Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Triennale in Matematica

Teorema di Alexandrov ed esistenza quasi ovunque delle derivate seconde di funzioni convesse

Relatore :
Prof. Roberto Monti

Laureando: Pietro Borsarini
Matricola: 2046384

Anno Accademico 2023/2024

20 settembre 2024

Indice

Introduzione

I primi risultati riguardanti le derivate seconde di funzioni convesse risalgono al 1935, quando H. Busemann e W. Feller ne mostrarono l'esistenza quasi ovunque (rispetto alla misura di Lebesgue in $\mathbb{R}^n$) nel caso di funzioni in due variabili. Nel 1939, A.D. Alexandrov non solo dimostrò questa proprietà nel caso generale delle funzioni in più variabili, ma arrivò anche a provare la differenziabilità quasi ovunque del gradiente di una funzione convessa: questo fatto assunse successivamente il nome di Teorema di Alexandrov. Nel corso dei decenni, numerosi matematici hanno affrontato il problema con diversi approcci, uno dei quali proposto da Yu. G. Reshetnyak in [1], facente utilizzo prevalentemente di strumenti classici di Teoria della Misura.
L'obiettivo del presente lavoro è quello di ripercorrere l'iter dimostrativo seguito da L.C. Evans e R. Gariepy in [3], a sua volta basato sui lavori di Yu. G. Reshetnyak. Nella prima parte, grazie al Teorema di ricoprimento di Besicovitch, mostreremo alcuni celebri risultati dell'Analisi Reale (Lemma di Radon-Nikodym, Teorema di Decomposizione di Lebesgue, Teorema di differenziazione di Lebesgue-Besicovitch) e introdurremo le idee fondamentali circa le misure a valori matriciali e la loro decomposzione polare.
Nella seconda parte, presenteremo alcune proprietà fondamentali delle funzioni convesse, quali la locale lipschitzianità e alcune stime locali (come (3.3.1), (3.3.2)), che saranno essenziali nel finale. Il cuore del lavoro risiede, però, nel Teorema 3.4: una funzione convessa ammette delle derivate seconde in senso debole che derivano da misure (con segno) di Radon, le quali costituiscono una matrice simmetrica di misure. La loro origine va ricercata nel Teorema di Rappresentazione di Riesz, che, tramite il Corollario 3.1, caratterizza i funzionali lineari non negativi attraverso delle misure di Radon. Dalla decomposizione di Lebesgue della matrice ottenuta, ricaviamo la 'Matrice Hessiana Generalizzata' $D^2 f$, che, con la dimostrazione del Teorema 4.1, proveremo essere l'Hessiana quasi ovunque della funzione convessa f.

Capitolo 1

Differenziazione di misure di Radon

1.1 Derivate di misure

In generale, indicheremo con $B(x,r)$ la palla chiusa in $\mathbb{R}^n$ di centro $x \in \mathbb{R}^n$ e raggio $r > 0$ e $\alpha(n) := \mathcal{L}^n(B(r))$, dove $\mathcal{L}^n$ è la misura di Lebesgue $n-$dimensionale.
Inoltre, a meno che non sia diversamente specificato, lavoreremo con misure μ, ν esterne: dunque, se ristette alle $\sigma-$algebre dei $\mu($o $\nu)-$misurabili, sono misure complete per il Teorema di Carathéodory. La completezza delle misure sarà un dettaglio fondamentale per i Teoremi 1.2 e 1.3.
In particolare, lavoreremo con misure di **Radon** su $\mathbb{R}^n$, ossia misure di Borel $\mu : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$ che siano anche:
(i) **Borel regolari**: comunque preso $A \subseteq \mathbb{R}^n$ esiste un boreliano B tale che $A \subseteq B$ e $\mu(A) = \mu(B)$;
(ii) **Finite sui compatti**: per ogni $K \subseteq \mathbb{R}^n$ compatto, $\mu(K) < +\infty$.

Definizione 1.1. Comunque preso $x \in \mathbb{R}^n$, definiamo

$$\overline{D}_\mu \nu(x) := \begin{cases} \limsup\limits_{r \to 0^+} \dfrac{\nu(B(x,r)}{\mu(B(x,r)} & \text{se } \mu(B(x,r)) > 0 \text{ per ogni } r > 0 \\ +\infty & \text{se esiste } r > 0 \text{ tale che } \mu(B(x,r)) = 0, \end{cases}$$

$$\underline{D}_\mu \nu(x) := \begin{cases} \liminf\limits_{r \to 0^+} \dfrac{\nu(B(x,r))}{\mu(B(x,r))} & \text{se } \mu(B(x,r)) > 0 \text{ per ogni } r > 0 \\ +\infty & \text{se esiste } r > 0 \text{ tale che } \mu(B(x,r)) = 0. \end{cases}$$

Definizione 1.2. Diciamo che ν è **differenziabile** rispetto a μ nel punto $x \in \mathbb{R}^n$ se $\overline{D}_\mu\nu(x) = \underline{D}_\mu\nu(x) < +\infty$ e definiamo $D_\mu\nu(x) := \overline{D}_\mu\nu(x) = \underline{D}_\mu\nu(x)$, detta **derivata** di ν rispetto a μ nel punto $x \in \mathbb{R}^n$.

Enunciamo ora due risultati che saranno fondamentali per giungere a una dimostrazione dei teoremi di differenziazione di fine capitolo. In particolare, il Teorema di ricoprimento di Besicovitch presenta una dimostrazione di natura geometrica che non fa però utilizzo di strumenti avanzati di teoria della misura. Le dimostrazioni, che omettiamo, si possono trovare in [3], pp. 30-37.

Teorema 1.1 (Teorema di ricoprimento di Besicovitch)**.** *Esiste una costante N_n, che dipende soltanto da n, per la quale vale:*
data $\mathcal{F}$ famiglia qualsiasi di palle chiuse in $\mathbb{R}^n$ (non degeneri) tale che

$$\sup\{\operatorname{diam}(B) \mid B \in \mathcal{F}\} < +\infty$$

e A l'insieme dei centri delle palle di $\mathcal{F}$, allora esistono $\mathcal{G}_1, ..., \mathcal{G}_{N_n} \subseteq \mathcal{F}$ tali che ciascuno dei $\mathcal{G}_i$ al variare di i in $\{1, ..., N_n\}$ è una famiglia numerabile di palle disgiunte di $\mathcal{F}$ e

$$A \subseteq \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{G}_i} B.$$

Corollario 1.1. *Siano μ misura di Borel su $\mathbb{R}^n$, $\mathcal{F}$ una qualsiasi famiglia di palle chiuse non degeneri, A l'insieme dei centri delle palle in $\mathcal{F}$ e supponiamo che $\mu(A) < +\infty$ e $\inf\{r \mid B(a,r) \in \mathcal{F}\} = 0$ per ogni $a \in A$. Allora, comunque preso un sottoinsieme aperto U di $\mathbb{R}^n$, esiste $\mathcal{G}$ sottofamiglia numerabile disgiunta di $\mathcal{F}$ tale che $\bigcup_{B \in \mathcal{G}} B \subseteq U$ e*

$$\mu\left((A \cap U) - \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

Il precedente corollario permette di dimostrare il lemma tecnico che segue, il quale rivestirà un ruolo chiave nell'ottenimento dei teoremi di differenziazione.

Lemma 1.1. *Dati $\alpha \in (0, +\infty)$ e $A \subseteq \mathbb{R}^n$, segue che:*
(i) Se $A \subseteq \{x \in \mathbb{R}^n \mid \underline{D}_\mu \nu(x) \leq \alpha\}$, allora $\nu(A) \leq \alpha\mu(A)$,
(ii) Se $A \subseteq \{x \in \mathbb{R}^n \mid \overline{D}_\mu \nu(x) \geq \alpha\}$, allora $\nu(A) \geq \alpha\mu(A)$.

Dimostrazione. (i) A meno di restringerci ad un compatto di $\mathbb{R}^n$, possiamo supporre, essendo ν e μ misure di Radon, che $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < +\infty$. Decomponendo poi $\mathbb{R}^n$ in un unione disgiunta di compatti, otteniamo il caso generale.
Dato A come nell'ipotesi, fissiamo $\epsilon > 0$, U aperto di $\mathbb{R}^n$ tale che $A \subseteq U$ e definiamo

$$\mathcal{F} := \{B \mid B = B(a, r), a \in A, r > 0, B \subseteq U, \nu(B) \leq (\alpha + \epsilon)\mu(B)\}.$$

Osserviamo che $\mathcal{F}$ è non vuoto e per ogni $a \in A$, poiché $\alpha < +\infty$ e $\underline{D}_\mu \nu(x) \leq \alpha$, comunque presi $\epsilon > 0$, $r_0 > 0$ esiste $r \in (0, r_0)$ tale che $\frac{\nu(B(a,r))}{\mu(B(a,r))} \leq \alpha + \epsilon$, da cui

$$\inf\{r > 0 | B(a, r) \in \mathcal{F}\} = 0.$$

Avendo supposto che $\nu(A) < +\infty$, otteniamo, dal Corollario 1.1, che esiste una sottofamiglia $\mathcal{G}$ di $\mathcal{F}$ numerabile e disgiunta tale che

$$\nu\Big(A - \bigcup_{B \in \mathcal{G}} B\Big) = 0.$$

Quindi

$$\begin{aligned}\nu(A) &\leq \nu\Big(\bigcup_{B \in \mathcal{G}} B\Big) + \nu\Big(A - \bigcup_{B \in \mathcal{G}} B\Big) = \sum_{B \in \mathcal{G}} \nu(B) \leq \sum_{B \in \mathcal{G}} (\alpha + \epsilon)\mu(B) \\ &= (\alpha + \epsilon)\mu\Big(\bigcup_{B \in \mathcal{G}} B\Big) \leq (\alpha + \epsilon)\mu(U).\end{aligned}$$

Per il Teorema di regolarità dall'esterno, si ha che

$$\nu(A) \leq (\alpha + \epsilon) \inf\{\mu(U) | A \subseteq U,\ U \text{ aperto}\} = (\alpha + \epsilon)\mu(A).$$

Infine, data l'arbitrarietà nella scelta di ϵ, concludiamo che $\nu(A) \leq \alpha\mu(A)$.
(ii) La dimostrazione è analoga a quella di (i). □

Osservazione 1.1. Il lemma precedente non richiede né la $\mu-$misurabilità né la $\nu-$misurabilità di A: ripercorrendo la dimostrazione si osserva che ciò discende dal fatto che la regolarità dall'esterno è valida per qualsiasi sottoinsieme di $\mathbb{R}^n$. Lo stesso non vale per la regolarità dall'interno.

Teorema 1.2. *Date μ, ν misure di Radon su $\mathbb{R}^n$, $D_\mu \nu$ esiste finita per $\mu-$q.o. $x \in \mathbb{R}^n$. Di più: la funzione $D_\mu \nu$ gode di $\mu-$misurabilità.*

Dimostrazione. : Suddividiamo la dimostrazione in tre passi:
Passo 1. Vogliamo mostrare che $D_\mu \nu(x)$ esiste finita $\mu-$q.o.
Possiamo suppore $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < +\infty$: se così non fosse, potremmo decomporre $\mathbb{R}^n$ in un'unione numerabile disgiunta di compatti e lavorare con μ e ν ristrette a ciascuno di essi.
Definiamo

$$I := \{x \in \mathbb{R}^n \mid \overline{D}_\mu \nu(x) = +\infty\}$$

e, comunque dati $0 < a < b \in \mathbb{R}$,

$$R(a,b) := \{x \in \mathbb{R}^n \mid \underline{D}_\mu \nu(x) < a < b < \overline{D}_\mu \nu(x) < +\infty\}.$$

Osserviamo che, poiché $I \subseteq \{x \in \mathbb{R}^n \mid \overline{D}_\mu \nu(x) \geq \alpha\}$ per ogni $\alpha > 0$, segue, per il Lemma 1.1, che $\nu(I) \geq \alpha \mu(I)$, cioè $\mu(I) \leq \frac{\nu(I)}{\alpha}$ per ogni $\alpha > 0$. Avendo che $\nu(I) < +\infty$, possiamo passare al limite per $\alpha \to +\infty$, ottenendo che $\mu(I) = 0$. Sempre in forza del Lemma 1.1, osserviamo che $\nu(R(a,b)) \leq a\mu(R(a,b))$ e $\nu(R(a,b)) \geq b\mu(R(a,b))$, da cui

$$\frac{\nu(R(a,b))}{b} \leq \frac{\nu(R(a,b))}{a} \leq \mu(R(a,b)) \leq \frac{\nu(R(a,b))}{b},$$

da cui chiaramente (essendo $\mu(R(a,b)) < +\infty$) $\mu(R(a,b)) = 0$ comunque presi $0 < a < b \in \mathbb{R}$. Osservato che

$$N := \{x \mid \underline{D}_\mu \nu(x) < \overline{D}_\mu \nu(x) < \infty\}) = \bigcup_{0<a<b\in\mathbb{Q}} R(a,b),$$

otteniamo

$$\mu(N) \leq \sum_{0<a<b\in\mathbb{Q}} \mu(R(a,b)) = 0,$$

concludendo che $D_\mu \nu$ esiste ed è finito $\mu-$q.o.
Passo 2. Vogliamo provare che, comunque presi $x \in \mathbb{R}^n$ e $r > 0$, valgono:
(a) $\limsup_{y\to x} \mu(B(y,r)) \leq \mu(B(x,r))$;
(b) $\limsup_{y\to x} \nu(B(y,r)) \leq \nu(B(x,r))$.
Mostriamo soltanto (a), dal momento che la dimostrazione di (b) è del tutto analoga.

Possiamo scegliere $\{y_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$ tale che $y_k \to x$ e definire $f := \chi_{B(x,r)}$, $f_k := \chi_{B(y_k,r)}$ per ogni $k \in \mathbb{N}$. Allora

$$\limsup_{k\to+\infty} f_k \leq f, \text{ossia } 1-f \leq \liminf_{k\to+\infty}(1-f_k) : \tag{1.2.1}$$

-se $y \notin B(x,r)$, allora $|y-x| > r$ e, poiché $y_k \to x$, esiste $\bar{k} \in \mathbb{N}$ tale che per ogni $k \geq \bar{k}$ si ha $|y_k - x| < |y-x| - r$, da cui per ogni $k \geq \bar{k}$ $|y - y_k| \geq |y-x| - |y_k - x| > r$, cioè $y \notin B(y_k, r)$. Quindi per ogni $k \geq \bar{k}$ $\;f_k(y) = 0 = f(y)$ e $0 = \limsup\limits_{k\to+\infty} f_k(y) \leq f(y) = 0$;
-se $y \in B(x,r)$, allora $f_k(y) \leq 1 = f(y)$ per ogni $k \in \mathbb{N}$, da cui $\limsup\limits_{k\to+\infty} f_k(y) \leq f(y)$.
Utilizziamo il Lemma di Fatou e otteniamo:

$$\mu(B(x,2r)) - \mu(B(x,r)) = \int_{B(x,2r)} (1-f)\ d\mu \overset{(1.2.1)}{\leq} \int_{B(x,2r)} \liminf_{k\to+\infty}(1-f_k)\ d\mu$$

$$\leq \liminf_{k\to+\infty} \int_{B(x,2r)} (1-f_k)\ d\mu = \liminf_{k\to+\infty}(\mu(B(x,2r)) - \mu(B(y_k,r))$$

$$\overset{(*)}{=} \mu(B(x,2r)) - \limsup_{k\to+\infty} \mu(B(y_k,r))$$

e, quindi, essendo che $\mu(B(x,2r)) < +\infty$ (in quanto μ è misura di Radon),

$$\limsup_{k\to+\infty} \mu(B(y_k,r)) \leq \mu(B(x,r)).$$

Data l'arbitrarietà nella scelta di $\{y_k\}_{k\in\mathbb{N}}$, possiamo concludere, per caratterizzazione sequenziale del limite, che $\limsup\limits_{y\to x} \mu(B(y,r)) \leq \mu(B(x,r))$.
Passo 3. Avendo mostrato al Passo 2 che, comunque preso $r > 0$, le funzioni $x \to \mu(B(x,r))$ e $x \to \nu(B(x,r))$ sono superiormente semicontinue, abbiamo che sono boreliane. Ha quindi senso definire, per $r > 0$, la funzione

$$f_r(x) := \begin{cases} \dfrac{\nu(B(x,r))}{\mu(B(x,r))} & \text{se } \mu(B(x,r)) > 0, \\ +\infty & \text{se } \mu(B(x,r)) = 0. \end{cases}$$

che è $\mu-$misurabile.
Quindi, poiché

$$\overline{D}_\mu \nu(x) = \limsup_{r\to 0^+} f_r(x) = \limsup_{k\to+\infty} f_{\frac{1}{k}}(x),$$

deduciamo che $\overline{D}_\mu\nu(x)$ è funzione $\mu-$misurabile, in quanto limite superiore di una successione di funzioni $\mu-$misurabili.
Per il Passo 1, $D_\mu\nu = \overline{D}_\mu\nu \quad \mu-$q.o. e μ, se pensata come misura sulla $\sigma-$algebra degli insiemi $\mu-$misurabili, è misura completa (Teorema di Carathéodory): quindi anche $D_\mu\nu$ è $\mu-$misurabile. □

Osservazione 1.2. Nella dimostrazione del Teorema 1.2:
-abbiamo usato in (*) che, date $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ tali che $a_n \to l \in \mathbb{R}$, valgono:

$$\liminf_{n\to+\infty}(a_n + b_n) = l + \liminf_{n\to+\infty} b_n \text{ e } \liminf_{n\to+\infty}(-a_n) = -\limsup_{n\to+\infty} a_n.$$

-Nel Passo 3, si è usato il fatto che una funzione superiormente semicontinua sia boreliana e ciò può essere provato mostrando, tramite la definizione di semicontinuità superiore, che gli insiemi di sottolivello (stretto) di f sono aperti, quindi boreliani.

1.2 Teoremi di Differenziazione

Ricordiamo che, date μ, ν misure di Radon su $\mathbb{R}^n$:
(i) $\nu \ll \mu$ se, comunque preso A $\mu-$misurabile tale che $\mu(A) = 0$, $\nu(A) = 0$.
(ii) $\mu \perp \nu$ se esiste $B \subseteq \mathbb{R}^n$ boreliano tale che $\mu(\mathbb{R}^n - B) = \nu(B) = 0$.

Teorema 1.3 (Lemma di Radon-Nikodym)**.**
Siano μ, ν misure di Radon su $\mathbb{R}^n$ e $\nu \ll \mu$. Allora

$$\nu(A) = \int_A D_\mu\nu \, d\mu$$

comunque preso A insieme $\mu-$misurabile.

Dimostrazione. Mostriamo che ogni A $\mu-$misurabile è anche $\nu-$misurabile. Comunque preso, infatti, A $\mu-$misurabile, poiché μ è misura di Radon, esiste un boreliano B che contenga A tale che $\mu(B - A) = 0$. Poiché $\nu \ll \mu$, $\nu(B - A) = 0$, da cui $B - A$ è $\nu-$misurabile e, quindi, lo è $A = B - (B - A)$.
Definiamo anche gli insiemi

$$Z := \{x \in \mathbb{R}^n \mid D_\mu\nu(x) = 0\},$$
$$I := \{x \in \mathbb{R}^n \mid D_\mu\nu(x) = +\infty\}.$$

Per il Teorema 1.2, $\mu(I) = 0$, da cui $\nu(I) = 0$. Inoltre, comunque preso $\alpha > 0$, si ha, per il Lemma 1.1, che $\nu(Z) \leq \alpha\mu(Z)$. Dunque, se $\mu(Z) < +\infty$, facendo tendere α a 0, si ottiene che $\nu(Z) = 0$. Se, invece, $\mu(Z) = +\infty$ è possibile decomporre $\mathbb{R}^n$ in un'unione numerabile di compatti e sfruttare il fatto che μ sia misura di Radon. In ogni caso, abbiamo dedotto che

$$\nu(I) = \int_I D_\mu\nu \; d\mu = 0 = \nu(Z) = \int_Z D_\mu\nu \; d\mu.$$

Sia $N := \{x \in \mathbb{R}^n \mid \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x)\}$: per il Teorema 1.2, $\mu(N) = 0$, quindi $\nu(N) = 0$.
Prendiamo A $\mu-$misurabile e $t \in (1, +\infty)$. Definiamo per $m \in \mathbb{Z}$ l'insieme

$$A_m := \{x \in A \mid t^m \leq D_\mu\nu(x) < t^{m+1}\}.$$

Dal momento che $D_\mu\nu$ è, per il Teorema 1.2, $\mu-$misurabile, si ha che per ogni $m \in \mathbb{Z}$ A_m è $\mu-$(e, dunque, $\nu-$)misurabile. Inoltre, per il Lemma 1.1, per ogni $m \in \mathbb{Z}$ valgono

$$\begin{aligned} \nu(A_m) &\leq t^{m+1}\mu(A_m) \\ \nu(A_m) &\geq t^m\mu(A_m). \end{aligned} \tag{1.3.1}$$

Poiché $t \in (1, +\infty)$, notiamo che $t^m \to +\infty$, $t^{-m} \to 0$ per $m \to +\infty$, da cui

$$A = \bigcup_{m=-\infty}^{+\infty} A_m \cup ((Z \cup I \cup N) \cap A).$$

Osservato che $\nu(Z \cup I \cup N) = \nu(Z) + \nu(I) + \nu(N) = 0$ e

$$\int_A D_\mu\nu \; d\mu = \int_{\bigcup_{m=-\infty}^{+\infty} A_m} D_\mu\nu \; d\mu = \sum_{m=-\infty}^{+\infty} \int_{A_m} D_\mu\nu \; d\mu,$$

si ha, per (1.3.1), che:

$$\begin{aligned} \nu(A) = \sum_{m=-\infty}^{+\infty} \nu(A_m) &\leq \sum_{m=-\infty}^{+\infty} t^{m+1}\mu(A_m) = \sum_{m=-\infty}^{+\infty} t\int_{A_m} t^m \; d\mu \\ &\leq \sum_{m=-\infty}^{+\infty} t\int_{A_m} D_\mu\nu \; d\mu = t\int_A D_\mu\nu \; d\mu. \end{aligned}$$

Analogamente, utilizzando (1.3.1), otteniamo:

$$\begin{aligned}\nu(A) = \sum_{m=-\infty}^{+\infty} \nu(A_m) \geq \sum_{m=-\infty}^{+\infty} t^m \mu(A_m) = \sum_{m=-\infty}^{+\infty} \frac{1}{t} \int_{A_m} t^{m+1}\, d\mu \\ \geq \sum_{m=-\infty}^{+\infty} \frac{1}{t} \int_{A_m} D_\mu \nu\, d\mu = \frac{1}{t} \int_A D_\mu \nu\, d\mu.\end{aligned}$$

Infine, per $t \to 1^+$, otteniamo l'uguaglianza dell'enunciato. □

Teorema 1.4 (Teorema di decomposizione di Lebesgue)**.** *Date μ e ν misure di Radon su $\mathbb{R}^n$:*
(i) Esistono ν_{ac}, ν_s misure di Radon su $\mathbb{R}^n$ tali che $\nu = \nu_{ac} + \nu_s$, $\nu_{ac} \ll \mu$ e $\nu_s \perp \mu$.
(ii) $D_\mu \nu = D_\mu \nu_{ac}$ e $D_\mu \nu_s = 0$ $\mu-$q.o. e, quindi, per ogni $A \subseteq \mathbb{R}^n$ boreliano,

$$\nu(A) = \int_A D_\mu \nu\, d\mu + \nu_s(A).$$

Dimostrazione. Come nel Teorema 1.2, possiamo supporre, senza perdita di generalità, che $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < +\infty$.
(i) Definiamo $\mathcal{E} := \{A \subseteq \mathbb{R}^n \mid A$ boreliano e $\mu(\mathbb{R}^n - A) = 0\}$: osserviamo che, dal momento che $\mu(\mathbb{R}^n - \mathbb{R}^n) = \mu(\oslash) = 0$, $\mathbb{R}^n \in \mathcal{E}$, da cui $\mathcal{E} \neq \oslash$ e, da $\nu(\mathbb{R}^n) < +\infty$, deduciamo che esiste $M < +\infty$ tale che $\inf_{A\in\mathcal{E}} \nu(A) = M$.
Dunque per ogni $k \in \mathbb{N}$ esiste $B_k \in \mathcal{E}$ tale che $\nu(B_k) < M + \frac{1}{k}$. Detto $B := \bigcap_{k\in\mathbb{N}} B_k$, notiamo che B è un boreliano e

$$\mu(\mathbb{R}^n - B) = \mu\left(\mathbb{R}^n - \bigcap_{k\in\mathbb{N}} B_k\right) = \mu\left(\bigcup_{k\in\mathbb{N}} (\mathbb{R}^n - B_k)\right) \leq \sum_{k\in\mathbb{N}} \mu(\mathbb{R}^n - B_k) = 0,$$

da cui $B \in \mathcal{E}$ e $\nu(B) = \inf_{A\in\mathcal{E}} \nu(A) = \min_{A\in\mathcal{E}} \nu(A)$. Definiamo

$$\nu_{ac} := \nu|_B,\ \nu_s := \nu|_{\mathbb{R}^n - B};$$

che sono misure di Radon su $\mathbb{R}^n$ (vedi Osservazione 1.3).
Dato $A \subseteq B$ Boreliano e $\mu(A) = 0$, allora $\nu(A) = 0$. Se, per assurdo, $\nu(A) > 0$, si avrebbe che

$$\mu(\mathbb{R}^n - (B - A)) = \mu(A) + \mu(\mathbb{R}^n - B) = 0,$$

da cui $B - A \in \mathcal{E}$; ma $\nu(B - A) = \nu(B) - \nu(A) < \nu(B) = \min_{E\in\mathcal{E}} \nu(E)$, che è assurdo. Segue che $\nu_{ac} \ll \mu$. Inoltre, essendo che $\mu(\mathbb{R}^n - B) = 0$, è vero che $\nu_s \perp \mu$.
(ii) Rimane da provare che $D_\mu \nu_s = 0$ $\mu-$q.o.. Per ogni $\alpha > 0$ definiamo

$$C_\alpha := \{x \in B \mid D_\mu \nu_s \geq \alpha\}$$

e, usando $\nu_s \perp \mu$ e il Lemma 1.1, otteniamo $\alpha\mu(C_\alpha) \leq \nu_s(C_\alpha) = 0$ per ogni $\alpha > 0$, da cui $\mu(C_\alpha) = 0$. Quindi

$$\mu(\{x \in \mathbb{R}^n \mid D_\mu \nu_s(x) > 0\}) = \mu\left(\bigcup_{n\geq 1} C_{\frac{1}{n}}\right) \leq \sum_{n\geq 1} \mu(C_{\frac{1}{n}}) = 0,$$

per cui $D_\mu \nu_s = 0$ $\mu-$q.o. e $D_\mu \nu = D_\mu \nu_s + D_\mu \nu_{ac} = D_\mu \nu_{ac}$ $\mu-$q.o.
Possiamo concludere che, comunque preso $A \subseteq \mathbb{R}^n$ boreliano, vale

$$\nu(A) = \nu_{ac}(A) + \nu_s(A) = \int_A D_\mu \nu_{ac}\, d\mu + \nu_s(A) = \int_A D_\mu \nu\, d\mu + \nu_s(A),$$

come volevamo dimostrare. $\square$

Osservazione 1.3. La finitezza sui compatti segue dal fatto che $\nu|_B, \nu|_{\mathbb{R}^n - B}$ siano restrizioni di ν, che è finita sui compatti. Inoltre, che le restrizioni siano ancora misure di Borel è facile da dedurre, visto che B e $\mathbb{R}^n - B$ sono Boreliani.
Rimane da vedere che le due misure siano Borel regolari.

Teorema 1.5 (Teorema di differenziazione di Lebesgue-Besicovitch)**.** *Data μ misura di Radon su $\mathbb{R}^n$ e $f \in L^1_{loc}(\mathbb{R}^n, \mu)$, allora, per $\mu-$q.o. $x \in \mathbb{R}^n$, vale:*

$$\lim_{r\to 0} ⨍_{B(x,r)} f\, d\mu = f(x). \tag{1.5.1}$$

*Chiameremo **punti di Lebesgue** di f quelli per cui vale (1.5.1).*

Dimostrazione. Ricordiamo che, per $0 < \mu(E) < +\infty$,

$$⨍_E f\, d\mu = \frac{1}{\mu(E)} \int_E f\, d\mu.$$

Decomponiamo $f := f^+ - f^-$ e definiamo per i boreliani $B \subseteq \mathbb{R}^n$ $\nu^\pm(B) := \int_B f^\pm d\mu$ e per ogni $A \subseteq \mathbb{R}^n$ diciamo che $\nu^\pm(A) = \inf\{\nu^\pm(B) \mid A \subseteq B, B \text{ boreliano}\}$.

Dal momento che $f \in L^1_{loc}(\mathbb{R}^n, \mu)$, si ha che ν^+, ν^- sono misure di Radon (positive) su $\mathbb{R}^n$ tali che $\nu^+, \nu^- \ll \mu$. Quindi, per il Teorema 1.3, comunque si prenda A $\mu-$misurabile, valgono:

$$\nu^{\pm}(A) = \int_A f^{\pm}\ d\mu = \int_A D_\mu \nu^{\pm}\ d\mu,$$

da cui $D_\mu \nu^{\pm} = f^{\pm}$ μ-q.o. Possiamo, dunque, concludere che

$$\begin{aligned}
\lim_{r\to 0} \fint_{B(x,r)} f\ d\mu &= \lim_{r\to 0} \frac{1}{\mu(B(x,r))} \left(\int_{B(x,r)} f^+\ d\mu - \int_{B(x,r)} f^-\ d\mu \right) \\
&= \lim_{r\to 0} \frac{\nu^+(B(x,r))}{\mu(B(x,r))} - \frac{\nu^-(B(x,r))}{\mu(B(x,r))} \\
&= D_\mu \nu^+(x) - D_\mu \nu^-(x) \\
&= f^+(x) - f^-(x) = f(x) \text{ per } \mu - \text{q.o. } x \in \mathbb{R}^n.
\end{aligned}$$

$\square$

Capitolo 2

Misure a valori matriciali

Indichiamo con $(M_m(\mathbb{R}), \|\cdot\|)$ lo spazio delle matrici quadrate di dimensione $m \in \mathbb{N}$ a entrate reali e norma $\|P\| = \|(p_{ij})_{i,j}\| = \sum_{i,j=1,...,m} |p_{ij}|$ per $P \in M_m(\mathbb{R})$. La scelta di questa norma per lo spazio delle matrici ha il fine di agevolare alcuni calcoli nelle proposizioni che seguiranno: la validità delle stesse è più ampia e gli argomenti utilizzati si possono riadattare anche alle altre norme (che sono equivalenti).

Definizione 2.1. Prendiamo ora una $\sigma-$algebra $\mathcal{M}$ di $\mathbb{R}^n$ e diciamo che $\mu : \mathcal{M} \to M_m(\mathbb{R})$, con $m \in \mathbb{N}$, è **misura a valori matriciali** se:
(i) $\mu(\oslash) = 0$;
(ii) $\mu\Big(\bigcup_{i\geq 1} A_i\Big) = \sum_{i\geq 1} \mu(A_i)$ comunque presa una famiglia disgiunta al più numerabile di $A_i \in \mathcal{M}$.

Inoltre, osserviamo che si può descrivere tale misura come $\mu = (\mu_{ij})_{i,j=1,...,m}$, dove $\mu_{ij} : \mathcal{M} \to \mathbb{R} \cup \{\pm\infty\}$ è misura con segno per ogni $i, j \in 1, ..., m$.
Se le misure μ_{ij} possono assumere valori infiniti, non tutte le definizioni e proprietà legate alle misure a valori matriciali continuano ad essere valide o ad assumere significato.
Nel seguito, supporremo quindi che tali misure siano finite, dal momento che quelle di nostro interesse (introdotte nel Teorema 3.4) saranno misure di Radon, dunque finite sui compatti, unico caso in cui ci troveremo a lavorare nella dimostrazione del Teorema 4.1 (ossia su palle chiuse).
Diremo che μ è misura di Radon, se lo sono tutte le sue componenti. In

particolare, $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{M}$. La trattazione può essere più generale, ma richiede l'estensione dei Teoremi 1.3,1.4 e 1.5.

Definizione 2.2. Data μ misura a valori matriciali, definiamo
$\|\mu\| : \mathcal{M} \to \mathbb{R}_{\geq 0}$ tale che per ogni $A \in \mathcal{M}$

$$\|\mu\|(A) := \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu(E)\|,$$

dove Π_A è la famiglia delle partizioni disgiunte (al più) numerabili di A.
Chiamiamo $\|\mu\|$ **variazione totale** di μ: si può mostrare che si tratta di una misura ed è di Radon se anche μ lo è.

Teorema 2.1 (Decomposizione polare)**.** *Data μ misura di Radon a valori matriciali, esiste $\sigma = (\sigma_{ij})_{i,j=1;\dots,m} : \mathbb{R}^n \to M_m(\mathbb{R})$, con σ_{ij} $\|\mu\|-$misurabili, tale che $\|\sigma\| = 1$ $\|\mu\|-q.o.$ e $d\mu = \sigma \ d \ \|\mu\|$.*

Dimostrazione. Vogliamo, innanzitutto, provare che $\mu \ll \|\mu\|$.
Fissato $A \in \mathcal{M}$, osserviamo che $\{A\}$ è partizione (al più) numerabile di A, da cui, per definizione di $\|\mu\|$, si ha

$$\|\mu(A)\| \leq \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu(E)\| = \|\mu\|(A). \tag{2.1.1}$$

Se le misure μ_{ij} sono misure di Radon, anche $\|\mu\|$ lo è. In particolare, vale il Teorema 1.3, cioè per ogni $i, j = 1, ..., m$ esiste $\sigma_{ij} \in L^1_{loc}(\mathbb{R}^n; \|\mu\|)$ tale che

$$d\mu_{ij} = \sigma_{ij} d\|\mu\|.$$

Definiamo $\sigma = (\sigma_{ij})_{i,j} : \mathbb{R}^n \to M_m(\mathbb{R})$, da cui $d\mu = \sigma d\|\mu\|$.
Dal Teorema 1.5, si ha, per (2.1.1), che

$$\begin{aligned}\|\sigma(x)\| &= \sum_{i,j=1}^m |\sigma_{ij}(x)| = \sum_{i,j=1}^m \lim_{r\to 0^+} \left| \fint_{B(x,r)} \sigma_{ij} d\|\mu\| \right| = \lim_{r\to 0^+} \sum_{i,j=1}^m \left| \frac{\mu_{ij}(B(x,r))}{\|\mu\|(B(x,r))} \right| \\ &= \lim_{r\to 0^+} \frac{\|\mu(B(x,r))\|}{\|\mu\|(B(x,r))} \leq \lim_{r\to 0^+} 1 = 1 \ per \ \|\mu\| - q.o. \ x \ in \ \mathbb{R}^n.\end{aligned}$$

D'altra parte, per ogni $r \in [0, 1)$, detto $A_r = \{x \in \mathbb{R}^n \mid \|\sigma(x)\| < r\}$, osserviamo che è $\|\mu\|-$misurabile, in quanto la funzione $\|\sigma(\cdot)\| = \sum_{i,j=1}^m |\sigma_{ij}(\cdot)|$

è misurabile perché somma delle $|\sigma_{ij}|$, che sono misurabili poiché lo sono le σ_{ij}. Quindi, comunque preso $\{E_k\}_{k\in\mathbb{N}} \in \Pi_{A_r}$,

$$\begin{aligned}\sum_{k\in\mathbb{N}}\|\mu(E_k)\| &= \sum_{k\in\mathbb{N}}\sum_{i,j=1}^{m}|\mu_{ij}(E_k)| = \sum_{k\in\mathbb{N}}\sum_{i,j=1}^{m}\left|\int_{E_k}\sigma_{ij}d\|\mu\|\right| \\ &\leq \sum_{k\in\mathbb{N}}\sum_{i,j=1}^{m}\int_{E_k}|\sigma_{ij}|d\|\mu\| = \sum_{k\in\mathbb{N}}\int_{E_k}\sum_{i,j=1}^{m}|\sigma_{ij}|d\|\mu\| \\ &= \sum_{k\in\mathbb{N}}\int_{E_k}\|\sigma\|d\|\mu\| \leq \sum_{k\in\mathbb{N}} r\|\mu\|(E_k) = r\|\mu\|(A_r),\end{aligned}$$

da cui estraendo il sup sulle famiglie $\{E_k\}$, si ottiene $\|\mu\|(A_r) \leq r\|\mu\|(A_r)$, cioè $\|\mu\|(A_r) = 0$ se $r < 1$ e, dunque, $\|\sigma\| = 1$ $\|\mu\|-$q.o. □

Osservazione 2.1. Non è restrittivo supporre che la funzione σ trovata sia di norma identicamente 1. Questa osservazione sarà cruciale per il teorema che segue:

Teorema 2.2. *Se $\mu : \mathcal{M} \to M_m(\mathbb{R})$ è misura (di Radon) a valori matriciali tale che esista $g \in L^1(M_m(\mathbb{R}); \mathcal{L}^n)$ per cui $\mu(E) := \int_E g\, d\mathcal{L}^n$ per ogni $E \in \mathcal{M}$, allora $\|\mu\|(E) = \int_E \|g\| d\mathcal{L}^n$ per ogni $E \in \mathcal{M}$.*

Dimostrazione. Sappiamo, per il Teorema 2.1, che esiste una funzione di norma unitaria $\sigma : \mathbb{R}^n \to M_m(\mathbb{R})$ tale che $d\mu = \sigma d\|\mu\|$ e con le proprietà di misurabilità dell'enunciato. Segue che $gd\mathcal{L}^n = d\mu = \sigma d\|\mu\|$, da cui per ogni $i, j \in \{1, ..., m\}$

$$g_{ij}d\mathcal{L}^n = \sigma_{ij}d\|\mu\|. \tag{2.2.1}$$

Per il Teorema 1.4, esiste $h \in L^1_{loc}(\mathcal{L}^n)$ tale che $d\|\mu\| = d\|\mu\|_s + hd\mathcal{L}^n$. Quindi

$$gd\mathcal{L}^n = \sigma d\|\mu\|_s + h\sigma d\mathcal{L}^n. \tag{2.2.2}$$

Osserviamo che $\sigma d\|\mu\|_s \ll d\|\mu\|_s$, che è ortogonale a $\mathcal{L}^n$, quindi è anch'essa ortogonale a $\mathcal{L}^n$. Per il Teorema 1.3, differenziando (2.2.2) rispetto a $\mathcal{L}^n$, otteniamo

$$g = D_{\mathcal{L}^n}\mu = \sigma D_{\mathcal{L}^n}\|\mu\|_s + h\sigma = h\sigma. \tag{2.2.3}$$

Inoltre, per (2.2.2), essendo $\sigma(x) \neq 0$ per ogni $x \in \mathbb{R}^n$, si ha che $\|\mu\|_s = 0$, da cui $\|\mu\| \ll \mathcal{L}^n$. Fissiamo $i, j \in \{1, ..., m\}$: detto

$$A_{ij} := \{x \in \mathbb{R}^n \mid \sigma_{ij}(x) g_{ij}(x) < 0\},$$

notiamo che è un boreliano di $\mathbb{R}^n$ e mostriamo che $\mathcal{L}^n(A_{ij}) = 0$.
Ora, supponiamo per assurdo che $\mathcal{L}^n(A_{ij}) > 0$: per ogni $k \in \mathbb{N}_0$ sia $A_{ij}^k := \{x \in A_{ij} \mid |g_{ij}(x)| > \frac{1}{k}\}$, che è boreliano: per continuità dal basso, $\lim_{k\to+\infty} \mathcal{L}^n(A_{ij}^k) = \mathcal{L}^n(A_{ij}) > 0$, cioè esiste $k \geq 1$ tale che $\mathcal{L}^n(A_{ij}^k) > 0$: possiamo supporre, senza perdita di generalità, che $\mathcal{L}^n(A_{ij}^{k,+}) > 0$, dove $A_{ij}^{k,+} := \{x \in A_{ij} \mid g_{ij}(x) > \frac{1}{k}\}$. Ma, per (2.2.1), otterremmo l'assurdo

$$\int_{A_{ij}^{k,+}} g_{ij} d\mathcal{L}^n \geq \frac{\mathcal{L}^n(A_{ij}^{k,+})}{k} > 0 \geq \int_{A_{ij}^{k,+}} \sigma_{ij} d\|\mu\| = \int_{A_{ij}^{k,+}} g_{ij} d\mathcal{L}^n.$$

Quindi, l'insieme $A := \bigcup_{i,j=1}^{m} A_{ij}$ ha misura di Lebesgue nulla. In particolare, per ogni $x \in A^c$ e per ogni $i, j \in \{1, ..., m\}$ $g_{ij}(x)\sigma_{ij}(x) \geq 0$: poiché $1 = \|\sigma(x)\| = \sum_{i,j=1}^{m} |\sigma_{ij}(x)|$, esisteranno indici $\bar{i}, \bar{j}$ tali che $\sigma_{\bar{i}\bar{j}}(x) \neq 0$, da cui $g_{\bar{i}\bar{j}}(x)\sigma_{\bar{i}\bar{j}}(x) \geq 0$ e, dunque, $h(x) = g_{\bar{i}\bar{j}}(x)\sigma_{\bar{i}\bar{j}}^{-1}(x) \geq 0$, cioè $h \geq 0 \quad \mathcal{L}^n-$q.o. Infine,

$$\|g\| = \|h\sigma\| = |h|\|\sigma\| = |h| :$$

ma $h \geq 0 \quad \mathcal{L}^n-$q.o., quindi $h = \|g\| \quad \mathcal{L}^n-$q.o., che dà la tesi. $\square$

Osservazione 2.2. In (2.2.3), abbiamo utilizzato il fatto che, date le misure di Radon μ, ν, η tali che $\mu \ll \nu, \nu \ll \eta$, vale la Regola della catena:

$$D_\eta \mu = D_\eta \nu \; D_\nu \mu.$$

Proponiamo ora alcune osservazioni che risulteranno utili nella dimostrazione del Teorema di Alexandrov.

Lemma 2.1. *Sia $\mu = (\mu^{ij})_{i,j}$ misura a valori matriciali (in $M_n(\mathbb{R})$). Per ogni $i, j \in \{1, ..., n\}$, consideriamo la decomposizione di Lebesgue di $\mu^{ij} = \mu_{ac}^{ij} + \mu_s^{ij}$ e definiamo le misure a valori matriciali $\mu_{ac} := (\mu_{ac}^{ij})_{i,j}$ e $\mu_s := (\mu_s^{ij})_{i,j}$. Allora*

$$\|\mu\| \leq \|\mu_{ac}\| + \|\mu_s\|.$$

Dimostrazione. Poiché $\mu = \mu_{ac} + \mu_s$, si ha per ogni $A \in \mathcal{M}$

$$\begin{aligned}
\|\mu\|(A) &= \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu(E)\| = \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu_{ac}(E) + \mu_s(E)\| \\
&\leq \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu_{ac}(E)\| + \|\mu_s(E)\| \\
&\leq \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu_{ac}(E)\| + \sup_{\pi \in \Pi_A} \sum_{E \in \pi} \|\mu_s(E)\| \\
&= \|\mu_{ac}\|(A) + \|\mu_s\|(A).
\end{aligned}$$

da cui la tesi. $\square$

Capitolo 3

Funzioni Convesse

Ricordiamo che una funzione $f : \mathbb{R}^n \to \mathbb{R}$ si dice **convessa** se, per ogni $x, y \in \mathbb{R}^n$ e $\lambda \in [0,1]$, vale

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Definiamo, inoltre, la funzione $\eta \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ tale che per ogni $x \in \mathbb{R}^n$

$$\eta(x) := \begin{cases} c\ e^{\frac{1}{|x|^2-1}} & \text{se } |x| < 1; \\ 0 & \text{se } |x| \geq 1, \end{cases}$$

dove la costante $c > 0$ è tale che $\displaystyle\int_{\mathbb{R}^n} \eta(x)\ dx = 1$.
Per ogni $\epsilon > 0$ e $x \in \mathbb{R}^n$ definiamo

$$\eta_\epsilon(x) := \frac{1}{\epsilon^n}\eta\Big(\frac{x}{\epsilon}\Big),$$

e chiamiamo $\{\eta_\epsilon\}_{\epsilon>0}$ **unità approsimata standard**. Presentiamo due teoremi cardine dell'Analisi Matematica (le cui dimostrazioni possono essere trovate in [3], pp.103-106, 59-64) e un lemma utili per le dimostrazioni che seguiranno:

Teorema 3.1 (Teorema di Rademacher)**.** *Una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ localmente lipschitziana è differenziabile $\mathcal{L}^n-$q.o.*

Teorema 3.2 (Teorema di rappresentazione di Riesz)**.** *Sia $L : \mathcal{C}_c(\mathbb{R}^n;\mathbb{R}^m) \to \mathbb{R}$ un funzionale lineare tale che per ogni compatto $K \subset \mathbb{R}^n$*

$$\sup\{L(f) \mid f \in \mathcal{C}_c(\mathbb{R}^n;\mathbb{R}^m), \mathrm{Supp} f \subseteq K, |f| \leq 1\} < +\infty.$$

Allora esistono una misura di Radon (positiva) μ e una funzione $\mu-$misurabile $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ tale che $|\sigma| = 1 \quad \mu-$q.o. e per ogni $f \in \mathcal{C}_c(\mathbb{R}^n; \mathbb{R}^m)$

$$L(f) = \int_{\mathbb{R}^n} f \cdot \sigma \, d\mu.$$

Lemma 3.1. *Sia $h : \mathbb{R}^n \to \mathbb{R}$ una funzione localmente lipschitziana e $|\cdot|$ la norma euclidea standard su $\mathbb{R}^n$. Allora per ogni $r > 0$ e $x, y \in B(r)$*

$$|h(x) - h(y)| \leq \operatorname*{ess\,sup}_{B(r)} |Dh||x - y|$$

Dimostrazione. (i) Poiché h è localmente lipschitziana, detta $L_{r,x}$ la costante di Lipschitz di h ristretta a $B(x, r)$, nei punti in cui Dh è definita (cioè $\mathcal{L}^n-$q.o.)

$$|Dh(y)| = |Dh(y) \cdot v| = \Big|\frac{\partial h}{\partial v}(y)\Big| = \lim_{t \to 0} \frac{|h(y + tv) - h(y)|}{|t|} \leq \frac{L_{r,x}|tv|}{|t|} = L_{r,x},$$

dove $v := \frac{Dh(y)}{|Dh(y)|}$ (se $Dh(y) = 0$ è banale), da cui, sapendo che Dh esiste $\mathcal{L}^n$-q.o. per il Teorema di Rademacher, $\operatorname*{ess\,sup}_{B(x,r)} |Dh| \leq L_{r,x} < +\infty$, cioè $Dh \in L^\infty(B(x, r))$, quindi $Dh \in L^1(B(x, r))$. Essendo h localmente lipschitziana, si può mostrare attraverso il Teorema di convergenza dominata che $\mathcal{L}^n - q.o.$ vale

$$Dh^\epsilon = Dh * \eta_\epsilon =: (Dh)^\epsilon,$$

dove $h^\epsilon := h * \eta_\epsilon$, per $\epsilon > 0$. Segue che $Dh^\epsilon \to Dh$ per $\epsilon \to 0^+$ nei punti di Lebesgue di Dh in $B(r)$. Inoltre, per ogni $x, y \in \overset{\circ}{B}(r)$ (cioè la palla aperta), sappiamo esiste $0 \leq \delta_{x,y} < 1$ tale che $x, y \in B(\delta_{x,y} r)$. Poiché $h^\epsilon \in \mathcal{C}^\infty(\mathbb{R}^n)$ per ogni $\epsilon > 0$ (si veda la proposizione 8.10 in [6]), per una conseguenza del Teorema del Valor Medio, si ha per ogni $\epsilon > 0$

$$|h^\epsilon(x) - h^\epsilon(y)| \leq \sup_{B(\delta_{x,y} r)} |Dh^\epsilon||x - y|.$$

Ma per ogni $z \in B(\delta_{x,y} r)$, se $\epsilon \in (0, (1 - \delta_{x,y})r)$,

$$\begin{aligned} |Dh^\epsilon(z)| &= \Big|\int_{\mathbb{R}^n} Dh(z - u)\eta_\epsilon(u) \, du\Big| \leq \int_{\mathbb{R}^n} |Dh(z - u)|\eta_\epsilon(u) \, du \\ &\leq \operatorname*{ess\,sup}_{B(\delta_{x,y} r + \epsilon)} |Dh| \leq \operatorname*{ess\,sup}_{B(r)} |Dh|, \end{aligned}$$

da cui, per gli stessi $\epsilon \in (0, (1-\delta_{x,y})r)$

$$\sup_{B(\delta_{x,y}r)} |Dh^\epsilon| \le \operatorname*{ess\,sup}_{B(r)} |Dh|.$$

Essendo h continua, facendo tendere $\epsilon \to 0^+$, otteniamo

$$|h(x) - h(y)| = \lim_{\epsilon \to 0^+} |h^\epsilon(x) - h^\epsilon(y)| \le \operatorname*{ess\,sup}_{B(r)} |Dh||x-y|.$$

Se $x, y \in B(r)$, poiché la chiusura di $\overset{\circ}{B}(r)$ è $B(r)$, esistono $\{x_k\}_{k\in\mathbb{N}}, \{y_k\}_{k\in\mathbb{N}} \subseteq \overset{\circ}{B}(r)$ tali che $x_k \to x, y_k \to y$ per $k \to +\infty$. Sapendo che per ogni $k \in \mathbb{N}$

$$|h(x_k) - h(y_k)| \le \operatorname*{ess\,sup}_{B(r)} |Dh||x_k - y_k|,$$

per $k \to +\infty$, poiché h è continua, si ha la tesi. $\square$

3.1 Proprietà delle Funzioni Convesse

Alcune delle proprietà delle funzioni convesse sono contenute nel seguente:

Teorema 3.3. *Data $f : \mathbb{R}^n \to \mathbb{R}$ convessa:*
(i) f è localmente limitata e localmente lipschitziana su $\mathbb{R}^n$.
(ii) Esiste $C > 0$, dipendente soltanto da n, tale che

$$\sup_{B(x,\frac{r}{2})} |f| \leq C \fint_{B(x,r)} |f(y)|\, dy; \tag{3.3.1}$$

$$\operatorname{ess\,sup}_{B(x,\frac{r}{2})} |Df| \leq \frac{C}{r} \fint_{B(x,r)} |f(y)|\, dy \tag{3.3.2}$$

comunque presa una palla $B(x,r)$ in $\mathbb{R}^n$.
(iii) Se $f \in \mathcal{C}^2(\mathbb{R}^n)$, allora $D^2 f(x) \geq 0$ per ogni $x \in \mathbb{R}^n$.

Dimostrazione. (i) Consideriamo i cubi del tipo $Q_L = [-L, L]^n$ e mostriamo che f è limitata su ciascuno di essi. Fissato $L > 0$, chiamiamo $\{v_i\}_{\{i=1,\dots,2^n\}}$ i vertici di Q_L e definiamo $M := \sup_i |f(v_i)|$. Sapendo che per ogni $x \in Q_L$ esistono $0 \leq \lambda_1, \dots, \lambda_{2^n} \leq 1$ tali che $\sum_{i=1}^{2^n} \lambda_i = 1$ e $x = \sum_{i=1}^{2^n} \lambda_i v_i$, per la disuguaglianza di Jensen,

$$f(x) = f\Big(\sum_{i=1}^{2^n} \lambda_i v_i\Big) \leq \sum_{i=1}^{2^n} \lambda_i f(v_i) \leq \sum_{i=1}^{2^n} \lambda_i M = M < +\infty$$

e, osservato che per ogni $x \in Q_L$

$$f(0) = f\Big(\frac{1}{2}x - \frac{1}{2}x\Big) \leq \frac{1}{2} f(x) + \frac{1}{2} f(-x),$$

otteniamo che

$$f(x) = 2f(0) - f(-x) \geq 2f(0) - M > -\infty,$$

da cui la locale limitatezza di f.
Consideriamo ora una palla $B(r)$ e fissiamo $x, y \in B(r)$. Sia $z := x + \mu(y - x) \in \partial B(2r)$ per un opportuno $\mu > 0$. In particolare, $\mu = \frac{|z-x|}{|y-x|}$ e $y = \frac{1}{\mu} z + \Big(1 - \frac{1}{\mu}\Big) x$, da cui, per definizione di funzione convessa,

$$\begin{aligned} f(y) &= f\Big(\frac{1}{\mu} z + \Big(1 - \frac{1}{\mu}\Big) x\Big) \leq \frac{1}{\mu} f(z) + \Big(1 - \frac{1}{\mu}\Big) f(x) \\ &\leq f(x) + \frac{|f(x) - f(z)|}{\mu} \leq f(x) + \frac{2M_r |y - x|}{|z - x|} \leq f(x) + \frac{2M_r}{r} |y - x|, \end{aligned}$$

dove $M_r := \sup\limits_{B(2r)} |f| < +\infty$. Analogamente $f(x) \leq f(y) + \frac{2M_r}{r}|y-z|$, da cui la locale lipschitzianità della funzione.
(ii) Supponiamo preliminarmente che $h \in \mathcal{C}^2(\mathbb{R}^n)$. Fissato $x \in \mathbb{R}^n$, comunque preso $y \in \mathbb{R}^n$, se $\lambda \in (0,1)$,

$$f(x + \lambda(y-x)) \leq f(x) + \lambda(f(y) - f(x)),$$

da cui

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \leq f(y) - f(x)$$

e, facendo tendere λ a 0^+, poiché $f \in \mathcal{C}^2(\mathbb{R}^n)$, si ha

$$Df(x) \cdot (y-x) = \frac{\partial f}{\partial (y-x)}(x) \leq f(y) - f(x). \tag{3.3.3}$$

Fissati $x \in \mathbb{R}^n$, $r > 0$ e $z \in B(x, \frac{r}{2})$, sappiamo, per (3.3.3), che per ogni $y \in \mathbb{R}^n$ vale

$$f(y) \geq f(z) + Df(z) \cdot (y-z),$$

da cui

$$\begin{aligned}
f(z) &= \fint_{B(z,\frac{r}{2})} f(z)\ dy \leq \fint_{B(z,\frac{r}{2})} f(y) - Df(z)\cdot(y-z)\ dy \\
&= \frac{1}{\mathcal{L}^n(B(z,\frac{r}{2}))} \int_{B(z,\frac{r}{2})} f(y)\ dy \leq \frac{1}{\alpha(n)(\frac{r}{2})^n} \int_{B(z,\frac{r}{2})} |f(y)|\ dy \\
&\leq \frac{\alpha(n)r^n}{\alpha(n)(\frac{r}{2})^n} \fint_{B(x,r)} |f(y)|\ dy = 2^n \fint_{B(x,r)} |f(y)|\ dy,
\end{aligned}$$

dove nel terzo passaggio si è utilizzato il fatto che

$$\int_{B(z,\frac{r}{2})} Df(z)\cdot(y-z)\ dy = 0.$$

Consideriamo ora una funzione $\zeta \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ tale che $0 \leq \zeta \leq 1$, $\mathrm{Supp}(\zeta) \subseteq B(x,r)$, $\zeta = 1$ su $B(x, \frac{r}{2})$ e $|D\zeta| \leq \frac{K}{r}$ per un'opportuna costante K positiva (indipendente da r). Una funzione siffatta può essere ottenuta per $r = 1$ tramite convoluzione dell'unità approssimata standard e la funzione gradino tra $B(x, \frac{1}{2})$ e $B(x,1)$ ed estesa al caso r generale considerando $\zeta(\frac{x}{r})$.
Sempre per (3.3.3), abbiamo

$$f(z) \geq f(y) + Df(y)\cdot(z-y),$$

per cui (C.S indica la Disuguaglianza di Cauchy-Schwarz):

$$
\begin{aligned}
&f(z)\fint_{B(x,r)} \zeta(y)\ dy \geq \fint_{B(x,r)} \zeta(y)(f(y) + Df(y)\cdot(z-y)\ dy \\
&= \fint_{B(x,r)} \zeta(y)f(y) + \sum_{i=1}^{n} \zeta(y)\frac{\partial f}{\partial y_i}(y)(z_i - y_i)\ dy \\
&= \fint_{B(x,r)} \zeta(y)f(y) + \sum_{i=1}^{n} \Big[\frac{\partial}{\partial y_i}(\zeta(y)f(y)(z_i-y_i)) - f(y)\frac{\partial}{\partial y_i}(\zeta(y)(z_i-y_i))\Big]\ dy \\
&= \fint_{B(x,r)} \zeta(y)f(y) - \sum_{i=1}^{n} f(y)\Big(\frac{\partial \zeta}{\partial y_i}(y)(z_i-y_i) - \zeta(y)\Big)\ dy \\
&\geq -(n+1)\fint_{B(x,r)} |f(y)||\zeta(y)|\ dy - \fint_{B(x,r)} |f(y)||D\zeta(y)\cdot(z-y)|\ dy \\
&\overset{C.S}{\geq} -(n+1)\fint_{B(x,r)} |f(y)|\ dy - \fint_{B(x,r)} |f(y)|\frac{K}{r}\cdot\frac{3r}{2}\ dy \\
&= -\Big(\frac{3K}{2} + n + 1\Big)\fint_{B(x,r)} |f(y)|\ dy.
\end{aligned}
$$

Ponendo quindi $C' := \max\Big\{\frac{3K}{2} + n + 1, 2^n\Big\}$ ed estraendo il sup su $B(x, \frac{r}{2})$, si ha la tesi. Dato ancora $z \in B(x, \frac{r}{2})$, definiamo

$$
S_z := \Big\{y \in \mathbb{R}^n \mid \frac{r}{4} \leq |y-z| \leq \frac{r}{2},\ Df(z)\cdot(y-z) \geq \frac{1}{2}|Df(z)||(y-z)|\Big\},
$$

che è intersezione di un cono con angolo di apertura $\frac{\pi}{3}$ e centro z e di una corona circolare di palle di centro z e raggi $\frac{r}{4}$ e $\frac{r}{2}$: in particolare contiene una palla di raggio $\frac{r}{8}$, da cui $\mathcal{L}^n(S_z) \geq \frac{\alpha(n)r^n}{8^n}$ uniformemente per $z \in B(x, \frac{r}{2})$.
Per (3.3.3), abbiamo che per ogni $y \in S_z$

$$
f(y) \geq f(z) + Df(z)\cdot(y-z) \geq f(z) + \frac{1}{2}|Df(z)||y-z| \geq f(z) + \frac{r}{8}|Df(z)|.
$$

Allora per ogni $z \in B(x, \frac{r}{2})$, sapendo che $S_z \subseteq B(x,r)$:

$$
\begin{aligned}
|Df(z)| &\leq \fint_{S_z} \frac{8}{r}|f(y)-f(z)|\ dy \leq \frac{8}{r\mathcal{L}^n(s_z)}\int_{S_z} |f(y)-f(z)|\ dy \\
&\leq \frac{8^{n+1}}{\alpha(n)r^{n+1}}\int_{B(x,r)} |f(y)-f(z)|\ dy \\
&\leq \Big(\frac{8}{r}\Big)^{n+1}\frac{\mathcal{L}^n(B(x,r))}{\alpha(n)}\fint_{B(x,r)} |f(y)-f(z)|\ dy
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{8^{n+1}}{r}\Big(|f(z)| + \fint_{B(x,r)} |f(y)|\, dy\Big)\\
&\leq \frac{8^{n+1}(C'+1)}{r}\fint_{B(x,r)} |f(y)|\, dy,
\end{aligned}
$$

da cui la tesi, ponendo $C := 8^{n+1}(C'+1)$, anche per la disuguaglianza (3.3.1).
Nel caso in cui f sia funzione soltanto convessa, sappiamo, per (i) e per il Teorema di Rademacher, che la funzione Df è definita q.o.
Sia $\{\eta_\epsilon\}_{\epsilon>0}$ l'unità approssimata standard e definiamo $f^\epsilon := f * \eta_\epsilon$. Sappiamo che $f^\epsilon \in \mathcal{C}^\infty(\mathbb{R}^n)$ e mostriamo che è anche convessa per ogni $\epsilon > 0$. Dati, infatti, $x, y \in \mathbb{R}^n$ e $\lambda \in [0,1]$, vale

$$
f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).
$$

Quindi

$$
\begin{aligned}
f^\epsilon(\lambda x + (1-\lambda)y) &= \int_{\mathbb{R}^n} \eta_\epsilon(z) f(\lambda x + (1-\lambda)y - z)\, dz\\
&= \int_{\mathbb{R}^n} \eta_\epsilon(z) f(\lambda(x-z) + (1-\lambda)(y-z))\, dz\\
&\leq \lambda \int_{\mathbb{R}^n} \eta_\epsilon(z) f(x-z)\, dz + (1-\lambda)\int_{\mathbb{R}^n} \eta_\epsilon(z) f(y-z)\, dz\\
&= \lambda f^\epsilon(x) + (1-\lambda) f^\epsilon(y-z).
\end{aligned}
$$

La costante C non dipende dalla funzione convessa scelta, dunque per ogni $\epsilon > 0$

$$
\sup_{B(x,\frac{r}{2})} |f^\epsilon| + r|Df^\epsilon| \leq C \fint_{B(x,r)} |f^\epsilon(y)|\, dy
$$

Essendo f limitata sulle palle (per (i)), $|f^\epsilon| \to |f|$, per $\epsilon \to 0^+$ uniformemente sulla palla $B(x, \frac{r}{2})$ (si veda 8.14c in [6]). Inoltre, per le considerazioni contenute nell'Osservazione 3.1, facendo il limite per $\epsilon \to 0^+$ e applicando il Teorema di Convergenza Dominata, otteniamo le stesse stime per f e Df, dunque la tesi nel caso generale.
(iii) Supponiamo ora che $f \in \mathcal{C}^2(\mathbb{R}^n)$: per la formula di Taylor con resto integrale, si ha per ogni $x, y \in \mathbb{R}^n$

$$
f(y) = f(x) + Df(x)\cdot(y-x) + (y-x)^t \cdot \int_0^1 (1-s)D^2 f(x+s(y-x))\, ds \cdot (y-x),
$$

da cui, per (3.3.1), per ogni $x, y \in \mathbb{R}^n$

$$(y-x)^t \cdot \int_0^1 (1-s) D^2 f(x+s(y-x))\; ds \cdot (y-x) \geq 0.$$

Fissati $x, v \in \mathbb{R}^n$ qualunque, posso imporre $y = x + tv$ e ricavare $(t > 0)$

$$v^t \cdot \int_0^1 (1-s) D^2 f(x+tsv)\; ds \cdot v \geq 0.$$

Per $t \to 0^+$, applicando il Teorema di convergenza dominata, si ottiene

$$v^t \cdot D^2 f(x) \cdot v \geq 0$$

e, dunque, data l'arbitrarietà nella scelta di v, la tesi. □

Osservazione 3.1. Per il Teorema 8.15 in [6], per $\epsilon \to 0^+$, $Df^\epsilon(y) \to Df(y)$ per y punto di Lebesgue di Df in $B(x,r)$ (infatti, per il Lemma 3.1, $Df \in L^1_{loc}(\mathbb{R}^n;\mathbb{R}^n)$ e $Df^\epsilon = (Df)^\epsilon$). Fissato $\delta > 0$, per definizione di ess sup, esiste y_δ punto di Lebesgue di Df in $B(x,r)$ tale che

$$|Df(y_\delta)| > \operatorname*{ess\,sup}_{B(x,r)} |Df| - \frac{\delta}{2}.$$

Poiché $Df^\epsilon(y_\delta) \to Df(y_\delta)$ per $\epsilon \to 0^+$, si ha che esiste $\epsilon_\delta > 0$ tale che per ogni $\epsilon \in (0, \epsilon_\delta)$ $|Df(y_\delta) - Df^\epsilon(y_\delta)| < \frac{\delta}{2}$, da cui

$$\sup_{B(x,r)} |Df^\epsilon| \geq |Df^\epsilon(y_\delta)| \geq |Df(y_\delta)| - |Df(y_\delta) - Df^\epsilon(y_\delta)| > \operatorname*{ess\,sup}_{B(x,r)} |Df| - \delta.$$

Segue che $\operatorname*{ess\,sup}_{B(x,r)} |Df| \leq \liminf_{\epsilon \to 0^+} \sup_{B(x,r)} |Df^\epsilon|$ e, quindi,

$$\begin{aligned} \operatorname*{ess\,sup}_{B(x,\frac{r}{2})} r|Df| + \sup_{B(x,\frac{r}{2})} |f| &\leq \liminf_{\epsilon \to 0^+} \sup_{B(x,\frac{r}{2})} r|Df^\epsilon| + \lim_{\epsilon \to 0^+} \sup_{B(x,\frac{r}{2})} |f^\epsilon| \\ &\leq \lim_{\epsilon \to 0^+} C \fint_{B(x,r)} |f^\epsilon|\; dy = C \fint_{B(x,r)} |f|\; dy. \end{aligned}$$

3.2 Derivate Seconde come Misure

Enunciamo ora (con una breve idea della dimostrazione) un corollario del Teorema di Rappresentazione di Riesz, che ci permetterà di descrivere le derivate seconde di funzioni convesse come misure:

Corollario 3.1 (Funzionali lineari non negativi)**.** *Sia $L : \mathcal{C}_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ lineare e non negativo (cioè se $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ è tale che $f \geq 0$, allora $L(f) \geq 0$). Dunque esiste una misura (positiva) di Radon μ su $\mathbb{R}^n$ tale che per ogni $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$*

$$L(f) = \int_{\mathbb{R}^n} f \; d\mu.$$

Dimostrazione. L'idea della dimostrazione sta nell'osservare che L può essere esteso a un funzionale lineare su $\mathcal{C}_c(\mathbb{R}^n)$ che soddisfi le ipotesi del Teorema di Riesz e, applicandolo, dedurre la tesi. □

Teorema 3.4 (Derivate Seconde come Misure)**.** *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione convessa. Allora esistono delle misure di Radon μ^{ij} tali che $\mu^{ij} = \mu^{ji}$ per ogni $i, j \in \{1, ..., n\}$ e per ogni $\phi \in \mathcal{C}_c^2(\mathbb{R}^n)$*

$$\int_{\mathbb{R}^n} f \phi_{x_i x_j} \; dx = \int_{\mathbb{R}^n} \phi d\mu^{ij}.$$

Dimostrazione. Fissiamo $v = (v_1, ..., v_n) \in \mathbb{R}^n$, $\{\eta_\epsilon\}_{\epsilon > 0}$ il mollificatore standard e $f^\epsilon := f * \eta_\epsilon$ per ogni $\epsilon > 0$. $f^\epsilon \in \mathcal{C}^\infty(\mathbb{R}^n)$ ed è convessa quindi $D^2 f^\epsilon \geq 0$. Allora per ogni $\phi \in \mathcal{C}_c^2(\mathbb{R}^n)$, con $\phi \geq 0$, si ha

$$\sum_{i,j=1}^n v_i v_j \int_{\mathbb{R}^n} f^\epsilon \phi_{x_i x_j} \; dx = \sum_{i,j=1}^n - \int_{\mathbb{R}^n} v_i v_j f^\epsilon_{x_i} \phi_{x_j} \; dx$$
$$= \sum_{i,j=1}^n \int_{\mathbb{R}^n} f^\epsilon_{x_i x_j} v_i v_j \phi \; dx = \int_{\mathbb{R}^n} v^t \cdot D^2 f^\epsilon \cdot v \; \phi \; dx \geq 0.$$

Facendo il limite per $\epsilon \to 0^+$, si ha che per ogni $\phi \in \mathcal{C}_c^2(\mathbb{R}^n)$

$$L_v(\phi) := \sum_{i,j=1}^n v_i v_j \int_{\mathbb{R}^n} f \phi_{x_i x_j} \; dx \geq 0.$$

Per il Corollario 3.1, per ogni $v \in \mathbb{R}^n$ esiste una misura di Radon (positiva) μ^v tale che $L_v(\phi) = \int_{\mathbb{R}^n} \phi \; d\mu^v$ per ogni $\phi \in \mathcal{C}_c^2(\mathbb{R}^n)$.

Imponiamo, per $i \in \{1, ..., n\}$, $\mu^{ii} := \mu^{e_i}$. Se, invece, $i \neq j$, consideriamo $v_{ij} = \frac{e_i + e_j}{\sqrt{2}}$. Poiché $\displaystyle\sum_{k,l=1}^{n} \phi_{x_k x_l} v_k v_l = \phi_{x_i x_j} + \frac{1}{2}(\phi_{x_i x_i} + \phi_{x_j x_j})$, otteniamo che

$$\begin{aligned}\int_{\mathbb{R}^n} f\phi_{x_i x_j}\, dx &= \sum_{k,l=1}^{n} v_k v_l \int_{\mathbb{R}^n} f\phi_{x_k x_l}\, dx - \frac{1}{2}\int_{\mathbb{R}^n} f\phi_{x_i x_i}\, dx - \frac{1}{2}\int_{\mathbb{R}^n} f\phi_{x_j x_j}\, dx \\ &= \int_{\mathbb{R}^n} \phi\, d\mu^{v_{ij}} - \frac{1}{2}\int_{\mathbb{R}^n} \phi\, d\mu^{ii} - \frac{1}{2}\int_{\mathbb{R}^n} \phi\, d\mu^{jj} = \int_{\mathbb{R}^n} \phi\, d\mu^{ij},\end{aligned}$$

dove $\mu^{ij} := \mu^{v_{ij}} - \frac{1}{2}\mu^{ii} - \frac{1}{2}\mu^{jj}$. □

Ad ogni funzione convessa $f : \mathbb{R}^n \to \mathbb{R}$ possiamo quindi associare la misura (di Radon) a valori matriciali

$$[D^2 f] := \begin{pmatrix} \mu^{11} & \mu^{12} & \cdots & \mu^{1n} \\ \mu^{21} & \mu^{22} & \cdots & \mu^{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu^{n1} & \mu^{n2} & \cdots & \mu^{nn} \end{pmatrix}.$$

Chiamiamo, per comodità, $\mu := [D^2 f]$. Per il Teorema 2.1 (di Decomposizione Polare), esiste una funzione $\sigma = (\sigma_{ij})_{i,j=1;...,n} : \mathbb{R}^n \to M_n(\mathbb{R})$, con σ_{ij} $\|\mu\|-$misurabili, tale che $\|\sigma\| = 1$ $\|\mu\|-$q.o. e $d\mu = \sigma\ d\ \|\mu\|$.
In particolare, per il Teorema 1.4, esistono per ogni $i, j \in \{1, ..., n\}$ delle misure $\mu_s^{ij}, \mu_{ac}^{ij}$ di Radon su $\mathbb{R}^n$ tali che $\mu_s^{ij} \perp \mathcal{L}^n$ e $\mu_{ac}^{ij} \ll \mathcal{L}^n$. Inoltre, per ogni $i, j \in \{1, ..., n\}$ esistono funzioni $f_{ij} \in L^1_{loc}(\mathbb{R}^n)$ tali che

$$d\mu_{ac}^{ij} = f_{ij} d\mathcal{L}^n.$$

Definiamo anche le matrici $[D^2 f]_s := (\mu_s^{ij})_{i,j=1,...,n}$ e $[D^2 f]_{ac} := (\mu_{ac}^{ij})_{i,j=1,...,n}$. Siamo ora in grado di costruire una matrice (simmetrica, come conseguenza della simmetria di μ, nota dal Teorema 3.4) di funzioni localmente integrabili. ossia le f_{ij}. Tale matrice, che indicheremo con $D^2 f := (f_{ij})_{i,j}$, si candida ad essere la matrice Hessiana della funzione f in quasi ogni punto di $\mathbb{R}^n$.

Capitolo 4

Teorema di Alexandrov

Mentre, grazie al Teorema 3.3 (punto (i)), abbiamo garanzia della locale lipschitzianità delle funzioni convesse, dunque dell'esistenza quasi ovunque del loro differenziale, non abbiamo ancora evidenza del fatto che la matrice $D^2 f$ rappresenti effettivamente l'Hessiana di f q.o. La conferma giunge dal

Teorema 4.1 (Teorema di Alexandrov)**.** *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione convessa. Allora f ammette derivate seconde $\mathcal{L}^n$-q.o. In particolare, per $\mathcal{L}^n-$q.o. $x \in \mathbb{R}^n$, se $y \to x$, vale*

$$\left| f(y) - f(x) - Df(x) \cdot (y-x) - \frac{1}{2}(y-x)^t \cdot D^2 f(x) \cdot (y-x) \right| = o(|y-x|^2).$$

Dimostrazione. Partiamo dall'enunciare alcuni fatti che utilizzeremo nel corso della dimostrazione. Per $\mathcal{L}^n-$q.o. $x \in \mathbb{R}^n$:

(a) $Df(x)$ esiste e $\lim_{r \to 0^+} \fint_{B(x,r)} |Df(y) - Df(x)| \; dy = 0$;

(b) $\lim_{r \to 0^+} \fint_{B(x,r)} \|D^2 f(y) - D^2 f(x)\| \; dy = 0$;

(c) $\lim_{r \to 0^+} \dfrac{\|[D^2 f]_s\|(B(x,r))}{r^n} = 0$.

Dimostrazione. (a) Il fatto che Df esista $\mathcal{L}^n$-q.o. segue dal Teorema di Rademacher, una volta dedotta la locale lipschitzianità di f nel punto (i) del Teorema 3.3. Inoltre, come osservato nel Lemma 3.1, $Df \in L^1_{loc}(\mathbb{R}^n; \mathbb{R}^n)$.
Imponendo che $|Df(x)| = \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(x) \right|$ (tutte le norme in $\mathbb{R}^n$ sono equivalenti) e applicando il Teorema 1.5 a ciascuna delle componenti di Df, che

sono L^1_{loc}, otteniamo la tesi.
(b) Per costruzione di $D^2 f$, le sue componenti $f_{ij} \in L^1_{loc}(\mathbb{R}^n)$. Utilizzando per $D^2 f$ la norma matriciale definita nel precedente capitolo (anche in questo caso, tutte le norme sono equivalenti), otteniamo, per additività del limite, la tesi, adoperando il Teorema 1.5.
Sappiamo che tutte le norme matriciali sono equivalenti: detta $\|\cdot\|_{op}$ la norma operatoriale su $M_n(\mathbb{R})$, esistono $\beta, \gamma > 0$ tali che

$$\beta\|\cdot\| \leq \|\cdot\|_{op} \leq \gamma\|\cdot\|.$$

(c) Osserviamo che è sufficiente ricondursi a mostrare che tale limite valga per $|\mu_s^{ij}|$ per ogni $i, j \in \{1, ..., n\}$. Attraverso la definzione di misura variazione totale si può notare come $|\mu_s^{ij}| \perp \mathcal{L}^n$ e concludere per il Teorema 1.4. $\square$

Osserviamo che ciascuna delle proprietà (a),(b) e (c) è vera per rispetivamente $x \in A^c, B^c, C^c$, con $\mathcal{L}^n(A) = \mathcal{L}^n(B) = \mathcal{L}^n(C) = 0$, ma non è detto che tali insiemi coincidano: poniamo, dunque, $N := A \cup B \cup C$ e da ora in poi considereremo $x \in N^c$. In particolare, non è restrittivo supporre $x = 0$. Sia $r > 0$ e consideriamo $f^\epsilon := \eta_\epsilon * f$, famiglia di funzioni convesse e $\mathcal{C}^\infty$. Fissato $y \in B(r)$, per la Formula di Taylor con resto integrale, vale

$$\begin{aligned} f^\epsilon(y) &= f^\epsilon(0) + Df^\epsilon(0) \cdot y + \int_0^1 (1-s) y^t \cdot D^2 f^\epsilon(sy) \cdot y \, ds \\ &= f^\epsilon(0) + Df^\epsilon(0) \cdot y + \frac{1}{2} y^t \cdot D^2 f(0) \cdot y \\ &+ \int_0^1 (1-s) y^t \cdot [D^2 f^\epsilon(sy) - D^2 f(0)] \cdot y \, ds. \end{aligned}$$

Fissiamo ora $\phi \in \mathcal{C}^2_c(B(r))$, $|\phi| \leq 1$. Dunque (per il Teorema di Fubini)

$$\begin{aligned} &\fint_{B(r)} \phi(y) \Big(f^\epsilon(y) - f^\epsilon(0) - Df^\epsilon(0) \cdot y - \frac{1}{2} y^t \cdot D^2 f(0) \cdot y \Big) \, dy \\ &= \int_0^1 (1-s) \fint_{B(r)} \phi(y) y^t \cdot [D^2 f^\epsilon(sy) - D^2 f(0)] \cdot y \, dy \, ds \end{aligned} \tag{4.1.1}$$

$$= \int_0^1 \frac{1-s}{s^2} \fint_{B(rs)} \phi\Big(\frac{z}{s}\Big) z^t \cdot [D^2 f^\epsilon(z) - D^2 f(0)] \cdot z \, dz \, ds.$$

Consideriamo ora

$$
\begin{aligned}
g_\epsilon(s) &:= \int_{B(rs)} \phi\Big(\frac{z}{s}\Big) z^t \cdot D^2 f^\epsilon(z) \cdot z \; dz \\
&= \int_{B(rs)} \phi\Big(\frac{z}{s}\Big) \sum_{i,j=1}^n z_i z_j \frac{\partial^2 f^\epsilon}{\partial z_i \partial z_j}(z) \; dz \\
&= \int_{B(rs)} \sum_{i,j=1}^n f^\epsilon(z) \frac{\partial^2}{\partial z_i \partial z_j} \Big(\phi\Big(\frac{z}{s}\Big) z_i z_j \Big) \; dz \\
&\to \int_{B(rs)} \sum_{i,j=1}^n f(z) \frac{\partial^2}{\partial z_i \partial z_j} \Big(\phi\Big(\frac{z}{s}\Big) z_i z_j \Big) \; dz \quad \text{per } \epsilon \to 0^+ \\
&= \sum_{i,j=1}^n \int_{B(rs)} \phi\Big(\frac{z}{s}\Big) z_i z_j \; d\mu^{ij} \\
&= \int_{B(rs)} \sum_{i,j=1}^n \phi\Big(\frac{z}{s}\Big) z_i z_j (f_{ij}(z) \; dz + d\mu_s^{ij}) \\
&= \int_{B(rs)} \phi\Big(\frac{z}{s}\Big) z^t \cdot D^2 f(z) \cdot z \; dz + \sum_{i,j=1}^n \int_{B(rs)} \phi\Big(\frac{z}{s}\Big) z_i z_j \; d\mu_s^{ij}.
\end{aligned}
$$

Ora, per il Teorema 1.3 e le osservazioni (b) e (c), per $\mathcal{L}^n-$q.o. $x \in \mathbb{R}^n$

$$
\lim_{\eta \to 0^+} \frac{\|[D^2 f]_s\|(B(x,\eta))}{\mathcal{L}^n(B(x,\eta))} = 0, \quad \lim_{\eta \to 0^+} \fint_{B(x,\eta)} \|D^2 f(y) - D^2 f(x)\| \; dy = 0
$$

Ancora una volta non è restrittivo supporre che sia vero anche per $x = 0$. Sapendo che $s \in [0,1]$, osserviamo che esistono $r_0, \epsilon_0 > 0$ tali che, se $0 < r < r_0, \; \epsilon < \epsilon_0$, allora

$$
\frac{\|[D^2 f]_s\|(B(rs+\epsilon))}{\mathcal{L}^n(B(rs+\epsilon))} + \fint_{B(rs+\epsilon)} \|D^2 f(y) - D^2 f(0)\| \; dy < 1
$$

Per il Lemma 2.1 ed il Teorema 2.2, otteniamo che

$$
\begin{aligned}
&\|D^2 f\|(B(rs+\epsilon)) \le \|[D^2 f]_s\|(B(rs+\epsilon)) + \|[D^2 f]_{ac}\|(B(rs+\epsilon)) \\
&= \alpha(n)(rs+\epsilon)^n \frac{\|[D^2 f]_s\|(B(rs+\epsilon))}{\mathcal{L}^n(B(rs+\epsilon))} + \int_{B(rs+\epsilon)} \|D^2 f(y)\| \; dy
\end{aligned}
$$

$$\le \alpha(n)(rs+\epsilon)^n\Big(\frac{\|[D^2f]_s\|(B(rs+\epsilon))}{\mathcal{L}^n(B(rs+\epsilon))}+⨍_{B(rs+\epsilon)}\|D^2f(y)-D^2f(0)\|\ dy+\|D^2f(0)\|\Big)$$
$$\le \alpha(n)(1+\|D^2f(0)\|)(rs+\epsilon)^n.$$

Per $0<r<r_0, 0<\epsilon<\epsilon_0$, siamo ora in grado di stimare

$$\begin{aligned}
\frac{g_\epsilon(s)}{s^{n+2}} &\overset{C.S.}{\le} \frac{\gamma}{s^{n+2}}\int_{B(rs)}\Big|\phi\Big(\frac{z}{s}\Big)\Big||z|^2\|D^2f^\epsilon(z)\|\ dz\\
&\le \frac{\gamma(rs)^2}{s^{n+2}}\int_{B(rs)}\sum_{i,j=1}^n\Big|\frac{\partial^2 f^\epsilon}{\partial z_i\partial z_j}(z)\Big|\ dz\\
&= \frac{\gamma r^2}{s^n}\int_{B(rs)}\sum_{i,j=1}^n\Big|\int_{\mathbb{R}^n}\frac{\partial^2\eta_\epsilon}{\partial z_i\partial z_j}(z-y)f(y)\ dy\Big|\ dz\\
&= \frac{\gamma r^2}{s^n}\int_{B(rs)}\sum_{i,j=1}^n\Big|\int_{\mathbb{R}^n}\eta_\epsilon(z-y)\ d\mu^{ij}(y)\Big|\ dz\\
&= \frac{\gamma r^2}{s^n}\int_{B(rs)}\Big\|\int_{\mathbb{R}^n}\eta_\epsilon(z-y)\ d[D^2f](y)\Big\|\ dz\\
&= \frac{\gamma r^2}{s^n}\int_{B(rs)}\Big\|\int_{\mathbb{R}^n}\eta_\epsilon(z-y)\sigma(y)\ d\|D^2f\|(y)\Big\|\ dz\\
&\le \frac{\gamma r^2}{s^n}\int_{B(rs)}\int_{\mathbb{R}^n}|\eta_\epsilon(z-y)|\|\sigma(y)\|\ d\|D^2f\|(y)\ dz\\
&\le \frac{\gamma r^2}{s^n}\int_{B(rs+\epsilon)}\int_{B(rs)\cap B(y,\epsilon)}\frac{c}{\epsilon^n}dz\ d\|D^2f\|(y)\\
&\le \frac{\gamma c r_0^2}{s^n\epsilon^n}\alpha(n)\min\{(rs)^n,\epsilon^n\}\|D^2f\|(B(rs+\epsilon))\\
&\le \gamma c r_0^2\alpha(n)^2(1+\|D^2f(0)\|)\Big(\frac{\min\{(rs)^n,\epsilon^n\}}{s^n\epsilon^n}(rs+\epsilon)^n\Big)\\
&\le \gamma c(1+\|D^2f(0)\|)\alpha(n)^2 2^n r_0^{n+2} =: C'.
\end{aligned}$$

Applichiamo il Teorema di Convergenza Dominata a (4.1.1) per $\epsilon\to 0^+$:

$$⨍_{B(r)}\phi(y)\Big(f(y)-f(0)-Df(0)\cdot y-\frac{1}{2}y^t\cdot D^2f(0)\cdot y\Big)\ dy$$
$$=\int_0^1\frac{1-s}{s^2}\Big(⨍_{B(rs)}\phi\Big(\frac{z}{s}\Big)z^t\cdot[D^2f(z)-D^2f(0)]\cdot z\ dz+\sum_{i,j=1}^n\phi\Big(\frac{z}{s}\Big)z_iz_j\ d\mu_s^{ij}\Big)\ ds$$

$$\leq \int_0^1 \gamma r^2 \fint_{B(rs)} \|D^2 f(z) - D^2 f(0)\| \, dz \, ds + \int_0^1 \frac{|1-s|}{s^2} \sum_{i,j=1}^n (rs)^2 |\mu_s^{ij}|(B(rs)) \, ds$$
$$\leq \gamma r^2 \fint_{B(r)} \|D^2 f(z) - D^2 f(0)\| dz + r^2 \int_0^1 \sum_{i,j=1}^n \frac{|\mu_s^{ij}|(B(rs))}{\mathcal{L}^n(B(rs))} \, ds$$
$$= r^2 o(1) = o(r^2)$$

per $r \to 0^+$, per (b) e l'osservazione fatta nel dimostare (c).
Prendendo il sup tra le funzioni ϕ costruite come sopra, deduciamo che, detta

$$h(y) := f(y) - f(0) - Df(0) \cdot y - \frac{1}{2} y^t \cdot D^2 f(0) \cdot y,$$

per $r \to 0^+$

$$\fint_{B(r)} |h(y)| \, dy = o(r^2). \tag{4.1.2}$$

Sia ora $\Lambda := \|D^2 f(0)\|_{op}$: vogliamo mostrare che $g := h + \frac{\Lambda}{2}|y|^2$ è convessa. Per farlo, osserviamo che è somma delle funzioni convesse $f(y) - f(0) - Df(0) \cdot y$ e $l(y) := \frac{\Lambda}{2}|y|^2 - \frac{1}{2} y^t \cdot D^2 f(0) \cdot y$. Infatti, per ogni $x, y \in \mathbb{R}^n$ e per ogni $\lambda, \mu \in [0,1]$ tali che $\lambda + \mu = 1$ è vero che

$$\begin{aligned}
&f(\lambda x + \mu y) - f(0) - Df(0) \cdot (\lambda x + \mu y) \\
&\leq \lambda f(x) + \mu f(y) - \lambda f(0) - \mu f(0) - \lambda Df(0) \cdot x - \mu Df(0) \cdot y \\
&= \lambda (f(x) - f(0) - Df(0) \cdot x) + \mu (f(y) - f(0) - Df(0) \cdot y),
\end{aligned}$$

mentre $l \in \mathcal{C}^2(\mathbb{R}^n)$ ammette Hessiana $Hl(y) = \Lambda \mathbb{I}_n - D^2 f(0) \geq 0$ per ogni $y \in \mathbb{R}^n$. Allora, per il Teorema 3.3 (punto (ii)), esiste $C > 0$ per cui

$$\begin{aligned}
\operatorname*{esssup}_{B(\frac{r}{2})} |Dh| &\leq \operatorname*{esssup}_{B(\frac{r}{2})} |Dg| + \sup_{B(\frac{r}{2})} \left| D\left(\frac{\Lambda}{2}|y|^2\right) \right| \leq \frac{C}{r} \fint_{B(r)} |g(y)| \, dy + \sup_{B(\frac{r}{2})} \Lambda |y| \\
&\leq \frac{C}{r} \fint_{B(r)} |h(y)| + \frac{\Lambda}{2}|y|^2 \, dy + \Lambda \frac{r}{2} \leq \frac{\bar{C}}{r} \fint_{B(r)} |h(y)| \, dy + \bar{C} r,
\end{aligned}$$

dove $\bar{C} := \max\{C, \frac{\Lambda}{2}(C+1)\}$.
Per concludere mostriamo che $\sup_{B(\frac{r}{2})} |h| = o(r^2)$ per $r \to 0^+$.
Fissiamo $\epsilon > 0$ e $\eta < 1$, $\eta^{\frac{1}{n}} < \frac{1}{2}$. Per (4.1.2), esiste $r_0 := r_0(\epsilon, \eta) > 0$ tale che

$$\mathcal{L}^n(\{z \in B(r) \mid |h(z)| \geq \epsilon r^2\}) \leq \int_{B(r)} \frac{|h(z)|}{\epsilon r^2} \, dz = o(r^n) < \eta \alpha(n) r^n$$

per $0 < r < r_0$. Per ogni $y \in B(\frac{r}{2})$ esiste z_y tale che $|h(z_y)| \leq \epsilon r^2$ e $|y - z_y| \leq \sigma := \eta^{\frac{1}{n}} r$. Se così non fosse, $B(y, \sigma) \subseteq \{z \in B(r) \mid |h(z)| \geq \epsilon r^2\}$, da cui

$$\mathcal{L}^n(\{z \in B(r) \mid |h(z)| \geq \epsilon r^2\}) \geq \mathcal{L}^n(B(y, \sigma)) = \eta \alpha(n) r^n,$$

che è assurdo. Segue che per ogni $y \in B(\frac{r}{2})$, poiché h è localmente lipschitziana (perché differenza di funzioni localmente lipschitziane), per il Lemma 3 e (4.1.2),

$$\begin{aligned} |h(y)| &\leq |h(z_y)| + |h(y) - h(z_y)| \leq \epsilon r^2 + \sigma \operatorname*{ess\,sup}_{B(r)} |Dh| \\ &\leq \epsilon r^2 + \sigma \bar{C} \Big(\frac{1}{r} \fint_{B(r)} |h(z)| \, dz + r \Big) \leq \epsilon r^2 + \eta^{\frac{1}{n}} r \bar{C}(2r) \\ &< 2\epsilon r^2, \end{aligned}$$

ponendo η tale che $2\bar{C}\eta^{\frac{1}{n}} < \epsilon$ e per $0 < r < r_0$, con r_0 opportuno (eventualmente minore di quello iniziale).
Dunque, per $r \to 0^+$, si ha che

$$\sup_{B(\frac{r}{2})} \Big| f(y) - f(0) - Df(0) \cdot y - \frac{1}{2} y^t \cdot D^2 f(0) \cdot y \Big| = o(r^2),$$

da cui la tesi nel caso $x = 0$. Si osservi che l'imposizione $x = 0$ non è restrittiva e che la tesi vale per ogni $x \in N^c$. $\square$

Osservazione 4.1. Si potrebbe tentare di caratterizzare i punti dell'insieme N: il punto di partenza sono ovviamente le Osservazioni (a),(b) e (c), che valgono rispettivamente per i punti di Lebesgue di Df, per quelli di $D^2 f$ e, infine, per l'intersezione di insiemi nulli per le $|\mu_s^{ij}|$, che sono ortogoanli a $\mathcal{L}^n$, dando appunto il complementare di un insieme di misura nulla (cioè N stesso).

Bibliografia

[1] Yu. G. Reshetnyak, *Generalized derivatives and differentiability almost everywhere*, Math. USSR-Sb. **4**, 293-302 (1968)

[2] V. Magnani, *Lipschitz continuity, Aleksandrov theorem and characterizations for H-convex functions*, Math. Ann., **334**, 199-233 (2006)

[3] L.C. Evans and R. Gariepy, *Measure Theory and Fine Properties of Functions*, CRC PRESS, Revised Edition, 2015. Pages 6-7, 10-11, 39-54, 266-276.

[4] G. Bianchi, A. Colesanti and C. Pucci, *On the second differentiability of convex surfaces*, Geometriae Dedicata, **60**, 39-48 (1996)

[5] L.C. Evans, *Partial Differential Equations*, American Mathematical Society, Second Edition, 2010

[6] G.B. Folland, *Real Analysis, Modern Techniques and Their Applications*, Wiley-Interscience, Second Edition, 1999

[7] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Series in Higher Mathematics, Third Edition, 1987, Pages 116-125