cioè y deve essere della forma

$$y = x - \nabla\Phi(x) = \nabla\left(\frac{1}{2}\,|\cdot|^2 - \Phi\right)(x) = \nabla\Psi(x).$$

Questo suggerisce di definire

$$y \coloneqq \nabla\Psi(x);$$

la funzione $c(\cdot; y) - \Phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ è convessa (infatti $\frac{1}{2}\,|\cdot|^2 - \Phi$ è concava come osservato in 3.1.3) e la scelta di y garantisce che ha gradiente nullo in x. Deduciamo che x è punto di minimo per la suddetta funzione.

Abbiamo verificato che π è concentrata sul grafico della funzione $\nabla\Psi$; infatti, π è supportata in $\mathrm{Graph}\,(\partial^c\Phi)$ e l'insieme dei punti

$$\{(x; y) \in \mathrm{Graph}\,(\partial^c\Phi) \;\mid\; \Phi \text{ è differenziabile in } x\}$$

ha misura π piena in $\mathbb{X} \times \mathbb{Y}$ (infatti è contenuto nella controimmagine secondo la mappa p_X dei punti di differenziabilità di Φ, che ha misura μ piena in $\mathbb{X}$). Per il lemma 3.1.2 vale che

$$\pi = (Id \times \nabla\Psi)_{\#}\mu.$$

2) Sia $T \coloneqq \nabla\psi$. Le nostre ipotesi garantiscono che $\nu = T_{\#}\mu \in \mathscr{P}_2(\mathbb{R}^n)$. Per il teorema 2.2.21, l'ottimalità della mappa di trasporto T è verificata se mostriamo che l'insieme

$$\Gamma \coloneqq \{(x; \nabla\psi(x)) \mid \psi \text{ è differenziabile in } x\}$$

è c-ciclicamente monotono. Siano $x_1, \ldots, x_N$ punti differenziabilità di ψ e σ una permutazione di $\{1; \ldots; N\}$. Dalla convessità di ψ deduciamo che per ogni $i \in \{1; \ldots; N\}$ vale che

$$< \nabla\psi(x_i), x_{\sigma(i)} - x_i > \leq \psi(x_{\sigma(i)}) - \psi(x_i);$$

segue che

$$\sum_{i=1}^{N} < \nabla\psi(x_i), x_{\sigma(i)} - x_i > \leq 0.$$

Possiamo equivalentemente scrivere che

$$\sum_{i=1}^{N} |\nabla\psi(x_i) - x_i|^2 \leq \sum_{i=1}^{N} \left|\nabla\psi(x_i) - x_{\sigma(i)}\right|^2,$$

che è la disuguaglianza desiderata. □

3.2 Sull'unicità della mappa di trasporto

Introduciamo la nozione di disintegrazione di una misura che vogliamo utilizzare per mostrare l'unicità della mappa di trasporto.

Definizione 3.2.1 (Disintegrazione).
Siano $\mathbb{X}, \mathbb{Z}$ spazi metrici, $\sigma \in \mathscr{P}(\mathbb{Z})$, $f : \mathbb{Z} \to \mathbb{X}$ una mappa boreliana. Poniamo $\theta := f_\# \sigma \in \mathscr{P}(\mathbb{X})$. Una famiglia $(\sigma_x)_{x\in\mathbb{X}} \subseteq \mathscr{P}(\mathbb{Z})$ è detta disintegrazione di σ rispetto alla mappa f se valgono le seguenti proprietà:

- per ogni $A \in \mathcal{B}(\mathbb{Z})$, la mappa $\mathbb{X} \ni x \to \sigma_x(A)$ è boreliana;
- per ogni $A \in \mathcal{B}(\mathbb{Z})$, vale che

$$\sigma(A) = \int_{\mathbb{X}} \sigma_x(A)\ d\theta;$$

- per θ-quasi ogni $x \in \mathbb{X}$ la probabilità σ_x è concentrata su $f^{-1}(x)$.

Nel contesto della definizione 3.2.1, è possibile mostrare l'esistenza di una famiglia $(\sigma_x)_x$ con le proprietà indicate se $\mathbb{X}, \mathbb{Z}$ sono spazi polacchi. In ogni caso, per i nostri scopi è sufficiente mostrare l'unicità di una disintegrazione.

Lemma 3.2.2 (Unicità della disintegrazione).
Siano $\mathbb{X}, \mathbb{Z}$ spazi metrici e supponiamo che $\mathbb{Z}$ sia separabile. Siano $\sigma \in \mathscr{P}(\mathbb{Z})$, $f : \mathbb{Z} \to \mathbb{X}$ una mappa boreliana e $\theta := f_\# \sigma$. Supponiamo che $(\sigma_x)_{x\in\mathbb{X}}, (\sigma'_x)_{x\in\mathbb{X}}$ siano due disintegrazioni di σ rispetto alla mappa f, come nella definizione 3.2.1. Allora esiste un insieme $N \in \mathcal{B}(\mathbb{X})$ di misura θ-nulla tale che per ogni $x \in \mathbb{X} \setminus N$ vale che $\sigma_x = \sigma'_x$.

Dimostrazione. Sia $(A_n)_n$ una famiglia numerabile di insiemi aperti in $\mathbb{Z}$ che genera la σ-algebra $\mathcal{B}(\mathbb{Z})$ ($\mathbb{Z}$ è separabile, quindi ha una base numerabile di aperti). Fissiamo $n \in \mathbb{N}$ e $B \in \mathcal{B}(\mathbb{X})$. Poniamo $A := A_n \cap f^{-1}(B) \in \mathcal{B}(\mathbb{Z})$. Per la seconda proprietà enunciata in 3.2.1, vale che

$$\int_B \sigma_x(A_n)\ d\theta = \int_B \sigma'_x(A_n)\ d\theta.$$

Essendo B un boreliano in $\mathbb{Z}$ arbitrario, deduciamo che esiste un insieme N_n boreliano in $\mathbb{X}$ di misura θ-nulla tale che per ogni $x \in \mathbb{X} \setminus N_n$ vale che $\sigma_x(A_n) = \sigma'_x(A_n)$. Essendo $(A_n)_n$ numerabile, possiamo invertire i quantificatori deducendo che esiste un insieme $N \subseteq \mathbb{X}$ boreliano di misura θ-nulla tale che per ogni per ogni $x \in \mathbb{X} \setminus N$ per ogni $n \in \mathbb{N}$ vale che $\sigma_x(A_n) = \sigma'_x(A_n)$. Fissato $x \in \mathbb{X} \setminus N$, abbiamo che σ_x e σ'_x sono misure di probabilità boreliane che coincidono su una base (numerabile) della topologia di $\mathbb{Z}$. Allora σ_x e σ'_x coincidono su $\mathcal{B}(\mathbb{Z})$. □

Teorema 3.2.3 (Unicità della mappa di trasporto)**.**
Siano $\mu \in \mathscr{P}_2(\mathbb{R}^n)$ e $\nu \in \mathscr{P}_2(\mathbb{R}^n)$. Supponiamo che μ sia assolutamente continua rispetto alla misura di Lebesgue; il problema di Kantorovich ha soluzione unica π. Per la precisione esiste una mappa di trasporto $T : \mathbb{R}^n \to \mathbb{R}^n$ univocamente determinata a meno di insiemi di misura μ-nulla tale che

$$\pi = (Id \times T)_{\#}\mu.$$

In particolare, anche il problema di Monge ha soluzione unica (a meno di insiemi di misura μ-nulla). Inoltre, T coincide μ-quasi ovunque con il gradiente di una funzione convessa $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ semicontinua inferiormente e finita μ-quasi ovunque.

Dimostrazione. Supponiamo che esistano due piani di trasporto ottimali π_1, π_2. Per il teorema di Brenier (vedi 3.1.6), esistono due funzioni convesse $\psi_1, \psi_2 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ semicontinue inferiormente e finite quasi ovunque tali che

$$\pi_1 = (Id \times \nabla\psi_1)_{\#}\mu, \;\; \pi_2 = (Id \times \nabla\psi_2)_{\#}\mu.$$

Poniamo $T_1 \coloneqq \nabla\psi_1$, $T_2 \coloneqq \nabla\psi_2$. Essendo il funzionale costo lineare rispetto alla misura e lo spazio $\Gamma(\mu;\nu)$ convesso, è immediato notare che anche

$$\pi \coloneqq \frac{1}{2}(\pi_1 + \pi_2)$$

è un piano di trasporto ottimo tra μ e ν. Per il teorema di Brenier (vedi 3.1.6), esiste una mappa $T : \mathbb{R}^n \to \mathbb{R}^n$ boreliana tale che

$$\pi = (Id \times T)_{\#}\mu.$$

Notiamo che per dedurre questo fatto bisogna veramente utilizzare il teorema di Brenier, infatti in generale non vale l'uguaglianza

$$\frac{1}{2}(Id \times T_1)_{\#}\mu + \frac{1}{2}(Id \times T_2)_{\#}\mu = \left(Id \times \frac{T_1 + T_2}{2}\right)_{\#}\mu.$$

Sia $p_X : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ la mappa di proiezione sul primo fattore; si nota immediatamente che

$$(p_X)_{\#}\pi_1 = (\pi_X)_{\#}\pi_2 = (p_X)_{\#}\pi = \mu.$$

Possiamo verificare che $\left(\delta_{(x;T_1(x))}\right)_{x\in\mathbb{R}^n}$ è la disintegrazione di π_1 rispetto alla mappa p_X. Infatti, fissato $A \in \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n)$ vale che

$$\delta_{(x;T_1(x))}(A) = \mathbb{1}_A(x;T_1(x));$$

quindi la mappa $\mathbb{R}^n \ni x \to \delta_{(x;T_1(x))}$ è boreliana. Analogamente, vale che

$$\begin{aligned}\pi_1(A) &= [(Id \times T_1)_{\#}\mu](A)\\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbb{1}_A \; d(Id \times T_1)_{\#}\mu\\ &= \int_{\mathbb{R}^n} \mathbb{1}_A(x;T_1(x)) \; d\mu.\\ &= \int_{\mathbb{R}^n} \delta_{(x;T_1(x))}(A) \; d\mu.\end{aligned}$$

Ragionando in maniera completamente analoga, possiamo verificare che $\left(\delta_{(x;T_2(x))}\right)_{x\in\mathbb{R}^d}$ è la disintegrazione di π_2 rispetto alla mappa p_X e che $\left(\delta_{(x;T(x))}\right)_{x\in\mathbb{R}^d}$ è la disintegrazione di π rispetto alla mappa p_X. Ricordando che $\pi = \frac{1}{2}(\pi_1+\pi_2)$, otteniamo immediatamente che anche

$$\left(\frac{1}{2}\delta_{(x;T_1(x))} + \frac{1}{2}\delta_{(x;T_2(x))}\right)_{x\in\mathbb{R}^d}$$

è una disintegrazione di π rispetto alla mappa p_X. Per il lemma 3.2.2, per μ-quasi ogni $x \in \mathbb{R}^d$ vale che

$$\frac{1}{2}\delta_{(x;T_1(x))} + \frac{1}{2}\delta_{(x;T_2(x))} = \delta_{(x;T(x))}.$$

É immediato dedurre che deve valere che $T_1(x) = T_2(x)$ per μ-quasi ogni $x \in \mathbb{R}^d$. A ben vedere, abbiamo mostrato l'unicità della soluzione del problema di Monge, cioè che esiste ed è unica (a meno di insiemi di misura μ-nulla) una mappa di trasporto ammissibile $T : \mathbb{R}^n \to \mathbb{R}^n$ boreliana che risolve il problema di Monge. $\square$

Concludiamo la trattazione con un semplice corollario.

Corollario 3.2.4. *Supponiamo che μ e ν siano assolutamente continue rispetto alla misura di Lebesgue. Detta $T^{\mu\to\nu}$ ($T^{\nu\to\mu}$ rispettivamente) l'unica mappa di trasporto ottimale da μ a ν (da ν a μ rispettivamente) vale che*

$$T^{\nu\to\mu} \circ T^{\mu\to\nu} = Id \quad \mu - \textit{quasi ovunque},$$

$$T^{\mu\to\nu} \circ T^{\nu\to\mu} = Id \quad \nu - \textit{quasi ovunque}.$$

Dimostrazione. Innanzitutto, notiamo che per il teorema 3.2.3 le mappe $T^{\mu\to\nu}$ e $T^{\nu\to\mu}$ sono univocamente determinate (a meno di insiemi di misura μ o ν nulla).

Introduciamo l'operatore di inversione $i : \mathbb{X} \times \mathbb{Y} \to \mathbb{Y} \times \mathbb{X}$ tale che

$$i(x;y) := (y;x).$$

Denotiamo con $\tilde{c}(y;x) := c(x;y)$. É immediato verificare che $i_\#$ manda $\Gamma(\mu;\nu)$ in $\Gamma(\nu;\mu)$ e $\Gamma_{0;c}(\mu;\nu)$ in $\Gamma_{0;\tilde{c}}(\nu;\mu)$. Nel nostro caso, $c = \tilde{c}$. Dette T ed S le uniche mappe di trasporto ottimo da μ a ν e da ν a μ rispettivamente, dall'unicità del piano di trasporto ottimo deduciamo che

$$(Id \times T)_\#\mu = (S \times Id)_\#\nu.$$

Sia $f : \mathbb{X} \times \mathbb{Y} \to [0 + \infty]$ una funzione boreliana; per la formula di cambio di variabili (vedi 1.1.7), vale che

$$\begin{aligned}\int_{\mathbb{R}^n} f(x;T(x))\ d\mu &= \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x;y)\ d(Id \times T)_\#\mu \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n} f(x;y)\ d(S \times Id)_\#\mu \\ &= \int_{\mathbb{R}^n} f(S(y);y)\ d\nu.\end{aligned}$$

Scegliendo $f(x;y) := |y - T(x)|$, si trova che

$$0 = \int_{\mathbb{R}^n} |y - T(S(y))|\ d\nu,$$

dunque $y = T(S(y))$ per ν-quasi ogni $y \in \mathbb{R}^n$. Analogamente, si trova che $S(T(x)) = x$ per μ-quasi ogni $x \in \mathbb{R}^n$. $\square$

Bibliografia

[1] Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Universitext. Springer New York, 2010.

[2] Dacorogna, B. *Direct methods in the calculus of variations. 2nd ed*, vol. 78. 01 2007.

[3] Rindler, F. *Introduction to the Modern Calculus of Variations.* 2015.

[4] Santambrogio, F. *Optimal Transport for Applied Mathematicians.* 2015.

[5] Villani, C. *Topics in Optimal Transportation Theory*, vol. 58. 01 2003.