UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Matematica
Corso di Laurea Triennale in Matematica

Trasporto ottimo. Problema di Monge-Kantorovich e teorema di Brenier

Laureando
Battistin Emanuele
1071650

Relatore
Prof. Monti Roberto

Anno Accademico 2016/2017
21 Luglio 2017

To my optimal girlfriend Camilla

Indice

Elenco dei simboli

$\mathbb{N}$	insieme dei numeri naturali	
$\mathbb{R}^n$	spazio euclideo di dimenisione n	
$\mathcal{L}^n$	misura di Lebesgue in dimensione n	
$x \cdot y$	prodotto scalare in $\mathbb{R}^n$	
$\langle x, y \rangle$	prodotto di dualità	
$\square$	conclude una dimostrazione	
※	conclude una dimostrazione per assurdo	
$\emptyset$	insieme vuoto	
$A \subset B$	A sottinsieme di B	
$\mathrm{Int}(A)$	interno di A	
$\overline{A}$	chiusura di A	
∂A	frontiera di A	
$B_r(x)$	palla di centro x e raggio r	
$C_c(X)$	funzioni continue a supporto compatto su X	
$C_b(X)$	funzioni continue limitate su X	
$T_{\#}\mu$	misura push-forward	pg.1
$\mathcal{T}(\mu, \nu)$	insieme delle mappe di trasporto	pg.1
$\Pi(\mu, \nu)$	insieme dei piani di trasporto	pg.3
$\mathrm{spt}(\mu)$	supporto di μ	pg.4
$\mathrm{spt}(\phi)$	supporto di ϕ	pg.5
$M(X)$	misure di Radon su X	pg.5
$\mu_h \overset{*}{\rightharpoonup} \mu$	convergenza debole* di misure	pg.5
$\mu \llcorner A$	misura ristretta ad A	pg.7
$\mathrm{Dom}(\phi)$	dominio di ϕ	pg.11
$\mathrm{epi}(\phi)$	epigrafico di ϕ	pg.11
ϕ^*	convessa coniugata di ϕ	pg.13
$\partial\phi(x)$	sottodifferenziale di ϕ in x	pg.16
$f_k \xrightharpoonup{L^1(X,\mu)} f$	convergenza debole in L^1	pg.25
$\mu \ll \nu$	misura assolutamente continua	pg.28

Introduzione

Work in progress

Capitolo 1

Problema di Monge-Kantorovich

1.1 Formulazione del problema di Monge

Siano μ, ν due misure di Borel finite su $\mathbb{R}^n$, $n \geq 1$, tali che $\mu(\mathbb{R}^n) = \nu(\mathbb{R}^n)$.

Definizione 1.1.1. (**Misura push-forward**)
Data una funzione $T : \mathbb{R}^n \to \mathbb{R}^m$ boreliana, cioè tale per cui se $B \subset \mathbb{R}^m$ è un insieme di Borel allora $T^{-1}(B) \subset \mathbb{R}^n$ è un insieme di Borel, resta associata una misura $T_{\#}\mu$, detta **misura push-forward**, o misura immagine, definita come $T_{\#}\mu(B) = \mu(T^{-1}(B))$, per ogni $B \subset \mathbb{R}^m$ insieme di Borel.

Definizione 1.1.2. (**Mappa di trasporto**)
Chiamamo una funzione $T : \mathbb{R}^n \to \mathbb{R}^n$ boreliana **mappa di trasporto** da μ a ν se $T_{\#}\mu = \nu$. Indichiamo con $\mathcal{T}(\mu, \nu)$ l'insieme di tutte le mappe che traportano μ su ν.

Consideriamo una "funzione costo", $c : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$ continua. Il problema di Monge consiste nello studio del problema di minimo

$$\min \left\{ \int_{\mathbb{R}^n} c(x, Tx) \, \mathrm{d}\mu(x) : T \in \mathcal{T}(\mu, \nu) \right\}.$$

Una volta impostato il problema, sorgono domande naturali circa l'effettiva esistenza di tale minimo ed eventuali modi di caratterizzare le mappe di trsporto che realizzano il minimo. La teoria è molto vasta e sviluppata in ambienti ben più generali e con ipotesi più deboli, e i risultati dipendono fortemente dalla struttura degli spazi considerati, dalle caratteristiche della funzione costo e in un qualche modo dalla "regolarità" delle misure scelte. Già il caso preso in considerazione in questa tesi risulta complicato da trattare con generalità.

Il problema originale posto da G. Monge considerava la funzione la distanza euclidea $c(x,y)=|x-y|$. Noi studeiremo $d(x,y)=\frac{1}{2}|x-y|^2$, riuscendo anche a dare una descrizione della soluzione, provando:

Teorema. (**Trasporto ottimo per costo quadratico, Brenier**)
Date μ, ν due misure di probabilità su $\mathbb{R}^n$ con momento quadratico finito, consideriamo il problema di Monge-Kantorovich associato alla funzione $c(x,y)=\frac{1}{2}|x-y|^2$.
Vale allora:

(i) (**Criterio di ottimalità di Knott-Smith**) $\pi \in \Pi(\mu,\nu)$ è ottimale se e solo se esiste una funzione ϕ convessa s.c.i. tale che $\operatorname{spt}(\pi) \subset \operatorname{gr}(\partial\phi)$, ovvero se per π-quasi ogni (x,y) vale $y \in \partial\phi(x)$.
Inoltre la coppia (ϕ,ϕ^*) minimizza il problema

$$\inf\left\{\int_{\mathbb{R}^n} \phi \,\mathrm{d}\mu + \int_{\mathbb{R}^n} \psi \,\mathrm{d}\nu : (\phi,\psi) \in \widetilde{\Phi} \text{ per ogni } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n\right\}.$$

(ii) (**Breiner**) Se $\mu \ll \mathcal{L}^n$, allora esiste un unico piano di trasporto ottimale π tale che $\pi = (\mathrm{Id} \times \nabla\phi)_{\#}\mu$, dove $\nabla\phi$ è l'unico gradiente di una funzione convessa definito μ-quasi ovunque tale che $\nabla\phi_{\#}\mu=\nu$.
Inoltre vale $\operatorname{spt}(\nu) = \overline{\nabla\phi(\operatorname{spt}(\mu))}$.

(iii) Se $\mu \ll \mathcal{L}^n$, allora $\nabla\phi$ è l'unica soluzione del problema

$$\int_{\mathbb{R}^n} |x-\nabla\phi(x)|^2 \,\mathrm{d}\mu(x) = \inf_{T \in \mathcal{T}(\mu,\nu)} \int_{\mathbb{R}^n} |x-T(x)|^2 \,\mathrm{d}\mu(x),$$

o equivalentemente,

$$\int_{\mathbb{R}^n} x \cdot \nabla\phi(x) \,\mathrm{d}\mu(x) = \sup_{T \in \mathcal{T}(\mu,\nu)} \int_{\mathbb{R}^n} x \cdot T(x) \,\mathrm{d}\mu(x).$$

(iv) Se $\mu \ll \mathcal{L}^n$ e $\nu \ll \mathcal{L}^n$, allora per μ-quasi ogni x e ν-quasi ogni y si ha

$$\nabla\phi^* \circ \nabla\phi(x) = x, \qquad \nabla\phi \circ \nabla\phi^*(y) = y,$$

e $\nabla\phi^*$ è l'unico gradiente di una funzione convessa definita ν-quasi ovunque che trasporta la misura ν su μ; inoltre è l'unica soluzione del problema di Monge per $T \in \mathcal{T}(\nu,\mu)$ con funzione costo quadratico.

In seguito presenteremo tutti gli strumenti e le definizioni necessarie per arrivare a provare questo risultato.

1.2 Riformulazione di Kantorovich

Vediamo ora una riformulazione del problema di Monge dovuta a L.V. Kantorovich. Consideriamo le proiezioni:

- $p^x(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ che agisce come $(x,y) \mapsto x$;
- $p^y(x,y) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ che agisce come $(x,y) \mapsto y$.

Possiamo ora dare le seguenti:

Definizione 1.2.1. (**Piano di trasporto**)
Siano μ, ν due misure di Borel, finite su $\mathbb{R}^n$, $n \geq 1$, tali che $\mu(\mathbb{R}^n) = \nu(\mathbb{R}^n)$. Chiamiamo π misura boreliana su $\mathbb{R}^n \times \mathbb{R}^n$ **piano di trasporto** se è un elemento dell'insieme

$$\Pi(\mu,\nu) = \left\{ \pi \text{ misura di Borel su } \mathbb{R}^n \times \mathbb{R}^n \,\middle|\, \begin{matrix} p^x_\# \pi = \mu \\ p^y_\# \pi = \nu \end{matrix} \right\}.$$

Equivalentemente, $\Pi(\mu,\nu)$ è l'insieme di tutte la misure π su $\mathbb{R}^n \times \mathbb{R}^n$ tali che per ogni $A, B \subset \mathbb{R}^n$ insiemi boreliani valgano $\pi(A \times \mathbb{R}^n) = \mu(A)$ e $\pi(\mathbb{R}^n \times B) = \nu(B)$.

Osservazione. L'insieme dei piani di traporto $\Pi(\mu,\nu)$ è non vuoto, contiene infatti la misura prodotto $\mu \otimes \nu \in \Pi(\mu,\nu)$.

Introduciamo il funzionale $I : \Pi(\mu,\nu) \to [0,+\infty]$, definito da

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\, \mathrm{d}\pi.$$

Il problema di Kantorovich chiede di trovare

$$\inf \{ I[\pi] : \pi \in \Pi(\mu,\nu) \}.$$

Osserviamo che tale riformulazione risulta vantaggiosa in quanto si richiede di cercare una misura anzichè una mappa.

Osservazione. Questa riformulazione non è equivalente al problema di Monge esposto nella sezione precedente, ma risulta invece essere una versione "rilassata": la classe di oggetti su cui si cerca l'estremo inferiore viene infatti ampliata. In particolare data una mappa di trasporto esiste un piano di trasporto associato, ma un dato un piano di trasporto può non essere legato ad una mappa di trasporto. Dettagliamo questa osservazione:

Consideriamo $T \in \mathcal{T}(\mu,\nu)$, e la funzione $\mathrm{Id} \times T : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$, $(\mathrm{Id} \times T)(x) \mapsto (x, Tx)$. La misura $\pi = (\mathrm{Id} \times T)_\# \mu$ risulta essere un elemento di $\Pi(\mu,\nu)$, infatti, scelti $A, B \subset \mathbb{R}^n$ due insiemi boreliani:

$$\begin{aligned} p^x_\# \pi(A) &= \pi\left((p^x)^{-1}(A) \right) = \pi(A \times \mathbb{R}^n) = \\ &= (\mathrm{Id} \times T)_\# \mu(A \times \mathbb{R}^n) = \mu\left((\mathrm{Id} \times T)^{-1}(A \times \mathbb{R}^n) \right) = \mu(A) \end{aligned}$$

e

$$p^y_{\#}\pi(B) = \pi\left(\left(p^y\right)^{-1}(B)\right) = \pi(\mathbb{R}^n \times B) = (\mathrm{Id} \times T)_{\#}\mu(\mathbb{R}^n \times B) =$$
$$= \mu\left(\left(\mathrm{Id} \times T\right)^{-1}(\mathbb{R}^n \times B)\right) = \mu\left(T^{-1}(B)\right) = \nu(B).$$

Se viceversa $\pi \in \Pi(\mu,\nu)$ è una misura tale che $\mathrm{spt}(\pi) =\subset \mathrm{gr}(T)$, dove con $\mathrm{spt}(\pi)$ si intende il più piccolo insieme chiuso $E \subset \mathbb{R}^n \times \mathbb{R}^n$ tale che $\pi(E^c) = 0$, per una fissata $T : \mathbb{R}^n \to \mathbb{R}^n$ mappa di Borel, allora si ha $\nu = T_{\#}\mu$, in quanto, scelto $A \subset \mathbb{R}^n$ insieme boreliano:

$$\begin{aligned}
T_{\#}\mu(A) &= \mu\left(T^{-1}(A)\right) = p^x_{\#}\pi\left(T^{-1}(A)\right) = \pi\left(\left(p^x\right)^{-1}\left(T^{-1}(A)\right)\right) = \\
&= \pi\left(T^{-1}(A) \times \mathbb{R}^n\right) = \pi(\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : y \in A, y = Tx\}) = \\
&= \pi(\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : y = Tx \in A, x \in \mathbb{R}^n\}) = \pi(\mathbb{R}^n \times A) = \nu(A).
\end{aligned}$$

Richiamiamo ora un risultato sulle misure push-forward:

Lemma 1.2.1. (**Cambio di variabile**)
Date $T : \mathbb{R}^n \to \mathbb{R}^m$, $f : \mathbb{R}^m \to [0,+\infty)$ funzioni boreliane, vale:

$$\int_{\mathbb{R}^m} f(y)\,\mathrm{d}T_{\#}\mu(y) = \int_{\mathbb{R}^n} f(T(x))\,\mathrm{d}\mu(x).$$

Dimostrazione. Iniziamo considerando il caso in cui $f(y) = \mathbb{1}_E(y)$ dove $E \subset \mathbb{R}^m$ è un insieme boreliano, otteniamo così dalla definizione

$$\int_{\mathbb{R}^m} f(y)\,\mathrm{d}T_{\#}\mu(y) = \int_{\mathbb{R}^m} \mathbb{1}_E(y)\,\mathrm{d}T_{\#}\mu(y) = \int_E \mathrm{d}\mu(T^{-1}\{y\}) = \int_{T^{-1}(E)} \mathrm{d}\mu(x) =$$
$$= \int_{\mathbb{R}^n} \mathbb{1}_{T^{-1}(E)}\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^n} \mathbb{1}_{\{x:T(x)\in E\}}\,\mathrm{d}\mu(x) = \int_{\mathbb{R}^n} f(T(x))\,\mathrm{d}\mu(x)$$

Dalla linearità dell'integrale segue che la formula continua a valere per tutte le funzioni a gradino, cioè funzioni che sono combinazione lineare di funzioni indicatrici. Inoltre il teorema della convergenza monotona ci permette di estendere la formula alle funzioni misurabili e positive, concludendo. □

Partendo dal funzionale del problema di Kantorovich, e applicando il lemma sopra:

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\,\mathrm{d}\pi(x,y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y)\,\mathrm{d}(\mathrm{Id} \times T)_{\#}\mu(x) = \int_{\mathbb{R}^n} c(x,Tx)\,\mathrm{d}\mu(x)$$

otteniamo il funzionale del problema di Monge. Dunque se il minimo del problema di Kantorovich è della forma $\pi = (\mathrm{Id} \times T)_{\#}\mu$, allora T realizza il minimo per il problema di Monge. Resta da provare l'effettiva esistenza di tale minimo, problema di cui ci occuperemo ora.

1.3 Esistenza del minimo di Kantorovich

Definizione 1.3.1. (**Misura di Radon**)
Una misura boreliana μ si dice **misura di Radon** se è Borel regolare, ovvero per ogni $A \subset \mathbb{R}^n$ esiste un insieme boreliano $B \supset A$ tale che $\mu(A) = \mu(B)$, e finita sui compatti, cioè per ogni $K \subset \mathbb{R}^n$ compatto vale $\mu(K) < +\infty$. Indicheremo con $M(X)$ l'insieme delle misure di Radon su X.

Definizione 1.3.2. (**Convergenza debole* per misure di Radon**)
Data $(\mu_h)_{h\in\mathbb{N}}$ successione di misure di Radon su $\mathbb{R}^n$, diciamo che μ_h **converge debole*** a μ, o equivalentemente $\mu_h \overset{*}{\rightharpoonup} \mu$, se

$$\lim_{h\to\infty} \int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}\mu_h = \int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}\mu \qquad \text{per ogni } \varphi \in C_c(\mathbb{R}^n).$$

Proposizione 1.3.1. (**Caraterizzazioni della convergenza debole***)
Data $(\mu_h)_{h\in\mathbb{N}}$ successione di misure di Radon su $\mathbb{R}^n$, sono equivalenti le seguenti affermazioni:

(*i*) $\mu_h \overset{*}{\rightharpoonup} \mu$;

(*ii*) per ogni $U \subseteq \mathbb{R}^n$, U aperto, vale $\mu(U) \leq \liminf\limits_{h\to\infty} \mu_h(U)$, e per ogni $K \subseteq \mathbb{R}^n$, K compatto vale $\limsup\limits_{h\to\infty} \mu_h(U) \leq \mu(K)$;

(*iii*) per ogni $B \subseteq \mathbb{R}^n$, B boreliano limitato tale che $\mu(\partial B) = 0$, vale il limite $\lim\limits_{h\to\infty} \mu_h(B) = \mu(B)$.

Dimostrazione. Mostriamo che (*i*) implica (*ii*): fissato $\epsilon > 0$, sia $U \subseteq \mathbb{R}^n$ un aperto e K un compatto contenuto in U. Fissiamo poi $\varphi \in C_c(\mathbb{R}^n)$ tale che $\mathrm{spt}(\varphi) = \overline{\{x \in \mathbb{R}^n : \varphi(x) \neq 0\}} \subset U$, $0 \leq \varphi \leq 1$, $\varphi \equiv 1$ su K. Allora:

$$\mu(K) \leq \int_{\mathbb{R}^n} \varphi\,\mathrm{d}\mu = \lim_{h\to\infty} \int_{\mathbb{R}^n} \varphi\,\mathrm{d}\mu_h \leq \liminf_{h\to\infty} \mu_h(U)$$

e dunque

$$\mu(U) = \sup\Big\{\mu(K) : K \text{ compatto}, K \subseteq U\Big\} \leq \liminf_{h\to\infty} \mu_h(U).$$

L'altra parte di (*ii*) si raggiunge in modo analogo.
Proviamo ora che (*ii*) implica (*iii*): sia B un boreliano limitato con $\mu(\partial B) = 0$. Avremo dunque:

$$\mu(B) = \mu(\mathrm{Int}(B)) \leq \liminf_{h\to\infty} \mu_h(\mathrm{Int}(B)) \leq \limsup_{h\to\infty} \mu_h(\overline{B}) \leq \mu(\overline{B}) = \mu(B).$$

Per concludere mostriamo che (*iii*) implica (*i*): fissiamo $\varepsilon > 0$ e $\varphi \in C_c(\mathbb{R}^n)$, $\varphi > 0$. Sia $R > 0$ tale che $\mathrm{spt}(\varphi) \subset B_0(R)$ e $\mu(\partial B_0(R)) = 0$. Scegliamo allora $0 = t_0 < t_1 < ... < t_N$ tali che $t_N = 2\|\varphi\|_{L^\infty}$, con $0 < t_i - t_{i-1} < \varepsilon$ e $\mu(\varphi^{-1}\{t_i)\} = 0$, per ogni $i = 1, ..., N$. Posto $B_i = \varphi^{-1}(t_{i-1}, t_i)$, vale quindi $\mu(\partial B_i) = 0$, per ogni $i = 1, ..., N$. Consideriamo allora

$$\sum_{i=2}^{N} t_{i-1}\mu_h(B_i) \leq \int_{\mathbb{R}^n} f d\mu_h \leq \sum_{i=2}^{N} t_i\mu_h(B_i) + t_1\mu_h(B_0(R))$$

e

$$\sum_{i=2}^{N} t_{i-1}\mu(B_i) \leq \int_{\mathbb{R}^n} f d\mu \leq \sum_{i=2}^{N} t_i\mu(B_i) + t_1\mu_h(B_0(R))$$

per cui da (iii) segue:

$$\limsup_{h\to\infty} \left| \int_{\mathbb{R}^n} \varphi \,\mathrm{d}\mu_h - \int_{\mathbb{R}^n} \varphi \,\mathrm{d}\mu \right| \leq 2\varepsilon \cdot \mu(B_0(R)).$$

□

Teorema 1.3.1. (**Compattezza debole per le misure di Radon**)
Sia $(\mu_h)_{h\in\mathbb{N}}$ successione di misure di Radon su $\mathbb{R}^n$ tali che $\sup_{h\in\mathbb{N}} \mu_h(K) < \infty$,per ogni $K \subset \mathbb{R}^n$ compatto. Allora esiste una sottosuccessione $(\mu_{h_j})_{j\in\mathbb{N}}$ ed una misura di Radon μ tali che $\mu_{h_j} \overset{*}{\rightharpoonup} \mu$.

Dimostrazione. Assumiamo per il momento che $M = \sup_{h\in\mathbb{N}} \mu_h(\mathbb{R}^n) < +\infty$. Sia $(\varphi_k)_{k\in\mathbb{N}}$ sia un sottoinsieme denso e numerabile di $C_c(\mathbb{R}^n)$. L'ipotesi aggiuntiva implica che la successione $\{\int_{\mathbb{R}^n} \varphi_1 d\mu_h\}_{h\in\mathbb{N}}$ sia limitata, dunque per il teorema di Bolzano-Weiestrass possiamo trovare una sottosuccessione $\{\mu_h^1\}_{h\in\mathbb{N}}$ e un valore $a_1 \in \mathbb{R}^n$ per cui

$$\int_{\mathbb{R}^n} \varphi_1 \,\mathrm{d}\mu_h^1 \to a_1.$$

Con un processo iterativo, selezioniamo una sottosuccessione di $\{\mu_h^k\}_{h\in\mathbb{N}}$ di $\{\mu_h^{k-1}\}_{h\in\mathbb{N}}$ e $a_k \in \mathbb{R}^n$ tali che

$$\int_{\mathbb{R}^n} \varphi_k \,\mathrm{d}\mu_h^k \to a_k.$$

Posto per brevità $\nu_h = \mu_h^h$, avremo

$$\int_{\mathbb{R}^n} \varphi_k \,\mathrm{d}\nu_h \to a_k, \qquad \text{per ogni } k \geq 1.$$

Definiamo allora l'operatore $L(\varphi_k) = a_k$, per cui vale $|L(\varphi_k)| \leq M\|\varphi_k)\|_{L^\infty}$, inoltre L può essere esteso in modo unico ad un funzionale limitato $\overline{L}$ su

$C_c(\mathbb{R}^n)$. Dal teorema di rappresentazione di Riesz segue l'esistenza di una misura di Radon μ su $\mathbb{R}^n$ tale per cui

$$\overline{L} = \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\mu, \text{ per ogni } \varphi \in C_c(\mathbb{R}^n)$$

Scelta dunque $\varphi \in C_c(\mathbb{R}^n)$, dalla densità di $(\varphi_k)_{k\in\mathbb{N}}$ discende l'esistenza di una sottosuccessione $(\varphi_i)_{i\in\mathbb{N}}$ che converge uniformemente a φ. Scelgo dunque $\varepsilon > 0$ e i sufficientemente grande tali che $\|\varphi - \varphi_i\| < \frac{\varepsilon}{4M}$. Fissiamo poi $H \in \mathbb{N}$ tale che per ogni $h > H$ vale

$$\left| \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\mu \right| < \frac{\varepsilon}{2}.$$

Allora per $h > H$ osserviamo:

$$\begin{aligned} &\left| \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\mu \right| = \\ &= \left| \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\nu_h + \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\mu + \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\mu - \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\mu \right| \leq \\ &\leq \left| \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\nu_h \right| + \left| \int_{\mathbb{R}^n} \varphi \, \mathrm{d}\mu - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\mu \right| + \left| \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\nu_h - \int_{\mathbb{R}^n} \varphi_i \, \mathrm{d}\mu \right| < \\ &\qquad < 2M \cdot \|\varphi - \varphi_i\| + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Questo conclude nel caso in cui $M < \infty$.
Se l'ipotesi aggiuntiva non sussiste, è sufficiente applicare il ragionamento precedente alle misure $\mu_k^l = \mu_k \llcorner B_0(l)$, dove $k, l = 1, 2, 3, ...$ e utilizzare poi un argomento di selezione diagonale. □

Diamo ancora una definizione:

Definizione 1.3.3. (**Misura con momento quadratico finito**)
Diciamo che una misura μ su $\mathbb{R}^n$ ha **momento quadratico finito** se

$$\int_{\mathbb{R}^n} |x|^2 \, \mathrm{d}\mu < +\infty.$$

Forti dei risultati appena ottenuti, dimostriamo il risultato principale di questa sezione:

Teorema 1.3.2. (**Esistenza del minimo di Kantorovitch**)
Siano μ, ν due misure di Radon con momento quadratico finito, tali che $\mu(\mathbb{R}^n) = \nu(\mathbb{R}^n) < +\infty$ e $c(x, y) = \frac{|x-y|^2}{2}$. Allora il problema di Kantorovitch

$$\inf_{\pi \in \Pi(\mu,\nu)} \left\{ I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) \, \mathrm{d}\pi \right\}$$

ammette minimo.

Dimostrazione. Essendo μ, ν misure finite, per ogni $\varepsilon > 0$ esiste $R > 0$ tale che $\mu(\mathbb{R}^n \setminus B_R(0)) < \varepsilon$ e $\nu(\mathbb{R}^n \setminus B_R(0)) < \varepsilon$. Poniamo per brevità $B_R(0) = B_R$. Avremo allora:

$$\begin{aligned}\pi((\mathbb{R}^n \times \mathbb{R}^n) \setminus (B_R \times B_R)) \leq \pi((\mathbb{R}^n \setminus B_R) \times \mathbb{R}^n)) + \pi(\mathbb{R}^n \times (\mathbb{R}^n \setminus B_R)) =\\ = \mu(\mathbb{R}^n \setminus B_R) + \nu(\mathbb{R}^n \setminus B_R) < 2\varepsilon.\end{aligned}$$

Consideriamo una successione minimizzante $(\pi_h)_{h\in\mathbb{N}}$, con $\pi_h \in \Pi(\mu,\nu)$, tale che

$$\lim_{h\to\infty} \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,\mathrm{d}\pi_h = m = \inf\{I[\pi] : \pi \in \Pi(\mu,\nu)\} < +\infty.$$

Consideriamo la successione di termine generale $\pi_h \llcorner (B_R \times B_R)$; per il teorema di compattezza delle misure di Radon, esiste una sottosuccessione $(\pi_j)_{j\in\mathbb{N}}$ tale che $\pi_j \overset{*}{\rightharpoonup} \pi^*$, con π^* misura su $(B_R \times B_R)$. Dunque vale:

$$\lim_{j\to\infty} \int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi\,\mathrm{d}\pi_j = \int_{\mathbb{R}^n\times\mathbb{R}^n} \varphi\,\mathrm{d}\pi^*$$

per ogni $\varphi \in C_c(B_R \times B_R)$. Applicando un procedimento di selezione diagonale, possiamo considerare funzioni $\varphi \in C_c(\mathbb{R}^n \times \mathbb{R}^n)$.
Mostriamo che anche $\pi^* \in \Pi(\mu,\nu)$: richiamando le caraterizzazioni della convergenza debole*, siano $A, B \subset \mathbb{R}^n$ insiemi boreliani con $\mu(\partial A) = \nu(\partial B) = 0$ abbiamo

$$\pi^*(A \times \mathbb{R}^n) = \lim_{j\to+\infty} \pi_j(A \times \mathbb{R}^n) = \lim_{j\to+\infty} \mu(A) = \mu(A).$$

Analogamente otteniamo anche $\pi^*(\mathbb{R}^n \times B) = \nu(B)$.
Consideriamo ora $(\varphi_k(x,y))_{k\in\mathbb{N}}$ successione monotona crescente di funzioni che converge a $c(x,y)$, $\varphi_k \in C_c(\mathbb{R}^n \times \mathbb{R}^n)$ per ogni $k \in \mathbb{N}$.

Calcoliamo allora:

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} c\,\mathrm{d}\pi_j &= \int_{B_R\times B_R} c\,\mathrm{d}\pi_j + \int_{(B_R\times B_R)^c} c\,\mathrm{d}\pi_j =\\ &= \int_{B_R\times B_R} (c-\varphi_k)\,\mathrm{d}\pi_j + \underbrace{\int_{B_R\times B_R} \varphi_k\,\mathrm{d}\pi_j}_{(i)} + \int_{(B_R\times B_R)^c} c\,\mathrm{d}\pi_j.\end{aligned}$$

Analizziamo (i): per $j \to \infty$ abbiamo

$$\int_{B_R\times B_R} \varphi_k\,\mathrm{d}\pi_j \to \int_{B_R\times B_R} \varphi_k\,\mathrm{d}\pi^*,$$

da cui segue:

$$\int_{B_R\times B_R}\varphi_k\,\mathrm{d}\pi^* = \int_{B_R\times B_R}(\varphi_k - c)\,\mathrm{d}\pi^* + \int_{B_R\times B_R} c\,\mathrm{d}\pi^* =$$
$$= \int_{B_R\times B_R}(\varphi_k - c)\,\mathrm{d}\pi^* + \int_{\mathbb{R}^n\times\mathbb{R}^n} c\,\mathrm{d}\pi^* - \underbrace{\int_{(B_R\times B_R)^c} c\,\mathrm{d}\pi^*}_{(ii)}.$$

Stimiamo allora (ii), dal momento che $|x-y|^2 \leq 2(|x|^2+|y|^2)$ e che le misure μ, ν hanno momento quadratico finito segue:

$$\int_{(B_R\times B_R)^c} c(x,y)\,\mathrm{d}\pi^* = \int_{(B_R\times B_R)^c}\frac{|x-y|^2}{2}\,\mathrm{d}\pi^* \leq$$
$$\leq \int_{(B_R\times B_R)^c}(|x|^2+|y|^2)\,\mathrm{d}\pi^* = \int_{B_R^c}|x|^2\,\mathrm{d}\mu + \int_{B_R^c}|y|^2\,\mathrm{d}\nu < 2\varepsilon.$$

Ricordando la convergenza delle funzioni φ_k a c, sostituendo troviamo:

$$\lim_{j\to\infty}\int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,\mathrm{d}\pi_j = \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)\,\mathrm{d}\pi^* = \inf\left\{I[\pi] : \pi\in\Pi(\mu,\nu)\right\}.$$

Dunque il problema di Kantorovich ammette minimo. □

1.4 Riformulazione duale di Kantorovich

Kantorovitch ha riformulato il problema anche in un modo duale. Data $c : \mathbb{R}^n \times \mathbb{R}^n \to [0,+\infty)$ funzione costo, definiamo

$$\Phi_c = \left\{(\phi,\psi)\in L^1(\mathbb{R}^n,\mu)\times L^1(\mathbb{R}^n,\nu) : \phi(x)+\psi(y)\leq c(x,y) \text{ per ogni } x,y\in\mathbb{R}^n\right\}$$

e il funzionale $J : \Phi_c \to \mathbb{R}$

$$J(\phi,\psi) = \int_{\mathbb{R}^n}\phi(x)\,\mathrm{d}\mu + \int_{\mathbb{R}^n}\psi(y)\,\mathrm{d}\nu.$$

Con questi dati Kantorovich considera il problema:

$$\sup\{J(\phi,\psi) : (\phi,\psi)\in\Phi_c\}.$$

Notiamo che se $\pi\in\Pi(\mu,\nu)$, allora

$$J(\phi,\psi) = \int_{\mathbb{R}^n}\phi(x)\,\mathrm{d}\mu + \int_{\mathbb{R}^n}\psi(y)\,\mathrm{d}\nu = \int_{\mathbb{R}^n\times\mathbb{R}^n}(\phi(x)+\psi(y))\,\mathrm{d}\pi,$$

ed essendo inoltre $\phi(x) + \psi(y) \leq c(x, y)$, vale

$$\sup_{(\phi,\psi)\in\Phi_c} J(\phi, \psi) \leq \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x, y)\, \mathrm{d}\pi = I[\pi],$$

quindi

$$\sup_{(\phi,\psi)\in\Phi_c} J(\phi, \psi) \leq \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Mostreremo nel Capitolo 3 che in realtà vale il simbolo di uguaglianza, prima dovremo però introdurre alcune nozioni di analisi convessa.

Capitolo 2

Elementi di analisi convessa

In questo capitolo richiamiamo alcune definizioni e introduciamo alcuni strumenti dell'analisi convessa, come la convessa coniugata di una funzione e un principio di dualità. Useremo una struttura generale, seguendo l'esposizione presente in [Br], che ci permetterà di dare una dimostrazione completa del teorema di dualità di Fechel-Rockafellar. Nella stessa sezione si è fatto riferimento anche a [Bo] per il lemma introduttivo. Nella terza parte introdurremo il sottodifferenziale di una funzione e analizzeremo il suo legame con la convessa coniugata. Dove non specificato diversamete ci porremo nel contesto di uno spazio vettoriale topologico normato E, siano E^* il suo duale e $\langle \cdot, \cdot \rangle$ il prodotto di dualità.

2.1 Richiami e fatti utili

Ricordiamo che una funzione $\phi : E \to]-\infty, +\infty]$, viene detta **semicontinua inferiormente** (s.c.i.) se per ogni $x \in E$ vale $\phi(x) \leq \liminf_{x \to y} \phi(y)$. Tale definizione è valida in ambienti più generali, è infatti sufficiente uno spazio topologico per poter dare tale definizione. Si osserva inoltre che una funzione continua è anche s.c.i..

Da questo momento in poi, supponiamo sempre $\phi \not\equiv +\infty$. Diciamo una funzione $\phi : E \to]-\infty, +\infty]$ **convessa** se per ogni $t \in [0, 1]$, per ogni $x, y \in E$ vale $\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)$. Una definizione equivalente di convessità può essere data nel modo seguente: una funzione f è convessa se e solo se il suo **epigrafico** $\text{epi}(f) = \{(x, \lambda) \in E \times \mathbb{R} : f(x) \leq \lambda\}$ è un insieme convesso. Non dimostreremo questa asserzione, tuttavia la useremo nella dimostrazione della Proposizione 2.2.1.
Ricordiamo inoltre che con **dominio** di una funzione ϕ intendiamo $\text{Dom}(\phi) = \{x \in E : \phi(x) < +\infty\}$, si noti in particolare che se $\phi : \mathbb{R}^n \to]-\infty, +\infty]$ è convessa la frontiera $\partial(\text{Dom}(\phi))$ è un insieme di misura $\mathcal{L}^n$-trascurabile.

Enunciamo ora un risultato che ci risulterà utile tra poco e successivamente per il Teorema di dualità di Fenchel-Rockafellar (Teorema2.3.1):

Proposizione 2.1.1. Sia $A \subset E$ un sottinsieme convesso con almeno un punto interno x_0. Allora per ogni $x \in \overline{A}$ vale $\{tx + (1-t)x_0 : t \in (0,1)\} \subset \text{Int}(A)$.

Dimostrazione. Fissati $x \in \overline{A}$ e $y \in \{tx + (1-t)x_0 : t \in (0,1)\}$ consideriamo l'omotetia f, con centro y e rapporto $\lambda < 0$ tale per cui valga $f(x_0) = x$. Sia $V \subset A$ intorno aperto di x_0, allora $f(V)$ è intorno di x e dunque contiene dei punti $f(z) \in A$. Vale

$$f(z) - y = \lambda(z - y) = \lambda(z - f(z)) + \lambda(f(z) - y),$$

da cui segue $y - f(z) = \frac{\lambda}{\lambda - 1}(z - f(z))$, cioè esiste un'omotetia g di centro $f(z)$ e rapporto $0 < \frac{\lambda}{\lambda-1} < 1$ tale che $g(z) = y$. In particolare $g(V)$ è un intorno di y contenuto in A, da questo segue la tesi. $\square$

Corollario. Dato un insieme A convesso, il suo interno $\text{Int}(A)$ è ancora un convesso. Inoltre se $\text{Int}(A) \neq \emptyset$ allora la chiusura di A coincide con la chiusura del suo interno.

Dimostrazione. Se $\text{Int}(A) \neq \emptyset$, ogni punto di $\overline{A}$ è punto di accumulazione di $\text{Int}(A)$. Mostriamo allora che ogni punto interno di $\overline{A}$ appartiene a $\text{Int}(A)$: sia x punto dell'interno di $\overline{A}$, per praticità sia $x = 0$; allora considerati V un intorno simmetrico di 0 contenuto in $\overline{A}$ e $y \in \text{Int}(A) \cap V$, abbiamo $-y \in \overline{A}$, dunque dalla Proposizione 2.1.1 abbiamo $0 \in \text{Int}(A)$ nel caso in cui $y \neq 0$. Nel caso $y = 0$ l'asserto è banale. $\square$

In particolare otteniamo:

Proposizione 2.1.2. Date ϕ, ψ funzioni s.c.i. e convesse tali che $\text{Int}(\text{Dom}(\phi)) = \text{Int}(\text{Dom}(\psi))$ e coincidenti su tale insieme, allora vale $\phi \equiv \psi$.

Dimostrazione. Vale $\text{Int}(\text{Dom}(\phi)) = \text{Int}(\text{Dom}(\psi))$, abbiamo allora

$$\text{Int}(\text{epi}(\phi)) = \text{Int}(\text{epi}(\psi)),$$

e per il corollario sopra vale

$$\overline{\text{epi}(\phi)} = \overline{\text{Int}(\text{epi}(\phi))} = \overline{\text{Int}(\text{epi}(\psi))} = \overline{\text{epi}(\psi)}.$$

Dunque vale $\overline{\text{epi}(\phi)} = \overline{\text{epi}(\psi)}$, che implica $\phi \equiv \psi$. $\square$

Ricordiamo che una funzione ϕ si dice **localmente lipschitziana**, se per ogni insieme compatto $K \subset E$ e per ogni $x, y \in K$ esiste una costante $L_K \geq 0$ tale che $|\phi(x) - \phi(y)| \leq L_K|x - y|$. Si osservi che, per definizione, una funzione convessa ϕ è localmente lipschitziana su $\text{Int}(\text{Dom}(\phi))$.
Enunciamo, senza dimostrarlo, il seguente:

Teorema 2.1.1. (**Rademacher**)
Siano $U \subset \mathbb{R}^n$ un insieme aperto e una funzione $\phi : U \to \mathbb{R}^m$, allora $\nabla\phi$ è ben definito e limitato, a meno di un insieme $\mathcal{L}^n$-trascurabile.

Per una dimostrazione si veda il Teorema 3.2 in [E].

2.2 Convessa coniugata

Definizione 2.2.1. (**Convessa coniugata**)
Data $\phi : E \to]-\infty, +\infty]$ si definisce **convessa coniugata** (o **trasformata di Fenchel**) di ϕ la funzione $\phi^* : E^* \to]-\infty, +\infty]$ definita da $\phi^*(f) = \sup_{x \in E}\{\langle f, x\rangle - \phi(x)\}$.

Proposizione 2.2.1. Data $\phi : E \to]-\infty, +\infty]$ sia ϕ^* la sua convessa coniugata. Valgono allora:

(*i*) ϕ^* è convessa e s.c.i..

(*ii*) Se ϕ è convessa, allora $\phi^* \not\equiv +\infty$

Dimostrazione. (*i*) È sufficiente osservare che fissato $x \in E$ la funzione $f \mapsto \langle f, x\rangle - \phi(x)$ risulta essere convessa e continua, dunque anche s.c.i.
(*ii*) Preliminarmente fissiamo $x_0 \in \mathrm{Dom}(\phi)$ e $\lambda_0 \in \mathbb{R}$ tale che $\lambda_0 < \phi(x_0)$. Consideriamo poi $A = \mathrm{epi}(\phi)$, insieme convesso e chiuso, e $B = \{(x_0, \lambda_0)\}$, insieme convesso e compatto. Dalla scelta di x_0 e λ_0 segue che $A \cap B = \emptyset$. Siamo dunque nelle ipotesi del Teorema di Hanh-Banach: applicandolo otteniamo l'esistenza di un iperpiano H che separa A e B, che diciamo avere equazione $\psi = \alpha$.
L'applicazione $x \to \psi((x, 0))$ è una forma lineare e continua, dunque $\psi((x, 0)) = \langle x, f\rangle$, con $f \in E^*$.
Posto allora $k = \psi((0, 1))$ abbiamo che $\psi((x, \lambda)) = \langle x, f\rangle + k\lambda$, per ogni $(x, \lambda) \in E \times \mathbb{R}$, allora imponendo le condizioni $\psi > \alpha$ su A e $\psi < \alpha$ su B, otteniamo

$$\begin{cases} \langle x, f\rangle + k\lambda > \alpha & \text{per ogni } (x, \lambda) \in \mathrm{epi}(\phi) \\ \langle x_0, f\rangle + k\lambda_0 < \alpha, \end{cases}$$

da cui segue anche $\langle x, f\rangle + k\phi(x) > \alpha$, per ogni $x \in \mathrm{Dom}(\phi)$, dunque in particolare

$$\langle x_0, f\rangle + k\lambda_0 < \alpha < \langle x_0, f\rangle + k\phi(x_0).$$

Ricordando la scelta di λ_0 fatta all'inizio concludiamo che $k > 0$, e per le considerazioni fatte sopra si ha

$$\left\langle x, -\frac{1}{k}f \right\rangle - \phi(x) < -\frac{\alpha}{k} < +\infty \qquad \text{per ogni } x \in \mathrm{Dom}(\phi)$$

e questo ci permette di concludere. $\square$

Possiamo ora considerare $\phi^{**} : E \to]-\infty, +\infty]$, che risulta definita da $\phi^{**}(x) = \sup_{f \in E^*} \langle f, x \rangle - \phi^*(f)$. In virtù dell'ipotesi che $\phi \not\equiv +\infty$, vale il teorema

Teorema 2.2.1. (**Fenchel-Moreau**)
Data $\phi : E \to]-\infty, +\infty]$ convessa e s.c.i., vale $\phi^{**} = \phi$.

Dimostrazione. Supponiamo preliminarmente che sia $\phi \geq 0$. Dalle definizioni segue che per ogni $x \in E$, per ogni $f \in E^*$ vale $\langle x, f \rangle \leq \phi(x) + \phi^*(f)$. Ragioniamo adesso per assurdo: supponiamo esista $x_0 \in E$ tale che $\phi^{**}(x_0) < \phi(x_0)$. Potrebbe anche essere $\phi(x_0) = +\infty$, ma per le ipotesi fatte si avrà sempre $\phi^{**}(x_0) < +\infty$. Consideriamo ora, come nella dimostrazione precedente, $A = epi(\phi)$ e $B = \{(x_0, \phi{*}{*}(x_0))\}$, e applichiamo la forma geometrica del teorema di Hanh-Banach, seguendo passo passo le medesime considerazioni fatte in precedenza troviamo

$$\begin{cases} \langle x, f \rangle + k\lambda > \alpha & \text{per ogni } (x, \lambda) \in \text{epi}(\phi) \\ \langle x_0, f \rangle + k\phi^{**}(x_0) < \alpha. \end{cases}$$

Dalla prima disequazione otteniamo $k > \frac{1}{\lambda} \frac{\alpha}{\langle x, f \rangle}$ e con $\lambda \to +\infty$ abbiamo $k \geq 0$. Scelto allora $\varepsilon > 0$, dalla positività di ϕ abbiamo

$$\langle x, f \rangle + (k + \varepsilon)\phi(x) \geq \alpha \qquad \text{per ogni } x \in \text{Dom}(\phi).$$

Allora abbiamo $\phi^*\Big(-\frac{f}{k+\varepsilon}\Big) \leq -\frac{\alpha}{k+\varepsilon}$ e dalla definizione di $\phi^{**}(x_0)$ segue

$$\phi^{**}(x_0) \geq \left\langle x_0, -\frac{f}{k+\varepsilon} \right\rangle - \phi^*(-\frac{f}{k+\varepsilon}) \geq \left\langle x_0, -\frac{f}{k+\varepsilon} \right\rangle + \frac{\alpha}{k+\varepsilon}.$$

Ma allora $\langle x_0, f \rangle + (k+\varepsilon)\phi^{**}(x_0) \geq \alpha$, per ogni $\varepsilon \geq 0$, ma questo contraddice la seconda disequazione del sistema. ※
Rimane ora da provare il caso generale: sia $f_0 \in \text{Dom}(\phi^*)$. Costruiamo allora la funzione $\overline{\phi}(x) = \phi(x) - \langle x, f_0 \rangle + \phi^*(f_0)$, la quale risulta essere convessa, s.c.i., $\overline{\phi} \not\equiv +\infty$ e $\overline{\phi} \geq 0$. Usando la prima parte concludiamo che $\overline{\phi} = \overline{\phi}^{**}$; calcolando $\overline{\phi}^*$ e $\overline{\phi}^{**}$ si trova

$$\begin{aligned} \overline{\phi}^*(f) &= \phi^*(f + f_0) - \phi^*(f_0) \\ \overline{\phi}^{**}(x) &= \phi^{**}(x) - \langle x, f_0 \rangle + \phi^*(f_0) \end{aligned}$$

Quindi

$$\phi(x) - \langle x, f_0 \rangle + \phi^*(f_0) = \overline{\phi}(x) = \overline{\phi}^{**}(x) = \phi^{**}(x) - \langle x, f_0 \rangle + \phi^*(f_0)$$

Confrontando primo e ultimo membro abbiamo la tesi. $\square$

Guardiamo adesso queste definizioni nel contesto di nostro interesse, ovvero in $\mathbb{R}^n$. Enunciamo il risultato per noi importante sotto forma di proposizione, la dimostrazione viene omessa in quanto conseguenza di fatti generali mostrati in questa sezione:

Proposizione 2.2.2. (**Dualità di Fenchel-Legendre**)
Sia $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, con $\phi \not\equiv +\infty$. Sono allora equivalenti:

(i) ϕ è convessa e s.c.i;

(ii) $\phi = \psi^*$ per qualche ψ, con $\psi \not\equiv +\infty$;

(iii) $\phi^{**} = \phi$.

2.3 Dualità di Fenchel-Rockafellar

Possiamo ora dimostrare:

Teorema 2.3.1. (**Dualità di Fenchel-Rockafellar**)
Siano E uno spazio vettoriale normato, E^* il suo duale topologico, Θ, Ξ : $E \to \mathbb{R} \cup \{+\infty\}$ due funzioni convesse, Θ^*, Ξ^* le loro convesse coniugate. Se esiste $x_0 \in E$ tale che $\Theta(x_0) < +\infty$, $\Xi(x_0) < +\infty$ e Θ è continua in x_0, allora

$$\inf_{x\in E} (\Theta(x) + \Xi(x)) = \sup_{f\in E^*} (-\Theta^*(-f) - \Xi^*(f)) = \max_{f\in E^*} (-\Theta^*(-f) - \Xi^*(f)) .$$

Dimostrazione. Poniamo anzitutto

$$\begin{aligned} a = &\inf_{x\in E} (\Theta(x) + \Xi(x)) , \\ b = &\sup_{f\in E^*} (-\Theta^*(-f) - \Xi^*(f)) . \end{aligned}$$

Risulta sempre verificata $b \le a$. Distinguiamo allora i casi $a \in \mathbb{R}$ e $a = -\infty$: nel secondo caso la conclusione è immediata. Sia dunque $a \in \mathbb{R}$; poniamo $C = \mathrm{epi}(\Theta)$ e osserviamo che $\mathrm{Int}(C) \neq \emptyset$, essendo Θ continua in x_0. Consideriamo gli insiemi $A = \mathrm{Int}(C)$ e $B = \{(x, \lambda) \in E \times \mathbb{R} : \lambda \le a - \Xi(x)\}$: entrambi sono convessi e non vuoti, proviamo inoltre che sono disgiunti: considerato $(x, \lambda) \in A$ abbiamo $\lambda > \Theta(x) \ge a - \Xi(x)$, dunque $(x, \lambda) \notin B$. Possiamo quindi applicare il teorema di Hahn-Banach nella sua forma geometrica e otteniamo l'esistenza di un iperpiano chiuso H che separa A e B con disuguaglianze non strette, quindi separa anche $\overline{A}$ e B, dove $\overline{A} = \overline{C}$.
Pertanto esistono $f \in E^*$, $k \in \mathbb{R}$ e $\alpha \in \mathbb{R}$ tali che H ha equazione $\Psi = \alpha$, con $\Psi(x, \lambda) = \langle f, x\rangle + k\lambda$. Ψ separa C e B in senso largo, dunque

$$\begin{cases} \langle f, x\rangle + k\lambda \ge \alpha & \text{per ogni } (x, \lambda) \in C \\ \langle f, x\rangle + k\lambda \le \alpha & \text{per ogni } (x, \lambda) \in B. \end{cases} \tag{$\star$}$$

Scegliendo $x = x_0$ e passando al limite $\lambda \to +\infty$ nella prima equazione $(\star)$ otteniamo $k \ge 0$. Mostriamo che deve essere $k > 0$.

Ragioniamo per assurdo, ricordando che $\Psi \neq 0$, cioè $||f|| + |k| \neq 0$, e supponiamo $k = 0$. Riscriviamo il sistema $(\star)$ ottenendo

$$\begin{cases} \langle f, x \rangle \geq \alpha & \text{per ogni } x \in \mathrm{Dom}(\Theta) \\ \langle f, x \rangle \leq \alpha & \text{per ogni } x \in \mathrm{Dom}(\Xi). \end{cases}$$

Sia $B_{x_0}(\varepsilon_0) \subset \mathrm{Dom}(\Theta)$ per ε_0 sufficientemente piccolo vale $\langle f, x_0 + \varepsilon_0 z \rangle \geq \alpha$ per ogni $z \in B_0(1)$, segue inoltre $\langle f, x_0 \rangle \geq \alpha + \varepsilon_0 \|f\|$. Ma dato che $x_0 \in \mathrm{Dom}(\Xi)$, deve valere anche $\langle f, x_0 \rangle \leq \alpha$. Deve allora essere $\|f\| = 0$, ma allora $\Psi = 0$. ※
Dunque vale $k > 0$, da $(\star)$ segue

$$\begin{cases} \Theta^*(-\frac{f}{k}) \leq -\frac{\alpha}{k} \\ \Xi^*(\frac{f}{k}) \leq \frac{\alpha}{k} - a, \end{cases}$$

sommando membro a membro e cambiando di segno abbiamo

$$-\Theta^* \left(-\frac{f}{k} \right) - \Xi^* \left(\frac{f}{k} \right) \geq a.$$

Dalla definizione di b si ha

$$-\Theta^* \left(-\frac{f}{k} \right) - \Xi^* \left(\frac{f}{k} \right) \leq b,$$

da cui si conclude che

$$a = b = -\Theta^* \left(-\frac{f}{k} \right) - \Xi^* \left(\frac{f}{k} \right)$$

□

2.4 Sottodifferenziale

Definizione 2.4.1. (**Sottodifferenziale**)
Data $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, diremo che x_0 sta nel **sottodifferenziale** $\partial\phi(x)$ se per ogni $y \in \mathbb{R}^n$ vale $\phi(y) \geq \phi(x) + \langle x_0, y - x \rangle$

Dalla sezione precedente, abbiamo che $\phi^*(y) = \sup_{x \in \mathbb{R}^n} x \cdot y - \phi(x)$, allora per ogni $x, y \in \mathbb{R}^n$ vale $x \cdot y \leq \phi(x) + \phi^*(y)$. Mettiamo ora in evidenza una proprietà che sarà utile in seguito, e che caratterizza i punti in cui vale la relazione di uguaglianza:

Proposizione 2.4.1. (**Caraterizzazione del sottodifferenziale**)
Data $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convessa e s.c.i., $\phi \not\equiv +\infty$, vale la relazione

$$x \cdot y = \phi(x) + \phi^*(y) \overset{(i)}{\Longleftrightarrow} y \in \partial\phi(x) \overset{(ii)}{\Longleftrightarrow} x \in \partial\phi^*(y)$$

Dimostrazione. i)Dalla definizione di convessa coniugata abbiamo $x \cdot y \geq \phi(x) + \phi^*(y)$, e sostituendo $x \cdot y \geq \phi(x) + y \cdot z - \phi(z)$, per ogni $z \in \mathbb{R}^n$. Ciò equivale a $\phi(z) \geq \phi(x) + y \cdot (z - x)$, per ogni $z \in \mathbb{R}^n$, ovvero $y \in \partial\phi(x)$. ii)La simmetria della definizione è una conseguenza della Proposizione 2.2.2. □

Enunciamo senza dimostrare i seguenti fatti, per le dimostrazioni si rimanda al Teorema 25.1 in [R]:

Proposizione 2.4.2. Data una funzione ϕ convessa valgono:

(i) per ogni $x \in \text{Int}(\text{Dom}(\phi))$, l'insieme $\partial\phi(x)$ è non vuoto;

(ii) ϕ è differenziabile nel punto x se e solo se $\partial\phi(x)$ contiene un solo elemento, che risulta essere $\nabla\phi(x)$.

Capitolo 3

Dimostrazione del Teorema di Brenier

3.1 Teorema di dualità

Come annunciato alla fine del Capitolo 1, enunciamo:

Teorema 3.1.1. (**Dualità di Kantorovich**)
Siano X, Y spazi metrici completi e separabili, μ,ν misure finite su X, Y rispettivamente, e sia $c : X \times Y \to [0, +\infty]$ una funzione costo s.c.i..
Consideriamo il problema di Kantorovich

$$\inf \left\{ I[\pi] = \int_{X \times Y} c(x, y) \, \mathrm{d}\pi : \pi \in \Pi(\mu, \nu) \right\}$$

e la sua riformulazione duale

$$\sup \left\{ J(\phi, \psi) = \int_X \phi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu : (\phi, \psi) \in \Phi_c \right\}.$$

Vale $\inf\limits_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup\limits_{(\phi,\psi) \in \Phi_c} J(\phi, \psi)$.

Nota. Dimostreremo questo teorema nel caso in cui gli spazi sono compatti e la funzione costo è continua. Il risultato che otterremo può essere generalizzato al teorema sopra enunciato con successive approssimazioni, tuttavia non presentiamo qui questo lungo ragionamento in quanto la "sostanza" del teorema è racchiusa nella dimostrazione che segue, si rimanda al Teorema 1.3 in [V] per l'esposizione completa dell'argomento.

Dimostrazione. (**Spazi compatti e funzione costo continua**)
Sia $E = C_b(X \times Y)$ l'insieme delle funzioni continue e limitate su $X \times Y$, fornito della norma $\| \cdot \|_\infty$. Dal teorema di rappresentazione di Riesz il duale topologico di E può essere identificato con lo spazio $E^* = M(X \times Y)$ delle

misure di Radon su $X \times Y$, con la norma della variazione totale. Fissata $u \in C_b(X \times Y)$, definiamo ora:

$$\Theta(u) = \begin{cases} 0 & \text{se per quasi ogni } (x,y) \in X \times Y \text{ vale } u(x,y) \geq -c(x,y) \\ +\infty & \text{altrove,} \end{cases}$$

$$\Xi(u) = \begin{cases} \int_X \phi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu & \text{se } u(x,y) = \phi(x) + \psi(y) \\ +\infty & \text{altrove.} \end{cases}$$

Notiamo che Ξ è ben definita: se $\phi(x) + \psi(y) = \widetilde{\phi}(x) + \widetilde{\psi}(y)$ per ogni x, y allora esiste $s \in \mathbb{R}$ tale per cui $\phi = \widetilde{\phi} + s$, $\psi = \widetilde{\psi} - s$, e quindi

$$\int_X \phi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu = \int_X \widetilde{\phi} \, \mathrm{d}\mu + \int_Y \widetilde{\psi} \, \mathrm{d}\nu.$$

Le ipotesi del teorema di dualità di Fenchel-Rockafellar sono soddisfatte per $x_0 = 1$, quindi

$$\inf_{x \in E} \{\Theta(x) + \Xi(x)\} = \sup_{f \in E^*} \{-\Theta^*(-f) - \Xi^*(f)\}. \tag{$*$}$$

Espandiamo ora i membri dell'uguaglianza $(*)$, partendo da quello di sinistra:

$$\inf \left\{ \int_X \phi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu : \phi(x) + \psi(y) \geq -c(x,y) \right\} = - \sup_{(\phi,\psi) \in \Phi_c} \{J(\phi,\psi)\}.$$

Ora, fissato $\pi \in M(X \times Y)$, valutiamo:

$$\Theta^*(-\pi) = \sup_{u \in C_b(X \times Y)} \left\{ - \int_{X \times Y} u(x,y) \, \mathrm{d}\pi : u(x,y) \geq -c(x,y) \text{ quasi ovunque} \right\} =$$
$$= \sup_{u \in C_b(X \times Y)} \left\{ \int_{X \times Y} u(x,y) \, \mathrm{d}\pi : u(x,y) \leq c(x,y) \text{ quasi ovunque} \right\}.$$

Se π non è una misura non negativa, allora esiste una funzione $v \in C_b(X \times Y)$ non positiva tale che $\int_{X \times Y} v \, \mathrm{d}\pi > 0$. Allora scelta $u = \lambda v$, con $\lambda \to +\infty$, l'estremo superiore risulta essere $+\infty$.
D'altra parte se π è non negativa allora l'estremo superiore è $\int_{X \times Y} c \, \mathrm{d}\pi$.
Allora

$$\Theta^*(-\pi) = \begin{cases} \int_{X \times Y} c(x,y) \, \mathrm{d}\pi & \text{se } \pi \in M_+(X \times Y) \\ +\infty & \text{altrove.} \end{cases}$$

Similmente

$$\Xi^*(\pi) = \begin{cases} 0 & \begin{array}{l} \text{se per ogni } (\phi,\psi) \in C_b(X) \times C_b(Y), \\ \int_{X \times Y} (\phi + \psi) \, \mathrm{d}\pi = \int_X \phi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu \end{array} \\ +\infty & \text{altrove.} \end{cases}$$

Sostituendo nel membro a destra in $(*)$ otteniamo:

$$
\begin{aligned}
&\sup_{f\in E^*}\{-\Theta^*(-f)-\Xi^*(f)\} = \\
&= \sup_{\pi\in M_+(X\times Y)}\left\{-\int_{X\times Y} c(x,y)\,\mathrm{d}\pi : \begin{array}{c}\text{per ogni } (\phi,\psi)\in C_b(X)\times C_b(Y),\\ \int_{X\times Y}(\phi+\psi)\,\mathrm{d}\pi = \int_X \phi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu\end{array}\right\} = \\
&= \sup_{\pi\in\Pi(\mu,\nu)}\left\{-\int_{X\times Y} c(x,y)\,\mathrm{d}\pi\right\} = \\
&= -\inf_{\pi\in\Pi(\mu,\nu)}\left\{\int_{X\times Y} c(x,y)\,\mathrm{d}\pi\right\}
\end{aligned}
$$

Quindi

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] = \sup_{(\phi,\psi)\in\Phi_c\cap C_b} J(\phi,\psi).$$

L'insieme su cui cerchiamo l'estremo superiore non è lo stesso presentato nella Sezione 1.4, per concludere mostriamo che

$$\sup_{(\phi,\psi)\in\Phi_c\cap C_b} J(\phi,\psi) \le \sup_{(\phi,\psi)\in\Phi_c\cap L^1} J(\phi,\psi) \le \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

La prima disuguaglianza discende da $C_b(X)\times C_b(Y)\subset L^1(\mu)\times L^1(\nu)$, concentriamoci quindi sulla seconda. Siano $(\phi,\psi)\in\Phi_c\cap L^1$ e $\pi\in\Pi(\mu,\nu)$, allora

$$J(\phi,\psi) = \int_X \phi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu = \int_{X\times Y}(\phi(x)+\psi(y))\,\mathrm{d}\pi(x,y).$$

Vale però π-quasi ovunque $\phi(x)+\psi(y)\le c(x,y)$: infatti scelti N_x , N_y insiemi misurabili tali che $\mu(N_x)=\nu(N_y)=0$ e la disuguaglianza valga per ogni $(x,y)\in N_x^c\times N_y^c$ possiamo scrivere $\pi(N_x\times Y)=\mu(N_x)=0=\nu(N_y)=\pi(X\times N_y)$, da cui si deduce $\pi((N_x^c\times N_y^c)^c)=0$. Dunque

$$\int_{X\times Y}(\phi(x)+\psi(y))\,\mathrm{d}\pi(x,y) \le \int_{X\times Y} c(x,y)\,\mathrm{d}\pi(x,y) = I[\pi]$$

e la tesi segue considerando l'estremo superiore a destra e l'estremo inferiore a sinistra. □

3.2 Lemma di doppia convessificazione

Da questo momento in poi fissiamo $c(x,y)=\frac{1}{2}|x-y|^2$, e consideriamo misure μ, ν di probabilità con momento quadratico finito. Poniamo per brevità

$$M_2 = \int_{\mathbb{R}^n}\frac{|x|^2}{2}\,\mathrm{d}\mu(x) + \int_{\mathbb{R}^n}\frac{|y|^2}{2}\,\mathrm{d}\mu(y) < +\infty.$$

La condizione $(\phi, \psi) \in \Phi_c$ diventa

$$\phi(x) + \psi(y) \leq \frac{1}{2}|x-y|^2,$$

e deve valere per μ-quasi ogni $x \in \mathbb{R}^n$ e per ν-quasi ogni $y \in \mathbb{R}^n$. Espandendo $c(x,y)$ e riarrangiando i termini troviamo

$$x \cdot y \leq \left(\frac{|x|^2}{2} - \phi(x)\right) + \left(\frac{|y|^2}{2} - \psi(y)\right).$$

Al fine di sfruttare la teoria delle funzioni coniugate convesse, introdotta nella Sezione 2.2, consideriamo le nuove incognite:

$$\widetilde{\phi}(x) = \frac{|x|^2}{2} - \phi(x), \qquad \widetilde{\psi}(y) = \frac{|y|^2}{2} - \psi(y).$$

Possiamo ora riscrivere i problema di Kantorovich come

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = M_2 - \sup_{\pi \in \Pi(\mu,\nu)} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} (x \cdot y) \,\mathrm{d}\pi(x,y) \right\},$$

e, posto $\widetilde{\Phi} = \{(\phi, \psi) \in L^1(\mu) \times L^1(\nu) : x \cdot y \leq \phi(x) + \psi(y)\}$, il problema duale come

$$\sup_{(\phi,\psi) \in \Phi_c} J(\phi, \psi) = M_2 - \inf_{(\phi,\psi) \in \widetilde{\Phi}} J(\phi, \psi).$$

Richiamando il teorema di dualità di Kantorovich (Sezione 3.1), e svolgendo semplici calcoli troviamo

$$\sup_{\pi \in \Pi(\mu,\nu)} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} (x \cdot y) \,\mathrm{d}\pi(x,y) \right\} = \inf_{(\phi,\psi) \in \widetilde{\Phi}} J(\phi, \psi).$$

Consideriamo ora $(\psi, \phi) \in \widetilde{\Phi}$: le funzioni soddisfano il vincolo $x \cdot y \leq \phi(x) + \psi(y)$, e richiamando quando visto nella Sezione 2.2 possiamo scrivere

$$\psi(y) \geq \sup_{x \in \mathbb{R}^n} x \cdot y - \phi(x) = \phi^*(x),$$

per ν-quasi ogni y, dunque abbiamo $J(\phi, \psi) \geq J(\phi, \phi^*)$, inoltre

$$\phi(x) \geq \sup_{y \in \mathbb{R}^n} x \cdot y - \phi^*(y) = \phi^{**}(x)$$

per μ-quasi ogni x, da cui segue $J(\phi, \phi^*) \geq J(\phi^{**}, \phi^*)$, ovvero

$$\inf_{(\phi,\psi) \in \widetilde{\Phi}} J(\phi, \psi) \geq \inf_{\phi \in L^1(\mu)} J(\phi^{**}, \phi^*).$$

Ottenuti questi risultati, proviamo ora il seguente:

Lemma 3.2.1. (**Doppia convessificazione**)
Siano μ, ν due misure di probabilità con supporti rispettivamente in $X, Y \subseteq \mathbb{R}^n$, tali che

$$M_2 = \int_X \frac{|x|^2}{2}\,\mathrm{d}\mu(x) + \int_Y \frac{|y|^2}{2}\,\mathrm{d}\nu(y) < +\infty.$$

Siano ϕ, ψ funzioni misurabili che prendono valori in $\mathbb{R}\cup\{+\infty\}$, consideriamo

$$\phi^*(y) = \sup_{x\in X} x\cdot y - \phi(x), \qquad \psi^*(x) = \sup_{y\in Y} x\cdot y - \psi(y).$$

Sia $\widetilde{\Phi} = \{(\phi,\psi) \in L^1(\mu)\times L^1(\nu) : x\cdot y \leq \phi(x) + \psi(y)\}$, e $(\psi_k,\phi_k)_{k\in\mathbb{N}}$ una successione di funzioni minimizzante per J su $\widetilde{\Phi}$. Allora:

(*i*) $(\psi_k,\phi_k)_{k\in\mathbb{N}}$ può essere modificata su insiemi di misura nulla per μ e ν rispettivamente, in modo che il vincolo $x\cdot y \leq \phi(x)+\psi(y)$ sia soddisfatto per ogni $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$ e senza modificare il valore di $J(\phi_k,\psi_k)$;

(*ii*) Esiste una successione reale $(a_k)_{k\in\mathbb{N}}$ tale che

$$(\overline{\phi}_k, \overline{\psi}_k) = (\phi_k^{**} + a_k, \phi_k^* - a_k)$$

è ancora una successione minimizzante per J su $\widetilde{\Phi}$, inoltre valgono le stime dal basso

$$\begin{aligned} \overline{\phi}_k(x) &\geq -\tfrac{|x|^2}{2} \quad \text{per ogni } x\in X,\\ \overline{\psi}_k(y) &\geq -\tfrac{|y|^2}{2} \quad \text{per ogni } y\in Y; \end{aligned}$$

e le "stime" dall'alto

$$\begin{aligned} \liminf_{k\to+\infty} \inf_{x\in X} \left(\tfrac{|x|^2}{2} + \overline{\phi}_k(x)\right) &\leq \inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi,\psi) + M_2,\\ \liminf_{k\to+\infty} \inf_{y\in Y} \left(\tfrac{|y|^2}{2} + \overline{\psi}_k(y)\right) &\leq \inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi,\psi) + M_2. \end{aligned}$$

Dimostrazione. (*i*) Data $(\psi_k,\phi_k)\in\widetilde{\Phi}$ esistono gli insiemi N_x, N_y tali che $\mu(N_x) = 0 = \nu(N_y)$ e tali che il vincolo $x\cdot y \leq \phi_k(x)+\psi_k(y)$ è soddisfatto per ogni $(x,y)\in N_x^c\times N_y^c$. Ridefiniamo ϕ_k in modo che assuma il valore $+\infty$ su N_x e ψ_k in modo che assuma il valore $+\infty$ su N_y: vale ancora $(\phi_k,\psi_k)\in\widetilde{\Phi}$ e il valore di $J(\phi_k,\psi_k)$ rimane invariato, avendo modificato le funzioni su insiemi μ-trascurabili e ν-trascurabili.
(*ii*) Sia $(\psi_k,\phi_k)_{k\in\mathbb{N}}$ una successione minimizzante per J. Sfruttando la prima parte, assumiamo che per ogni $x\in X$, $y\in Y$ valga $\phi_k(x)\geq x\cdot y - \psi_k(y)$, allora, dato che $\psi_k \not\equiv +\infty$, deduciamo che ϕ_k è limitata per ogni $x\in X$ da

una qualche funzione affine, sia tale funzione $x \cdot y_0 + b_0$, per qualche $y_0 \in Y$. Allora

$$\phi_k^*(y_0) = \sup_{x \in X} x \cdot y_0 - \phi_k(x) \leq -b_0$$

ed in particolare $\phi_k^* \not\equiv +\infty$.
Inoltre ϕ_k^* è limitata dal basso da una funzione affine, questo implica che

$$a_k = \inf_{y \in Y} \left(\phi_k^*(y) + \frac{|y|^2}{2} \right)$$

è finito. Scegliamo ora $(\overline{\phi}_k, \overline{\psi}_k) = (\phi_k^{**} + a_k, \phi_k^* - a_k)$, osservando in particolare che $\overline{\phi}_k = (\overline{\psi}_k)^*$. Per costruzione abbiamo

$$\inf_{y \in Y} \left(\overline{\psi}_k(y) + \frac{|y|^2}{2} \right) = 0.$$

Calcoliamo

$$\overline{\phi}_k(x) + \frac{|x|^2}{2} = (\overline{\psi}_k)^*(x) + \frac{|x|^2}{2} = \sup_{y \in Y} \left(x \cdot y - \overline{\psi}_k(y) + \frac{|x|^2}{2} \right) \geq$$
$$\geq \sup_{y \in Y} \left(-\frac{|y|^2}{2} - \overline{\psi}_k(y) \right) = -\inf_{y \in Y} \left(\frac{|y|^2}{2} + \overline{\psi}_k(y) \right) = 0,$$

dunque la coppia $(\overline{\phi}_k, \overline{\psi}_k)$ soddisfa le stime dal basso enunciate.
Mostriamo ora l'integrabilità di $\overline{\phi}_k$ e $\overline{\psi}_k$, ricordando che per quanto visto in precedenza

$$J(\overline{\phi}_k, \overline{\psi}_k) = J(\phi_k^{**}, \phi_k^*) \leq J(\phi_k, \psi_k) < +\infty.$$

Vale quindi

$$\int_X \left(\frac{|x|^2}{2} + \overline{\phi}_k(x) \right) \mathrm{d}\mu(x) + \int_Y \left(\frac{|y|^2}{2} + \overline{\psi}_k(y) \right) \mathrm{d}\nu(y) < +\infty,$$

ed essendo entrambe le integrande non negative, deduciamo che $\overline{\phi}_k(x) \in L^1(\mu)$ e $\overline{\psi}_k(y) \in L^1(\nu)$. Quindi $(\overline{\phi}_k, \overline{\psi}_k) \in \widetilde{\Phi}$, e per l'osservazione la sucessione è minimizzante.
Restano da provare le "stime" dal basso:

$$J(\overline{\phi}_k, \overline{\psi}_k) + M_2 = \int_X \left(\frac{|x|^2}{2} + \overline{\phi}_k(x) \right) \mathrm{d}\mu(x) + \int_Y \left(\frac{|y|^2}{2} + \overline{\psi}_k(y) \right) \mathrm{d}\nu(y)$$
$$\geq \inf_{x \in X} \left(\frac{|x|^2}{2} + \overline{\phi}_k(x) \right) + \inf_{y \in Y} \left(\frac{|y|^2}{2} + \overline{\psi}_k(y) \right).$$

Le quantità tra parentersi sono non negative, quindi segue che

$$\liminf_{k \to +\infty} \inf_{x \in X} \left(\frac{|x|^2}{2} + \overline{\phi}_k(x) \right) \leq \liminf_{k \to +\infty} J(\overline{\phi}_k, \overline{\psi}_k) + M_2 = \inf_{(\phi, \psi) \in \widetilde{\Phi}} J(\phi, \psi) + M_2,$$

e l'asserto analogo, questo conclude. □

Corollario. Scegliendo $X = Y = \mathbb{R}^n$, allora ϕ^* e ψ^* coincidono con le convesse coniugate di ϕ e ψ rispettivamente, e vale

$$\inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi,\psi) = \inf_{\phi\in L^1(\mu)} J(\phi^{**},\phi^*).$$

Dimostrazione. Si applicano le definizioni e si usa il lemma precedente. $\square$

Prima di passare al teorema che ci permetterà di scrivere il minimo come coppia di funzioni convesse coniugate abbiamo bisogno di alcuni risulati:

Definizione 3.2.1. (**Convergenza debole in** L^1)
Dati uno spazio misurato (X,μ) e $(f_k)_{k\in\mathbb{N}}$ una successione in $L^1(X,\mu)$, diciamo che f_k **converge debolelmente** a f in $L^1(X,\mu)$, o $f_k \xrightarrow{L^1(X,\mu)} f$, se per ogni $g \in L^\infty(X,\mu)$ vale

$$\lim_{k\to+\infty}\int_X f_k g \,\mathrm{d}\mu = \int_X fg\,\mathrm{d}\mu.$$

Definizione 3.2.2. (**Equiintegrabilità**)
Data $\mathcal{F} \subset L^1(X,\mu)$, diciamo che $\mathcal{F}$ è **equiintegrabile** se valgono:

(i) per ogni $\varepsilon > 0$ esiste un insieme μ-misurabile A di misura finita tale che vale

$$\int_{X\smallsetminus A} |f|\,\mathrm{d}\mu < \varepsilon \qquad \text{per ogni } f \in \mathcal{F};$$

(ii) per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni insieme μ-misurabile E con $\mu(E) < \delta$ vale

$$\int_E |f|\,\mathrm{d}\mu < \varepsilon \qquad \text{per ogni } f \in \mathcal{F}.$$

Osservazione. Si nota in partcolare che se la misura μ è finita allora il punto (i) della definizione precedente è sempre soddisfatto, è sufficiente considerare $A = X$.

Enunciamo ora, senza dare una dimostrazione:

Teorema 3.2.1. (**Dunford -Pettis**)
Sia $\mathcal{F} \subset L^1(X,\mu)$ una famiglia limitata ed equiintegrabile, allora $\mathcal{F}$ è debolmente sequenzialmente compatta, ovvero esistono una successione $(f_k)_{k\in\mathbb{N}} \subset \mathcal{F}$ e una funzione $f \in \mathcal{F}$ tali che $f_k \xrightarrow{L^1(X,\mu)} f$.

Per una dimostrazione del teorema di Dunford-Pettis si può consultare ad esempio [AFP], Teorema 1.38. Ottenuti questi risultati possiamo provare:

Teorema 3.2.2. Siano μ, ν misure di probabilità su $\mathbb{R}^n$, con momento quadratico finito e $\widetilde{\Phi} = \{(\phi, \psi) \in L^1(\mu) \times L^1(\nu) : x \cdot y \leq \phi(x) + \psi(y)\}$. Allora esiste una coppia (ϕ, ϕ^*) di funzioni s.c.i. convesse coniugate su $\mathbb{R}^n$, tali che

$$\inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi, \psi) = J(\phi, \phi^*)$$

Dimostrazione. Iniziamo modificando il vincolo $x \cdot y \leq \phi(x) + \psi(y)$, in modo da avere un vincolo non negativo:

$$\phi(x) + \psi(y) + \frac{|x|^2}{2} + \frac{|y|^2}{2} \geq x \cdot y + \frac{|x|^2}{2} + \frac{|y|^2}{2};$$

$$\left(\phi(x) + \frac{|x|^2}{2}\right) + \left(\psi(y) + \frac{|y|^2}{2}\right) \geq \frac{|x+y|^2}{2} \geq 0.$$

Sia ora $(\phi_k, \psi_k)_{k\in\mathbb{N}}$ una successione minimizzante per J, per il Lemma 3.2.1 e quanto detto sopra troviamo

$$0 \leq \phi_k(x) + \frac{|x|^2}{2}, \qquad 0 \leq \psi_k(y) + \frac{|y|^2}{2}$$

e

$$\left(\phi_k(x) + \frac{|x|^2}{2}\right) + \left(\psi_k(y) + \frac{|y|^2}{2}\right) \geq \frac{|x+y|^2}{2} \geq 0.$$

Le funzioni tra parentesi sono limitate, mostriamo ora che sono anche equiintegrabili: per $l \in \mathbb{N}$ definiamo le successioni troncate

$$\begin{aligned} \phi_k^{(l)}(x) + \tfrac{|x|^2}{2} &= \min\left\{\phi_k(x) + \tfrac{|x|^2}{2}, l\right\}, \\ \psi_k^{(l)}(y) + \tfrac{|y|^2}{2} &= \min\left\{\psi_k(x) + \tfrac{|y|^2}{2}, l\right\}. \end{aligned}$$

Notiamo che

$$\begin{cases} 0 \leq \phi_k^{(l)}(x) + \frac{|x|^2}{2} \leq l, \\ 0 \leq \psi_k^{(l)}(y) + \frac{|y|^2}{2} \leq l; \end{cases} \tag{$\bullet$}$$

e

$$\begin{cases} 0 \leq \phi_k^{(1)}(x) \leq \phi_k^{(2)}(x) \leq \cdots \leq \phi_k^{(l)}(x) \leq \ldots, \\ 0 \leq \psi_k^{(1)}(y) \leq \psi_k^{(2)}(y) \leq \cdots \leq \psi_k^{(l)}(x) \leq \ldots; \end{cases} \tag{i.a}$$

$$J(\phi_k^{(l)}, \psi_k^{(l)}) \leq J(\phi_k, \psi_k); \tag{ii.a}$$

$$\phi_k^{(l)}(x) + \psi_k^{(l)}(y) \geq \min\left\{\frac{|x+y|^2}{2}, l\right\} - \left(\frac{|x|^2}{2} + \frac{|y|^2}{2}\right). \tag{iii.a}$$

Da $(\bullet)$ deduciamo che per ogni l vale

$$\phi_k^{(l)}(x) = -\frac{|x|^2}{2} + O(l),$$

essendo $\frac{|x|^2}{2}$ una funzione fissata in $L^1(\mu)$, abbiamo che $(\phi_k^{(l)})_{k\in\mathbb{N}}$ ammette una sottosuccesione debolmente convergente per $k \to +\infty$ in $L^1(\mu)$, sia $\phi^{(l)}$ il limite di tale sottosuccessione. Usiamo un procedimento analogo con $(\psi_k^{(l)})_{k\in\mathbb{N}}$, ottenendo debole convergenza a $\psi^{(l)} \in L^1(\nu)$. Usando ora un criterio di selezione diagonale estraiamo una sottosuccessione di indici $k \in \mathbb{N}$ tali per cui la convergenza valga per ogni $l \in \mathbb{N}$, otteniamo allora

$$\begin{cases} 0 \le \phi^{(1)}(x) \le \phi^{(2)}(x) \le \cdots \le \phi^{(l)}(x) \le \dots, \\ 0 \le \psi^{(1)}(y) \le \psi^{(2)}(y) \le \cdots \le \psi^{(l)}(x) \le \dots; \end{cases} \tag{i.b}$$

$$J(\phi^{(l)}, \psi^{(l)}) = \lim_{k\to+\infty} J(\phi_k^{(l)}, \psi_k^{(l)}) \le \\ \le \liminf_{k\to+\infty} J(\phi_k, \psi_k) = \inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi, \psi); \tag{ii.b}$$

$$\phi^{(l)}(x) + \psi^{(l)}(y) \ge \min\left\{\frac{|x+y|^2}{2}, l\right\} - \left(\frac{|x|^2}{2} + \frac{|y|^2}{2}\right). \tag{iii.b}$$

Le successioni $(\phi^{(1)})_{l\in\mathbb{N}}$ e $(\psi^{(1)})_{l\in\mathbb{N}}$ sono limitate in L^1, non decrescenti e limitate inferiormente da funzioni L^1 fissate, questo ci permette di applicare il teorema della convergenza monotona, e dedurre l'esistenza di

$$\widehat{\phi} = \lim_{l\to+\infty} \phi^{(l)},$$
$$\widehat{\psi} = \lim_{l\to+\infty} \psi^{(l)};$$

dove $\widehat{\phi}$ e $\widehat{\psi}$ sono funzioni in L^1 definite quasi ovunque che soddisfano

$$J(\widehat{\phi}, \widehat{\psi}) = \lim_{l\to+\infty} J(\phi^{(l)}, \psi^{(l)}) \le \inf_{(\phi,\psi)\in\widetilde{\Phi}} J(\phi, \psi).$$

Richiamiamo ora $(111.b)$,

$$\phi^{(l)}(x) + \psi^{(l)}(y) \ge \min\left\{\frac{|x+y|^2}{2}, l\right\} - \left(\frac{|x|^2}{2} + \frac{|y|^2}{2}\right)$$

e facciamo tendere $l \to +\infty$, ottenendo

$$\widehat{\phi}(x) + \widehat{\psi}(y) \ge \frac{|x+y|^2}{2} - \left(\frac{|x|^2}{2} + \frac{|y|^2}{2}\right),$$

cioè

$$\widehat{\phi}(x) + \widehat{\psi}(y) \ge x \cdot y$$

dunque $(\widehat{\phi}, \widehat{\psi}) \in \widetilde{\Phi}$, ed è una coppia minimizzante.
Applicando il Lemma 3.2.1 si conclude. $\square$

3.3 Teorema di Brenier

Richiamiamo preliminarmente:

Definizione 3.3.1. (**Assoluta continuità di misure**)
Date μ, ν due misure, diciamo μ è **assolutamente continua** rispetto a ν, scriviamo $\mu \ll \nu$, se per ogni A tale che $\nu(A) = 0$ vale $\mu(A) = 0$.

Siamo ora pronti per dimostrare:

Teorema 3.3.1. (**Trasporto ottimo per costo quadratico, Brenier**)
Date μ, ν due misure di probabilità su $\mathbb{R}^n$ con momento quadratico finito, consideriamo il problema di Monge-Kantorovich associato alla funzione $c(x,y) = \frac{1}{2}|x-y|^2$.
Vale allora:

(*i*) (**Criterio di ottimalità di Knott-Smith**) $\pi \in \Pi(\mu,\nu)$ è ottimale se e solo se esiste una funzione ϕ convessa s.c.i. tale che $\mathrm{spt}(\pi) \subset \mathrm{gr}(\partial\phi)$, ovvero se per π-quasi ogni (x,y) vale $y \in \partial\phi(x)$.
Inoltre la coppia (ϕ, ϕ^*) minimizza il problema

$$\inf\left\{\int_{\mathbb{R}^n} \phi \,\mathrm{d}\mu + \int_{\mathbb{R}^n} \psi \,\mathrm{d}\nu : (\phi,\psi) \in \widetilde{\Phi} \text{ per ogni } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n\right\}.$$

(*ii*) (**Breiner**) Se $\mu \ll \mathcal{L}^n$, allora esiste un unico piano di trasporto ottimale π tale che $\pi = (\mathrm{Id} \times \nabla\phi)_{\#}\mu$, dove $\nabla\phi$ è l'unico gradiente di una funzione convessa definito μ-quasi ovunque tale che $\nabla\phi_{\#}\mu = \nu$.

(*iii*) Con le ipotesi del punto (*ii*) si ha che $\nabla\phi$ è l'unica soluzione del problema di Monge con costo quadratico:

$$\int_{\mathbb{R}^n} |x - \nabla\phi(x)|^2 \,\mathrm{d}\mu(x) = \inf_{T \in \mathcal{T}(\mu,\nu)} \int_{\mathbb{R}^n} |x - T(x)|^2 \,\mathrm{d}\mu(x),$$

o equivalentemente,

$$\int_{\mathbb{R}^n} x \cdot \nabla\phi(x) \,\mathrm{d}\mu(x) = \sup_{T \in \mathcal{T}(\mu,\nu)} \int_{\mathbb{R}^n} x \cdot T(x) \,\mathrm{d}\mu(x).$$

(*iv*) Se $\mu \ll \mathcal{L}^n$ e $\nu \ll \mathcal{L}^n$, allora per μ-quasi ogni x e ν-quasi ogni y si ha

$$\nabla\phi^* \circ \nabla\phi(x) = x, \qquad \nabla\phi \circ \nabla\phi^*(y) = y,$$

e $\nabla\phi^*$ è l'unico gradiente di una funzione convessa definita ν-quasi ovunque che trasporta la misura ν su μ; inoltre è l'unica soluzione del problema di Monge per $T \in \mathcal{T}(\nu,\mu)$ con funzione costo quadratico.

Dimostrazione. (i) La condizione ottimalità per π è già stata presentata nell'Osservazione della Sezione 1.2. Abbiamo provato l'esistenza di un minimo per il problema di Kantorovitch (Teorema 1.3.2), che sappiamo essere realizzato da una coppia di funzioni convesse coniugate (Teorema 3.2.2), allora per considerazioni fatte in precedenza abbiamo

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^n}\phi\,\mathrm{d}\mu+\int_{\mathbb{R}^n}\phi^*\,\mathrm{d}\nu=\int_{\mathbb{R}^n\times\mathbb{R}^n}(\phi+\phi^*)\,\mathrm{d}\pi(x,y),$$

ovvero

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(\phi(x)+\phi^*(y)-x\cdot y)\,\mathrm{d}\pi(x,y)=0.$$

La funzione integranda deve essere nulla per π-quasi ogni (x,y), dunque per la caraterizzazione del sottodifferenziale (Proposizione 2.4.1) vale $y\in\partial\phi(x)$ per π-quasi ogni (x,y).
Data invece $\pi\in\Pi(\mu,\nu)$ tale che $y\in\partial\phi(x)$ per π-quasi ogni (x,y) abbiamo, con le stesse argomentazioni di sopra,

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^n}\phi\,\mathrm{d}\mu+\int_{\mathbb{R}^n}\phi^*\,\mathrm{d}\nu,$$

dunque il Teorema di Dualità (Teorema 3.1.1) assicura la minimalità della coppia (ϕ,ϕ^*).
(ii) Supponiamo ora $\mu\ll\mathcal{L}^n$. Vale $\phi\in L^1(\mu)$, quindi ϕ è finita μ-quasi ovunque, ovvero $\mu(\mathrm{Dom}(\phi))=1$, inoltre $\partial\,\mathrm{Dom}(\phi)$ è un insieme $\mathcal{L}^n$-trascurabile e dunque anche μ-trascurabile; allora vale $\mu(\mathrm{Int}(\mathrm{Dom}(\phi)))=1$. Su $\mathrm{Int}(\mathrm{Dom}(\phi))$ l'insieme di non differenziabilità risulta essere μ-trascurabile per il teorema di Rademacher, quindi per μ-quasi ogni $x\in\mathbb{R}^n$ il sottogradiente $\partial\phi(x)$ risulta essere il singoletto $\{\nabla\phi(x)\}$. Questo è vero per μ-quasi ogni $x\in\mathbb{R}^n$ e quindi anche per π-quasi ogni $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$, otteniamo allora $y=\nabla\phi(x)$ per π-quasi ogni $(x,y)\in\mathbb{R}^n\times\mathbb{R}^n$. Sappiamo dal punto precedente che il piano di trasporto associato è della forma $(\mathrm{Id}\times\nabla\phi)_{\#}\mu$, per qualche funzione convessa ϕ tale che $\nabla\phi_{\#}\mu=\nu$. Mostriamo ora che tale piano di trasporto è unico: consideriamo $\overline{\phi}$ una funzione convessa tale che $\nabla\overline{\phi}_{\#}\mu=\nu$: sappiamo che $(\mathrm{Id}\times\nabla\phi)_{\#}\mu$ è un piano di trasporto ottimo, sia $(\overline{\phi},\overline{\phi}^*)$ una coppia una coppia che realizza il massimo nel problema duale, come (ϕ,ϕ^*). Allora

$$\int_{\mathbb{R}^n}\overline{\phi}\,\mathrm{d}\mu+\int_{\mathbb{R}^n}\overline{\phi}^*\,\mathrm{d}\nu=\int_{\mathbb{R}^n}\phi\,\mathrm{d}\mu+\int_{\mathbb{R}^n}\phi^*\,\mathrm{d}\nu.$$

Sia $\pi=(\mathrm{Id}\times\nabla\phi)_{\#}\mu$, allora

$$\int_{\mathbb{R}^n\times\mathbb{R}^n}(\overline{\phi}+\overline{\phi}^*)\,\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}(\phi+\phi^*)\,\mathrm{d}\pi(x,y)=\int_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)\,\mathrm{d}\pi(x,y),$$

che vale a dire, ricordando la definizione di π,

$$\int_{\mathbb{R}^n}(\overline{\phi}(x)+\overline{\phi}^*(\nabla\phi(x)))\,\mathrm{d}\mu(x)=\int_{\mathbb{R}^n}(x\cdot\nabla\phi(x))\,\mathrm{d}\mu(x),$$

quindi

$$\int_{\mathbb{R}^n} (\overline{\phi}(x) + \overline{\phi}^*(\nabla\phi(x)) - x \cdot \nabla\phi(x))\,\mathrm{d}\mu(x) = 0.$$

La funzione integranda è non negativa, quindi deve essere nulla ν-quasi ovunque, quindi per la Proposizione 2.4.1 abbiamo $\nabla\phi(x) \in \partial\overline{\phi}(x)$ per μ-quasi ogni x, ed essendo $\overline{\phi}$ differenziabile deve essere $\nabla\phi(x) = \nabla\overline{\phi}(x)$ per μ-quasi ogni x. Dunque abbiamo, a meno di un insieme μ-trascurabile, $\nabla\phi = \nabla\overline{\phi}$, e in questo modo abbiamo mostrato sia l'unicità della soluzione del problema di Monge-Kantorovich sia l'unicità del gradiente della funzione convessa ϕ.
(iii) Discende dai punti precedenti.
(iv) Vale $y = \nabla\phi(x)$ π-quasi ovunque, quindi $x \in \partial\phi^*(y)$, ed essendo ϕ^* finita ν-quasi ovunque risulta differenziabile ν-quasi ovunque. Dunque vale $x = \nabla\phi^*(y) = \nabla\phi^*(\nabla\phi(x))$ per π-quasi ogni (x, y). Considerando primo e ultimo termine troviamo che vale $x = \nabla\phi^*(\nabla\phi(x))$ per μ-quasi ogni x. Analogamente otteniamo l'altra parte. $\square$

Bibliografia

[V] Villani C.: *Topics in optimal transportation*, American Mathematical Society, 2003

[M] Monti R.: *Introduzione al Calcolo delle Variazioni*, *Manuscripta*, 2017

[E] Evans L.C., Gariepy R.F.: *Measure Theory and Fine Properties of Functions. Revised Edition*, CRC Press, 2015

[Br] Brezis H.: *Analisi funzionale. Teoria e applicazioni*, Liguori, 1986

[Bo] Bourbaki N.: *Topological Vector Spaces*, Springer, 1987

[R] Rockafellar R.T.: *Convex Analysis*, Princeton University Press, 1972

[AFP] Ambrosio L., Fusco N., Pallara D.: *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, 1999

[F] Folland G.B.: *Real Analysis: Modern Techniques and Their Applications*, Wiley-Interscience, 1999