Università degli Studi di Milano

Corso di laurea in Matematica
e in Matematica per le applicazioni

Analisi Convessa

APPUNTI NON UFFICIALI

anno accademico 2007/08

Questi appunti riportano lo schema delle lezioni del Corso di Analisi Convessa per il Corso di laurea specialistica in Matematica. Testi di riferimento per il corso:

R.T. Rockafellar, Convex Analysis*. Princeton University Press.*
A.W. Roberts and D.E. Varberg, Convex Functions*, New York Academic Press.*
S.R. Lay, Convex Set and their Applications*, Malabar, Florida Krieger Publishing Company.*
J.R. Giles, Convex Analysis with Application in the Differentiation of Convex Functions*, Boston Pitman Advanced Publishing Program.*

Indice

Capitolo 1

Proprietà topologiche degli insiemi convessi

L'analisi convessa è quella particolare branca dell'analisi che si occupa dello studio degli insiemi e delle funzioni convesse. L'analisi convessa, per certi versi, è molto vicina alle tematiche dell'analisi funzionale lineare svolte nell'omonimo corso: frequentemente, infatti, i risultati dell'analisi convessa sono le prime, naturali estensioni a casi non lineari di proprietà già dimostrate ed analizzate nel caso lineare.

1.1 Definizioni e proprietà di base

Non si può non iniziare con la definizione di insieme convesso e dei suoi parenti più intimi.

Definizione 1.1 (Insieme convesso/affine/lineare) Siano X spazio vettoriale e $A \subseteq X$. Diciamo che A è:

- *convesso* se per ogni $x, y \in A$ e $\lambda \in [0,1]$ si ha $(1-\lambda)x + \lambda y \in A$, ovvero A contiene il segmento che ha per estremi x e y.
- *affine* se per ogni $x, y \in A$ e $\lambda \in \mathbb{R}$ si ha $(1-\lambda)x + \lambda y \in A$, ovvero A contiene la retta per x e y.
- *lineare* se è un sottospazio vettoriale di X.

Osservazione È semplice osservare che intersezione di una famiglia di insiemi convessi (affini, lineari) è ancora un insieme convesso (affine, lineare).

La definizione che segue serve per dare un minimo di nomenclatura agli oggetti su cui investigheremo.

Definizione 1.2 (Combinazione lineare/lineare affine/lineare convessa) Siano X spazio vettoriale e $A \subseteq X$. Chiamiamo

- *combinazione lineare* di elementi di A ogni somma della forma $\sum_{i=1}^{n} \lambda_i x_i$, con $x_i \in A$ e $\lambda_i \in \mathbb{R}$.

- *combinazione lineare affine* di elementi di A ogni somma della forma $\sum_{i=1}^{n} \lambda_i x_i$, con $x_i \in A$, $\lambda_i \in \mathbb{R}$ e $\sum_{i=1}^{n} \lambda_i = 1$.
- *combinazione lineare convessa* di elementi di A ogni somma della forma $\sum_{i=1}^{n} \lambda_i x_i$, con $x_i \in A$, $\lambda_i \geq 0$ e $\sum_{i=1}^{n} \lambda_i = 1$.

Naturalmente i nomi usati in precedenza per definire classi di combinazioni di elementi, non sono casuali ma sono legati dal seguente teorema.

Teorema 1.1 (di Jensen) Siano X spazio vettoriale e $A \subseteq X$. A è convesso (affine, lineare) se e solo se ogni combinazione lineare convessa (affine, lineare) di elementi di A appartiene ad A.

Dimostrazione Ci occuperemo solo del caso di A convesso, gli altri casi sono simili. Una delle due implicazioni è ovvia, dimostreremo solo l'altra. Sia quindi A convesso; si procede per induzione su n, il numero di elementi di cui si fa combinazione.

$(n = 2)$ Ovvio;

$(n \rightsquigarrow (n+1))$ Siano $\{x_1\}_{i=1}^{n+1} \subseteq A$, $\{\lambda_i\}_{i=1}^{n+1} \subseteq \mathbb{R}^+$ tali che $\sum_{i=1}^{n+1} \lambda_i = 1$. Supponiamo che $\lambda_{n+1} \neq 1$, altrimenti è banale, in tal caso

$$\sum_{i=1}^{n+1} \lambda_i x_i = \lambda_{n+1} x_{n+1} + \sum_{i=1}^{n} \lambda_i x_i = \lambda_{n+1} x_{n+1} + (1 - \lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_{n+1}} x_i$$

osservando che $\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} = 1$ e $\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i \in A$ per ipotesi d'induzione, si ha la tesi. $\square$

È interessante osservare come gli insiemi affini e gli insiemi lineari siano profondamente collegati.

Proposizione 1.1 Siano X spazio vettoriale e $A \subseteq X$. Sono tra loro equivalenti:

1. $\forall a, b \in A$ vale $A - a = A - b$ e tale insieme è lineare;
2. A è un insieme affine;
3. A è un traslato di un insieme lineare.

Dimostrazione Dimostriamo solo un'implicazione, le restanti vengono lasciate al lettore.

$(2. \Rightarrow 1.)$ Ovviamente $A - a = (A + (b - a)) - b$, se dimostriamo che $A = A + (b - a)$ abbiamo la tesi. Dimostriamo le due inclusioni: sia $x \in A + (b - a)$ allora

$$x = y + (b - a) \in A$$

essendo combinazione lineare affine di elementi di A. Quindi $A + (b - a) \subseteq A$. Sia ora $x \in A$ si ha

$$x = \underbrace{(x - b + a)}_{\in A} + b - a \in A + (b - a).$$

La verifica che $A - a$ è lineare viene lasciata come facile esercizio. $\square$

Ovviamente non tutti gli insiemi di uno spazio vettoriale risultano essere convessi (affini, lineari), è quindi utile definire cosa intendiamo per involucro convesso (affine, lineare).

Definizione 1.3 (Involucro convesso/affine/lineare) Siano X spazio vettoriale e $E \subseteq X$. Chiamiamo *involucro convesso* (*affine*, *lineare*) di E il più piccolo insieme convesso (affine, lineare) contenente E, ovvero l'intersezione di tutti gli insiemi convessi (affini, lineari) contenenti E. Tale insieme verrà indicato con

$$\mathrm{conv}(E) \qquad (\mathrm{aff}(E),\ \mathrm{span}(E)).$$

La seguente proposizione consente di esprimere l'involucro convesso (affine, lineare) di un insieme in termini dei soli elementi dell'insieme stesso.

Proposizione 1.2 Siano X spazio vettoriale e $E \subseteq X$. L'involucro convesso (affine, lineare) di E è l'insieme delle combinazioni convesse (affini, lineari) di elementi di E.

Dimostrazione Dimostreremo la proposizione solo nel caso dell'involucro convesso, le altre dimostrazioni sono simili. Sia C l'insieme delle combinazioni lineari convesse di elementi di E. È ovvio che

$$C \subseteq \mathrm{conv}(E).$$

Per avere la tesi è sufficiente mostrare, siccome C contiene E, che C è convesso. Siano

$$\sum_{i=1}^{n} \lambda_i x_i, \sum_{j=1}^{m} \mu_j y_j \in C \qquad \alpha, \beta \geq 0,\ \alpha + \beta = 1$$

allora

$$\alpha\left(\sum_{i=1}^{n} \lambda_i x_i\right) + \beta\left(\sum_{j=1}^{m} \mu_j y_j\right) = \sum_{i=1}^{n} (\alpha\lambda_i) x_i + \sum_{j=1}^{m} (\beta\mu_j) y_j;$$

è semplice osservare che $\alpha\lambda_i \geq 0$ e $\beta\mu_j \geq 0$ per ogni i, j e inoltre

$$\sum_{i=1}^{n} \alpha\lambda_i + \sum_{j=1}^{m} \beta\mu_j = \alpha \underbrace{\sum_{i=1}^{n} \lambda_i}_{=1} + \beta \underbrace{\sum_{j=1}^{m} \mu_j}_{=1} = \alpha + \beta = 1.$$

Abbiamo quindi che C è convesso. $\square$

Vogliamo ridurre ancora di più l'insieme su cui fare le combinazioni lineari convesse in modo da renderlo maggiormente maneggiabile, per fare ciò dobbiamo supporre di essere in dimensione finita.

Teorema 1.2 (di Carathéodory) Siano X spazio vettoriale di dimensione finita d e $E \subseteq X$. Ogni elemento di $\mathrm{conv}(E)$ è combinazione lineare di al più $d+1$ elementi distinti di E.

Dimostrazione Siano $v_0, \ldots, v_n \in E$ con $n > d$, altrimenti non c'è niente da dimostrare, e consideriamo

$$x = \sum_{i=0}^{n} \lambda_i v_i \qquad \sum_{i=0}^{n} \lambda_i = 1,\ \lambda_i \geq 0.$$

Senza perdita di generalità trasliamo l'insieme E in modo tale che $0 \in E$ e supponiamo che $v_0 = 0$. Siccome $n > d$, si ha che i v_i sono linearmente dipendenti e quindi esistono $\{\mu_j\}_{j=1}^{n} \subseteq \mathbb{R}$ non tutti nulli tali che

$$\sum_{j=1}^{n} \mu_j v_j = 0.$$

Senza perdita di generalità possiamo supporre che $\sum_{j=1}^n \mu_j \geq 0$. Per ogni $t > 0$ vale

$$x = \sum_{i=1}^n \lambda_i v_i - t \sum_{j=1}^n \mu_j v_j = \sum_{i=1}^n (\lambda_i - t\mu_i) v_i,$$

vorremo che questa sia una parte di una combinazione convessa e quindi vogliamo $(\lambda_i - t\mu_i) \geq 0$. Osserviamo che se $\mu_i \leq 0$ non ci sono problemi, quindi basta porre[1]

$$t = \frac{\lambda_k}{\mu_k} = \min \left\{ \frac{\lambda_i}{\mu_i} \,:\, 1 \leq i \leq n,\ \mu_i > 0 \right\}.$$

Per tale t si ha

$$x = \sum_{\substack{i=1 \\ i \neq k}}^n (\lambda_i - t\mu_i) v_i.$$

Ora useremo per la prima volta un trucco che ritroveremo anche in seguito. È semplice osservare che

$$s := \sum_{\substack{i=1 \\ i \neq k}}^n (\lambda_i - t\mu_i) = \sum_{i=1}^n (\lambda_i - t\mu_i) = \underbrace{\sum_{i=1}^n \lambda_i}_{\leq 1} \underbrace{-t \sum_{i=1}^n \mu_i}_{\leq 0} \leq 1$$

quindi

$$x = \sum_{\substack{i=1 \\ i \neq k}}^n (\lambda_i - t\mu_i) v_i + (1-s) \cdot 0$$

e questa è una combinazione lineare di n elementi non più $n+1$. Si replica questo processo fino a giungere a $d+1$ elementi.[2] □

È utile dare un concetto di dimensione per insiemi convessi, bisogna però tenere presente che in generale il seguente concetto di dimensione non coincide con la dimensione topologica.

Definizione 1.4 (Dimensione) Siano X spazio vettoriale e $E \subseteq X$. Se E è affine allora la *dimensione* di E è

$$\dim(E) = \dim(E - e) \qquad \forall e \in E.$$

In generale definiamo *dimensione* di E come segue:

$$\dim(E) = \dim(\operatorname{aff}(E)).$$

Quindi un semplice corollario del teorema di Carathéodory è il seguente.

Corollario Siano X spazio vettoriale e $E \subseteq X$ sottoinsieme di dimensione finita d. Vale

$$\operatorname{conv}(E) = \left\{ \sum_{i=1}^{d+1} \lambda_i v_i \,:\, v_i \in E,\ \lambda_i \geq 0,\ \sum_{i=1}^{d+1} \lambda_i = 1 \right\}.$$

---

[1] L'insieme $\left\{ \frac{\lambda_i}{\mu_i} \,:\, 1 \leq i \leq n,\ \mu_i > 0 \right\}$ è non vuoto siccome abbiamo supposto $\sum_{j=1}^n \mu_j \geq 0$ e i μ_j non sono tutti nulli.

[2] Non si può scendere sotto $d+1$ in generale, si consideri per esempio $\mathbb{R}$ l'involucro convesso dei punti 0 e 1.

Lasciamo un paio di esercizi di base sugli insiemi convessi.

Esercizio 1.1.1 Siano X spazio vettoriale e $A \subseteq X$. A è convesso se e solo se per ogni $\alpha, \beta > 0$ vale $\alpha A + \beta A = (\alpha + \beta)A$.

Esercizio 1.1.2 Siano X uno spazio normato, $\alpha_i, r_i \geq 0$, $x_i \in X$ con $i = 1, \ldots, n$. Dimostrare che

$$\sum_{i=1}^{n} \alpha_i \, \mathcal{B}(x_i, r_i) = \mathcal{B}\left(\sum_{i=1}^{n} \alpha_i x_i, \sum_{i=1}^{n} \alpha_i r_i\right).$$

Ricordando che $\mathcal{B}(x, r) = \{y \in X \, : \, \|y - x\| \leq r\}$.

1.2 Compattezza

Lasciamo ora la pura algebra e iniziamo a studiare alcune proprietà topologiche degli insiemi convessi, in particolare in questa sezione ci occuperemo di trovare condizioni sufficienti affinché l'involucro convesso di un insieme sia compatto.

Chiamiamo il seguente risultato corollario siccome è diretta conseguenza del teorema di Carathéodory.

Corollario Siano X spazio vettoriale topologico e $K \subseteq X$. Se K è compatto e finito-dimensionale allora $\operatorname{conv}(K)$ è compatto.

Dimostrazione Sia $\dim K = d$ e definiamo

$$\Lambda = \left\{ \lambda = (\lambda_1, \ldots, \lambda_{d+1}) \in \mathbb{R}^{d+1} \, : \, \lambda_i \geq 0, \, \sum_{i=1}^{d+1} \lambda_i = 1 \right\},$$

che è un compatto di $\mathbb{R}^{d+1}$. Vogliamo ora che $\operatorname{conv}(K)$ sia immagine continua di un compatto, per fare ciò definiamo

$$\begin{aligned} F : \Lambda \times X^{d+1} &\longrightarrow X \\ (\lambda, v_1, \ldots, v_{d+1}) &\longrightarrow \sum_{i=1}^{d+1} \lambda_i v_i \end{aligned}$$

Tale funzione è continua e $\Lambda \times K^{d+1}$ è compatto, quindi $F(\Lambda \times K^{d+1}) = \operatorname{conv}(K)$ è compatto. $\square$

Osservazione L'ipotesi $\dim K < +\infty$ è essenziale. Si consideri $X = \ell^2$ e il compatto

$$K = \left\{ \frac{1}{n} e_n \right\}_{n \in \mathbb{N}} \cup \{0\}.$$

Sappiamo che se $x \in \operatorname{conv}(K)$, allora x è a supporto finito. Consideriamo $\lambda_i > 0$ tali che $\sum_{i=1}^{+\infty} \lambda_i = 1$ e $y_n \in K$ definito come

$$y_n = \sum_{i=1}^{n} \frac{\lambda_i}{i} e_i + \left(1 - \sum_{j=1}^{n} \lambda_j\right) \cdot 0.$$

è semplice osservare che

$$y_n \xrightarrow{\|\cdot\|} y = \sum_{i=1}^{+\infty} \frac{\lambda_i}{i} e_i.$$

Ma $y \notin \operatorname{conv}(K)$ siccome non è a supporto finito.

Vogliamo adesso considerare la possibilità di convessificare l'unione di un numero finito di insiemi convessi, in particolare ci interesserà quando tale convesso risulti compatto o chiuso.

Proposizione 1.3 Siano X spazio vettoriale e $C_1, \dots, C_n \subseteq X$ convessi. Valgono

1. $\operatorname{conv}\left(\bigcup_{i=1}^n C_i\right) = \{\sum_{i=1}^n \lambda_i x_i \;:\; x_i \in C_i,\; \lambda_i \geq 0,\; \sum_{i=1}^n \lambda_i = 1\}$;
2. se X è uno spazio vettoriale topologico e tutti i C_i sono compatti, allora $\operatorname{conv}\left(\bigcup_{i=1}^n C_i\right)$ è compatto;
3. se X è uno spazio vettoriale topologico e C_i è compatto per $i = 1, \dots, n-1$, mentre C_n è chiuso e limitato, allora $\operatorname{conv}\left(\bigcup_{i=1}^n C_i\right)$ è chiuso.

Dimostrazione Poniamo $C = \operatorname{conv}\left(\bigcup_{i=1}^n C_i\right)$.

1. Se $x \in C$ allora x è combinazione convessa di elementi di $\bigcup_{i=1}^n C_i$. Fissiamo k e consideriamo la parte riguardante C_k
$$\sum_{j \in F_k} \lambda_j y_j \tag{$\star$}$$
con $y_j \in C_k$. Poniamo $s_k = \sum_{j \in F_k} \lambda_j$ ed osserviamo che se $s_k \neq 0$, altrimenti la parte riguardante C_k può essere elisa, è possibile riscrivere $(\star)$ come una combinazione convessa di elementi di C_k moltiplicata per un elemento positivo nel seguente modo
$$\sum_{j \in F_k} \lambda_j y_j = s_k \sum_{j \in F_k} \frac{\lambda_j}{s_k} y_j.$$
Osservando che $\sum_{k=1}^n s_k = 1$ e iterando questo processo per ogni $k = 1, \dots, n$ si ha la tesi.
2. Definiamo
$$\Lambda = \left\{ \lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n \;:\; \lambda_i \geq 0,\; \sum_{i=1}^n \lambda_i = 1 \right\},$$
che è un compatto di $\mathbb{R}^n$. Vogliamo ora che C sia immagine continua di un compatto, per fare ciò definiamo
$$\begin{aligned} F : \Lambda \times X^n &\longrightarrow X \\ (\lambda, x_1, \dots, x_n) &\longrightarrow \sum_{i=1}^n \lambda_i x_i \end{aligned}$$
Tale funzione è continua e $\Lambda \times C_1 \times \cdots \times C_n$ è compatto, quindi
$$F(\Lambda \times C_1 \times \cdots \times C_n) = C$$
è compatto.
3. Sia $\{c_k\} \subseteq C$ una successione tale che $c_k \to z$. Sappiamo che ogni c_k è esprimibile come combinazione convessa nel seguente modo
$$c_k = \sum_{i=1}^n \lambda_i^{(k)} x_i^{(k)} \qquad x_i^{(k)} \in C_i.$$

Per compattezza sappiamo che passando ad un'opportuna sottosuccessione è possibile supporre che

$$\lambda_i^{(k)} \xrightarrow{k\to+\infty} \lambda_i \in [0,1] \qquad i = 1, \ldots, n$$
$$x_i^{(k)} \xrightarrow{k\to+\infty} x_i \in C_i \qquad i = 1, \ldots, n-1$$

Distinguiamo ora due casi

($\lambda_n = 0$) Siccome $x_n^{(k)}$ è limitata allora $\lambda_n^{(k)} x_n^{(k)} \to 0$. Quindi

$$z = \lim_{k\to+\infty} \sum_{i=1}^{n-1} \lambda_i^{(k)} x_i^{(k)}.$$

Per la chiusura di $\operatorname{conv}\left(\bigcup_{i=1}^{n-1} C_i\right)$ si ha

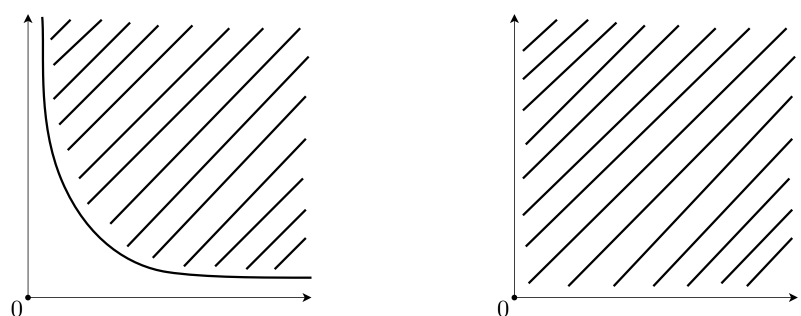
$$z \in \operatorname{conv}\left(\bigcup_{i=1}^{n-1} C_i\right) \subseteq C.$$

($\lambda_n > 0$) Per il teorema della permanenza del segno esiste k_0 tale che per ogni $k \geq k_0$ vale $\lambda_n^{(k)} > 0$. Quindi

$$x_n^{(k)} = \frac{1}{\lambda_n^{(k)}}\left(c_k - \sum_{i=1}^{n-1} \lambda_i^{(k)} x_i^{(k)}\right) \xrightarrow{k\to+\infty} \frac{1}{\lambda_n}\left(z - \sum_{i=1}^{n-1} \lambda_i x_i\right) = x_n \in C_n.$$

In definitiva $z = \sum_{i=1}^{n} \lambda_i x_i$, ovvero $z \in C$. $\square$

Osservazione È utile osservare che in generale l'involucro convesso di un chiuso non è chiuso. Si consideri il seguente esempio:



S — involucro convesso di S

Esercizio 1.2.1 Siano X spazio vettoriale topologico e $A \subseteq X$. Dimostrare che se A è aperto allora $\operatorname{conv}(A)$ è aperto.

Passiamo ora allo studio delle proprietà di compattezza della chiusura dell'involucro convesso di un insieme in spazi di Banach.

Teorema 1.3 Siano X normato e $E \subseteq X$ totalmente limitato (=precompatto). Valgono

1. $\operatorname{conv}(E)$ è totalmente limitato;

2. se X è uno spazio di Banach allora $\overline{\mathrm{conv}}(E)$ è compatto.

Dimostrazione 2. deriva da 1. ricordando che in spazi metrici compatto è equivalente a totalmente limitato e completo. Dimostriamo quindi 1. Fissiamo $\varepsilon > 0$ e consideriamo una ε-net finita associata $\{x_i\}_{i=1}^n$. Facendo i conti il compatto

$$\mathrm{conv}\left(\{x_i\}_{i=1}^n\right)$$

risulta essere un ε-net, in generale non finita, per $\mathrm{conv}(E)$. Per compattezza possiamo estrarre una ε-net finita di elementi di $\mathrm{conv}\left(\{x_i\}_{i=1}^n\right)$ e questa risulta essere un 2ε-net per $\mathrm{conv}(E)$. □

Corollario Siano X uno spazio di Banach e $K \subseteq X$ compatto. Si ha che $\overline{\mathrm{conv}}(K)$ è compatto.

Osservazione Se X non è uno spazio di Banach esiste sicuramente un insieme compatto E tale che $\overline{\mathrm{conv}}(E)$ non sia compatto. Si consideri in X una successione di Cauchy non convergente $\{y_n\}_{n=1}^\infty$. Consideriamo la successione $\{z_n\}$, definita come

$$z_i = y_{i+1} - y_i.$$

Tale successione ammette una sottosuccessione $\{z_{n_i}\}$ tale che

$$\|z_{n_i}\| = \|y_{n_{i+1}} - y_{n_i}\| < \frac{1}{2^{2i}}.$$

Tale sottosuccessione può essere costruita con un processo di diagonalizzazione: $\{y_n\}$ è di Cauchy e quindi $\{z_n\}$ ammette una sottosuccessione $\{z_{1,n}\}$ costituita da elementi minori di $\frac{1}{2}$ in norma. La $\{z_{1,n}\}$ continua a tendere a zero e quindi è possibile estrarre una sottosuccessione $\{z_{2,n}\}$ in cui ogni $z_{2,n}$ è minore di $\frac{1}{2^2}$ in norma. Questa, a sua volta, continuerà a tendere a zero e così via. Procedendo così abbiamo costruito una successione di sottosuccessioni $\{z_{m,n}\}$ tali che $\|z_{m,n}\| < \frac{1}{2^{2m}}$. A questo punto costruiamo la nostra candidata sottosuccessione estraendo la diagonale delle $\{z_{m,n}\}$. La $\{y_{n_i}\}$ continua ad essere di Cauchy e non convergente in X, per l'unicità del limite nel completamento di X. Indichiamo ora la successione $\{y_{n_i}\}$ con $\{y_n\}$ e sia $\{z_n\}$ la successione dei $z_i = y_{i+1} - y_i$, con $\|z_i\| < \frac{1}{2^{2i}}$. Osserviamo che $\sum_{i=1}^n z_i = y_n$, quindi la serie $\sum_{n=1}^\infty z_n$ non converge.
Siano dunque

$$\{x_n\} = 2^n z_n \quad \text{e} \quad \{a^n\} = \frac{1}{2^n}.$$

Per come abbiamo costruito $\{z_n\}$, la successione $\{x_n\}$ converge a zero; infatti

$$\|x_n\| = \|2^n z_n\| < \frac{1}{2^n},$$

e la serie $\sum_{i=1}^\infty a^n x_n = \sum_{i=1}^\infty z_n$ non converge. Consideriamo $K = \{x_n\} \cup \{0\}$. K è compatto siccome è costituito da una successione convergente unita al suo limite. Sia ora $C = \overline{\mathrm{conv}}(K)$ la chiusura convessa di K. Osserviamo che

$$\sum_{n=1}^{+\infty} a^n = \sum_{n=1}^{+\infty} \frac{1}{2^n} = 1,$$

quindi per ogni $j \geq 1$

$$\sum_{n=1}^{j} a^n x_n = \sum_{n=1}^{j} \frac{1}{2^n} x_n + \left(1 - \sum_{n=1}^{j} \frac{1}{2^n}\right) \cdot 0$$

è una combinazione convessa di elementi di K e dunque

$$y_j = \sum_{n=1}^{j} a^n x_n \in \mathrm{conv}(K) \qquad \forall j \geq 1.$$

Ora se C fosse compatto dovrebbe essere completo, ma la successione $\{y_j\}$ è una successione di Cauchy non convergente, assurdo.

Dal punto di vista dell'analisi funzionale rivestono un'importanza vitale la topologia debole e la topologia debole*. Siamo quindi interessati a condizioni che ci assicurino la compattezza della chiusura dell'involucro convesso di un insieme in tali topologie.

Proposizione 1.4 Siano X normato e $E \subseteq X^*$. Se E è limitato allora $\overline{\mathrm{conv}}^{w^*}(E)$ è w^*-compatto.

Dimostrazione Dalla limitatezza di E si ha che $\overline{\mathrm{conv}}^{w^*}(E)$ è limitato. Abbiamo dunque che $\overline{\mathrm{conv}}^{w^*}(E)$ è w^*-chiuso e limitato, ovvero w^*-compatto. □

È naturale chiedersi se un simile teorema valga anche per la topologia debole. La risposta è affermativa ma ovviamente non possiamo sperare in una dimostrazione simile a quella fatta per spazi di Banach, né simile a quella fatta per la topologia debole*. In realtà il teorema che andiamo ad enunciare ha una dimostrazione molto complessa, ma noi la semplificheremo utilizzando il seguente teorema.

Teorema 1.4 (di James) Siano X spazio di Banach e $C \subseteq X$. Se C è convesso e chiuso, allora sono tra loro equivalenti:

1. C è w-compatto;
2. ogni $x^* \in X^*$ assume il suo massimo su C.

Di questo teorema ci guarderemo bene dal fornirne una dimostrazione, in compenso come facile corollario otteniamo il seguente.

Teorema 1.5 (di Krein–Šmulyan) Siano X spazio di Banach e $K \subseteq X$. Se K è w-compatto allora $\overline{\mathrm{conv}}(K)$ è w-compatto.[3]

Dimostrazione Sia $x^* \in X^*$, per continuità sappiamo che esiste $y_0 \in K$ tale che

$$x^*(y_0) = \max x^*(K).$$

L'equazione $x^*(x) = x^*(y_0)$ definisce un iperpiano in X. K è contenuto in uno dei due semispazi definiti da tale iperpiano, ma anche $\overline{\mathrm{conv}}(K)$ lo è. Necessariamente si ha $\max x^*(K) = \max x^*(\overline{\mathrm{conv}}(K))$. La tesi segue dal teorema di James. □

1.3 Interno relativo e algebrico

Passiamo ora allo studio delle proprietà riguardanti l'interno degli insiemi convessi. Incominciamo con la seguente semplice osservazione.

[3]Ricordiamo che per insiemi convessi la chiusura debole e la chiusura forte coincidono.

Osservazione Siano X spazio vettoriale topologico e $C \subseteq X$ convesso. Se $x \in \operatorname{int} C$ e $y \in C$ allora
$$[x, y) \subseteq \operatorname{int} C.$$

I seguenti teoremi permetteranno di comprendere come, dal punto di vista finito-dimensionale, gli insiemi convessi godano di ottime proprietà riguardanti l'interno.

Lemma Siano X normato di dimensione finita n e $C \subseteq X$ sottoinsieme convesso. Se $\operatorname{aff}(C) = X$ allora C ha punti interni.

Dimostrazione L'idea è quella di mostrare che C contiene n punti linearmente indipendenti (un poliedro) e dimostrare che l'involucro convesso di tali punti ha interno non vuoto. Rigorosamente, senza perdita di generalità possiamo supporre che $0 \in C$, dal fatto che
$$\operatorname{span}(C) = \operatorname{aff}(C) = X$$
possiamo dire che esistono $\{c_1, \ldots, c_n\} \subseteq C$ che siano una base per X. Consideriamo ora la seguente mappa $T : X \to \mathbb{R}^n$ definita come segue
$$T\left(\sum_{i=1}^{n} \alpha_i c_i\right) = (\alpha_1, \ldots, \alpha_n),$$
è semplice osservare che T è un omeomorfismo lineare. Sapendo che[4]
$$\left\{\sum_{i=1}^{n} \lambda_i c_i \,:\, \lambda_i > 0 \ \sum_{i=1}^{n} \lambda_i < 1\right\} \subseteq C$$
da cui
$$\left\{(\lambda_1, \ldots, \lambda_n) \,:\, \lambda_i > 0 \ \sum_{i=1}^{n} \lambda_i < 1\right\} \subseteq T(C).$$
Quindi $T(C)$ contiene un aperto non vuoto, ma sfruttando il fatto che T è un omeomorfismo lineare allora anche C contiene un aperto non vuoto. □

Con la seguente definizione possiamo arrivare ad un teorema molto importante riguardante i convessi finito-dimensionali.

Definizione 1.5 (Interno relativo) Siano X spazio vettoriale topologico e $E \subseteq X$. Si chiama *interno relativo* di E l'interno di E nel più piccolo insieme affine che lo contiene, ovvero
$$\operatorname{ri}(E) = \operatorname{int}_{\operatorname{aff}(E)}(E).$$

Corollario Siano X spazio vettoriale topologico e $C \subseteq X$ convesso. Se C ha dimensione finita, allora $\operatorname{ri}(C) \neq \emptyset$.

Sappiamo che in dimensione finita l'interno di un insieme convesso coincide con l'interno della sua chiusura, è naturale chiedersi se un tale risultato può valere ancora in dimensione infinita. Purtroppo la risposta è in generale no, ma il seguente teorema ci fornisce delle condizioni sufficienti.

Teorema 1.6 Siano X spazio vettoriale topologico e $C \subseteq X$ convesso. Si ha

[4] Questo sempre sfruttando la possibilità di scaricare ciò che manca per raggiungere una combinazione convessa sullo zero.

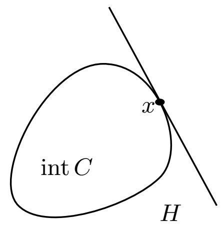
1. se $\operatorname{int} C \neq \emptyset$ allora $\operatorname{int} C = \operatorname{int} \overline{C}$;

2. se $\dim C < +\infty$ allora $\operatorname{int} C = \operatorname{int} \overline{C}$;

3. se X è localmente convesso e $\operatorname{int} C \neq \emptyset$ allora $\overline{C} = \overline{\operatorname{int} C}$

Dimostrazione

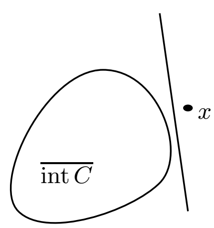
1. Dal fatto che $C \subseteq \overline{C}$ segue $\operatorname{int} C \subseteq \operatorname{int} \overline{C}$. Sia ora $x \notin \operatorname{int} C$



per Hahn–Banach esiste un semispazio chiuso H contenente $\operatorname{int} C$ e che abbia sul bordo x. Per l'osservazione di inizio sezione $C \subseteq H$, quindi $\overline{C} \subseteq H$, in conclusione $\operatorname{int} \overline{C} \subseteq \operatorname{int} H$. Abbiamo quindi la tesi, ovvero $x \notin \operatorname{int} \overline{C}$.

2. Se $\operatorname{int} C \neq \emptyset$ possiamo applicare il punto precedente. Supponiamo che $\operatorname{int} C = \emptyset$ e per assurdo $x \in \operatorname{int} \overline{C}$. Siccome siamo in dimensione finita $\operatorname{aff}(C)$ è chiuso e quindi
$$\operatorname{aff}(C) = \operatorname{aff}(\overline{C}).$$
Dal fatto che $\operatorname{int} \overline{C} \neq \emptyset$ si ottiene che C è denso in una bolla. Quindi $\operatorname{aff}(C)$ è denso in X e, per la chiusura, $\operatorname{aff}(C) = X$. Per il lemma $\operatorname{int} C \neq \emptyset$, abbiamo ottenuto l'assurdo.

3. Dal fatto che $\operatorname{int} C \subseteq C$ segue $\overline{\operatorname{int} C} \subseteq \overline{C}$. Sia $x \notin \overline{\operatorname{int} C}$ e separiamo con Hahn–Banach.



Per assurdo supponiamo $x \in \overline{C}$, per il primo punto sappiamo che $\operatorname{int} \overline{C} = \operatorname{int} C$. Sia $y \in \operatorname{int} C \subseteq \overline{\operatorname{int} C}$, per l'osservazione di inizio sezione
$$[y, x) \subseteq \operatorname{int} \overline{C} = \operatorname{int} C \subseteq \overline{\operatorname{int} C}.$$
Da qui l'assurdo per la separazione forte. □

Osservazione È interessante osservare che, se $\dim X = +\infty$, nel punto 1. del teorema sia imprescindibile l'ipotesi $\operatorname{int} C \neq \emptyset$. Infatti in dimensione infinita esiste almeno un funzionale lineare discontinuo $l : X \to \mathbb{R}$, posto $C = l^{-1}(0)$, esso è convesso, denso e ha interno vuoto, ma $\operatorname{int} \overline{C} = X$.

Passiamo ora a studiare le proprietà dell'interno algebrico.

Definizione 1.6 (Interno algebrico) Siano X spazio vettoriale e $E \subseteq X$. L'*interno algebrico* di E è il seguente insieme

$$\text{a-int}(E) = \{x \in E : \forall v \in X\ \exists \delta > 0 \text{ tale che } [x, x + \delta v) \subseteq E\}.$$

Equivalentemente, $x \in \text{a-int}\, E$ se e solo se, per ogni retta $r \subseteq X$ passante per x, si ha $x \in \text{ri}(E \cap r)$.

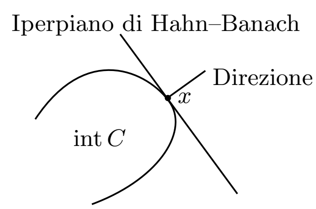
È ovvio che $\text{int}\, E \subseteq \text{a-int}(E)$ è naturale chiedersi quando tali insiemi coincidono. Il seguente teorema ci fornisce alcune condizioni sufficienti.

Teorema 1.7 Siano X spazio vettoriale topologico e $C \subseteq X$ convesso. L'uguaglianza $\text{int}\, C = \text{a-int}(C)$ vale in ciascuno dei seguenti casi:

1. $\text{int}\, C \neq \emptyset$;
2. $\dim C < +\infty$;
3. se X è uno spazio di Banach e C è un F_σ.

Dimostrazione

1. Sia $x \in C \setminus \text{int}\, C$, usiamo Hahn–Banach per separare x da $\text{int}\, C$. Dal disegno[5] si evince che $x \notin \text{a-int}\, C$, basta infatti considerare una direzione nel semispazio che non contiene C.



2. Se $\text{a-int}\, C = \emptyset$ abbiamo la tesi. Sia quindi $x \in \text{a-int}\, C$, senza perdita di generalità possiamo supporre che $x = 0$. In tal caso sia $\{e_i\}_{i \in I}$ una base di Hamel per X, allora $[0, \delta_i e_i) \subseteq C$ per opportuni δ_i. Da ciò segue che $X = \text{aff}(C)$ e dunque $\dim X = \dim C < +\infty$, quindi $\text{int}\, C \neq \emptyset$.
3. Senza perdita di generalità possiamo supporre che $0 \in \text{a-int}\, C$. Per definizione di interno algebrico

$$X = \bigcup_{t \geq 0} tC = \bigcup_{n \in \mathbb{N}} nC = \bigcup_{n \in \mathbb{N}} n \left(\bigcup_{k \in \mathbb{N}} F_k \right)$$

ove la seconda uguaglianza è dovuta alla convessità di C e la terza al fatto che C è un F_σ, quindi supporremo che ognuno degli F_k sia chiuso. Per il teorema di Baire esistono k, n tali che $\text{int}\, nF_k \neq \emptyset$. Da ciò $\text{int}\, C \neq \emptyset$. □

[5]Non è esattamente la cosa più rigorosa, ma permette di risparmiarsi una dimostrazione formale noiosa.

Capitolo 2

Proprietà topologiche delle funzioni convesse

L'idea di questo capitolo, come è possibile evincere dal titolo, sarà quella di analizzare le proprietà topologiche delle funzioni convesse, in particolare sotto quali ipotesi sia possibile ottenere la continuità. È ovviamente imprescindibile la definizione di funzione convessa.

Definizione 2.1 (Funzione convessa) Siano X spazio vettoriale, $C \subseteq X$ convesso e $f : C \to [-\infty, +\infty]$. f si dice *convessa* su C se l'*epigrafico* di f, ovvero l'insieme

$$\operatorname{epi}(f) = \{(x,t) \in X \times \mathbb{R} : x \in C,\ f(x) \leq t\},$$

è convesso.

Esempio La funzione

$$f(x) = \begin{cases} (\operatorname{sgn}(x))(+\infty) & x \in \mathbb{R} \setminus \{0\} \\ 1 & x = 0 \end{cases}$$

è convessa su $\mathbb{R}$.

Esercizio 2.0.1 Siano X spazio vettoriale, $C \subseteq X$ convesso e $f : C \to [-\infty, +\infty]$. f è convessa su C se e solo se $\forall x, y \in C$ e $\forall \lambda \in [0,1]$ si ha

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

laddove l'espressione sulla destra abbia senso, ovvero non ci siano forme d'indecisione del tipo somme di infiniti di segno discorde.

Il seguente teorema è un risultato importante, ma dalla dimostrazione non impossibile.[1]

Teorema 2.1 (di Jensen) Siano X spazio vettoriale, $C \subseteq X$ sottoinsieme convesso e $f : C \to [-\infty, +\infty]$. f è convessa se e solo se per ogni $n \in \mathbb{N}$, $x_1, \ldots, x_n \in C$ e $\lambda_1, \ldots, \lambda_n \geq 0$ con $\sum_{i=1}^n \lambda_i = 1$ vale

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

[1] La dimostrazione viene lasciata per esercizio, come suggerimento su consiglia di procedere per induzione su n.

2.1 Continuità

Vogliamo giungere ad un teorema che ricorda molto quello che caratterizza la continuità dei funzionali lineari. Ma per fare ciò abbiamo bisogno di un congruo numero di lemmi preliminari.

Lemma 1 Siano X spazio vettoriale, $C \subseteq X$ convesso e simmetrico rispetto all'origine. Se $f : C \to \mathbb{R}$ è convessa e superiormente limitata allora f è limitata su C. Sarà inoltre valida la stima

$$\|f\|_\infty \leq \max\{|\sup f(C)|, |2f(0) - \sup f(C)|\}.$$

Dimostrazione Sia $x \in C$, possiamo scrivere $0 = \frac{1}{2}x + \frac{1}{2}(-x)$ e per convessità

$$f(0) \leq \frac{1}{2}f(x) + \frac{1}{2}f(-x) \leq \frac{1}{2}f(x) + \frac{1}{2}\sup f(C).$$

Da cui $f(x) \geq 2f(0) - \sup f(C)$. □

Lemma 2 Siano X normato, $D \subseteq C \subseteq X$ convessi e $f : C \to \mathbb{R}$ convessa e limitata. Se

$$D + r\,\mathcal{B}_X \subseteq C$$

per un r opportuno allora f è lipschitziana su D. Si otterrà inoltre la stima $\mathrm{Lip}(f) \leq \frac{2\|f\|_\infty}{r}$. ove $\mathrm{Lip}(f)$ indica la costante di Lipschitz di f.

Dimostrazione Siano $x, y \in D$ con $x \neq y$. Consideriamo $z = y + r\frac{y-x}{\|y-x\|} \in C$. Vogliamo scrivere y come combinazione convessa di elementi di C, e con un po' di conti si ottiene

$$y = \frac{\|y-x\|}{\|y-x\|+r}z + \frac{r}{\|y-x\|+r}x.$$

Stimiamo f in y

$$f(y) \leq \frac{\|y-x\|}{\|y-x\|+r}f(z) + \frac{r}{\|y-x\|+r}f(x)$$

e facendo due conti

$$f(y) - f(x) \leq \frac{f(z) - f(y)}{r}\|y-x\| \leq \frac{2\|f\|_\infty}{r}\|y-x\|.$$

Ripetendo lo stesso ragionamento per x si ottiene la lipschitzianità, più precisamente

$$|f(x) - f(y)| \leq \frac{2\|f\|_\infty}{r}\|x-y\|.$$

□

Corollario Siano X normato, $r > 0$, $\varepsilon \in (0, r)$, $x_0 \in X$ e

$$f : \mathcal{B}(x_0, r) \to \mathbb{R}.$$

Se f è convessa e limitata superiormente allora f è lipschitziana su $\mathcal{B}(x_0, r-\varepsilon)$.

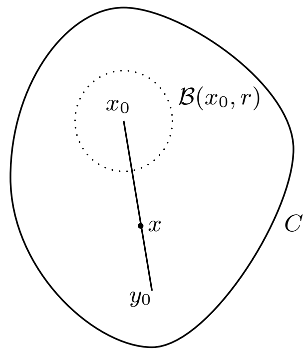
Dimostrazione f è convessa e limitata superiormente su un insieme simmetrico, quindi per li lemma 1 f è limitata e applicando il lemma 2 f risulta essere lipschitziana su ogni bolla più piccola contenuta in $\mathcal{B}(x_0, r)$. Inoltre vale la stima

$$\mathrm{Lip}(f) \leq \phi(f(x_0), \sup f(\mathcal{B}(x_0, r)), \varepsilon)$$

con ϕ funzione opportuna. □

Lemma 3 Siano X normato, $C \subseteq X$ aperto convesso e $f : C \to \mathbb{R}$ convessa. Se f è limitata superiormente in un intorno di qualche $x_0 \in C$, allora f è localmente limitata superiormente su C (e quindi localmente lipschitziana).

Dimostrazione Il disegno permette una maggiore comprensione di ciò che faremo.



f è limitata superiormente su $\mathcal{B}(x_0, r) \subseteq C$. Sia $x \in C$ arbitrario. Siccome C è aperto, esiste $y_0 \in C$ tale che $x \in (x_0, y_0)$. Quindi

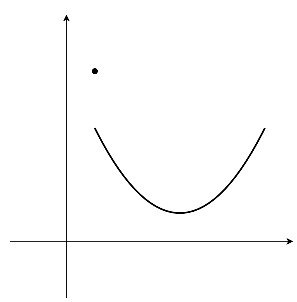
$$x \in \operatorname{int}(\operatorname{conv}(\{y_0\} \cup \mathcal{B}(x_0, r))) \subseteq C,$$

ovvero esiste $\rho > 0$ tale che $\mathcal{B}(x, \rho) \subseteq \operatorname{conv}(\{y_0\} \cup \mathcal{B}(x_0, r))$. Sia $z \in \mathcal{B}(x, \rho)$, sappiamo che esistono $\lambda \in [0, 1]$ e $w \in \mathcal{B}(x_0, r)$ tali che $z = (1 - \lambda)y_0 + \lambda w$ e quindi

$$f(z) \leq (1 - \lambda) f(y_0) + \lambda f(w) \leq f(y_0) + \sup f(\mathcal{B}(x_0, r)). \qquad \square$$

Osservazione Vogliamo ora arrivare al teorema che caratterizza la sezione, ma per fare ciò sono utili due piccole osservazioni.

- Se C non è aperto una funzione convessa può non essere continua sebbene sia limitata superiormente, si consideri per esempio il seguente diagramma



- Esistono funzioni convesse, definite su tutto lo spazio, che non sono continue. Ad esempio un funzionale lineare discontinuo. Vedremo in seguito che in dimensione finita ciò non può accadere.

Teorema 2.2 Siano X normato, $C \subseteq X$ aperto convesso e $f : C \to \mathbb{R}$ convessa. Sono tra loro equivalenti

1. f è localmente lipschitziana su C;
2. f è continua su C;
3. f è continua in un punto di C;

4. f è localmente limitata su C;
5. esiste $A \subseteq C$ aperto non vuoto tale che
$$\sup_{x \in A} f(x) < +\infty.$$

Dimostrazione La dimostrazione avviene con il seguente schema:
$$1. \Longrightarrow 2. \Longrightarrow 3. \Longrightarrow 5. \overset{\text{Lemma 3}}{\Longrightarrow} 4. \overset{\text{Corollario}}{\Longrightarrow} 1.$$
Osserviamo inoltre che, con i lemmi fin'ora enunciati, risultano di facile dimostrazione anche le implicazioni 2. $\Rightarrow$ 4., 3. $\Rightarrow$ 5. e 4. $\Rightarrow$ 5. □

Vediamo ora cosa accade in dimensione finita.

Teorema 2.3 Siano X spazio vettoriale topologico, $C \subseteq X$ convesso e finito-dimensionale e $f : C \to \mathbb{R}$. Se f è convessa allora f è continua sull'interno relativo di C.

Dimostrazione Basta limitare il nostro studio a $\operatorname{aff}(C)$, quindi senza perdita di generalità possiamo supporre che $\dim X = n$ finita e $\operatorname{int} C \neq \emptyset$. Vogliamo mostrare che f è localmente limitata in un punto. Sia $x_0 \in C$, sappiamo che in dimensione finita ogni punto ammette un intorno poliedrale, quindi esistono $\{y_1, \ldots, y_n\} \subseteq C$ tali che
$$x_0 \in U = \operatorname{int}(\operatorname{conv}(\{y_1, \ldots, y_n\})).$$
Sia $u \in U$ tale che $u = \sum_{i=1}^{n} \lambda_i y_i$ combinazione lineare convessa, per il teorema di Jensen si ha
$$f(u) \leq \sum_{i=1}^{n} \lambda_i f(y_i) \leq \max_{x \in \{y_1, \ldots, y_n\}} f(x).$$
Abbiamo quindi ottenuto che f è localmente limitata su U. □

Osservazione In particolare una funzione convessa finita su un aperto $C \subseteq \mathbb{R}^d$ è continua su C.

2.2 Semicontinuità e convessità

Ricordiamo le definizioni di funzioni semicontinue inferormente e superiormente, alcune loro proprietà di base; in seguito esploreremo le correlazioni tra semicontinuità e continuità per una funzione convessa.

Definizione 2.2 (Funzione semicontinua) Siano X spazio topologico di Hausdorff, $x_0 \in X$ e $f : X \to [-\infty, +\infty]$. Diciamo che

- f è *semicontinua inferiormente* in x_0 se
$$\forall \alpha \in \mathbb{R} \text{ per cui } f(x_0) > \alpha \ \exists U \in \mathcal{U}(x_0) \text{ tale che } f(x) > \alpha \ \forall x \in U;$$
- f è *semicontinua superiormente* in x_0 se
$$\forall \alpha \in \mathbb{R} \text{ per cui } f(x_0) < \alpha \ \exists U \in \mathcal{U}(x_0) \text{ tale che } f(x) < \alpha \ \forall x \in U.$$

Spesso si utilizzano gli acronimi derivanti dall'inglese *lsc* (da *lower semicontinous*) per indicare le funzioni semicontinue inferiormente e *usc* (da *upper semicontinous*) per indicare le funzioni semicontinue superiormente.

È importante il seguente teorema di caratterizzazione.

Teorema 2.4 Siano X uno spazio topologico di Hausdorff e $f : X \to [-\infty, +\infty]$. Sono tra loro equivalenti:

1. f è inferiormente semicontinua su X;
2. per ogni $\alpha \in \mathbb{R}$ l'insieme $\{x \in X \,:\, f(x) > \alpha\}$ è aperto in X;
3. per ogni $\alpha \in \mathbb{R}$ l'insieme $\{x \in X \,:\, f(x) \leq \alpha\}$ è chiuso in X;
4. $\mathrm{epi}(f)$ è chiuso in $X \times \mathbb{R}$;
5. per ogni $x_0 \in X$ vale
$$f(x_0) \leq \liminf_{x \to x_0} f(x);$$
6. (per X metrizzabile) per ogni $\{x_n\} \subseteq X$ con $x_n \to x_0 \in X$ vale
$$f(x_0) \leq \liminf_{n \to +\infty} f(x_n).$$

Osservazione Le seguenti osservazioni, anche se basilari, sono importanti per ciò che segue.

- f è semicontinua inferiormente in x_0 se e solo se $-f$ è semicontinua superiormente in x_0.
- Se f è continua in x_0 allora f è sia inferiormente semicontinua sia superiormente semicontinua in x_0.
- Sia $\{f_\alpha\}_{\alpha \in A}$ una famiglia di funzioni semicontinue inferiormente (convesse). Definita
$$g(x) = \sup_{\alpha \in A} f_\alpha(x),$$
essa è semicontinua inferiormente (convessa). Basta osservare che
$$(x,t) \in \mathrm{epi}(g) \Leftrightarrow g(x) \leq t \Leftrightarrow f_\alpha(x) \leq t \;\forall \alpha \in A \Leftrightarrow (x,t) \in \bigcap_{\alpha \in A} \mathrm{epi}(f_\alpha)$$
e quindi $\mathrm{epi}(g) = \bigcap_{\alpha \in A} \mathrm{epi}(f_\alpha)$.
- Siccome $\mathcal{B}(0,r) = \|\cdot\|^{-1}([0,r])$ sono w-chiuse si ha che $\|\cdot\|$ è w-inferiormente semicontinua.

Il seguente teorema è sicuramente importante e permette di osservare come le funzioni inferiormente semicontinue si comportano su un compatto.

Teorema 2.5 Siano X spazio topologico compatto e di Hausdorff e
$$f : X \to (-\infty, +\infty].$$
Se f è semicontinua inferiormente allora f assume minimo su X.

Dimostrazione Il caso $f \equiv +\infty$ è banale. Sia $\beta = \inf f(X) < +\infty$. Esiste una successione $\{\beta_n\} \subseteq \mathbb{R}$ strettamente decrescente tale che $\beta_n \to \beta$. Consideriamo

$$C_n = \{x \in X \,:\, f(x) \leq \beta_n\}$$

essi sono chiusi in un compatto e quindi compatti, inscatolati, e non vuoti. Allora

$$\{x \in X \,:\, f(x) \leq \beta\} = \bigcap_{n=1}^{+\infty} C_n \neq \emptyset$$

cioè l'estremo inferiore viene assunto. □

Passiamo ora allo studio della proprietà di convessità e di semicontinuità.

Corollario Siano X spazio normato, $C \subseteq X$ aperto convesso e $f : C \to \mathbb{R}$. Se f è semicontinua superiormente in un punto di C, allora f è continua su C.

Dimostrazione Se f è semicontinua superiormente allora f è limitata superiormente in un aperto non vuoto. Da ciò la tesi. □

Proposizione 2.1 Siano X spazio di Banach, C aperto convesso e $f : C \to \mathbb{R}$ convessa. Se f è semicontinua inferiormente allora f è continua su C.

Dimostrazione Ovviamente, posti $F_n = \{x \in C \,:\, f(x) \leq n\}$,

$$C = \bigcup_{n=1}^{+\infty} F_n.$$

Essendo C uno spazio di Baire (aperto in uno spazio metrico completo) esiste k tale che $\mathrm{int}_C(F_k) = \mathrm{int}(F_k) \neq \emptyset$. Quindi f è limitata superiormente su un aperto non vuoto. □

2.3 Famiglie di funzioni convesse

Ci occupiamo ora dello studio di famiglie di funzioni convesse. In particolare vogliamo ottenere un risultato simile al teorema di Ascoli–Arzelà. Il seguente teorema può essere visto come una generalizzazione del teorema di Banach–Steinhaus.

Teorema 2.6 Siano X spazio di Banach, $C \subseteq X$ aperto convesso e $\mathcal{F}$ una famiglia di funzioni convesse e continue su C. Se $\mathcal{F}$ è una famiglia di funzioni puntualmente limitate su C allora $\mathcal{F}$ è una famiglia di funzioni localmente equilimitate e localmente equilipschitziane.

Dimostrazione Consideriamo

$$g(x) = \sup_{f \in \mathcal{F}} f(x),$$

essa è convessa e inferiormente semicontinua, quindi g è continua su C. La continuità di una funzione convessa equivale alla locale limitatezza, ovvero

$$\forall x \in C \;\exists r_x > 0 \;\exists \eta_x \in \mathbb{R} \text{ tale che } g \leq \eta_x \text{ su } \mathcal{B}(x, r_x).$$

Questo implica in particolare che per ogni $f \in \mathcal{F}$ vale $f \leq \eta_x$ su $\mathcal{B}(x, r_x)$, quindi $\mathcal{F}$ è una famiglia di funzioni localmente equilimitate. Tali funzioni sono quindi lipschitziane su ogni bolla più piccola contenuta in $\mathcal{B}(x, r_x)$, in particolare

$$\forall x \in X \;\exists L_x > 0 \text{ tale che ogni } f \in \mathcal{F} \text{ è } L_x\text{-lipschitziana su } \mathcal{B}\left(x, \frac{r_x}{2}\right).$$

In definitiva $\mathcal{F}$ è una famiglia di funzioni localmente equilipschitziane. □

Corollario Siano X spazio di Banach, $C \subseteq X$ aperto convesso e $f_n : C \to \mathbb{R}$ una successione di funzioni convesse continue. Se $\{f_n\}$ converge puntualmente a g, ovvero

$$\forall x \in X \quad f_n(x) \xrightarrow{n\to+\infty} g(x),$$

allora g è convessa e continua e la convergenza è uniforme sui compatti $K \subseteq C$.

Dimostrazione Applichiamo il teorema precedente alla famiglia di funzioni $\{f_n\}$, in questo modo la famiglia risulta essere localmente equilimitata e localmente equilipschitziana. Di conseguenza g è localmente equilimitata e convessa,[2] e quindi continua su C. Sia ora $K \subseteq C$ compatto, la famiglia di funzioni $\{f_n\} \subseteq \mathcal{C}^0(K)$ è equilipschitziana, quindi equicontinua e equilimitata[3] e per il teorema di Ascoli–Arzelà si ha che $\{f_n\} \subseteq \mathcal{C}^0(K)$ è precompatto. Supponiamo per assurdo che f_n non converga uniformemente su K a g, ovvero

$$\exists\{f_{n_k}\} \subseteq f_n \ \exists\delta > 0 \text{ tale che } \|f_n - g\|_{\mathcal{C}(K)} > \delta$$

Ma $\{f_{n_k}\}$ è precompatto e quindi ammette una sottosuccessione $\{f_{n_{k_j}}\}$ convergente in $\mathcal{C}^0(K)$ e per convergenza puntuale il limite deve necessariamente essere g, ovvero

$$\|f_{n_{k_j}} - g\|_{\mathcal{C}^0(K)} \xrightarrow{k\to+\infty} 0.$$

Ma ciò è assurdo. □

Corollario Siano $C \subseteq \mathbb{R}^d$ aperto convesso e $\{f_n\}$ una successione di funzioni convesse su C. Se $\{f_n\}$ è puntualmente limitata allora esiste $\{f_{n_k}\} \subseteq \{f_n\}$ che converge localmente uniformemente a una funzione convessa.

Dimostrazione Siamo in dimensione finita, quindi le f_n sono continue, ovvero localmente equilipschtziane e localmente equilimitate. Consideriamo una successione di compatti invasori di C

$$K_1 \subseteq K_2 \subseteq K_3 \subseteq \cdots$$

con in aggiunta[4]

$$C = \bigcup_{n=1}^{+\infty} K_n = \bigcup_{n=1}^{+\infty} \operatorname{int}(K_n).$$

$\{f_n\}$ è precompatta in ogni $\mathcal{C}^0(K_i)$, per il ragionamento fatto nel precedente corollario. Usiamo un procedimento diagonale: sia $\{f_n^{(1)}\} \subseteq \{f_n\}$ una sottosuccessione convergente in $\mathcal{C}^0(K_1)$, $\{f_n^{(2)}\} \subseteq \{f_n^{(1)}\}$ una sottosuccessione convergente in $\mathcal{C}^0(K_2)$ e cosí via. Consideriamo la successione $\{f_k^{(k)}\}$, questa è la sottosuccessione che cercavamo. La convergenza di $\{f_k^{(k)}\}$ segue mostrando che è di Cauchy, con un gioco di disuguaglianze triangolari, su ognuno dei K_i; sfruttando poi il fatto che tali K_i invadono C segue la convergenza localmente uniforme. □

2.4 Funzioni convesse notevoli

In questa sezione vogliamo studiare alcuni esempi importanti di funzioni convesse, che in seguito utilizzeremo.

[2]Questo è ovvio per definizione.

[3]Basta considerare il ricoprimento dato dagli intorni su cui si ha locale equilipschitzianità ed estrarre un sottoricoprimento finito.

[4]L'ultima richiesta è realizzabile siccome siamo in dimensione finita.

2.4.1 Funzione indicatrice

Siano X spazio vettoriale e $C \subseteq X$, la *funzione indicatrice di* C è definita come:

$$\delta_C(x) = \begin{cases} 0 & x \in C \\ +\infty & x \notin C \end{cases}$$

Facciamo un elenco delle proprietà di tale funzione:

- $\operatorname{epi}(\delta_C) = C \times [0, +\infty)$;
- δ_C è convessa se e solo se C è convesso;
- δ_C è semicontinua inferiormente se e solo se C è chiuso.

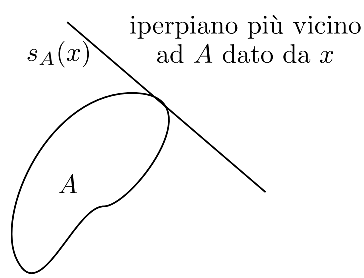
2.4.2 Funzione di supporto

Siano $A \subseteq X^*$ e $B \subseteq X$ non vuoti, la *funzione di supporto* di tali insiemi è definita rispettivamente come:

$$s_A(x) = \sup_{a^* \in A} a^*(x) \qquad \sigma_B(x^*) = \sup_{b \in B} x^*(b)$$

Facciamo un elenco delle proprietà di tali funzioni:

- s_A e σ_B sono convesse e semicontinue inferiormente, siccome estremi superiori di famiglie di funzioni convesse e semicontinue inferiormente;
- s_A è w-semicontinua inferiormente e σ_B è w^*-semicontinua inferiormente;
- geometricamente



per questo

$$s_A = s_{\operatorname{conv}(A)} = s_{\overline{\operatorname{conv}}(A)} = s_{\overline{\operatorname{conv}}^{w^*}(A)} \qquad \sigma_B = \sigma_{\overline{\operatorname{conv}}(B)}.$$

2.4.3 Funzione di distanza

Siano (X, d) spazio metrico e $A \subseteq X$ non vuoto. Definiamo *funzione di distanza da* A la seguente

$$\mathrm{d}_A(x) = \operatorname{dist}(x, A) = \inf_{a \in A} \mathrm{d}(x, a).$$

Facciamo un elenco delle proprietà di tale funzione:

- d_A è 1-lipschitziana, siccome estremo inferiore di una famiglia di funzioni 1-lipschitziane;
- vale $\mathrm{d}_A = \mathrm{d}_{\overline{A}}$;

- se X è normato e $C \subseteq X$ è convesso, allora d_C è convessa.

  Dimostrazione Siano $x, y \in X$, $\lambda \in (0,1)$ e poniamo $z = (1-\lambda)x + \lambda y$. Fissato $\varepsilon > 0$ esistono $c, d \in C$ tali che

  $$\|x - c\| \leq \mathrm{d}_C(x) + \varepsilon \qquad \|y - d\| \leq \mathrm{d}_C(y) + \varepsilon,$$

  poniamo $e = (1-\lambda)c + \lambda d$. Si ha dunque

  $$\begin{aligned} \mathrm{d}_C(z) &\leq \|z - e\| \leq (1-\lambda)\|x - c\| + \lambda\|y - d\| \leq \\ &\leq (1-\lambda)\,\mathrm{d}_C(x) + (1-\lambda)\varepsilon + \lambda\,\mathrm{d}_C(y) + \lambda\varepsilon = \\ &= (1-\lambda)\,\mathrm{d}_C(x) + \lambda\,\mathrm{d}_C(y) + \varepsilon \end{aligned}$$

  Per l'arbitrarietà di ε si ottiene la tesi. $\square$

- se $D \subseteq X$ tale che $X \setminus D$ è convesso, allora d_D è concava su $X \setminus D$.

  Dimostrazione Senza perdita di generalità possiamo supporre D chiuso, osserviamo che

  $$\mathrm{d}_D(x) = \max\{r > 0 \,:\, \mathrm{int}(\mathcal{B}(x,r)) \cap D = \emptyset\}.$$

  Sia $z = (1-\lambda)x + \lambda y$ e $r_z = (1-\lambda)\,\mathrm{d}_D(x) + \lambda\,\mathrm{d}_D(y)$, abbiamo che

  $$\mathrm{int}(\mathcal{B}(z, r_z)) = (1-\lambda)\,\mathrm{int}(\mathcal{B}(x, \mathrm{d}_D(x))) + \lambda\,\mathrm{int}(\mathcal{B}(y, \mathrm{d}_D(y))) \subseteq X \setminus D.$$

  Quindi $\mathrm{d}_D(z) \geq r_z$. $\square$

2.4.4 Funzionale di Minkowski

Siano X spazio vettoriale e $C \subseteq X$ convesso tale che $0 \in C$. Chiamiamo *funzionale di Minkowski di C* la seguente funzione

$$\mathrm{p}_C(x) = \inf\{t > 0 \,:\, x \in tC\} = \inf\left\{t > 0 \,:\, \frac{x}{t} \in C\right\}.$$

Facciamo un elenco delle proprietà di tale funzione:

- $\mathrm{p}(0) = 0$;

- se $x \neq 0$ allora $\mathrm{p}_C(x) = 0$ se e solo se C contiene tutta la semiretta passante per x che fuoriesce dall'origine.

  Dimostrazione $\mathrm{p}_C(x) = 0 \iff \forall t > 0\ \frac{x}{t} \in C \iff \forall s > 0\ sx \in C.$ $\square$

- $\mathrm{p}_C(x) < +\infty$ se e solo se esiste $\delta > 0$ tale che $[0, \delta x] \subseteq C$. In particolare p_C è finito per ogni $x \in X$ se e solo se $0 \in \text{a-int}\, C$.

  Dimostrazione $\mathrm{p}_C(x) < +\infty \iff \exists t > 0\ \frac{x}{t} \in C \iff \exists s > 0\ sx \in C.$ $\square$

- p_C è positivamente omogeneo, ovvero per ogni $\alpha > 0$ vale $p_C(\alpha x) = \alpha\,\mathrm{p}_C(x)$;

- p_C è subadditivo, ovvero $\mathrm{p}_C(x + y) \leq \mathrm{p}_C(x) + \mathrm{p}_C(y)$.

Dimostrazione Per ogni $\varepsilon > 0$ esistono $t_x, t_y > 0$ tali che

$$t_x \le \mathrm{p}_C(x) + \varepsilon \qquad t_y \le \mathrm{p}_C(y) + \varepsilon,$$

quindi $\frac{x}{t_x} \in C$ e $\frac{y}{t_y} \in C$. Ora

$$\frac{x+y}{t_x+t_y} = \frac{t_x}{t_x+t_y}\frac{x}{t_x} + \frac{t_y}{t_x+t_y}\frac{y}{t_y} \in C$$

e quindi

$$\mathrm{p}_C(x+y) \le t_x + t_y = \mathrm{p}_C(x) + \mathrm{p}_C(y) + 2\varepsilon$$

e per l'arbitrarietà di ε si ha la tesi. □

- p_C è sublineare, siccome subadditiva e positivamente omogenea. In particolare, p_C è convessa.
- vale $\{x \in X : \mathrm{p}_C(x) < 1\} \subseteq C \subseteq \{x \in X : \mathrm{p}_C(x) \le 1\}$.

  Dimostrazione Se $\mathrm{p}_C(x) < 1$ allora $\frac{x}{1} \in C$. Se $x \in C$ allora $\frac{x}{1} \in C$ e quindi $\mathrm{p}_C(x) \le 1$. □

- riguardo la continuità del funzionale di Minkowski abbiamo la seguente proposizione.

  Proposizione 2.2 Siano X normato e $C \subseteq X$ convesso. Sono tra loro equivalenti

  1. p_C è continuo e finito;
  2. p_C è continuo nell'origine;
  3. $0 \in \operatorname{int} C$;
  4. p_C è lipschitziano e finito.

  Dimostrazione Sono ovvie 1. ⇔ 2. siccome p_C è convessa.

  (4. ⇒ 1.) Ovvio.

  (1. ⇒ 3.) Segue da $0 \in \{x \in X : \mathrm{p}_C(x) < 1\} \subseteq C$, infatti $\{x \in X : \mathrm{p}_C(x) < 1\}$ è aperto.

  (3. ⇒ 4.) Esiste $\delta > 0$ tale che $\delta\,\mathcal{B}_X \subseteq C$. Se $x \neq 0$, allora

  $$\mathrm{p}_C(x) = \mathrm{p}_C\left(\frac{\|x\|}{\delta}\frac{\delta x}{\|x\|}\right) = \frac{\|x\|}{\delta}\,\mathrm{p}_C\underbrace{\left(\frac{\delta x}{\|x\|}\right)}_{\in \delta\,\mathcal{B}_X \subseteq C} \le \frac{1}{\delta}\|x\|,$$

  ovvero p_C è lipschitziana nell'origine. Ora sfruttando la subadditività, per ogni $x, y \in X$

  $$\mathrm{p}_C(x) - \mathrm{p}_C(y) = \mathrm{p}_C(y + (x-y)) - \mathrm{p}_C(y) \le \mathrm{p}_C(y) + \mathrm{p}_C(x-y) - \mathrm{p}_C(y) = $$
  $$= \mathrm{p}_C(x-y) \le \frac{1}{\delta}\|x-y\|.$$

  La disuguaglianza per $\mathrm{p}_C(y) - \mathrm{p}_C(x)$ è analoga. □

- se X è normato, C è chiuso se e solo se $C = \{x \in X : \mathrm{p}_C(x) \le 1\}$.

Dimostrazione Per continuità $\{x \in X : \mathrm{p}_C(x) \leq 1\}$ è chiuso e per omogeneità

$$\overline{\{x \in X : \mathrm{p}_C(x) < 1\}} = \{x \in X : \mathrm{p}_C(x) \leq 1\}.$$

La tesi segue dal fatto che $\overline{\{x \in X : \mathrm{p}_C(x) < 1\}} \subseteq C \subseteq \{x \in X : \mathrm{p}_C(x) \leq 1\}$.□

- se X è normato, C è aperto se e solo se $C = \{x \in X : \mathrm{p}_C(x) < 1\}$;
- se C e D sono convessi contenenti l'origine e $C \subseteq D$, allora $\mathrm{p}_D \leq \mathrm{p}_C$;
- se X è normato, allora $\mathrm{p}_{\mathcal{B}_X}(x) = \|x\|$;
- se X è normato e $r > 0$, allora $\mathrm{p}_{r\,\mathcal{B}_X}(x) = \frac{1}{r}\|x\|$.

  Dimostrazione

  $$\mathrm{p}_{r\,\mathcal{B}_X}(x) = \inf\{t > 0 : x \in tr\,\mathcal{B}_x\} = \frac{1}{r}\inf\{rt > 0 : x \in rt\,\mathcal{B}_x\} = \frac{1}{r}\,\mathrm{p}_{\mathcal{B}_X}(x).$$

  □

- combinando queste ultime tre proprietà otteniamo che se X è normato e C è un convesso tale che $r\,\mathcal{B}_X \subseteq C \subseteq R\,\mathcal{B}_X$ allora

  $$\frac{1}{R}\|\cdot\| \leq \mathrm{p}_C \leq \frac{1}{r}\|\cdot\|.$$

Vediamo un'applicazione interessante del funzionale di Minkowski.

Teorema 2.7 Siano X normato, $C, D \subseteq X$ convessi, chiusi e limitati, $x_0 \in \operatorname{int} C$ e $y_0 \in \operatorname{int} D$. Esiste un omeomorfismo $F : C \to D$ tale che $F(x_0) = y_0$.

Dimostrazione Senza perdita di generalità possiamo supporre che $x_0 = y_0 = 0$. Definiamo

$$\begin{cases} F(x) = \frac{\mathrm{p}_C(x)}{\mathrm{p}_D(x)}x & x \in C \smallsetminus \{0\} \\ F(0) = 0 \end{cases}$$

Ricordiamo che per chiusura $x \in C$ se e solo se $\mathrm{p}_C(x) \leq 1$ e $F(x) \in D$ se e solo se $\mathrm{p}_D(F(x)) \leq 1$. Si ha

$$\mathrm{p}_D(F(x)) = \mathrm{p}_D\left(\frac{\mathrm{p}_C(x)}{\mathrm{p}_D(x)}x\right) = \frac{\mathrm{p}_C(x)}{\mathrm{p}_D(x)}\,\mathrm{p}_D(x) = \mathrm{p}_C(x).$$

Quindi $x \in C$ se e solo se $F(x) \in D$. Sappiamo già che il funzionale di Minkowski è lipschitziano in tutti i punti ma non possiamo dire nulla sull'origine, basta controllare la continuità in tale punto

$$\|F(x)\| = \frac{\mathrm{p}_C(x)}{\mathrm{p}_D(x)}\|x\| \xrightarrow{\|x\|\to 0} 0,$$

questo perché $\mathrm{p}_C(x) \xrightarrow{\|x\|\to 0} 0$ e $\frac{\|x\|}{\mathrm{p}_D(x)}$ è limitato. Di conseguenza F è continua, dimostrare che l'inversa

$$\begin{cases} G(y) = \frac{\mathrm{p}_D(y)}{\mathrm{p}_C(y)}y & y \in D \smallsetminus \{0\} \\ G(0) = 0 \end{cases}$$

è continua è semplice. □

Capitolo 3

Teorema di Helly e simili

Ci occuperemo ora di trovare condizioni sufficienti che consentono di affermare che una famiglia di insiemi convessi ha intersezione non vuota. Incominciamo con il definire l'oggetto del nostro studio.

Definizione 3.1 (Famiglia centrata) Siano X insieme e $\mathcal{F}$ una famiglia di sottoinsiemi di X. $\mathcal{F}$ è detta *centrata* se per ogni $\mathcal{F}_0 \subseteq \mathcal{F}$ con $\mathrm{card}(\mathcal{F}_0) < +\infty$ vale

$$\bigcap \mathcal{F}_0 \neq \emptyset.$$

Diciamo che la famiglia $\mathcal{F}$ è *k-centrata* se per ogni $\mathcal{F}_0 \subseteq \mathcal{F}$ con $\mathrm{card}(\mathcal{F}_0) \leq k$ vale

$$\bigcap \mathcal{F}_0 \neq \emptyset.$$

Proposizione 3.1 Sia X spazio topologico di Hausdorff e $\mathcal{F}$ una famiglia centrata di chiusi. Se esiste $\mathcal{F}_0 \subseteq \mathcal{F}$ con $\mathrm{card}(\mathcal{F}_0) < +\infty$ tale che $\bigcap \mathcal{F}_0$ è compatta, allora $\bigcap \mathcal{F} \neq \emptyset$.

Dimostrazione Sia $X_0 = \bigcap \mathcal{F}_0$ e consideriamo

$$\widetilde{\mathcal{F}} = \{F \cap X_0 \,:\, F \in \mathcal{F}\}.$$

Tale famiglia è una famiglia centrata di compatti e quindi $\bigcap \widetilde{\mathcal{F}} \neq \emptyset$. Infatti se per assurdo $\bigcap \widetilde{\mathcal{F}} = \emptyset$ allora $X_0 = \bigcup \widetilde{\mathcal{F}}^{\complement}$, ove

$$\widetilde{\mathcal{F}}^{\complement} = \left\{F^{\complement} \,:\, F \in \widetilde{\mathcal{F}}\right\}.$$

Ma tali insiemi costituiscono un rivestimento aperto di X_0, possiamo quindi estrarne un sottoricoprimento finito $\left\{F_i^{\complement}\right\}_{i=1}^{n}$, ma

$$\bigcap_{i=1}^{n} F_i = \emptyset$$

il che è assurdo. $\square$

Enunciamo ora il teorema principale di questa sezione.

Teorema 3.1 (di Helly) Sia $\mathcal{F}$ una famiglia finita di insiemi convessi di $\mathbb{R}^d$. Se $\mathcal{F}$ è $(d+1)$-centrata allora

$$\bigcap \mathcal{F} \neq \emptyset.$$

Dimostrazione Sia $\mathcal{F} = \{C_1, \ldots, C_n\}$ e procediamo per induzione su n.

$(n \leq d+1)$ Ovvio;

$((n-1) \rightsquigarrow n)$ Sia $n > d+1$ e supponiamo che l'ipotesi d'induzione valga per tutte le famiglie con meno di n convessi, ovvero

$$\forall k = 1, \ldots, n \text{ esiste } x_k \in \bigcap_{\substack{i=1 \\ i \neq k}}^{n} C_i.$$

Consideriamo l'applicazione lineare $T : \mathbb{R}^n \to \mathbb{R}^{d+1}$ definita come segue

$$T(\lambda_1, \ldots, \lambda_n) = \left(\sum_{i=1}^{n} \lambda_i x_i, \sum_{i=1}^{n} \lambda_i \right),$$

essendo $n > d+1$ il nucleo di T è non banale. Consideriamo

$$P_\lambda = \{i \,:\, \lambda_i > 0\} \qquad N_\lambda = \{i \,:\, \lambda_i \leq 0\}$$

e osserviamo che necessariamente se $\lambda \in \ker T \setminus \{0\}$, allora $P_\lambda, N_\lambda \neq \emptyset$. Per un tale λ si ha quindi

$$\sum_{i \in P_\lambda} \lambda_i x_i = \sum_{j \in N_\lambda} |\lambda_j| x_j \qquad \sum_{i \in P_\lambda} \lambda_i = \sum_{j \in N_\lambda} |\lambda_j| =: s > 0.$$

Posto

$$z = \sum_{i \in P_\lambda} \frac{\lambda_i}{s} x_i = \sum_{j \in N_\lambda} \frac{|\lambda_j|}{s} x_j,$$

esse sono due combinazioni lineari convesse, la prima a elementi in $\bigcap_{i \in P_\lambda} C_i$ e la seconda ad elementi in $\bigcap_{j \in N_\lambda} C_j$. Di conseguenza

$$z \in \left(\bigcap_{i \in P_\lambda} C_i \right) \cap \left(\bigcap_{j \in N_\lambda} C_j \right) = \bigcap \mathcal{F}.$$

□

Ora unendo la proposizione precedente e il teorema di Helly otteniamo il seguente.

Teorema 3.2 (di Helly "bis") Sia $\mathcal{F}$ una famiglia $(d+1)$-centrata di convessi chiusi di $\mathbb{R}^d$. Se esiste una sottofamiglia $\mathcal{F}_0$ finita tale che $\bigcap \mathcal{F}_0$ sia compatta allora

$$\bigcap \mathcal{F} \neq \emptyset.$$

Osservazione Vogliamo fare un paio di commenti sulle ipotesi del teorema di Helly.

- Non è possibile richiedere che la famiglia sia solo d-centrata. Si consideri per esempio in $\mathbb{R}^2$ un triangolo, i suoi lati costituiscono una famiglia di convessi 2-centrata, ma la loro intersezione è vuota.
- La finitezza della famiglia è indispensabile. Se in $\mathbb{R}$ prendiamo in considerazione la famiglia $\mathcal{F} = \{[n, +\infty)\}_{n=1}^{+\infty}$, tale famiglia è 2-centrata, ma ha intersezione vuota.[1]

Passiamo ora ad analizzare le conseguenze del teorema di Helly, la prima è il seguente teorema conosciuto anche come *teorema dello spiedino*.

[1] In realtà questo risulta essere un controesempio anche alla proposizione 3.1.

Teorema 3.3 (della trasversale comune) Sia $\{I_\alpha\}_{\alpha\in A}$ una famiglia di segmenti compatti a due a due paralleli nel piano. Se ogni tre di essi sono intersecati da una retta, allora esiste una retta che li interseca tutti.

Dimostrazione Senza perdita di generalità possiamo supporre che

$$I_\alpha = \{u_\alpha\} \times [a_\alpha, b_\alpha],$$

ovvero i segmenti sono tutti verticali. Ci sono due casi:

1. esistono $\alpha \neq \beta$ tali che $u_\alpha = u_\beta$ e $I_\alpha \cap I_\beta = \emptyset$. In tal caso tutti i segmenti devono stare necessariamente sulla stessa retta verticale;

2. senza perdita di generalità possiamo semplificare la famiglia supponendo che se $u_\alpha = u_\beta$ allora $I_\alpha \cap I_\beta \neq \emptyset$, ovvero sopra ogni punto di $\mathbb{R}$ o non ci sono intervalli o c'è né uno solo. Questa supposizione assicura che dati 3 segmenti è possibile scegliere una retta non verticale d'intersezione.[2] Una retta del tipo $y = mx + q$ interseca l'intervallo I_α se e solo se

   $$a_\alpha \leq mu_\alpha + q \leq b_\alpha.$$

   Consideriamo

   $$C_\alpha = \big\{(m, q) \in \mathbb{R}^2 \,:\, a_\alpha \leq mu_\alpha + q \leq b_\alpha\big\}.$$

   La famiglia $\{C_\alpha\}$ è una famiglia di chiusi 3-centrata e se $\alpha \neq \beta$ allora $C_\alpha \cap C_\beta$ è compatto. Applicando il teorema di Helly si ottiene che

   $$\bigcap C_\alpha \neq \emptyset. \qquad \square$$

Il teorema della trasversale comune fornisce delle interessanti applicazioni, ci permetteremo di enunciare il seguente corollario a titolo di esempio.[3]

Corollario Siano $I \subseteq \mathbb{R}$ intervallo non banale e $f, g : I \to \mathbb{R}$ tali che $f \leq g$. Sono tra loro equivalenti

1. esiste $a : I \to \mathbb{R}$ affine[4] tale che $f \leq a \leq g$;

2. per ogni $x, y \in I$ e $\lambda \in (0, 1)$ devono valere

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

Dimostrazione

(1. ⇒ 2.) Questa implicazione è un semplice conto

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

L'altra disuguaglianza si ottiene in maniera simile.

[2]Ciò segue dal teorema di Helly, infatti sulla retta verticale per u_α esiste un punto comune a tutti i segmenti lì definiti. Basta far "partite" le rette da tale punto.

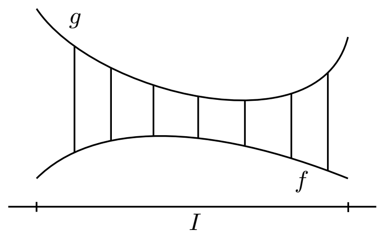
[3]Ricordiamo che dati $[a_1, b_1], [a_2, b_2] \subseteq \mathbb{R}$, $[a_1, b_1] \cap [a_2, b_2] \neq \emptyset$ se e solo se $a_1 \leq b_2$ e $a_2 \leq b_1$.

[4]Ovvero tale che per ogni $\lambda \in [0, 1]$ e $x, y \in I$ si ha $a((1-\lambda)x + \lambda y) = (1-\lambda)a(x) + \lambda a(y)$.

(2. $\Rightarrow$ 1.) Consideriamo la famiglia di intervalli

$$\mathcal{F} = \{\{x\} \times [f(x), g(x)]\}_{x \in I}.$$

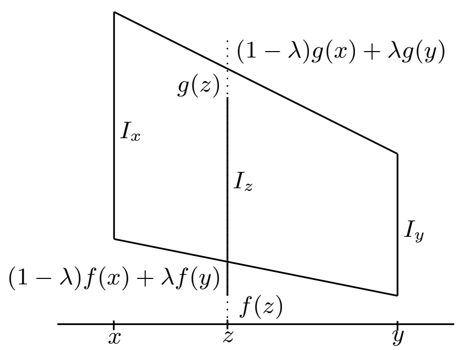
Il disegno rende bene l'idea.



Consideriamo ora $x < z < y$, esiste $\lambda \in (0, 1)$ tale che $z = (1 - \lambda)x + \lambda y$. Esiste una retta per I_x, I_y, I_z se e solo se

$$I_z \cap \{z\} \times [(1 - \lambda)f(x) + \lambda f(y), (1 - \lambda)g(x) + \lambda g(y)] \neq \emptyset,$$

lo si capisce bene dal seguente disegno.



Ma ciò è vero per l'assunzione. La tesi si ottiene per il teorema della trasversale comune. $\square$

Osservazione Si osservi che la condizione 2. del corollario è automaticamente soddisfatta se una delle funzioni f, g è convessa e l'altra concava (in qualsiasi ordine).

Il seguente teorema è un'applicazione del teorema di Helly, la dimostrazione viene lasciata per esercizio, con qualche suggerimento.

Teorema 3.4 (di Klee) Siano $\mathcal{F}$ una famiglia di convessi compatti di $\mathbb{R}^d$ e $K \subseteq \mathbb{R}^d$ convesso e compatto. Se per ogni $\mathcal{F}_0 \subseteq \mathcal{F}$ con $\mathrm{card}(\mathcal{F}_0) \leq d + 1$ si ha che $\bigcap \mathcal{F}_0$ contiene (è contenuta in) un traslato di K, allora $\bigcap \mathcal{F}$ contiene (è contenuta in) un traslato di K.

Dimostrazione L'idea è quella di considerare la famiglia $\widetilde{\mathcal{F}} = \left\{\widetilde{C}\right\}_{C \in \mathcal{F}}$ ove

$$\widetilde{C} = \left\{x \in \mathbb{R}^d \,:\, x + K \subseteq C\right\},$$

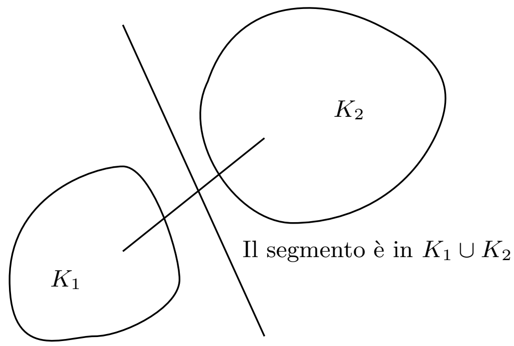
e applicare ad essa il teorema di Helly. $\square$

Concludiamo questa sezione con un teorema molto importante che costituisce una generalizzazione del teorema di Helly, e alcuni suoi corollari.

Teorema 3.5 (di Klee–Berge) Siano X spazio vettoriale topologico localmente convesso, $n \geq 2$ e $\{K_1, \dots, K_n\}$ una famiglia $(n-1)$-centrata di insiemi convessi e compatti di X. Se $\bigcup_{i=1}^n K_i$ è convessa, allora $\bigcap_{i=1}^n K_i$ è non vuota.

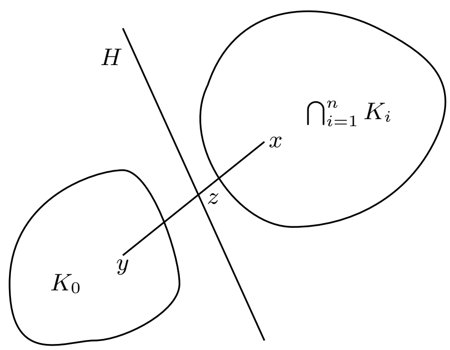
Dimostrazione Facciamo induzione su n.

$(n = 2)$ La famiglia è 1-centrata. Se $K_1 \cap K_2 = \emptyset$ allora, per Hahn–Banach, esiste un iperpiano che li separa fortemente.



Ovviamente si ha l'assurdo dal fatto che il segmento è contenuto in $K_1 \cup K_2$, ma tali insiemi sono separati fortemente.

$(n \rightsquigarrow n+1)$ Consideriamo una famiglia n-centrata $\{K_0, \dots, K_n\}$ che soddisfi le ipotesi. Se per assurdo i convessi e compatti K_0 e $\bigcap_{i=1}^n K_i$ sono disgiunti, separiamo con Hahn–Banach. Il seguente disegno permette di avere una maggiore comprensione di ciò che faremo, H rappresenta l'iperpiano separatore che ci fornisce Hahn–Banach.



Consideriamo $L_i = K_i \cap H$ con $i = 1, \dots, n$. Vogliamo mostrare che tale famiglia di compatti è $(n-1)$-centrata, per fare ciò possiamo, senza perdita di generalità mostrare solo che

$$L_1 \cap \dots \cap L_{n-1} \neq \emptyset.$$

Sia x come in figura e $y \in K_0 \cap \dots \cap K_{n-1}$, esiste

$$z \in (x, y) \cap H \subseteq (K_1 \cap \dots \cap K_{n-1}) \cap H = L_1 \cap \dots \cap L_{n-1}.$$

Ora per sfruttare l'ipotesi d'induzione basta mostrare che $\bigcup_{i=1}^n L_i$ è convessa, ma

$$\bigcup_{i=1}^n L_i = \bigcup_{i=1}^n (K_i \cap H) = \bigcup_{i=0}^n (K_i \cap H) = \left(\bigcup_{i=0}^n K_i \right) \cap H,$$

ove la prima uguaglianza sfrutta il fatto che $K_0 \cap H = \emptyset$. Per ipotesi d'induzione $\bigcap_{i=1}^n L_i \neq \emptyset$, ma d'altra parte

$$\emptyset \neq \bigcap_{i=1}^n L_i = \bigcap_{i=1}^n (K_i \cap H) = \left(\bigcap_{i=1}^n K_i \right) \cap H,$$

il che è assurdo per separazione. □

Corollario Siano X normato finito-dimensionale e $\{C_1, \dots, C_n\}$ una famiglia $(n-1)$-centrata di convessi chiusi di X. Se $\bigcup_{i=1}^n C_i$ è convessa, allora $\bigcap_{i=1}^n C_i$ è non vuota.

Dimostrazione Per ogni k consideriamo un elemento

$$x_k \in \bigcap_{\substack{i=1 \\ i \neq k}}^n C_i,$$

e $r > \max \|x_k\|$. Applichiamo ora il teorema di Klee–Berge alla famiglia di compatti $\{K_i\}_{i=1}^n$ definiti come $K_i = C_i \cap \mathcal{B}(0, r)$. □

Corollario Siano X normato, $A \subseteq X$ chiuso convesso e $\{C_1, \dots, C_n\}$ una famiglia di convessi chiusi di X tali che $A \subseteq \bigcup_{i=1}^n C_i$ e

$$A \cap \left(\bigcap_{\substack{i=1 \\ i \neq k}}^n C_i \right) \neq \emptyset. \qquad \forall k = 1, \dots, n$$

Se A è compatto o X è finito-dimensionale, allora

$$A \cap \left(\bigcap_{i=1}^n C_i \right) \neq \emptyset.$$

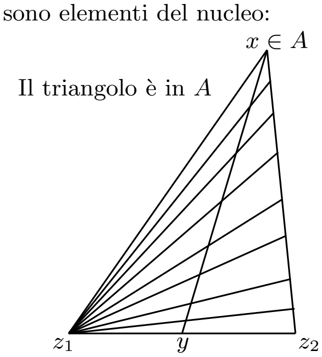
3.1 Una generalizzazione agli stellati

Vogliamo generalizzare il teorema di Klee–Berge per una nuova classe di insiemi che andiamo ora a definire.

Definizione 3.2 (Insieme stellato) Siano X spazio vettoriale e $A \subseteq X$. A si dice *stellato* se esiste $z \in A$ tale che

$$[z, x] \subseteq A \qquad \forall x \in A.$$

Osservazione L'insieme dei punti che rendono stellato un insieme A prende il nome di *nucleo di A*. Tale insieme è convesso, l'idea per dimostrare tale fatto si basa sul seguente disegno, ove z_1, z_2 sono elementi del nucleo:



Teorema 3.6 Siano X spazio vettoriale e $\{C_1, \ldots, C_n\}$ una famiglia $(n-1)$-centrata di convessi di X. Se l'intersezione di ogni C_i con una retta è un insieme chiuso nella retta, e

$$\bigcup_{i=1}^{n} C_i$$

è stellata, allora $\bigcap_{i=1}^{n} C_i \neq \emptyset$.

Dimostrazione Consideriamo l'insieme $D = \mathrm{conv}\{x_1, \ldots, x_n, \overline{z}\}$, dove

$$x_k \in \bigcap_{\substack{i=1 \\ i \neq k}}^{n} C_i$$

e $\overline{z}$ è un punto del nucleo di $\bigcup_{i=1}^{n} C_i$. Gli insiemi $K_i = C_i \cap D$ sono contenuti in $\mathrm{span}(D)$, che è finito-dimensionale, e soddisfano le stesse ipotesi dei C_i.[5] Inoltre, dalla chiusura per rette di K_i si può dedurre che K_i è compatto in $\mathrm{aff}(K_i)$.[6] Di conseguenza, è sufficiente dimostrare il teorema per $X = \mathbb{R}^d$ e tutti i C_i compatti. Si procede per induzione rispetto a n.

$(n = 2)$ Segue dal fatto che $C_1 \cup C_2$ è connessa e quindi i due insiemi non possono essere disgiunti.

$(n \rightsquigarrow (n+1))$ Procediamo per assurdo. Sia $\{C_0, \ldots, C_n\}$ una famiglia n-centrata di convessi compatti di $\mathbb{R}^d$ con intersezione vuota e unione stellata. Sia $\overline{z}$ un punto del nucleo di $\bigcup_{i=0}^{n} C_i$. Possiamo supporre che $\overline{z} \in C_0$. Denotiamo

$$P = \bigcap_{i=1}^{n} C_i.$$

Per il teorema di Hahn-Banach, esiste un iperpiano H separante C_0 e P e disgiunto da essi. Come nella dimostrazione del teorema di Klee–Berge, si dimostra che gli insiemi $L_i = C_i \cap H$ formano una famiglia $(n-1)$-centrata di insiemi compatti convessi. Per utilizzare l'ipotesi d'induzione, e ottenere un assurdo simile a quello ottenuto nella dimotrazione del teorema di Klee–Berge, rimane da dimostrare che l'insieme $\bigcup_{i=1}^{n} L_i$ è stellato. Sia $p \in P$ e denotiamo con $\overline{y}$ il punto di intersezione del segmento $[p, \overline{z}]$ con H. Sia

$$x \in \bigcup_{i=1}^{n} L_i = H \cap \left(\bigcup_{i=1}^{n} C_i \right).$$

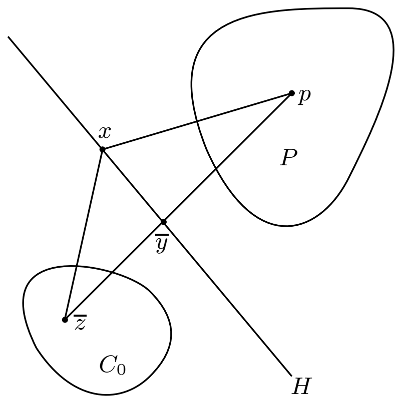
Si ha che $[p, x] \subseteq \bigcup_{i=1}^{n} C_i \subseteq \bigcup_{i=0}^{n} C_i$. Per come abbiamo scelto $\overline{z}$ si ha

$$[u, \overline{z}] \subseteq \bigcup_{i=0}^{n} C_i \qquad \forall u \in [p, x].$$

Il seguente disegno permette di avere una maggiore comprensione delle nostre azioni.

[5] La dimostrazione di tale fatto è semplice, basta ricordare che D è chiuso in $\mathrm{span}(D)$ e che l'intersezione tra uno stellato e un convesso è stellata.

[6] Infatti se supponiamo che esiste $x \in \overline{K}_i \setminus K_i$, siccome $\dim K_i < +\infty$ segue che $\mathrm{ri}(K_i) \neq \emptyset$. Fissato $u_i \in \mathrm{ri}(K_i) = \mathrm{ri}(\overline{K}_i)$ sappiamo che $[u, x) \subseteq \mathrm{ri}(K_i) = \mathrm{ri}(\overline{K}_i)$. Ma allora esiste una retta l tale che $x \notin l$, ma $[u, x) \subseteq l \cap K_i$. Assurdo.



Ne segue che il triangolo conv $\{p, x, \overline{z}\}$ è contenuto in $\bigcup_{i=0}^n C_i$. Di conseguenza

$$[\overline{y}, x] \subseteq H \cap \left(\bigcup_{i=0}^n C_i \right) = H \cap \left(\bigcup_{i=1}^n C_i \right) = \bigcup_{i=1}^n L_i.$$

Abbiamo dimostrato che quest'ultimo insieme è stellato rispetto al punto $\overline{y}$. □

Questa generalizzazione permette di ottenere come corollario il seguente recente teorema.[7]

Teorema 3.7 (di Breen) Sia $\mathcal{F}$ una famiglia di convessi compatti di $\mathbb{R}^d$ con $d \geq 1$. Se per ogni $\mathcal{F}_0 \subseteq \mathcal{F}$ con $\text{card}(\mathcal{F}_0) \leq d + 1$ si ha $\bigcup \mathcal{F}_0$ è stellata, allora $\bigcap \mathcal{F} \neq \emptyset$.

Dimostrazione La famiglia $\mathcal{F}$ è 2-centrata, infatti per ogni coppia di insiemi $A, B \in \mathcal{F}$ si ha, per ipotesi, $A \cup B$ stellata e in particolare connessa. Applicando ora il teorema precedente otteniamo che la famiglia $\mathcal{F}$ è 3-centrata. Si procede di questo passo fino a giungere al fatto che $\mathcal{F}$ è $(d+1)$-centrata. Si applica ora il teorema di Helly e si ottiene che $\bigcap \mathcal{F} \neq \emptyset$. □

[7]Si confronti M. Breen, *Proc. Amer. Math. Soc*, 108 (1990), 817–820.

Capitolo 4

Medie integrali e disuguaglianza di Jensen

Conosciamo già la disuguaglianza di Jensen in versione discreta, la quale afferma che se f è una funzione convessa su un insieme C convesso, allora

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i).$$

Osserviamo che se poniamo $\mu = \sum_{i=1}^{n} \lambda_i \delta_{x_i}$, dove δ_{x_i} indica la misura di Dirac concentrata in x_i, allora possiamo rileggere la disuguaglianza di Jensen nel seguente modo

$$f\left(\int x\, d\mu\right) \leq \int f\, d\mu.$$

Vogliamo dare un senso a tutto ciò, in particolare vogliamo poter cambiare misura. Per fare ciò abbiamo bisogno di alcuni prerequisiti.

4.1 Funzioni affini

Vogliamo analizzare in questa sezione alcune proprietà utili delle funzioni affini. Ovviamente è d'obbligo una definizione.

Definizione 4.1 (Funzione affine) Siano X spazio vettoriale e $a : X \to \mathbb{R}$. a si sice *affine* se a è contemporaneamente convessa e concava, ovvero $-a$ è convessa.

Esercizio 4.1.1 Siano X spazio vettoriale e $a : X \to \mathbb{R}$. Sono tra loro equivalenti:

1. a è affine;
2. per ogni $x, y \in X$ e $\lambda \in [0, 1]$ vale $a((1-\lambda)x + \lambda y) = (1-\lambda)a(x) + \lambda a(y)$;
3. esistono unici $l : X \to \mathbb{R}$ lineare e $\beta \in \mathbb{R}$ tali che $a(x) = l(x) + \beta$.

La seguente proposizione riveste una certa importanza.

Proposizione 4.1 Siano X spazio vettoriale topologico, $Z = X \times \mathbb{R}$, $x_0 \in X$, $t_1, t_2 \in \mathbb{R}$ e

$$z_i = (x_0, t_i) \qquad i = 1, 2.$$

Se esiste $z^* \in Z^*$ che separa i punti z_1 e z_2, allora ogni iperpiano del tipo $(z^*)^{-1}(\alpha)$ coincide con il grafico di una funzione affine $a : X \to \mathbb{R}$.

Dimostrazione Sappiamo che $(X \times \mathbb{R})^* = X^* \times \mathbb{R}^*$, quindi $z^* = (x^*, \lambda)$ e

$$z^*(x,t) = x^*(x) + \lambda t.$$

Siccome z^* separa i punti z_1 e z_2 abbiamo

$$0 \neq z^*(z_1 - z_2) = x^*(x_0 - x_0) + \lambda(t_1 - t_2) = \lambda(t_1 - t_2)$$

e dunque $\lambda \neq 0$. Ciò significa

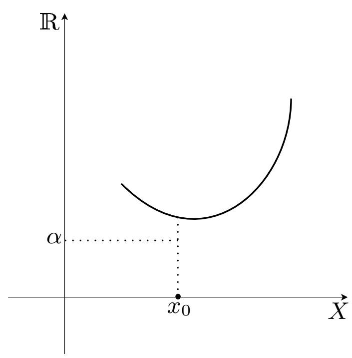
$$(x,t) \in (z^*)^{-1}(\alpha) \;\Leftrightarrow\; x^*(x) + \lambda t = \alpha \;\Leftrightarrow\; t = \frac{\alpha - x^*(x)}{\lambda} =: a(x). \qquad \square$$

Il risultato che andiamo ora ad enunciare ci verrà utile in seguito per dimostrare la disuguaglianza integrale di Jensen, lo chiameremo per tale motivo lemma.

Lemma 4 Siano X spazio vettoriale topologico localmente convesso, $C \subseteq X$ convesso, $x_0 \in C$, $f : C \to \mathbb{R}$ convessa e semicontinua inferiormente e $\alpha \in \mathbb{R}$. Se $\alpha < f(x_0)$ allora esiste $a : X \to \mathbb{R}$ affine e continua tale che

$$a(x_0) > \alpha \qquad a(x) < f(x) \quad \forall x \in C.$$

Dimostrazione Fissiamo le idee con un diagramma.



Fissiamo $\varepsilon > 0$ tale che $f(x_0) > \alpha + \varepsilon$ e consideriamo

$$A = \{x \in C \;:\; f(x) > \alpha + \varepsilon\}.$$

A è aperto in C per semicontinuità inferiore e contiene x_0. Esiste dunque V intorno convesso di x_0 tale che $V \subseteq A$. Abbiamo che

$$\mathrm{epi}(f) \cap [V \times (-\infty, \alpha + \varepsilon)] = \emptyset,$$

entrambi sono due convessi non vuoti e, in particolare, il secondo è aperto. Usiamo il teorema di Hahn–Banach e separiamo con un iperpiano, che per la proposizione precedente risulta essere il grafico di una funzione affine, ovvero esiste $\widetilde{a} : X \to \mathbb{R}$ affine continua tale che

$$\forall x \in V \quad \widetilde{a}(x) \geq \alpha + \varepsilon \qquad\qquad \forall x \in C \quad \widetilde{a}(x) \leq f(x).$$

Se definiamo $a(x) = \widetilde{a}(x) - \frac{\varepsilon}{2}$ abbiamo la tesi. $\square$

4.2 Proprietà di base dell'integrale di Pettis

Ci serviranno, in particolar modo, alcuni concetti riguardanti l'integrazione di funzioni a valori in spazi normati. Incominciamo quindi con il definire cosa si intende per integrale a valori in uno spazio normato.

Definizione 4.2 (Integrale di Pettis) Siano (Ω, Σ, μ) spazio di misura, X spazio vettoriale topologico localmente convesso e $F : \Omega \longrightarrow X$ misurabile. Diciamo F è *Pettis integrabile*, o *(P)-integrabile*, e che $x \in X$ è l'*integrale di Pettis di F su* Ω e scriviamo

$$\int_{\Omega}^{(P)} F \, d\mu = x$$

se per ogni $y^* \in X^*$ vale $\int_{\Omega} y^* \circ F \, d\mu = y^*(x)$.

Osservazione F è (P)-integrabile se e solo se per ogni $y^* \in X^*$ si ha $y^* \circ F \in \mathrm{L}^1(\mu)$ e l'applicazione lineare

$$y^* \mapsto \int_{\Omega} y^* \circ F \, d\mu$$

è w^*-continua.

4.2.1 Baricentri di misure

Passiamo ora a definire l'argomento che entrerà in gioco maggiormente.

Definizione 4.3 (Baricentro) Siano X spazio vettoriale topologico localmente convesso e μ una misura boreliana di probabilità su X, brevemente $\mu \in \mathcal{M}_1(X)$.[1] Diciamo che $x_\mu \in X$ è un *baricentro per* μ se

$$x_\mu = \int_X^{(P)} x \, d\mu$$

Osservazione

- Una misura non può avere più di un baricentro.[2]
- Una misura μ ammette baricentro se e solo se ogni $y^* \in X^*$ è μ-integrabile e l'applicazione lineare

  $$y^* \mapsto \int_X y^* \, d\mu$$

  è w^*-continua.

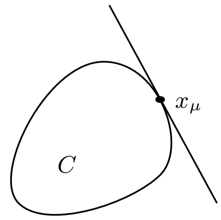
È naturale chiedersi dove si trovi il baricentro di un insieme. La seguente proposizione fornisce una risposta a tale domanda.

Proposizione 4.2 Siano X spazio vettoriale topologico localmente convesso, $C \subseteq X$ convesso e $\mu \in \mathcal{M}_1(C)$. Se C è aperto o chiuso e esiste il baricentro x_μ, allora $x_\mu \in C$.

[1] Il simbolo $\mathcal{M}_1(B)$, per B boreliano, indica l'insieme delle misure di probabilità μ su X concentrate in B, ovvero tali che $\mu(X \setminus B) = 0$.

[2] Segue facilmente dal teorema di Hahn–Banach.

Dimostrazione Faremo la dimostrazione solo nel caso di C aperto, quella per C chiuso risulta essere simile.[3] Supponiamo per assurdo che $x_\mu \notin C$, usiamo il teorema di Hahn–Banach per separare.



Esiste quindi $y^* \in X^*$ tale che

$$y^*(x_\mu) > y^*(x) \qquad \forall x \in C.$$

Da ciò si ha subito l'assurdo

$$0 < \int_C (y^*(x_\mu) - y^*(x))\, d\mu(x) = y^*(x_\mu) - \underbrace{\int_C y^*(x)\, d\mu(x)}_{=y^*(x_\mu)} = 0.$$

□

A questo punto ci si chiede se c'è una classe di insiemi per i quali si ha certezza riguardo l'esistenza del baricentro.

Teorema 4.1 Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso e $\mu \in \mathcal{M}_1(K)$. Se K è compatto, allora μ ammette baricentro.

Dimostrazione Vogliamo

$$\bigcap_{y^* \in X^* \smallsetminus \{0\}} \left\{ x \in K \,:\, y^*(x) = \int_K y^*\, d\mu \right\} \neq \emptyset.$$

Ognuno di tali insiemi è compatto, essendo intersezione di un iperpiano con un compatto, risulta quindi naturale cercare di dimostrare che tale famiglia è centrata. Fissiamo $y_1^*, \ldots, y_n^* \in X^* \smallsetminus \{0\}$ e consideriamo l'applicazione lineare e continua $T : X \to \mathbb{R}^n$ definita come segue

$$T(x) = (y_1^*(x), \ldots, y_n^*(x)).$$

Dunque $T(K) \subseteq \mathbb{R}^n$ è un convesso compatto. Consideriamo

$$\xi = \left(\int_K y_1^*\, d\mu, \ldots, \int_K y_n^*\, d\mu \right),$$

se $\xi \in T(K)$ allora la famiglia risulta centrata. Supponiamo per assurdo che $\xi \notin T(K)$, esiste quindi $\eta^* \in \mathbb{R}^n$ tale che

$$\langle \eta^*, \xi \rangle > \max \eta^*(T(K)).$$

Si ottiene che

$$\langle \eta^*, \xi \rangle = \sum_{i=1}^n \eta_i \int_K y_i^*\, d\mu = \int_K \left(\sum_{i=1}^n \eta_i y_i^* \right) d\mu \leq \sup_{x \in K} \left(\sum_{i=1}^n \eta_i y_i^*(x) \right),$$

e

$$\max \eta^*(T(K)) = \sup \eta^*(T(K)) = \sup_{x \in K} \langle \eta^*, T(x) \rangle = \sup_{x \in K} \left(\sum_{i=1}^n \eta_i y_i^*(x) \right).$$

Ma ciò è assurdo. □

[3] In realtà la dimostrazione che daremo funziona anche per quegli insiemi C che siano boreliani e *e-convessi*, ovvero siano intersezione di semispazi aperti.

Corollario Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso e $\mu \in \mathcal{M}_1(K)$. Se K è w-compatto, allora ammette baricentro.

4.3 Immagine di una misura

Risulterà utile anche il concetto di immagine di una misura. Siano (Ω, Σ, μ) spazio di probabilità, (T, τ) spazio topologico e $g : \Omega \to T$ misurabile.[4] Detto $\mathcal{B}_\tau$ l'insieme dei boreliani della topologia τ, poniamo

$$\nu : \mathcal{B}_\tau \longrightarrow [0, 1]$$

definita come $\nu(B) = \mu(g^{-1}(B))$. ν risulta essere una misura di probabilità su $(\tau, \mathcal{B}_\tau)$ detta *misura immagine di* μ.

Vogliamo definire l'integrale di una funzione $s : T \to \mathbb{R}$. Supponiamo, innanzitutto, che s sia semplice e boreliana, ovvero

$$s = \sum_{i=1}^{n} \alpha_i \chi_{B_i}$$

con B_i boreliani. Si ha

$$\int_T s\, d\nu = \sum_{i=1}^{n} \alpha_i \nu(B_i) = \sum_{i=1}^{n} \alpha_i \mu(g^{-1}(B_i)) = \int_\Omega s \circ g\, d\mu.$$

Per le funzioni positive si procede nel modo abituale approssimandole con una successione di funzioni semplici. Quindi

$$f \in \mathrm{L}^1(\nu) \Longleftrightarrow f \circ g \in \mathrm{L}^1(\mu)$$

e vale $\int_T f\, d\nu = \int_\Omega f \circ g\, d\mu$.

4.4 Disuguaglianze di Jensen e applicazioni

Possiamo finalmente enunciare il teorema tanto agognato.

Teorema 4.2 (Disuguaglianza integrale di Jensen) Siano X spazio vettoriale topologico localmente convesso, C convesso boreliano, $\mu \in \mathcal{M}_1(C)$, $f : C \to \mathbb{R}$ convessa e semicontinua inferiormente. Se esiste il baricentro $x_\mu \in C$, allora

1. esiste $\int_C f\, d\mu \in (-\infty, +\infty]$;
2. $f(x_\mu) \leq \int_C f\, d\mu$.

Dimostrazione

1. Gli insiemi $\{x \in C : f(x) \leq \alpha\}$ sono chiusi per semicontinuità inferiore, quindi f è boreliana. Per il lemma 4 esiste $a : X \to \mathbb{R}$ affine continua tale che $f \geq a$. Sappiamo che
$$a(x) = y^*(x) + \beta,$$
ma $y^* \in \mathrm{L}^1(\mu)$, per via dell'esistenza del baricentro, e $\beta \in \mathrm{L}^1(\mu)$ perché μ è una misura di probabilità. Dunque $a \in \mathrm{L}^1(\mu)$ e l'integrale della parte negativa di f non può essere infinito.

[4] Nel senso che retroimmagine di un boreliano tramite g è un insieme misurabile.

2. Per assurdo supponiamo che
$$f(x_\mu) > \int_C f\,d\mu.$$
Per il lemma 4 esiste $a : X \to \mathbb{R}$ affine e continua tale che
$$a(x_\mu) > \int_C f\,d\mu \qquad a(x) < f(x) \quad \forall x \in C.$$
Sappiamo inoltre che $a(x) = y^*(x) + \beta$. Abbiamo il seguente assurdo
$$\int_C f\,d\mu > \int_C a\,d\mu = \int_C y^*\,d\mu + \beta = y^*(x_\mu) + \beta = a(x_\mu) = \int_C f\,d\mu. \qquad \square$$

Vediamo ora alcuni corollari notevoli.

Corollario (Disuguaglianza di Hermite–Hadamard) Se $f : [a,b] \to \mathbb{R}$ è convessa e continua, allora
$$f\left(\frac{a+b}{2}\right) \overset{1.}{\leq} \frac{1}{b-a}\int_a^b f(t)\,dt \overset{2.}{\leq} \frac{f(a)+f(b)}{2}.$$

Dimostrazione

1. Definiamo $d\mu = \frac{dt}{b-a}$ e osserviamo che $\mu \in \mathcal{M}_1([a,b])$. Calcoliamo il baricentro di tale misura
$$x_\mu = \int_a^b x\,d\mu = \frac{1}{b-a}\int_a^b t\,dt = \frac{a+b}{2}.$$
Applicando Jensen si ottiene questa prima disuguaglianza.
2. Questa disuguaglianza è un conto coadiuvato da un opportuno cambio di variabili.
$$\frac{1}{b-a}\int_a^b f(t)\,dt = \begin{bmatrix} t = (1-\lambda)a + \lambda b \\ dt = (b-a)\,d\lambda \end{bmatrix} = \int_0^1 f((1-\lambda)a + \lambda b)\,d\lambda \leq$$
$$\leq f(a)\int_0^1 (1-\lambda)\,d\lambda + f(b)\int_0^1 \lambda\,d\lambda = \frac{f(a)+f(b)}{2} \qquad \square$$

Corollario (Seconda disuguaglianza integrale di Jensen) Siano $I \subseteq \mathbb{R}$ intervallo, (Ω, Σ, μ) spazio di probabilità e $g : \Omega \to I$ misurabile tale che $g \in \mathrm{L}^1(\mu)$. Se $f : I \to \mathbb{R}$ è convessa e continua su I allora
$$f\left(\int_\Omega g\,d\mu\right) \leq \int_\Omega f \circ g\,d\mu.$$

Dimostrazione Consideriamo ν l'immagine della misura μ tramite g, e osserviamo che anch'essa è una misura di probabilità su I. Infatti
$$\nu(I) = \int_I d\nu = \int_\Omega 1 \circ g\,d\mu = \int_\Omega d\mu = \mu(\Omega) = 1$$
e quindi $\nu \in \mathcal{M}_1(I)$. Per la disuguaglianza di Jensen (se esiste il baricentro)
$$f(x_\nu) \leq \int_I f\,d\nu = \int_\Omega f \circ g\,d\mu.$$
Dal fatto
$$x_\nu = \int_I x\,d\nu = \int_\Omega g\,d\mu,$$
segue la tesi. $\square$

Corollario Siano X spazio di Banach, $C \subseteq X$ aperto o chiuso, convesso e limitato, $\{x_n\}_{n=1}^{+\infty} \subseteq C$ e $\{\lambda_n\}_{n=1}^{+\infty} \subseteq [0, +\infty)$ tale che

$$\sum_{n=1}^{+\infty} \lambda_n = 1.$$

Se $f : C \to \mathbb{R}$ è convessa e semicontinua inferiormente, allora

1. esiste $x_0 = \sum_{n=1}^{+\infty} \lambda_n x_n \in C$;
2. $f\Big(\sum_{n=1}^{+\infty} \lambda_n x_n\Big) \leq \sum_{n=1}^{+\infty} \lambda_n f(x_n)$.

Dimostrazione È evidente che 2. è una diretta applicazione della disuguaglianza di Jensen. Dimostreremo solo 1. Per limitatezza

$$\sum_{n=1}^{+\infty} \|\lambda_n x_n\| \leq M \sum_{n=1}^{+\infty} \lambda_n < +\infty,$$

ovvero la serie converge assolutamente. Mostriamo ora che $x_0 = x_\mu$ con $\mu = \sum_{n=1}^{+\infty} \lambda_n \delta_{x_n}$, la quale è una misura di probabilità su C. Calcoliamone il baricentro

$$\int_C y^* \, d\mu = \int_{\{x_n\}_{n=1}^{+\infty}} y^* \, d\mu =$$

y^* è limitato su C e quindi integrabile

$$= \lim_{N \to +\infty} \int_{\{x_n\}_{n=1}^{N}} y^* \, d\mu = \lim_{N \to +\infty} \sum_{n=1}^{N} \lambda_n y^*(x_n) = y^*(x_0)$$

Ove l'ultimo passaggio è stato possibile per linearità e continuità. □

Esempio Vogliamo applicare la seconda disuguaglianza di Jensen alla seguente situazione: $I = (0, +\infty)$, $f(x) = \log(x)$. Dati quindi uno spazio di probabilità (Ω, Σ, μ) e una funzione $g : \Omega \to I$ tale che $g \in \mathrm{L}^1(\mu)$, vale

$$\log\left(\int_\Omega g \, d\mu\right) \geq \int_\Omega \log(g) \, d\mu.$$

Nel caso particolare in cui $\Omega = [a, b]$ si ottiene

$$\log\left(\frac{1}{b-a} \int_a^b g(t) \, dt\right) \geq \frac{1}{b-a} \int_a^b \log(g(t)) \, dt.$$

Capitolo 5

Punti estremi e teorema di Krein–Milman

Lo scopo di questo capitolo è quello di studiare la rappresentabilità degli insiemi convessi, in particolare ci interessa conoscere una sorta di insieme minimale tra tutti quelli il cui involucro convesso coincida con l'insieme di partenza. È immancabile una definizione di base per iniziare il nostro studio.

Definizione 5.1 (Punto estremo) Siano X spazio vettoriale, $C \subseteq X$ convesso e $x \in C$. x si dice *punto estremo per* C se quando

$$x = \frac{u+v}{2} \qquad u, v \in C$$

allora $u = v = x$, ovvero se x non è interno a nessun segmento non banale con estremi in C. L'insieme dei punti estremi per C si indica con $\operatorname{ext} C$.

È interessante conoscere che struttura topologica abbiano i punti estremi di un convesso, la seguente proposizione ci fornisce una risposta.

Proposizione 5.1 Siano X uno spazio vettoriale topologico e $K \subseteq X$ convesso. Se K è compatto e metrizzabile, allora $\operatorname{ext} K$ è un G_δ.

Dimostrazione Mostreremo che il complementare in K di $\operatorname{ext} K$ è un F_σ. È semplice osservare che

$$K \smallsetminus \operatorname{ext} K = \left\{ \frac{u+v}{2} \ : \ u, v \in K \ u \neq v \right\} = \bigcup_{n \in \mathbb{N}} \left\{ \frac{u+v}{2} \ : \ u, v \in K \ \mathrm{d}(u,v) \geq \frac{1}{n} \right\},$$

e se dimostriamo che tali insiemi sono chiusi abbiamo concluso. Facciamo vedere la chiusura per successioni: siano

$$\frac{u_k + v_k}{2} \xrightarrow{k \to +\infty} x \qquad \mathrm{d}(u_k, v_k) \geq \varepsilon,$$

per compattezza estraiamo due sottosuccessioni $\{u_{k_i}\} \subseteq \{u_k\}$ e $\{v_{k_i}\} \subseteq \{v_k\}$ tali che

$$u_{k_i} \xrightarrow{i \to +\infty} u \qquad v_{k_i} \xrightarrow{i \to +\infty} v$$

e utilizzando la continuità delle operazioni e della distanza otteniamo

$$x = \frac{u+v}{2} \qquad \mathrm{d}(u,v) \geq \varepsilon.$$

□

Osservazione

- Se $X = \mathbb{R}^2$ e C è convesso e compatto allora $\operatorname{ext} C$ è chiuso. Infatti se $\dim C = 0$ allora C è un punto, $\dim C = 1$ allora C è un segmento. Supponiamo che $\dim C = 2$, e togliamo a C i punti non estremi del bordo, allora abbiamo tolto un aperto.
- In generale anche se C è compatto non è detto che $\operatorname{ext} C$ sia chiuso. Si consideri per esempio un segmento tangente alla bolla di $\mathbb{R}^3$. Se consideriamo C l'involucro convesso di tale segmento e della bolla, risulta che $\operatorname{ext} C$ non è chiuso poiché $\operatorname{ext} C$ contiene il meridiano perpendicolare al segmento nel punto di tangenza, ma non quest'ultimo.
- Sia X spazio vettoriale topologico e $K \subseteq X$ convesso e compatto. Se esiste una successione $\{y_n^*\} \subseteq X^*$ che separa i punti di K, allora K è metrizzabile.[1]

5.1 Ambito finito-dimensionale

Il seguente teorema risolve, in ambito finito-dimensionale il problema che ci eravamo posti inizialmente.

Teorema 5.1 Siano X spazio vettoriale topologico, $K \subseteq X$ convesso, compatto e finito-dimensionale e $A \subseteq K$. Sono tra loro equivalenti:

1. $K = \operatorname{conv}(A)$;
2. $\operatorname{ext} K \subseteq A$.

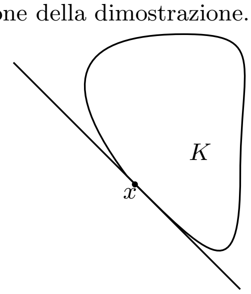
Dimostrazione Prima di iniziare la dimostrazione vogliamo far presente che in generale l'implicazione 2. $\Rightarrow$ 1. prende il nome di *teorema di Minkowski.*

(2. $\Rightarrow$ 1.) Basta mostrare che $K = \operatorname{conv}(\operatorname{ext}(K))$, per fare ciò procediamo per induzione su $d = \dim K$.

($d = 1$) Ovvio.

($d \rightsquigarrow (d+1)$) Sia $\dim K = d+1$, senza perdita di generalità possiamo supporre che $K \subseteq \mathbb{R}^{d+1}$ e $\operatorname{int} K \neq \emptyset$. Dobbiamo analizzare due casi.

- Sia $x \in \partial K$, per il teorema di Hahn–Banach esiste un iperpiano H che separa x da $\operatorname{int} K$ e tale che $x \in H$. Il seguente disegno risulta utile per la comprensione della dimostrazione.



[1] L'idea per dimostrare tale fatto è quella di mostrare che la seguente funzione

$$\mathrm{d}(x, y) = \sum_{n=1}^{+\infty} \alpha_n |y_n^*(x-y)|,$$

con $\alpha_n > 0$ che faccia convergere la serie, è una distanza su K.

Sia $K_1 = K \cap H$, sappiamo che $\dim K_1 \le d$ e quindi $x \in \operatorname{conv}(\operatorname{ext}(K_1))$. Se mostriamo che $\operatorname{ext} K_1 \subseteq \operatorname{ext} K$ abbiamo concluso. Se $y \in \operatorname{ext} K_1$ tale che

$$y = \frac{u+v}{2} \qquad u, v \in K,$$

allora necessariamente $u, v \in H$, altrimenti il segmento $[u, v]$ non sarebbe contenuto interamente in K, ma allora $u, v \in K_1$ e quindi $u = v = y$.

- Sia $x \in \operatorname{int} K$. Cosideriamo una retta r passate per x e poniamo $\{a, b\} = r \cap \partial K$. Dunque x è combinazione convessa di a e b che per il punto precedente sono combinazione convessa di punti estremi. In definitiva $x \in \operatorname{conv}(\operatorname{ext} K)$.

(1. $\Rightarrow$ 2.) Sia $x \in \operatorname{ext} K$, per ipotesi

$$x = \sum_{i=1}^{n} \lambda_i a_i$$

combinazione convessa di elementi di A. Possiamo supporre che n sia il più piccolo intero che permetta di scrivere x in quel modo. Ci sono 2 possibilità:

- Se $n = 1$ allora $x = a_1 \in A$ e abbiamo la tesi.
- Se $n > 1$ allora

$$x = \lambda_1 \underbrace{a_1}_{\in K} + (1 - \lambda_1) \underbrace{\sum_{i=2}^{n} \frac{\lambda_i}{1 - \lambda_1} a_i}_{\in K}.$$

Ma dal fatto che x è estremo per K discende che $x = a_1 \in A$. □

Vediamo ora un'applicazione interessante di tale risultato.

Corollario (Principio di massimo) Siano X spazio vettoriale topologico, K convesso compatto e finito-dimensionale contenuto in X e $f : K \to \mathbb{R}$ convessa. Se f è semicontinua superiormente, allora f assume massimo in un qualche punto estremo di K.

Dimostrazione Sappiamo che se f è semicontinua superiormente su un compatto allora esiste $x_0 \in K$ tale che

$$f(x_0) = \max f(K).$$

Per il teorema precedente $x_0 = \sum_{i=1}^{n} \lambda_i y_i$ combinazione convessa con $y_i \in \operatorname{ext} K$. Se fosse $f(y_i) < \max f(K)$ per ogni i si avrebbe

$$f(x_0) \le \sum_{i=1}^{n} \lambda_i f(y_i) < \max f(K),$$

il che è assurdo. □

5.2 Ambito infinito-dimensionale

Il seguente è il teorema che sta alla base della teoria dei punti estremi in spazi infinito-dimensionali.[2]

[2]Ricordiamo che la dimostrazione classica del teorema di Krein–Milman non ricorre a tale teorema, ma esso risulta molto comodo anche per altri utilizzi.

Teorema 5.2 Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso, compatto e non vuoto e $U \subseteq K$ convesso aperto relativamente a K. Se $\operatorname{ext} K \subseteq U$ allora $U = K$.

Dimostrazione Supponiamo per assurdo che $U \neq K$, e consideriamo la famiglia

$$\mathcal{V} = \{V \subseteq K \,:\, V \text{ convesso e aperto, } U \subseteq V,\ V \neq K\},$$

ovviamente $\mathcal{V} \neq \emptyset$ siccome $U \in \mathcal{V}$. Sia ora $\mathcal{V}_1 \subseteq \mathcal{V}$ una catena, posto

$$V_1 = \bigcup \mathcal{V}_1,$$

abbiamo che $K \smallsetminus V_1 = \bigcap_{W \in \mathcal{V}_1} (K \smallsetminus W) \neq \emptyset$, essendo una famiglia centrata di compatti. Dunque $V_1 \in \mathcal{V}$ e, per il lemma di Zorn, esiste $V \in \mathcal{V}$ massimale. Fissiamo $x \in V$, $\lambda \in (0,1)$ e consideriamo

$$W_{x,\lambda} = \{y \in K \,:\, (1-\lambda)x + \lambda y \in V\}.$$

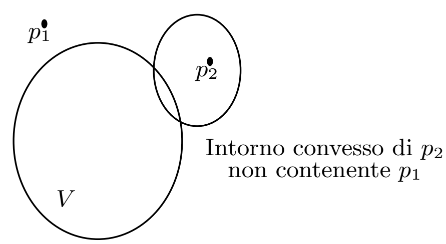
Vediamo alcune proprietà di $W_{x,\lambda}$:

- $W_{x,\lambda}$ è convesso, essendo V convesso. Infatti se $y, z \in W_{x,\lambda}$ e $\mu \in [0,1]$ allora
$$(1-\lambda)x + \lambda((1-\mu)y + \mu z) = (1-\mu)(\underbrace{(1-\lambda)x + \lambda y}_{\in V}) + \mu(\underbrace{(1-\lambda)x + \lambda z}_{\in V}) \in V.$$
- $W_{x,\lambda}$ è aperto per continuità delle operazioni di somma e prodotto.
- $V \subseteq W_{x,\lambda}$, per convessità di V.
- $\overline{V} \subseteq W_{x,\lambda}$. Infatti se $y \in \overline{V}$ allora $(1-\lambda)x + \lambda y \in [x, y) \subseteq V$ e quindi $y \in W_{x,\lambda}$.
- Dal fatto che $\overline{V} \neq V$ (per la connessione di K) e dalla massimalità di V si ha che $W_{x,\lambda} = K$.

Abbiamo appena mostrato che

$$K = \{y \in K \,:\, (1-\lambda)x + \lambda y \in V\} \qquad \forall x \in V\ \forall \lambda \in (0,1).$$

Ciò dimostra che se $x \in V$ e $y \in K$ allora $[x, y) \subseteq V$. Quindi per ogni A aperto convesso di K si ha che $A \cup V$ è convesso e per massimalità $A \cup V = K$. Se per assurdo esistono $p_1, p_2 \in K \smallsetminus V$ distinti si avrebbe



ma la loro unione non è K. Necessariamente $K \smallsetminus V = \{z_0\}$. Dimostriamo che z_0 è un punto estremo per K, se

$$z_0 = \frac{u+v}{2} \qquad u, v \in K$$

e $u, v \neq z_0$, necessariamente $u, v \in V$ e $z_0 \in V$ il che non può accadere. Ma V contiene U il quale a sua volta contiene $\operatorname{ext} K$, e quindi $z_0 \in V$. Assurdo. $\square$

Corollario (Teorema di Krein–Milman) Siano X spazio vettoriale topologico localmente convesso e $K \subseteq X$ convesso e compatto. Vale

$$K = \overline{\operatorname{conv}}(\operatorname{ext}(K)).$$

Dimostrazione Sappiamo già che $K_0 = \overline{\mathrm{conv}}(\mathrm{ext}(K)) \subseteq K$. Supponiamo per assurdo che esista $x_0 \in K \smallsetminus K_0$, utilizziamo il teorema di Hahn–Banach per separare K_0 da x_0, ovvero esiste $y^* \in X^*$ tale che

$$y^*(x_0) > \sup y^*(K_0).$$

Consideriamo

$$U = \{x \in K \,:\, y^*(x) < y^*(x_0)\},$$

tale insieme è un convesso aperto di K e contiene $K_0 \supseteq \mathrm{ext}\, K$. Per il teorema precedente $U = K$, ma ciò è assurdo siccome $x_0 \notin U$. □

Osservazione Facciamo alcune osservazioni su questo importante teorema.

- Innanzitutto osserviamo che non è possibile eliminare la chiusura dall'involucro convesso. Se consideriamo $X = c_0$ e $X^* = \ell^1$ per il teorema di Krein–Milman sappiamo
$$\mathcal{B}_{\ell^1} = \overline{\mathrm{conv}}^{w^*}(\mathrm{ext}\, \mathcal{B}_{\ell^1}),$$
ma $\mathcal{B}_{\ell^1} \neq \mathrm{conv}(\mathrm{ext}(\mathcal{B}_{\ell^1}))$. Si può infatti mostrare che $\mathrm{ext}\, \mathcal{B}_{\ell^1} = \{\pm e_n\}_{n=1}^{+\infty}$, dunque $\mathrm{conv}(\mathrm{ext}(\mathcal{B}_{\ell^1}))$ è costituito solo da elementi a supporto finito.
- Un'applicazione classica del teorema di Krein–Milman è quella di mostrare che alcuni spazi non possono essere isomorfi a nessun duale. Se X è uno spazio di Banach, il teorema di Banach–Alaoglu asserisce che $\mathcal{B}_{X^*}$ è w^*-compatta e dunque per il teorema di Krein–Milman
$$\mathcal{B}_{X^*} = \overline{\mathrm{conv}}^{w^*}(\mathrm{ext}\, \mathcal{B}_{X^*}).$$
Dunque spazi come c_0, $\mathcal{C}^0([0,1])$ e $\mathrm{L}^1([0,1])$ non possono essere isomorfi a un duale, infatti è possibile mostrare che
$$\mathrm{ext}\, \mathcal{B}_{c_0} = \emptyset \qquad \mathrm{ext}\, \mathcal{B}_{\mathcal{C}^0([0,1])} = \{\pm 1\} \qquad \mathrm{ext}\, \mathcal{B}_{\mathrm{L}^1([0,1])} = \emptyset.$$

Vogliamo ora ottenere un principio di massimo simile a quello ottenuto in dimensione finita.

Corollario (Principio di massimo di Bauer) Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ compatto e convesso e $f : K \to \mathbb{R}$ convessa. Se f è semicontinua superiormente allora f assume massimo in un punto estremo di K.

Dimostrazione Consideriamo

$$U = \{x \in K \,:\, f(x) < \max f(K)\},$$

tale insieme è convesso e aperto per semicontinuità superiore. Se il teorema fosse falso si avrebbe che $\mathrm{ext}\, K \subseteq U$ e quindi $U = K$ il che è assurdo. □

Ci verranno utili le seguenti proprietà degli spazi vettoriali topologici localmente convessi. Dato X spazio vettoriale topologico localmente convesso e $A \subseteq X$

- esiste una base di $\mathcal{U}(0)$ fatta di intorni convessi simmetrici.

  Dimostrazione Siccome X è localmente convesso esiste una base $\{V_i\}_{i \in I}$ di $\mathcal{U}(0)$ fatta di intorni convessi. Per avere la tesi è sufficiente considerare
$$\{(-V_i) \cap V_i\}_{i \in I}.$$
□

- per ogni $U \in \mathcal{U}(0)$ esiste $V \in \mathcal{U}(0)$ tale che $\overline{V} \subseteq U$.

  Dimostrazione L'applicazione
  $$(x, y) \mapsto x + y$$
  è continua nell'origine, quindi fissato $U \in \mathcal{U}(0)$ esiste $V \in \mathcal{U}(0)$ simmetrico tale che $V + V \subseteq U$. Sia $x \in \overline{V}$, sappiamo che $(x + V) \cap V \neq \emptyset$, ovvero esistono $v_1, v_2 \in V$ tali che
  $$x + v_1 = v_2.$$
  Quindi $x = v_2 + (-v_1) \in V + V \subseteq U$. □

- se A è relativamente compatto, ovvero $\overline{A}$ è compatto, allora per ogni $U \in \mathcal{U}(0)$ esiste $A_0 \subseteq A$ finito tale che $A \subseteq A_0 + U$.

  Dimostrazione Sia V un intorno simmetrico dell'origine contenuto in U. Consideriamo $\{a + V : a \in A\}$, vogliamo mostrare che è ricoprimento aperto di $\overline{A}$. Ma dal fatto che se $x \in \overline{A}$ allora
  $$(x + V) \cap A \neq \emptyset$$
  segue che esistono $v \in V$ e $a \in A$ tali che $x + v = a$, quindi $x \in a + V$. Estraiamo un sottoricoprimento finito e gli elementi di tale sottoricoprimento danno A_0.□

- se $\mathcal{B}$ è una base di $\mathcal{U}(0)$ fatta di insiemi simmetrici e per ogni $V \in \mathcal{B}$ si ha $x \in A + V$ allora $x \in \overline{A}$.

  Dimostrazione Per assurdo se $x \notin \overline{A}$ allora esiste $V \in \mathcal{B}$ tale che
  $$(x + V) \cap A = \emptyset.$$
  Ma ciò significa che $x \notin A + V$. □

Teorema 5.3 (di Milman) Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso e compatto e $A \subseteq K$. Se $K = \overline{\mathrm{conv}}(A)$ allora $\operatorname{ext} K \subseteq \overline{A}$.

Dimostrazione Siano $x \in \operatorname{ext} K$ e $\mathcal{B}$ una base di $\mathcal{U}(0)$ fatta da intorni chiusi, convessi e simmetrici. Sappiamo che A è relativamente compatto, quindi per ogni $V \in \mathcal{B}$ esistono $a_1, \ldots, a_n \in A$ tali che
$$A \subseteq \bigcup_{i=1}^{n} (a_i + V).$$
Poniamo
$$K_i = \overline{\mathrm{conv}}(A \cap (a_i + V)) \subseteq K \cap (a_i + V) \qquad i = 1, \ldots, n$$
tali insiemi soddisfano
$$K = \overline{\mathrm{conv}}(A) = \overline{\mathrm{conv}}\left(\bigcup_{i=1}^{n} K_i\right) = \mathrm{conv}\left(\bigcup_{i=1}^{n} K_i\right),$$
dove nell'ultima uguaglianza si sfrutta il fatto che i K_i sono compatti e convessi. Quindi $x = \sum_{i=1}^{n} \lambda_i u_i$ combinazione convessa con $u_i \in K_i$. Siccome x è punto estremo $x = u_i$ per un certo i, ovvero
$$x \in K_i \subseteq a_i + V \subseteq A + V.$$
Abbiamo mostrato che dato $V \in \mathcal{B}$ si ha $x \in A + V$, da cui $x \in \overline{A}$. □

Otteniamo dunque un teorema simile a quello enunciato per la dimensione finita.

Corollario Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso e compatto e $A \subseteq K$. Sono tra loro equivalenti

1. $K = \overline{\mathrm{conv}}(A)$;
2. $\mathrm{ext}\, K \subseteq \overline{A}$.

5.3 Intermezzo topologico: le net

Questa sezione costituisce solo un ripasso delle proprietà delle net o successioni generalizzate, che saranno molto utili in seguito.

Definizione 5.2 (Insieme diretto) Sia $(I, \leq)$ insieme parzialmente ordinato. I è detto *insieme diretto* o *insieme diretto in su* (dall'inglese *upwards directed*) se

$$\forall \alpha, \beta \in I \ \exists \gamma \in I \text{ tale che } \alpha \leq \gamma \text{ e } \beta \leq \gamma.$$

Esempi Facciamo un paio di esempi di insiemi diretti notevoli.

- Le partizioni di $[a, b]$ e diciamo che $P \leq Q$ se ogni punto di P è punto di Q.
- Fissato un punto x_0 in uno spazio topologico, possiamo considerare $I = \mathcal{U}(x_0)$ con
$$U_1 \leq U_2 \Longleftrightarrow U_2 \subseteq U_1.$$

Definizione 5.3 (Net) Siano (X, τ) spazio topologico di Hausdorff e I insieme diretto. Una *net* in X è una funzione $\varphi : I \to X$. Spesso una net è indicata come $\{x_\alpha\}_{\alpha \in I}$ ove $x_\alpha = \varphi(\alpha)$.

Definizione 5.4 (Convergenza) Siano (X, τ) spazio topologico di Hausdorff, I insieme diretto e $\{x_\alpha\}_{\alpha \in I}$ net. Diciamo che x_α *converge* a $y \in X$ se

$$\forall V \in \mathcal{U}(y) \ \exists \alpha_0 \in I \text{ tale che } \forall \alpha \geq \alpha_0 \ x_\alpha \in V.$$

Esempio Consideriamo f Riemann-integrabile su $[a, b]$ e, come insieme diretto, l'insieme I delle partizioni su $[a, b]$. Se $P \in I$ definiamo

$$x_P = \sum_{i=1}^{n} M_i (x_i - x_{i-1}),$$

ove $M_i = \sup f([x_{i-1}, x_i])$. Si può mostrare che

$$x_P \longrightarrow \int_a^b f.$$

Definizione 5.5 (Subnet) Siano (X, τ) spazio topologico di Hausdorff, I e J insiemi diretti e $\{x_\alpha\}_{\alpha \in I}$, $\{y_\beta\}_{\beta \in J}$ net. Diciamo che la net $\{y_\beta\}_{\beta \in J}$ è una *subnet* di $\{x_\alpha\}_{\alpha \in I}$ se esiste un'applicazione $\varphi : J \to I$ non decrescente e iniettiva tale che

$$y_\beta = x_{\varphi(\beta)}.$$

Esempio Consideriamo la successione $\{n^2\}_{n\in\mathbb{N}}$, si ha che

$$\{n^2 + 2nm + m^2\}_{(m,n)\in\mathbb{N}\times\mathbb{N}}$$

è una sua subnet. L'applicazione che mostra ciò è $\varphi(m,n) = m + n$, con le notazioni della definizione abbiamo posto $J = \mathbb{N}\times\mathbb{N}$ con ordine lessicografico e $I = \mathbb{N}$.

Enunciamo ora in un teorema, senza dimostrazione, tutte le proprietà di base delle net che ci verranno utili in seguito.

Teorema 5.4 Siano X, Y spazi topologici di Hausdorff e $A \subseteq X$. Valgono

1. $x \in \overline{A}$ se e solo se esiste $\{x_\alpha\}_{\alpha\in I} \subseteq A$ net tale che $x_\alpha \to x$;
2. un'applicazione $F : X \to Y$ è continua se e solo se per ogni $\{x_\alpha\}_{\alpha\in I} \subseteq X$ net tale che $x_\alpha \to \overline{x}$ si ha che
$$F(x_\alpha) \longrightarrow F(\overline{x});$$
3. A è compatto se e solo se ogni net $\{x_\alpha\}_{\alpha\in I} \subseteq A$ ammette una subnet $\{x_\beta\}_{\beta\in I}$ convergente a qualche elemento di A.

Concludiamo questo intermezzo topologico facendo notare, con degli esempi concreti, che le successioni non sono sufficienti per definire quali siano i chiusi in uno spazio topologico in generale. Se T è uno spazio topologico e $A \subseteq T$ indichiamo con

$$A^s = \left\{x \in T \,:\, \exists\{y_n\}_{n\in\mathbb{N}} \subseteq A \text{ con } y_n \to x\right\}$$

la *chiusura sequenziale* di A. Ovviamente $A \subseteq A^s \subseteq \overline{A}$, e se T è metrizzabile $A^s = \overline{A}$.

Esempi

- Sia $T = \Gamma \cup \{\infty\}$ con Γ non numerabile. Su Γ poniamo la topologia discreta, mentre diciamo che U è un intorno di ∞ se $\infty \in U$ e $T \smallsetminus U$ è al più numerabile. Ovviamente $\infty \in \overline{\Gamma}$ ma $\infty \notin \Gamma^s$, infatti se $\{y_n\} \in \Gamma$ allora $\{y_n\}^\complement$ è un intorno di ∞ privo di tale successione.
- Il seguente esempio è dovuto a Von Neumann e serve a mostrare che, in generale $A^s \neq (A^s)^s$.[3] Sia $T = \ell^2$ dotato della topologia debole e
$$A = \{e_m + me_n \,:\, m, n \in \mathbb{N},\ m < n\}.$$
Per m fissato
$$e_m + me_n \xrightarrow[n\to+\infty]{w} e_m,$$
di conseguenza $e_m \in A^s$ e $0 \in (A^s)^s$. Vogliamo mostrare che $0 \notin A^s$. Supponiamo che esista $\{a_k\} \subseteq A$ che converga debolmente a 0, sappiamo
$$a_k = e_{m_k} + m_k e_{n_k}.$$
Siccome a_k è debolmente convergente allora è limitata in norma e quindi $\{m_k\}$ è limitata. Abbiamo ora due casi
  - Se n_k è limitata, allora $\{a_k\}$ è finita e dunque a_k è definitivamente costante, ma ciò è assurdo siccome a_k converge debolmente a 0 e $0 \notin A$.

[3] Quindi la chiusura sequenziale non è nemmeno un operatore di chiusura.

– Se n_k è illimitata possiamo, passando ad un opportuna successione, supporre che $n_k \to +\infty$, ma allora

$$m_k e_{n_k} \xrightarrow{w} 0$$

e dunque $a_k \to e_{\overline{m}} \neq 0$.

Ovvero la chiusura successionale può non essere chiusa per successioni.

- Sia $T = \ell^1$ con la topologia debole e $A = \mathcal{S}_{\ell^1}$. Tale spazio gode della *proprietà di Schur*,[4] dunque $A = A^s$, ma $A \neq \overline{A}$.

Esercizio 5.3.1 Si consideri $T = \ell^2$ dotato della topologia debole e

$$B = \{e_m + \sqrt{m} e_n \,:\, m, n \in \mathbb{N},\ n < m\}.$$

Mostrare che $B = B^s$, ma $0 \in \overline{B} \smallsetminus B$.[5]

5.4 Una nuova luce per Krein–Milman

Vogliamo ottenere un teorema, in ambito infinito-dimensionale, che può essere visto come analogo al fatto che in dimensione finita gli elementi di un compatto possono essere scritti come combinazione convessa di punti estremi. Ricordiamo che se K è un compatto di Hausdorff e μ una misura boreliana, finita e positiva su K, definito

$$F(u) = \int_K u\, d\mu \qquad \forall u \in \mathcal{C}^0(K)$$

si ha $F \in (\mathcal{C}^0(K))^*_+$. Il *teorema di Riesz* afferma che vale il viceversa, ovvero per ogni $F \in (\mathcal{C}^0(K))^*_+$ esiste un'unica μ misura boreliana, finita e positiva su K tale che

$$F(u) = \int_K u\, d\mu \qquad \forall u \in \mathcal{C}^0(K),$$

inoltre

$$\|F\| = \sup_{\|u\|_\infty \leq 1} F(u) = F(1) = \int_K d\mu = \mu(K).$$

Possiamo enunciare il teorema promesso ad inizio sezione.

Teorema 5.5 (Forma integrale del teorema di Krein–Milman) Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso e compatto. Fissato $x \in K$ esiste $\mu \in \mathcal{M}_1(\overline{\operatorname{ext} K})$ tale che

$$x = x_\mu.$$

Dimostrazione Dal teorema di Krein–Milman sappiamo che $K = \overline{\operatorname{conv}}(\operatorname{ext} K)$, quindi se $x \in K$ esiste una net $\{x_\alpha\} \subseteq \operatorname{conv}(\operatorname{ext} K)$ tale che $x_\alpha \longrightarrow x$. Da cui

$$x_\alpha = \sum_{i=1}^{n_\alpha} \lambda_i^{(\alpha)} u_i^{(\alpha)}$$

[4] Ovvero una successione è debolmente convergente se e solo se è fortemente convergente.

[5] Si ricordi che se $y \in \ell^2$ allora $\liminf \sqrt{m}|y^m| = 0$. Infatti se definitivamente $\sqrt{m}|y^m| \geq \delta$ allora $|y^m| \geq \frac{\delta}{\sqrt{m}}$, ma $\sum \frac{\delta}{m}$ non converge.

combinazione convessa di elementi di $\operatorname{ext} K$. Da ciò si ottiene che $x_\alpha = x_{\mu_\alpha}$ con μ_α definita come

$$\mu_\alpha = \sum_{i=1}^{n_\alpha} \lambda_i^{(\alpha)} \delta_{u_i^{(\alpha)}},$$

ove δ_y indica la misura di Dirac concentrata in y. Sappiamo che $\overline{\operatorname{ext} K}$ è compatto e $\mu_\alpha \in \mathcal{M}_1(\overline{\operatorname{ext} K})$, possiamo pensare che

$$\mu_\alpha \in (\mathcal{C}^0(\overline{\operatorname{ext} K}))^*_+ \qquad \|\mu_\alpha\| = 1,$$

ma $\mathcal{B}_{(\mathcal{C}^0(\overline{\operatorname{ext} K}))^*}$ è w^*-compatta per il teorema di Banach–Alaoglu, estraiamo una subnet convergente (che chiameremo ancora μ_α) tale che

$$\mu_\alpha \xrightarrow{w^*} F \in \mathcal{B}_{(\mathcal{C}^0(\overline{\operatorname{ext} K}))^*}.$$

Per il teorema della permanenza del segno $F \in (\mathcal{C}^0(\overline{\operatorname{ext} K}))^*_+$ e per il teorema di Riesz esiste una misura μ boreliana, positiva e finita tale che

$$F(u) = \int_{\overline{\operatorname{ext} K}} u \, d\mu \qquad \forall u \in \mathcal{C}^0(\overline{\operatorname{ext} K}),$$

inoltre per convergenza debole*

$$\mu(\overline{\operatorname{ext} K}) = \mu(1) = \lim_\alpha \mu_\alpha(1) = \lim_\alpha \mu_\alpha(\overline{\operatorname{ext} K}) = 1$$

dunque $\mu \in \mathcal{M}_1(\overline{\operatorname{ext} K})$. Mostriamo che x è il baricentro di tale misura, sia $y^* \in X^*$ e consideriamo $y^*_{|\overline{\operatorname{ext} K}} \in \mathcal{C}^0(\overline{\operatorname{ext} K})$, innanzitutto

$$y^*(x_\alpha) \longrightarrow y^*(x)$$

per continuità e

$$y^*(x_\alpha) = \int_{\overline{\operatorname{ext} K}} y^* \, d\mu_\alpha = \mu_\alpha(y^*_{|\overline{\operatorname{ext} K}}) \longrightarrow \mu(y^*_{|\overline{\operatorname{ext} K}}) = \int_{\overline{\operatorname{ext} K}} y^* \, d\mu.$$

Dall'unicità del limite segue la tesi. $\square$

Risulta naturale chiedersi quando sia possibile ottenere una misura che sia concentrata solo sui punti estremi e non sulla loro chiusura. Il seguente teorema, di cui omettiamo la dimostrazione, risponde a tale domanda.

Teorema 5.6 (di Choquet) Siano X spazio vettoriale topologico localmente convesso, $K \subseteq X$ convesso, compatto e metrizzabile. Fissato $x \in K$ esiste $\mu \in \mathcal{M}_1(K)$ concentrata su $\operatorname{ext} K$ tale che

$$x = x_\mu.$$

Capitolo 6

Minimizzazione di funzioni convesse

Lo scopo di questo capitolo è quello di affrontare il problema dell'approssimazione. Il seguente teorema fornisce delle condizioni sufficienti affinché una funzione inferiormente semicontinua (nella topologia debole*) assume minimo su un chiuso.

Teorema 6.1 Siano X normato, $C \subseteq X^*$ non vuoto e w^*-chiuso e

$$f : C \to (-\infty, +\infty]$$

w^*-semicontinua inferiormente. Se inoltre C è illimitato supponiamo che f sia *coercitiva*, ovvero

$$\lim_{\substack{\|x\| \to +\infty \\ x \in C}} f(x) = +\infty.$$

f ammette minimo su C.

Dimostrazione Il teorema è ovvio se C è limitato, infatti C risulta essere w^*-compatto, sia quindi C illimitato. Se $f \equiv +\infty$ la tesi è ovvia. Supponiamo $f \not\equiv +\infty$ e fissiamo $a \in \mathbb{R}$ tale che $a > \inf f(C)$. Per coercitività esiste $r > 0$ tale che $f(x) > a$ se $x \in C$ e $\|x\| > r$. Quindi

$$\inf f(C) = \inf f(C \cap r\,\mathcal{B}_{x^*}) = \min f(C \cap r\,\mathcal{B}_{x^*}) = \min f(C)$$

sfruttando il fatto che $C \cap \mathcal{B}_{X^*}$ è w^*-compatto e f è w^*-semicontinua inferiormente. $\square$

Vediamo come applicare questo teorema alle funzioni convesse.

Corollario Siano X spazio di Banach riflessivo, $C \subseteq X$ chiuso, convesso e non vuoto e $f : C \to (-\infty, +\infty]$ convessa e semicontinua inferiormente. Se inoltre C è illimitato supponiamo che f sia coercitiva. f ammette minimo su C.

Dimostrazione Supponiamo che $X = Y^*$ (dove $Y = X^*$), sappiamo che

$$w_X = \sigma(X, X^*) = \sigma(Y^*, Y) = w^*_{Y^*}$$

Inoltre C è w-chiuso e $\{x \in C : f(x) \leq \alpha\}$ sono chiusi e convessi, quindi w-chiusi. Da ciò risulta f w-inferiormente semicontinua e per ottenere la tesi basta applicare il teorema precedente. $\square$

Osservazione Dalla dimostrazione appare evidente che l'ipotesi di convessità della funzione sia eccessiva. Basta infatti richiedere molto meno, più precisamente è sufficiente che per ogni $\alpha > 0$ gli insiemi $\{x \in C : f(x) \leq \alpha\}$ siano convessi, ovvero che la funzione f sia *quasi-convessa*.

Esercizio 6.0.1 Si provi che f è quasi-convessa su C se e solo se

$$\forall x, y \in C \ \forall \lambda \in [0,1] \quad f((1-\lambda)x + \lambda y) \leq \max\{f(x), f(y)\}.$$

Sfruttando questo risultato si mostri che la funzione $\|\cdot\|^a$ e convessa se e solo se $a \geq 1$, è quasi-convessa se e solo se $a > 0$.

6.1 Punti più vicini

Iniziamo questa sezione con alcune definizioni.

Definizione 6.1 (Punti più vicini) Siano (X, d) spazio metrico, $A \subseteq X$ non vuoto e $x \in X$. L'*insieme dei punti di A più vicini ad x* è

$$P_A(x) = \{a \in A : \mathrm{d}(x, a) = \mathrm{dist}(x, A)\}.$$

Questa definizione ne porta automaticamente altre con sé.

Definizione 6.2 (Insieme prossiminale/di unicità/di Chebyshev) Siano (X, d) spazio metrico, $A \subseteq X$ non vuoto. A si dice

- *prossiminale* se per ogni $x \in X$ si ha $P_A(x) \neq \emptyset$;
- *di unicità* se per ogni $x \in X$ si ha $\mathrm{card}\, P_A(x) \leq 1$;
- *di Chebyshev* se A è sia prossiminale sia di unicità.

Osservazione Le seguenti sono osservazioni di base.

- Se $x \in A$ allora $P_A(x) = \{x\}$.
- Se A è compatto allora A è prossiminale.
- Se A è prossiminale allora A è chiuso.

Il seguente teorema fornisce una caratterizzazione della riflessività in termini di insiemi prossiminali.

Teorema 6.2 Siano X spazio di Banach. Sono tra loro equivalenti

1. X è riflessivo;
2. ogni $C \subseteq X$ convesso, chiuso e non vuoto è prossiminale;
3. ogni iperpiano chiuso $H \subseteq X$ è prossiminale.

Dimostrazione

(1. $\Rightarrow$ 2.) Consideriamo per $x \notin C$ la funzione $f : C \to \mathbb{R}$ definita come

$$f(u) = \|u - x\|,$$

essa è convessa, continua e coercitiva quindi ammette minimo su C.

(2. $\Rightarrow$ 3.) Ovvio.

(3. $\Rightarrow$ 1.) Se X non fosse riflessivo, per il teorema di James esisterebbe un funzionale lineare continuo $x^* \in \mathcal{S}_{X^*}$ che non assume la norma. Consideriamo l'iperpiano $H = (x^*)^{-1}(1)$, abbiamo che $\mathcal{B}_X \cap H = \emptyset$, ma

$$(1+\varepsilon)\,\mathcal{B}_X \cap H \neq \emptyset \qquad \forall \varepsilon > 0.$$

Quindi $\operatorname{dist}(0, H) = 1$, ma $P_H(0) = \emptyset$. □

Arriviamo ad un'altra caratterizzazione degli spazi finito-dimensionali.

Teorema 6.3 Sia X normato. Ogni insieme chiuso e non vuoto in X è prossiminale se e solo se $\dim X < +\infty$.

Dimostrazione Se $\dim X = +\infty$ sappiamo che esiste una successione $\{x_n\}$ δ-separata sulla sfera. Consideriamo

$$E = \left\{ \left(1 + \frac{1}{n}\right) x_n \,:\, n \in \mathbb{N} \right\},$$

tale insieme è chiuso e $E \cap \mathcal{B}_X = \emptyset$, ma per ogni $\varepsilon > 0$ vale $E \cap (1+\varepsilon)\,\mathcal{B}_X \neq \emptyset$, dunque $\operatorname{dist}(0, E) = 1$ ma $P_E(0) = \emptyset$. L'altra implicazione è lasciata per esercizio. □

Enunciamo come conclusione il seguente teorema.

Teorema 6.4 Sia X normato. Sono tra loro equivalenti

1. $\mathcal{B}_X$ è strettamente convessa, ovvero $\operatorname{ext}\mathcal{B}_X = \mathcal{S}_X$;
2. ogni insieme convesso di X è di unicità;
3. ogni retta di X è di Chebyshev.

Dimostrazione

(1. $\Rightarrow$ 2.) Sia $C \subseteq X$ convesso e supponiamo per assurdo che esiste $x \in X$ tale che $\{c_1, c_2\} \subseteq P_C(x)$ con $c_1 \neq c_2$. Ma per stretta convesstà della bolla vale

$$\left\| \frac{c_1 + c_2}{2} - x \right\| < \|c_1 - x\|,$$

ovvero $c_1 \notin P_C(x)$. Assurdo.

(2. $\Rightarrow$ 3.) Per ipotesi ogni retta $r \subseteq X$ è di unicità. Ovviamente se $x \in r$ allora $P_r(x) = \{x\}$. Sia ora $x \notin r$ e supponiamo che

$$\operatorname{dist}(x, r) < d.$$

Consideramo $r \cap \mathcal{B}(x, d)$, tale insieme è compatto e la funzione $\mathrm{d}(x, \cdot)$ è inferiormente semicontinua su tale insieme e dunque assume minimo.

(3. $\Rightarrow$ 1.) Supponiamo per assurdo che esista $x \in \mathcal{S}_X \setminus \operatorname{ext}\mathcal{B}_X$, sappiamo che

$$x = \frac{y+z}{2} \qquad y, z \in \operatorname{ext}\mathcal{B}_X \subseteq \mathcal{S}_X \quad y \neq z.$$

Consideriamo ora la retta r passante per y e z. Vogliamo mostrare che $\operatorname{dist}(0, r) = 1$, supponiamo per assurdo che $\operatorname{dist}(0, r) < 1$. La retta r è di Chebyshev e quindi esiste $w \in r \cap \operatorname{int}\mathcal{B}_X$ tale che

$$\mathrm{d}(0, w) = \operatorname{dist}(0, r).$$

Ora, a meno di scambiare y con z, sappiamo che

$$x \in [w, z) \subseteq \operatorname{int}\mathcal{B}_X$$

il che è assurdo siccome $x \in \mathcal{S}_X$. Dunque $\operatorname{dist}(0, r) = 1$, ma allora $\{y, z\} \subseteq P_r(0)$ ma ciò è assurdo perché r è di Chebyshev. □

Osservazione Quindi se X è di Banach ogni sottoinsieme chiuso, convesso e non vuoto di X è di Chebyshev se e solo se X è riflessivo e $\mathcal{B}_X$ è strettamente convessa.

6.2 Centri di Chebyshev

Se (X, d) è uno spazio metrico e $A \subseteq X$ un insieme limitato e non vuoto, possiamo considerare

$$r_A(x) = \sup_{a \in A} \mathrm{d}(x, a).$$

Definizione 6.3 (Centro di Chebyshev) Siano (X, d) spazio metrico e $A \subseteq X$ limitato e non vuoto. Un punto $x_0 \in X$ è detto *centro di Chebyshev per A* se

$$r_A(x_0) = \min r_A(X).$$

Osservazione È noto che esistono insiemi di tre punti in $\mathcal{C}^0([0,1])$ privi di centri di Chebyshev. Un insieme può avere più di un centro di Chebyshev, ciò non accade se lo spazio è di Hilbert.

Il seguente teorema fornisce una condizione sufficiente all'esistenza di un centro di Chebyshev.

Teorema 6.5 In uno spazio di Banach duale X^*, ogni insieme limitato e non vuoto $A \subseteq X^*$ ammette un centro di Chebyshev.

Dimostrazione Sappiamo

$$r_A(x^*) = \sup_{a^* \in A} \|x^* - a^*\|,$$

ovvero r_A è estremo superiore di funzioni convesse e w^*-inferiormente semicontinue. Si può mostrare facilmente che a seguito della limitatezza di A, r_A è coercitiva e quindi r_A ammette minimo su X^*. □

Corollario Siano X spazio di Banach. Se X è 1-completato in X^{**}, ovvero se esiste $P : X^{**} \to X$ proiezione lineare con $\|P\| = 1$, allora ogni insieme limitato e non vuoto $A \subseteq X$ ammette centro di Chebyshev.

Dimostrazione Per il teorema precedente esiste $x_0^{**} \in X^{**}$ centro di Chebyshev di A in X^{**}, ovvero

$$r_A(x_0^{**}) \le r_A(x^{**}) \qquad \forall x^{**} \in X^{**}.$$

Calcoliamo

$$r_A(P(x_0^{**})) = \sup_{a \in A} \|P(x_0^{**}) - a\| =$$

sfruttando il fatto che $P(a) = a$,

$$= \sup_{a \in A} \|P(x_0^{**} - a)\| \le \sup_{a \in A} \|x_0^{**} - a\| = r_A(x_0^{**}) \le r_A(x) \qquad \forall x \in X$$

Dunque $P(x_0^{**})$ è un centro di Chebyshev per A in X. □

Osservazione Non è difficile mostare che se X è un duale allora X è 1-completato in X^{**}. Ma bisogna osservare che $\mathrm{L}^1([0,1])$, pur non essendo isomorfo ad un duale, è 1-completato nel suo biduale.

Capitolo 7

Differenziabilità

Il seguente capitolo costituisce un primo approccio al problema della differenziazione delle funzioni convesse.

7.1 Differenziabilità di funzioni convesse in $\mathbb{R}$

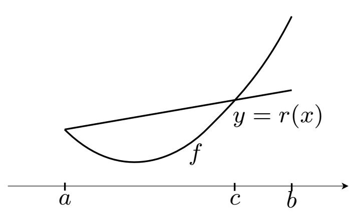
Incominciamo con un'introduzione riguardante la differenziazione di funzioni convesse in $\mathbb{R}$ analizzando le proprietà di base, che in seguito saranno utili per costruire una teoria della differenziazione generale. In tutta la sezione $I \subseteq \mathbb{R}$ intervallo con interno non vuoto, indichiamo con

$$a = \inf I \qquad b = \sup I$$

e infine sia $f : I \to \mathbb{R}$ convessa. Elencheremo ora una serie di proprietà che ha la funzione f in questa situazione.

- f è continua su I se e solo se f è inferiormente semicontinua su $\{a, b\} \cap I$.[1]

  Dimostrazione Sappiamo già che abbiamo continuità nei punti interni per il teorema 2.2. Supponiamo $a \in I$ e $c \in \operatorname{int} I$. Si consideri il seguente disegno



  dove $r(x)$ è la retta per $(a, f(a))$ e $(c, f(c))$. Sappiamo che

$$f(a) \leq \liminf_{x\to a^+} f(x) \leq \limsup_{x\to a^+} f(x) \leq \limsup_{x\to a^+} r(x) = r(a) = f(a),$$

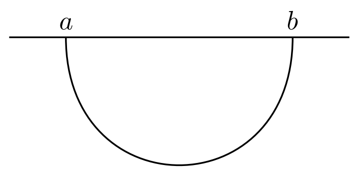
  di conseguenza

$$\liminf_{x\to a^+} f(x) = \limsup_{x\to a^+} f(x) = f(a).$$

  Analogo per b. □

[1] Tale fatto non vale più già in $\mathbb{R}^2$.

- Definiamo
$$R(x,y) = \frac{f(y)-f(x)}{y-x} \qquad y \neq x$$
Se $x_1, x_2, x_3 \in I$ con $x_1 < x_2 < x_3$ si ha[2]
$$R(x_1,x_2) \leq R(x_1,x_3) \leq R(x_2,x_3).$$
Quindi se $x \in I \setminus \{b\}$ esiste $f'_+(x) = \lim_{h\to 0^+} R(x,x+h)$ e per monotonia di R si ha
$$f'_+(x) = \inf_{\substack{h>0\\ x+h\in I}} R(x,x+h).$$
In modo analogo se $x \in I \setminus \{a\}$ esiste $f'_-(x) = \lim_{h\to 0^-} R(x,x+h)$ e per monotonia di R si ha
$$f'_-(x) = \sup_{\substack{h<0\\ x+h\in I}} R(x,x+h).$$
- Se $x \in \operatorname{int} I$ allora $f'_-(x) \leq f'_+(x)$ e se $x_1 < x_2$ allora $f'_+(x_1) \leq f'_-(x_2)$. Quindi $f'_\pm(x)$ sono monotone non decrescenti su $I \setminus \{a\}$ e $I \setminus \{b\}$.
- Se $x \in \operatorname{int} I$ allora $f'_\pm(x) \in \mathbb{R}$. Invece
$$f'_+(a) \in [-\infty,+\infty) \qquad f'_-(b) \in (-\infty,+\infty].$$
Esistoon funzioni f tali che $f'_+(a) = -\infty$ e $f'_-(b) = +\infty$ come mostrato nel seguente disegno.



- f'_+ è continua da destra e f'_- è contniua da sinistra.

Dimostrazione Analizziamo il caso $f'_+(x_0) \in \mathbb{R}$, gli altri casi si studiano in modo simile. Senza perdita di generalità possiamo supporre $x_0 = 0$ e $f(0) = f'_+(0) = 0$. Supponiamo per assurdo che
$$\lim_{x\to 0^+} f'_+(x) = \inf_{\substack{x>0\\ x\in I}} f'_+(x) = \alpha > 0,$$
siccome $f'_+(0) = 0$ esiste h tale che
$$R(0,h) = m < \alpha.$$
Fissato $x \in (0,h)$ si ha $0 \leq f(x) \leq mx$. Per ogni $t > x$ otteniamo
$$f(t) \geq f(x) + f'_+(x)(t-x) \geq f'_+(x)(t-x),$$
da cui
$$\frac{f(t)}{t-x} \geq f'_+(x) \Longrightarrow \inf_{\substack{x>0\\ t\neq x}} \frac{f(t)}{t-x} = \frac{f(t)}{t} \geq \inf_{x>0} f'_+(x) = \alpha$$
ma ciò è assurdo per $t = h$. □

[2] Questo è semplice usando la seguente formula magica
$$x_2 = \frac{x_2-x_1}{x_3-x_1}x_3 + \frac{x_3-x_2}{x_3-x_1}x_1.$$

- Per ogni $x_0 \in I \setminus \{b\}$ vale
$$f'_+(x_0) = \lim_{x\to x_0^+} f'_-(x) = \lim_{x\to x_0^+} f'_+(x).$$
Ciò segue dal fatto che se $x > x_0$ allora $f'_+(x_0) \le f'_-(x) \le f'_+(x)$. Ovviamente se $x_0 \in I \setminus \{a\}$ vale
$$f'_-(x_0) = \lim_{x\to x_0^-} f'_+(x) = \lim_{x\to x_0^-} f'_-(x).$$

- Chiamiamo *subdifferenziale di f in x_0* l'insieme dei coefficienti angolari delle rette passanti per il punto del grafico di ascissa x_0 e minoranti la funzione f su I, ovvero
$$\partial f(x_0) = \{m \in \mathbb{R} \,:\, f(x) \ge f(x_0) + m(x - x_0)\ \forall x \in I\}.$$
Gli elementi di $\partial f(x_0)$ sono detti *subgradienti*. Le seguenti sono alcune proprietà di base del subdifferenziale:
  - se f è differenziabile in x_0 allora card$(\partial f(x_0))$=1;
  - se $x_0 \in \operatorname{int} I$ allora $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$;
  - si ha $\partial f(a) = (-\infty, f'_+(a)]$ e $\partial f(b) = [f'_-(b), +\infty)$;[3]
  - siano $\{x_n\} \subseteq I$ e per ogni n scegliamo $m_n \in \partial f(x_n)$. Se $x_n \to x_0^\pm$ allora
  $$m_n \longrightarrow f'_\pm(x_0);$$
  - se $x_0 \in \operatorname{int} I$ sono tra loro equivalenti:
    1. f è differenziabile in x_0;
    2. $f'_+(x_0) = f'_-(x_0)$;
    3. $\operatorname{card}(\partial f(x_0)) = 1$;
    4. f'_+ è continua da sinistra in x_0 (f'_- è continua da destra in x_0);
    5. $\partial f(x_0) = \{m_0\}$ e ∂f è continua in x_0, ovvero se $x_n \to x_0$ e $m_n \in \partial f(x_n)$ allora
    $$m_n \xrightarrow{n\to+\infty} m_0.$$

- Se $D_1 = \{x \in \operatorname{int} I \,:\, f \text{ differenziabile in } x\}$ e $N_1 = \operatorname{int} I \setminus D_1$ allora N_1 è al più numerabile.

  Dimostrazione Ad ogni $x \in N_1$ possiamo associare l'intervallo
  $$J_x = (f'_-(x), f'_+(x)).$$
  Per monotonia tali intervalli sono a due a due disgiunti. Se consideriamo un razionale in ognuno di essi abbiamo la tesi. □

- Se f è continua su $[\alpha, \beta] \subseteq I$ allora
$$f(\beta) - f(\alpha) = \int_\alpha^\beta f'_+(t)\,dt = \int_\alpha^\beta f'_-(t)\,dt = \int_\alpha^\beta f'(t)\,dt.$$

[3]In particolare se $x_0 \in \operatorname{int} I$ allora $\partial f(x_0) \neq \emptyset$. Invece $\partial f(a) \neq \emptyset$ se e solo se $f'_+(a) \in \mathbb{R}$ e $\partial f(b) \neq \emptyset$ se e solo se $f'_-(b) \in \mathbb{R}$.

Dimostrazione Sia $\gamma \in (\alpha, \beta)$. Su $[\alpha + \frac{1}{n}, \gamma]$ sappiamo che f'_+ è monotona crescente e finita, quindi Riemann-integrabile. Inoltre f è lipschitziana e dunque assolutamente continua e

$$f(\gamma) - f\left(\alpha + \frac{1}{n}\right) = \int_{\alpha + \frac{1}{n}}^{\gamma} f'.$$

Per continuità

$$f(\gamma) - f(\alpha) = \int_{\alpha}^{\gamma} f'$$

e dunque l'integrale esiste finito. Facendo un ragionamento analogo per β si ottiene la tesi. $\square$

Concludiamo questa sezione con un teorema che caratterizza quando una funzione convessa è due volte derivabile in un punto.

Definizione 7.1 (Derivabilità secondo Peano) Siano $I \subseteq \mathbb{R}$, $f : I \to \mathbb{R}$ e $x_0 \in \operatorname{int} I$. Diciamo che f è *due volte derivabile secondo Peano in* x_0 se f è derivabile in x_0 e esiste $\Delta \in \mathbb{R}$ tale che

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \Delta\frac{h^2}{2} + o(h^2) \qquad h \to 0.$$

Osservazione Si ha praticamente subito che se f è convessa e esiste Δ allora $\Delta \geq 0$.

Dimostrazione Sappiamo

$$\Delta = \frac{2}{h}\left(\frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0)\right) + 2\frac{o(h^2)}{h^2} \qquad h \to 0^+.$$

Per la convessità sappiamo che

$$\left(\frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0)\right) \geq 0.$$

Si ha $\Delta \geq o(1) \to 0$ e dunque $\Delta \geq 0$. Analogo per $h \to 0^-$. $\square$

Teorema 7.1 Siano $I \subseteq \mathbb{R}$, $f : I \to \mathbb{R}$ convessa e $x_0 \in \operatorname{int} I$. Sono tra loro equivalenti

1. f è due volte derivabile secondo Peano in x_0;
2. f'_+ è derivabile in x_0 e $(f'_+)'(x_0) = \Delta$;
3. f'_- è derivabile in x_0 e $(f'_-)'(x_0) = \Delta$;
4. Vale
$$\Delta = \lim_{\substack{x \to x_0 \\ x \in D_1}} \frac{f'(x) - f'(x_0)}{x - x_0}.$$

Dimostrazione Osserviamo che 2. $\Rightarrow$ 4. e 3. $\Rightarrow$ 4. sono immediati e che 1. $\Rightarrow$ 3. si ottiene in maniera analoga a 1. $\Rightarrow$ 2.

(4. $\Rightarrow$ 1.) Sappiamo che

$$f'(x_0+h) = f'(x_0) + \Delta h + \omega(h)$$

dove $\omega(h) = o(h)$ per $h \to 0$. Si ha dunque

$$f(x_0+h) - f(x_0) = \int_0^h f'(x_0+t)\,dt = \int_0^h (f'(x_0) + \Delta t + \omega(t))\,dt =$$
$$= f'(x_0)h + \Delta\frac{h^2}{2} + \int_0^h \omega(t)\,dt.$$

Per il teorema di De l'Hôpital otteniamo

$$\lim_{h\to 0} \frac{\int_0^h \omega(t)\,dt}{h^2} = \lim_{h\to 0}\frac{\omega(h)}{2h} = 0.$$

e quindi $\int_0^h \omega(t)\,dt = o(h^2)$, da cui la tesi.[4]

(1. $\Rightarrow$ 2.) Senza perdita di generalità possiamo supporre che $x_0 = 0$ e $f(0) = f'(0) = 0$. Sotto tali ipotesi risulta

$$f(h) = \Delta\frac{h^2}{2} + o(h^2).$$

Sia ora $h > 0$, vogliamo stimare $\limsup_{h\to 0^+} \frac{f'(h)}{h}$. Sia $\varepsilon > 0$

$$\limsup_{h\to 0^+}\frac{f'(h)}{h} \le \limsup_{h\to 0^+}\frac{R(h, h+\varepsilon h)}{h} = \limsup_{h\to 0^+}\frac{1}{\varepsilon h^2}(f(h+\varepsilon h) - f(h)) =$$
$$= \limsup_{h\to 0^+}\frac{1}{\varepsilon h^2}\left(\frac{\Delta}{2}h^2(1+2\varepsilon+\varepsilon^2) - \frac{\Delta}{2}h^2 + o(h^2(1+\varepsilon)^2) + o(h^2)\right) =$$
$$= \limsup_{h\to 0^+}\left(\frac{\Delta}{2\varepsilon}(2\varepsilon+\varepsilon^2) + o(1)\right) = \Delta\left(1+\frac{\varepsilon}{2}\right)$$

e per l'arbitrarietà di ε si ottiene $\limsup_{h\to 0^+}\frac{f'(h)}{h} \le \Delta$. Si procede in maniera analoga per stimare[5] $\liminf_{h\to 0^+}\frac{f'(h)}{h}$ e si ottiene $\liminf_{h\to 0^+}\frac{f'(h)}{h} \ge \Delta$. □

Osservazione È possibile dimostrare che se f è convessa su I allora f è quasi ovunque (rispetto alla misura di Lebesgue) derivabile due volte secondo Peano su I.

7.2 Alcune nozioni di calcolo differenziale

Vogliamo dare senso al concetto di differenziabilità per funzioni tra due spazi normati. Ricordiamo che, se X, Y sono spazi normati, indichiamo con

$$\mathcal{L}(X,Y) = \{L : X \to Y \ :\ L \text{ continuo e lineare}\}.$$

Inoltre dati $A \subseteq X$ aperto non vuoto, $F : A \to Y$, $a \in A$ e $v \in X$ indichiamo

$$F'(a,v) = \lim_{t\to 0}\frac{F(a+tv) - F(a)}{t}.$$

Si possono dare due definizioni differenti di differenziabilità, che sono le seguenti.

[4] Si osservi che ω ha tutte le proprietà buone di f', siccome differisce da quest'ultima per una funziona lineare, e quindi è in particolare assolutamente continua.

[5] Osservando che per $\varepsilon > 0$ si ha

$$\liminf_{h\to 0^+}\frac{f'(h)}{h} \ge \liminf_{h\to 0^+}\frac{R(h-\varepsilon h, h)}{h}.$$

Definizione 7.2 (Gâteaux differenziabilità) Siano X, Y spazi normati $A \subseteq X$ aperto non vuoto, $F : A \to Y$, $a \in A$. Diciamo che F è *Gâteaux differenziabile in* a o, più semplicemente, *G-differenziabile in* a se esiste $L \in \mathcal{L}(X, Y)$ tale che per ogni $v \in X$ vale

$$F'(a, v) = Lv.$$

Definizione 7.3 (Fréchet differenziabilità) Siano X, Y spazi normati $A \subseteq X$ aperto non vuoto, $F : A \to Y$, $a \in A$. Diciamo che F è *Fréchet differenziabile in* a o, più semplicemente, *F-differenziabile in* a se esiste $L \in \mathcal{L}(X, Y)$ tale che

$$\lim_{\|h\| \to 0} \frac{F(a+h) - F(a) - Lh}{\|h\|} = 0.$$

Osservazione

- F è G-differenziabile in a se e solo se $F'(a, \cdot) \in \mathcal{L}(X, Y)$, ovvero se e solo se esiste $L \in \mathcal{L}(X, Y)$ tale che per ogni $v \in X$

  $$F(a + tv) = F(a) + t \cdot Lv + o(t) \qquad t \to 0$$

  con $o(t)$ che può dipendere da v.
- F è F-differenziabile in a se e solo se esiste $L \in \mathcal{L}(X, Y)$ tale che

  $$F(a + h) = F(a) + Lh + o(\|h\|) \qquad \|h\| \to 0.$$

- In entrambi i casi si denota $F'(a) := L$.

Esercizio 7.2.1 Dimostrare le seguenti affermazioni:

1. Se F è F-differenziabile in a allora F è G-differenziabile in a. In generale non vale il viceversa.
2. Se F è G-differenziabile in a allora per ogni $v \in X$ esiste $F'(a, v)$. In generale non vale il viceversa.
3. Se F è F-differenziabile in a allora F è continua in a. In generale non vale il viceversa.
4. Se F è G-differenziabile in a non è detto che F sia continua in a.

La seguente proposizione riveste una certa importanza, ma preferiamo lasciare la dimostrazione come esercizio per il lettore.

Proposizione 7.1 Siano X, Y spazi normati $A \subseteq X$ aperto non vuoto, $a \in A$ e $F : A \to Y$. Sono tra loro equivalenti

1. F è F-differenziabile in a;
2. F è G-differenziabile in a e il limite

   $$\lim_{t \to 0} \frac{F(a + tv) - F(a)}{t} = F'(a)v$$

   è uniforme rispetto a $v \in \mathcal{S}_X$.

Possiamo ora enunciare un lemma che riveste sicuramente un'importanza elevata in ambito finito-dimensionale.

Lemma 5 Siano X, Y spazi normati $A \subseteq X$ aperto non vuoto, $F : A \to Y$, $a \in A$. Se X è finito-dimensionale e F è lipschitziana in un intorno di a allora F è F-differenziabile in a se e solo se F è G-differenziabile in a.

Dimostrazione Se F è F-differenziabile in a è immediato mostrare che F è G-differenziabile in a. Supponiamo per assurdo che F non sia F-differenziabile da ciò segue che

$$\lim_{t\to 0} \frac{F(a+tv)-F(a)}{t}$$

non è uniforme in $v \in \mathcal{S}_X$ e quindi esistono $t_n \to 0$, $v_n \in \mathcal{S}_X$ e $\delta > 0$ tale che

$$\left\| F'(a)v_n - \frac{F(a+t_nv_n)-F(a)}{t_n} \right\| > \delta.$$

Passando ad una sottosuccessione possiamo supporre che $v_n \to v$, sfruttano il fatto che $\mathcal{S}_X$ è compatta, quindi

$$\begin{aligned}
\left\| F'(a)v - \frac{F(a+t_nv)-F(a)}{t_n} \right\| &\geq \\
\geq \delta - \left\| \frac{F(a+t_nv_n)-F(a+t_nv)}{t_n} + F'(a)v - F'(a)v_n \right\| &\geq \\
\geq \delta - \frac{\|F(a+t_nv_n)-F(a+t_nv)\|}{|t_n|} - \|F'(a)(v-v_n)\| &\geq \\
\geq \delta - \frac{L_F\|t_n(v_n-v)\|}{|t_n|} - \|F'(a)\|\|(v-v_n)\| &=
\end{aligned}$$

dove L_F è la costante di Lipschitz di F

$$= \delta - (L_F + \|F'(a)\|)\|v_n - v\| > \frac{\delta}{2}$$

definitivamente. Ma ciò è assurdo infatti in tal caso $F'(a)$ non sarebbe il G-differenziale di F in a. $\square$

Enunciamo per motivi di tipo culturale i seguenti teoremi.

Teorema 7.2 (Differenziabilità della funzione composta) Siano X, Y, Z normati, $A \subseteq X$ e $B \subseteq Y$ aperti,

$$A \xrightarrow{F} B \xrightarrow{G} Z,$$

$a \in A$ e G F-differenziabile in $F(a)$. Poniamo $H = G \circ F$.

1. Se esiste $F'(a, v)$ allora esiste $H'(a, v)$ e
$$H'(a,v) = G'(F(a))[F'(a,v)].$$
2. Se F è F-differenziabile (G-differenziabile) allora H è F-differenziabile (G-differenziabile) e vale
$$H'(a) = G'(F(a)) \circ F'(a).$$

Teorema 7.3 (Differenziabilità della mappa inversa) Siano X, Y spazi di Banach,[6] $A \subseteq X$ aperto e $F : A \to Y$. Se F è F-differenziabile in $a \in A$ con $F'(a)$ invertibile,[7] allora

[6] Questa ipotesi è dovuta al fatto che nella dimostrazione si vuole utilizzare il teorema di Banach–Caccioppoli.
[7] Per il teorema della mappa aperta $(F'(a))^{-1} \in \mathcal{L}(Y, X)$.

1. esiste $U \in \mathcal{U}(a)$ aperto tale che $F_{|_U}$ sia iniettiva;
2. $F(a) \in \operatorname{int} F(U)$;
3. $F^{-1} : \operatorname{int}(F(U)) \to U$ è F-differenziabile in $F(a)$ e vale
$$(F^{-1})'(F(a)) = (F'(a))^{-1}.$$

Osservazione Per quanto riguarda la G-differenziabilità non esiste un teorema simile.

Vediamo come applicare questi concetti alle funzioni convesse. Siano X normato, $A \subseteq X$ aperto convesso, $a \in A$ e $f : A \to \mathbb{R}$ convessa e continua. Valgono le seguenti proprietà:

- per ogni $v \in X$ esiste finito
$$f'_+(a, v) = \inf_{\substack{t>0 \\ a+tv}} \frac{f(a + tv) - f(a)}{t}.$$

- $f'_-(a, v) = -f'_+(a, -v)$.

  Dimostrazione È un semplice conto
$$f'_-(a, v) = \lim_{t\to 0^-} \frac{f(a + tv) - f(a)}{t} = [\ s = -t\ ] =$$
$$= \lim_{s\to 0^+} \frac{f(a + s(-v)) - f(a)}{-s} = -f'_+(a, -v). \qquad \square$$

- $f'_+(a, \cdot)$ è sublineare (in particolare convessa) e lipschitziana.

  Dimostrazione La dimostrazione che $f'_+(a, \cdot)$ è positivamente omogenea è semplice. Dimostriamo la subadditività, siano $u, v \in X$
$$f'_+(a, u + v) = \lim_{t\to 0^+} \frac{f(a + tu + tv) - f(a)}{t} = \begin{bmatrix} a + tu + tv = \\ = \frac{1}{2}(a + 2tu) + \frac{1}{2}(a + 2tv) \end{bmatrix} \leq$$
  per convessità
$$\leq \lim_{t\to 0^+} \frac{\frac{1}{2}f(a + 2tu) + \frac{1}{2}f(a + 2tv) - \frac{1}{2}f(a) - \frac{1}{2}f(a)}{t} =$$
$$= \lim_{t\to 0^+} \left(\frac{f(a + 2tu) - f(a)}{2t} + \frac{f(a + 2tv) - f(a)}{2t} \right) = f'_+(a, u) + f'_+(a, v).$$
  Passiamo alla lipschitzianità. Per la definizione
$$f'_+(a, v) \leq \frac{f(a + v) - f(a)}{1} = f(a + v) - f(a),$$
  dunque $f'_+(a, \cdot)$ è limitata superiormente in un opportuno intorno dell'origine. Per proprietà delle funzioni convesse $f'_+(a, \cdot)$ è lipschitziana in $r\,\mathcal{B}_X$ con r opportuno. Ora dimostriamo che in realtà è globalmente lipschitziana: siano $u, v \in X$, sappiamo che esiste $\lambda > 0$ tale che $\lambda u, \lambda v \in r\,\mathcal{B}_X$. Ma allora
$$\left|f'_+(a, u) - f'_+(a, v)\right| = \frac{1}{\lambda}\left|f'_+(a, \lambda u) - f'_+(a, \lambda v)\right| \leq \frac{L}{\lambda}\|\lambda u - \lambda v\| = L\|u - v\|.$$
$\square$

Riportiamo ora i risultati principali riguardanti le funzioni convesse in questa sezione. Per fare ciò abbiamo bisogno di un piccolo teorema.

Teorema 7.4 Siano Z spazio vettoriale e $p : Z \to \mathbb{R}$ sublineare. Valgono

1. per ogni $v \in Z$ si ha $-p(-v) \leq p(v)$;
2. $V = \{v \in Z \, : \, -p(-v) = p(v)\}$ è un sottospazio vettoriale di Z;
3. $p_{|V}$ è lineare.

Dimostrazione L'affermazione 3. è semplice da dimostrare e viene lasciata per esercizio.

1. $0 = p(0) = p(v + (-v)) \leq p(v) + p(-v)$ da cui la tesi.
2. Innazitutto sia $v \in V$ e $\lambda \in \mathbb{R}$, vogliamo mostrare che $\lambda x \in V$. Se $\lambda \geq 0$, per sublinearità, abbiamo la tesi. Sia $\lambda < 0$ abbiamo

$$-p(-\lambda v) = -p(|\lambda| v) = p(-|\lambda| v) = p(\lambda v).$$

Siano ora $u, v \in V$

$$p(u+v) \leq p(u) + p(v) = -p(-u) - p(-v) \leq -p(-u-v) \leq p(u+v),$$

si ha l'uguaglianza ovunque e quindi la tesi. □

Corollario Siano X normato, $A \subseteq X$ aperto convesso, $a \in A$ e $f : A \to \mathbb{R}$ convessa e continua. L'insieme

$$V = \{v \in X \, : \, \text{esiste } f'(a, v)\}$$

è un sottospazio vettoriale chiuso di X e $f_{|a+V}$ è G-differenziabile in a.

Osservazione In particolare se $V = X$ allora f è G-differenziabile in a.

Corollario ("Calculus student's delight") Sia $f : \mathbb{R}^n \to \mathbb{R}$ convessa e continua. Se esistono $\frac{\partial f}{\partial x_i}(a)$ rispetto a tutte le i, allora f è F-differenziabile in a.

Dimostrazione Per il precedente corollario f è G-differenziabile e per il fatto che le funzioni convesse e continue sono localmente lipschitziane, applicando il lemma 5 si ha la tesi. □

7.3 Il subdifferenziale

Abbiamo già definito la nozione di subdifferenziale per funzioni in $\mathbb{R}$, vogliamo ora generalizzare tale concetto e studiarlo più approfonditamente.

Definizione 7.4 (Subdifferenziale) Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $a \in A$. Il *subdifferenziale di f in a* è l'insieme dei funzionali lineari continui che definiscono un funzionale di supporto all'epigrafico della funzione f in $(a, f(a))$, ovvero

$$\partial f(a) = \{x^* \in X^* \, : \, f(x) \geq f(a) + x^*(x-a) \ \forall x \in A\}.$$

Un elemento di $\partial f(a)$ è detto *subgradiente di f in a*.

Osservazione Seppure apparentemente la definizione sembri dipendere da A, in realtà è possibile dimostrare che se $f = g$ in un opportuno intorno di a allora

$$\partial f(a) = \partial g(a),$$

ovvero la definizione è solo locale.

Teorema 7.5 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $a \in A$. Sono tra loro equivalenti:

1. $x^* \in \partial f(a)$;
2. per ogni $v \in X$ si ha $x^*(v) \leq f'_+(a, v)$;
3. per ogni $v \in X$ si ha $-f'_+(a, -v) \leq x^*(v) \leq f'_+(a, v)$.

Dimostrazione Daremo solo un'idea di questa dimostrazione senza entrare nei particolari.

(1. $\Rightarrow$ 3.) Sia $x^* \in \partial f(a)$ e calcoliamo $f'_+(a, v)$ ricordando la definizione di subdifferenziale

$$f'_+(a, v) = \lim_{t \to 0^+} \frac{f(a + tv) - f(a)}{t} \geq \lim_{t \to 0^+} \frac{x^*(tv)}{t} = x^*(v).$$

(3. $\Rightarrow$ 2.) Ovvio.

(2. $\Rightarrow$ 3.) È sufficiente scrivere 2. con $-v$ in luogo di v e si ottiene 3. moltiplicando per -1 l'equazione ottenuta

$$x^*(-v) \leq f'_+(a, -v).$$

(3. $\Rightarrow$ 1.) Dimostreremo la conversa. Siano $x^* \notin \partial f(a)$, allora esiste $v \in A$ tale che $x = a + v$ e

$$f(a + v) < f(a) + x^*(v).$$

Calcoliamo $f'_+(a, v)$, si ha

$$f'_+(a, v) = \lim_{t \to 0^+} \frac{f(a + tv) - f(a)}{t} = \inf_{\substack{t > 0 \\ a + tv \in A}} \frac{f(a + tv) - f(a)}{t} \leq$$
$$\leq f(a + v) - f(a) < x^*(v)$$

e da ciò segue $\neg 2$. □

7.3.1 Proprietà topologiche

Vogliamo ora mostrare alcune proprietà topologiche di $\partial f(a)$, in particolare vogliamo mostrare che è w^*-compatto, convesso e non vuoto. Per fare ciò abbiamo bisogno di due lemmi.

Lemma 6 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua, $r > 0$ e $E \subseteq A$. Se per ogni $x \in E$ vale

$$\partial f(x) \cap r\,\mathcal{B}_{X^*} \neq \emptyset,$$

allora f è r-lipschitziana su E.

Dimostrazione Siano $x_1, x_2 \in E$ e $x_i^* \in \partial f(x_i)$ con $\|x_i^*\| \leq r$ per $i = 1, 2$. Per definizione $f(x_2) \geq f(x_1) + x_1^*(x_2 - x_1)$. Da cui

$$f(x_1) - f(x_2) \leq x_1^*(x_1 - x_2) \leq \|x_1^*\| \cdot \|x_1 - x_2\| \leq r\|x_1 - x_2\|.$$

In modo analogo $f(x_2) - f(x_1) \leq r\|x_2 - x_1\|$ e dunque

$$|f(x_2) - f(x_1)| \leq r\|x_2 - x_1\|. \qquad \square$$

Lemma 7 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $G \subseteq A$ aperto. Se f è r-lipschitziana su G, allora per ogni $x \in G$ si ha

$$\partial f(x) \subseteq r\,\mathcal{B}_{X^*}\,.$$

Dimostrazione Siano $x \in G$ e $\delta > 0$ tale che $x + \delta\,\mathcal{B}_{X^*} \subseteq G$. Fissiamo $x^* \in \partial f(x)$ e calcoliamo

$$\|x^*\| = \sup x^*(\mathcal{B}_X) = \frac{1}{\delta} \sup x^*(\delta\,\mathcal{B}_X) = \frac{1}{\delta} \sup_{\|h\| \leq \delta} \{x^*((x+h) - x)\} \leq$$

per la definizione di subdifferenziale

$$\leq \frac{1}{\delta} \sup_{\|h\| \leq \delta} \{f(x+h) - f(x)\} \leq$$

per la lipschitzianità

$$\leq \frac{1}{\delta} \sup_{\|h\| \leq \delta} \{r\|h\|\} \leq r. \qquad \square$$

Di conseguenza una funzione convessa continua definita su un aperto ha subdifferenziale localmente limitato, detto in altre parole la mappa multivoca ∂f è localmente limitata in A, ovvero

$$\forall x \in A\ \exists U \in \mathcal{U}(x) \text{ tale che } \partial f = \bigcup_{y \in U} \partial f(y) \text{ limitato.}$$

In particolare $\partial f(a)$ è limitato. Enunciamo ora nel seguente corollario le proprietà del subdifferenziale di una funzione convessa continua.

Corollario Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $a \in A$. L'insieme $\partial f(a)$ è non vuoto, convesso e w^*-compatto.

Dimostrazione La convessità è una semplice conseguenza della definizione. Il fatto che non sia vuoto segue dal teorema di Hahn–Banach, infatti esiste un iperpiano che separa $(a, f(a))$ da $\operatorname{int}(\operatorname{epi}(f))$. Sappiamo già che $\partial f(a)$ è limitato, basta ora mostrare che è w^*-chiuso. Fissato $x \in A$ l'insieme

$$H_x = \{x^* \in X^* \,:\, x^*(x - a) \leq f(x) - f(a)\}$$

è un semispazio w^*-chiuso. È semplice osservare che

$$\partial f(a) = \bigcap_{x \in A} H_x$$

e dunque $\partial f(a)$ è w^*-compatto. $\square$

Un'ultima proprietà interessante del subdifferenziale è enunciata nella seguente proposizione

Proposizione 7.2 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua. La mappa multivoca $\partial f : A \to 2^{X^*}$ è un operatore monotono, ovvero

$$\forall x, y \in A \ \forall x^* \in \partial f(x) \ \forall y^* \in \partial f(y) \text{ vale } (x^* - y^*)(x - y) \geq 0.$$

Dimostrazione La dimostrazione consiste in un semplice conto. Siano $x, y \in A$, $x^* \in \partial f(x)$ e $y^* \in \partial f(y)$, per definizione di subdifferenziale abbiamo le disuguaglianze

$$f(y) - f(x) \geq x^*(y - x) \qquad f(x) - f(y) \geq y^*(x - y) = -y^*(y - x)$$

Sommando tali disuguaglianze e moltiplicando per -1 si ha la tesi. □

7.3.2 Continuità

Vogliamo ora definire il concetto di mappa multivoca continua e applicare tali definizioni al subdifferenziale in modo da osservare che tipo di continutà è possibile ottenere.

Definizione 7.5 (Mappa multivoca semicontinua) Siano Y, Z spazi topologici, $y_0 \in Y$ e $\phi : Y \to 2^Z$ mappa multivoca a valori non vuoti. Diciamo che

- ϕ è *superiormente semicontinua in* y_0 se per ogni $W \subseteq Z$ aperto tale che $\phi(y_0) \subseteq W$ esiste $V \in \mathcal{U}(y_0)$ tale che per ogni $y \in V$ si ha $\phi(y) \subseteq W$;
- ϕ è *inferiormente semicontinua in* y_0 se per ogni $W \subseteq Z$ aperto tale che $\phi(y_0) \cap W \neq \emptyset$ esiste $V \in \mathcal{U}(y_0)$ tale che per ogni $y \in V$ si ha $\phi(y) \cap W \neq \emptyset$.

Osservazione Se ϕ è univoca allora le due definizioni coincidono e sono la continuità di ϕ.

Esercizio 7.3.1 Siano Y, Z spazi topologici, $y_0 \in Y$ e $\phi : Y \to 2^Z$ mappa multivoca a valori non vuoti. Posto $\phi^{-1}(G) = \{y \in Y \ : \ \phi(y) \cap G \neq \emptyset\}$ dimostrare che

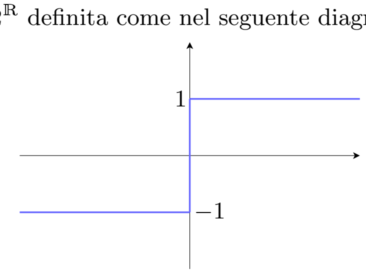
1. ϕ è superiormente semicontinua in Y se e solo se per ogni chiuso $C \subseteq Z$ si ha che $\phi^{-1}(C)$ è chiuso;
2. ϕ è inferiormente semicontinua in Y se e solo se per ogni aperto $A \subseteq Z$ si ha che $\phi^{-1}(A)$ è aperto.

Osservazione Per le mappe multivoche "complementazione e controimmagine non commutano", ovvero non è detto che nell'esercizio precedente, nel punto 1. le stesse cose accadano per gli aperti e, nel punto 2. accadano per i chiusi.

Facciamo alcuni esempi in modo da apprezzare le definizioni appena fornite.

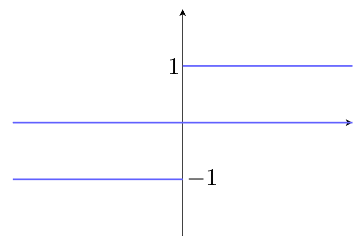
Esempi

- Si consideri $\phi : \mathbb{R} \to 2^{\mathbb{R}}$ definita come nel seguente diagramma.



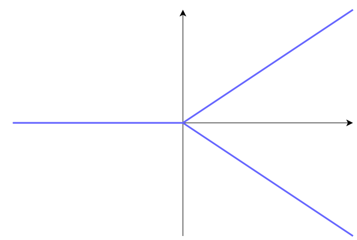
Tale mappa è multivoca solo nell'origine. ϕ è superiormente semicontinua su $\mathbb{R}$, ma non è inferiormente semicontinua nell'origine. È semplice inoltre osservare che tale mappa è il subdifferenziale del modulo, ovvero $\phi = \partial|\cdot|$.

- Si consideri $\phi : \mathbb{R} \to 2^{\mathbb{R}}$ definita come nel seguente diagramma.



ϕ è inferiormente semicontinua su $\mathbb{R}$, ma non è superiormente semicontinua nell'origine, infatti la retroimmagine di $\left(-\frac{1}{2}, \frac{1}{2}\right)$ contiene $\phi(0)$ ma nessun'altra immagine di punti "vicini" all'origine.

- Si consideri $\phi : \mathbb{R} \to 2^{\mathbb{R}}$ definita come nel seguente diagramma.



ϕ è sia inferiormente semicontinua sia superiormente semicontinua su $\mathbb{R}$, quindi ϕ è una mappa multivoca continua.

È importante il concetto di selezione di una mappa multivoca.

Definizione 7.6 (Selezione) Siano Y, Z spazi topologici e $\phi : Y \to 2^Z$ mappa multivoca a valori non vuoti. Una mappa $\varphi : Y \to Z$ è una *selezione per* ϕ se

$$\varphi(y) \in \phi(y) \qquad \forall y \in Y.$$

Una domanda naturale da porsi è quando sia possibile garantire l'esitenza di una selezione continua. Dagli esempi fatti è evidente che la presenza di semicontinutà superiore non è basterebbe, ma il seguente teorema ci rassicura nel caso di semicontinuità inferiore.

Teorema 7.6 (di Michael) Siano T spazio topologico, X spazio di Banach e $\phi : T \to 2^X$ mappa multivoca a valori non vuoti, convessi e chiusi. Se ϕ è semicontinua inferiormente e T è compatto o metrizzabile (o più in generale paracompatto), allora ϕ ammette una selezione continua.

Vogliamo ora arrivare a capire che carattere di continuità possiede il subdifferenziale, per fare ciò abbiamo bisogno di un lemma.

Lemma Siano X normato e $\{x_\alpha\} \subseteq X$, $\{x_\alpha^*\} \subseteq X^*$ due net tali che

$$x_\alpha \xrightarrow{\|\cdot\|} x \qquad x_\alpha^* \xrightarrow{w^*} x^*.$$

Se $\{x_\alpha^*\}$ è limitata, allora $x_\alpha^*(x_\alpha) \to x^*(x)$.

Dimostrazione È un semplice conto

$$|x_\alpha^*(x_\alpha) - x^*(x)| = |(x_\alpha^* - x^*)(x) + x_\alpha^*(x_\alpha - x)| \le |(x_\alpha^* - x^*)(x)| + M\|x_\alpha - x\|$$

e per ipotesi $|(x_\alpha^* - x^*)(x)| \to 0$ e $\|x_\alpha - x\| \to 0$. Da ciò la tesi. □

Teorema 7.7 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua. $\partial f : A \to 2^{X^*}$ è $(\|\cdot\|\text{-}w^*)$-superiormente semicontinua su A.

Dimostrazione Supponiamo per assurdo che non sia $(\|\cdot\|\text{-}w^*)$-superiormente semicontinua in $a \in A$, ovvero esiste $W \subseteq X^*$ w^*-aperto con $\partial f(a) \subseteq W$ tale che per ogni $V \in \mathcal{U}(a)$ esiste $x_V \in V$ tale che

$$\partial f(x_V) \nsubseteq W.$$

Sia $V_n = \operatorname{int}(\mathcal{B}(a, \frac{1}{n}))$ e consideriamo $x_n \in V_n$, sappiamo che $x_n \to a$ e esiste $x_n^* \in \partial f(x_n) \smallsetminus W$. In un intorno di a ∂f è limitato, quindi sia $\{x_n\}$ sia $\{x_n^*\}$ sono limitate. Per il teorema di Banach–Alaoglu esiste $\{x_{n_\alpha}^*\}$ subnet di $\{x_n^*\}$ tale che

$$x_{n_\alpha}^* \xrightarrow{w^*} x^* \in X^*.$$

Ma sappiamo che

$$f(y) \ge f(x_{n_\alpha}) + x_{n_\alpha}^*(y - x_{n_\alpha})$$

passando al limite e sfruttano il lemma precedente e la continuità di f

$$f(y) \ge f(a) + x^*(y - a),$$

da cui $x^* \in \partial f(a) \subseteq W$. Ma $\{x_{n_\alpha}^*\} \subseteq X^* \smallsetminus W$ che è w^*-chiuso e quindi $x^* \in X^* \smallsetminus W$. Assurdo. □

Osservazione Se ∂f è univoco allora risulta essere una funzione continua da un insieme con la topologia della norma ad un insieme con la topologia debole*.

Veniamo ora a delle carattrerizzazioni della G-differenziabilità e della F-differenziabilità in termini di continuità del subdifferenziale e delle sue selezioni.

Teorema 7.8 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $a \in A$. Sono tra loro equivalenti

1. f è G-differenziabile in a;
2. $\operatorname{card}(\partial f(a)) = 1$;
3. $\operatorname{card}(\partial f(a)) = 1$ e ∂f è $(\|\cdot\|\text{-}w^*)$-superiormente semicontinua in a;
4. esiste $\varphi : A \to X^*$ selezione di ∂f tale che φ è $(\|\cdot\|\text{-}w^*)$-continua in a.

Dimostrazione

(1. $\Rightarrow$ 2.) Supponiamo per assurdo che esistano $x^*, y^* \in \partial f(a)$ distinti, ovvero esiste $v \in \mathcal{S}_X$ tale che $x^*(v) < y^*(v)$. Ma

$$-f'_+(a, -v) \le x^*(v) < y^*(v) \le f'_+(a, v),$$

ovvero, nella direzione v, la derivata destra e sinistra non coincidono. Dunque f non è G-differenziabile in a, il che è assurdo.

(2. $\Rightarrow$ 3.) È il teorema precedente.

(3. $\Rightarrow$ 4.) 3. dice che tutte le selezioni di ∂f sono $(\|\cdot\|\text{-}w^*)$-continue in a.

(4. $\Rightarrow$ 1.) Sia $t > 0$ e $v \in X$. Per la definizione di subdifferenziale

$$f(a+tv) - f(a) \le (\varphi(a+tv))[tv].$$

Da cui

$$\begin{aligned} \frac{f(a+tv)-f(a)}{t} - (\varphi(a))[v] &= \frac{f(a+tv)-f(a)-(\varphi(a))[tv]}{t} \le \\ \le \frac{(\varphi(a+tv))[tv]-(\varphi(a))[tv]}{t} &= (\varphi(a+tv)-\varphi(a))[v] \xrightarrow{t\to 0^+} 0 \end{aligned}$$

Abbiamo mostrato che per ogni $v \in X$

$$f'_+(a,v) = (\varphi(a))[v] \qquad -f'_+(a,-v) = -(\varphi(a))[-v] = (\varphi(a))[v]$$

ovvero $f'(a,v) = (\varphi(a))[v]$ che è equivalente alla G-differenziablilità in a. □

Teorema 7.9 Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua e $a \in A$. Sono tra loro equivalenti

1. f è F-differenziabile in a;
2. $\mathrm{card}(\partial f(a)) = 1$ e ∂f è $(\|\cdot\|$-$\|\cdot\|)$-superiormente semicontinua in a;
3. esiste $\varphi : A \to X^*$ selezione di ∂f che è $(\|\cdot\|$-$\|\cdot\|)$-continua in a.

Dimostrazione Osserviamo anzitutto che la condizione 2. equivale a richiedere che se $x_n \to a$ e $x_n^* \in \partial f(x_n)$ allora $x_n^* \xrightarrow{\|\cdot\|} x_0^* \in \partial f(a)$.
(1. $\Rightarrow$ 2.) Sappiamo già che f è G-differenziabile in a e per il teorema precedente

$$\partial f(a) = \{x_0^*\}$$

e tale elemento è il differenziale di Fréchet di f in a, ovvero

$$0 \le f(a+h) - f(a) - x_0^*(h) \le \omega(\|h\|) \qquad (\star)$$

con $\frac{\omega(t)}{t} \xrightarrow{t\to 0^+} 0$. Sia $x_n \to a$ e $x_n^* \in \partial f(x_n)$, siccome ∂f è localmente limitato allora esiste M tale che $\|x_n^*\| \le M$ per tutti gli n. Bisogna ora fare alcuni conti

$$\begin{aligned} (x_n^* - x_0^*)(h) &= -x_0^*(h) + x_n^*(x_n + h - x_n) = \\ &= -x_0^*(h) + x_n^*(a+h-x_n) + x_n^*(x_n - a) \le \end{aligned}$$

per la $(\star)$ si ha

$$\le \omega(\|h\|) - f(a+h) + f(a) + x_n^*(a+h-x_n) + x_n^*(x_n-a) \le$$

per la definizione di subdifferenziale e per limitatezza

$$\le \omega(\|h\|) - f(a+h) + f(a) + f(a+h) - f(x_n) + M\|x_n - a\|.$$

Dunque si ottiene

$$\begin{aligned} \|x_n^* - x_0^*\| &= \sup_{\|h\|=r} (x_n^* - x_0^*)\left(\frac{h}{r}\right) \le \\ &\le \sup_{\|h\|=r} \left\{ \frac{\omega(r)}{r} + \frac{f(a)-f(x_n)}{r} + M\frac{\|x_n-a\|}{r} \right\}, \end{aligned}$$

possiamo togliere l'estremo superiore succome tutto è indipendente da h e passando al limite superiore ottenendo

$$\limsup_{n\to+\infty} \|x_n^* - x_0^*\| \leq \limsup_{n\to+\infty} \left\{ \frac{\omega(r)}{r} + \frac{f(a) - f(x_n)}{r} + M\frac{\|x_n - a\|}{r} \right\} = \frac{\omega(r)}{r}$$

per ogni $r > 0$. Da cui $\limsup \|x_n^* - x_0^*\| = 0$ e dunque $x_n^* \xrightarrow{\|\cdot\|} x_0^*$.

(2. $\Rightarrow$ 3.) Ovvio.

(3. $\Rightarrow$ 1.) Consideriamo

$$0 \leq f(a+h) - f(a) - (\varphi(a))[h] \leq (\varphi(a+h))[h] - (\varphi(a))[h] \leq$$

passando alle norme

$$\leq \|\varphi(a+h) - \varphi(a)\|\|h\| = o(1)\|h\| = o(\|h\|)$$

per $h \to 0$. Quindi f è F-differenziabile in a. □

7.4 I teoremi di Zajíček e di Preiss–Zajíček

Ci interessa ora studiare dove una funzione convessa non è differenziabile, in particolare ci interesserà capire che peso topologico (nel senso di Baire) hanno i punti di non differenziabilità. Abbiamo prima bisogno di una piccola, ma importante, proposizione riguardante l'estensione di funzioni lipschitziane.

Proposizione 7.3 (Estensione di mappe lipschitziane) Siano (X, d) spazio metrico, $E \subseteq X$ e $\varphi : E \to \mathbb{R}$ mappa r-lipschitziana. La mappa

$$\widehat{\varphi}(x) = \inf \{\varphi(y) + r\,\mathrm{d}(x,y) \ : \ y \in E\}$$

è un'estensione r-lipschitziana di φ su X. Inoltre se X è normato, E è convesso e φ è convessa su E allora $\widehat{\varphi}$ è convessa su E.

Dimostrazione $\widehat{\varphi}$ è estremo inferiore di mappe r-lipschitziane e quindi è r-lipschitziana. Inoltre se $x \in E$ allora

$$\widehat{\varphi}(x) \leq \varphi(x) + r\,\mathrm{d}(x,x) \leq \varphi(y) + r\,\mathrm{d}(x,y)$$

ove la seconda disuguaglianza è dovuta alla lipschitzianità. Passando agli estremi inferiori si ottiene $\widehat{\varphi}(x) \leq \varphi(x) \leq \widehat{\varphi}(x)$, ovvero $\widehat{\varphi}_{|E} = \varphi$. Facciamo ora la seconda parte. Siano $x_1, x_2 \in E$ e $\alpha_1, \alpha_2 \geq 0$ tali che $\alpha_1 + \alpha_2 = 1$, per ogni $\varepsilon > 0$ siano $y_i \in E$ tale che

$$\varphi(y_i) + r\|x_i - y_i\| \leq \widehat{\varphi}(x_i) + \varepsilon \qquad i = 1, 2.$$

E ora il solito conto

$$\begin{aligned}\widehat{\varphi}(\alpha_1 x_1 + \alpha_2 x_2) &\leq \varphi(\alpha_1 y_1 + \alpha_2 y_2) + r\|\alpha_1 x_1 + \alpha_2 x_2 - \alpha_1 y_1 - \alpha_2 y_2\| \leq \\ &\leq \alpha_1\varphi(x_1) + \alpha_2\varphi(x_2) + \alpha_1 r\|x_1 - y_1\| + \alpha_2 r\|x_2 - y_2\| \leq \\ &\leq \alpha_1\widehat{\varphi}(x_1) + \alpha_2\widehat{\varphi}(x_2) + \varepsilon\end{aligned}$$

e dall'arbitrarietà di ε si ha la tesi. □

Ancora un paio di definizioni preliminari.

Definizione 7.7 (Ipersuperficie lipschitziana) Siano X normato e $L \subseteq X$. L si dice *ipersuperficie lipschitziana* se esistono H sottospazio chiuso di X di codimensione 1 e $\varphi : H \to \mathbb{R}$ mappa lipschitziana tale che

$$L = \{u + \varphi(u)v \,:\, u \in H\}$$

con $v \in \mathcal{S}_X \smallsetminus H$ opportuno.

Osservazione Se L è un'ipersuperficie lipschitziana allora L è mai denso. Addirittura se $X = \mathbb{R}^d$ si ottiene che L ha misura nulla secondo Lebesgue (in pratica segue dal teorema di Fubini).

Possiamo finalmente enunciare il teorema che abbiamo promesso ad inizio sezione.

Teorema 7.10 (di Zajíček) Siano X normato separabile e $T : X \to 2^{X^*}$ operatore monotono. L'insieme

$$M(T) = \{x \in X \,:\, \mathrm{card}(T(x)) > 1\}$$

è contenuto in un'unione al più numerabile di ipersuperfici lipschitziane.

Dimostrazione L'idea della dimostrazione è quella di decomporre in maniera furba l'insieme $M(T)$. Per ogni $x \in M(T)$ esistono $a_x^*, b_x^* \in T(x)$ distinti. Sia $D \subseteq \mathcal{S}_X$ denso e numerabile, sappiamo che

- esiste $v_x \in D$ tale che $a_x^*(v_x) < b_x^*(v_x)$;
- esistono $\alpha_x, \beta_x \in \mathbb{Q}$ tali che $a_x^*(v_x) < \alpha_x < \beta_x < b_x^*(v_x)$;
- esiste $n_x \in \mathbb{N}$ tale che $\|a_x^*\| \le n_x$ e $\|b_x^*\| \le n_x$.

Se definiamo

$$A(v, \alpha, \beta, n) = \{x \in M(T) \,:\, v_x = v,\ \alpha_x = \alpha,\ \beta_x = \beta,\ n_x = n\},$$

abbiamo

$$M(T) = \bigcup_{\substack{v \in D, n \in \mathbb{N} \\ \alpha, \beta \in \mathbb{Q}, \alpha \neq \beta}} A(v, \alpha, \beta, n).$$

Per semplicità indichiamo con A uno degli insiemi $A(v, \alpha, \beta, n)$ sopra definiti, vogliamo mostrare che tale insieme è contenuto in un'ipersuperficie lipschitziana. Sia $v^* \in \mathcal{S}_{X^*}$ tale che $v^*(v) = \|v\|$ e $H = \ker(v^*)$. Fissati $x, y \in A$ sicuramente

$$x = u_x + t_x v \qquad y = u_y + t_y v$$

con $u_x, u_y \in H$ e $t_x, t_y \in \mathbb{R}$. Per monotonia

$$\begin{aligned} 0 \le (a_y^* - b_x^*)(y - x) = (a_y^* - b_x^*)(u_y - u_x) + (a_y^* - b_x^*)(v) \cdot (t_y - t_x) < \\ < (a_y^* - b_x^*)(u_y - u_x) + (\alpha - \beta)(t_y - t_x). \end{aligned}$$

Quindi

$$(t_y - t_x) < \frac{\|a_y^* - b_x^*\| \cdot \|u_y - u_x\|}{\beta - \alpha} \le \frac{2n}{\beta - \alpha}\|u_y - u_x\|.$$

Scambiando poi x con y si ottiene la stessa disuguaglianza con il modulo. La funzione $\varphi(u_x) = t_x$ è ben data e lipschitziana da un sottoinsieme di H a valori in $\mathbb{R}$, possiamo estenderla a tutto H alla mappa $\widehat{\varphi} : H \to \mathbb{R}$. Quindi

$$A \subseteq \{u + \widehat{\varphi}(u)v \,:\, u \in H\},$$

ma ciò è la tesi. □

Vediamo cosa comporta il teorema di Zajíček riguardo il subdifferenziale. Se f è convessa continua su $A \subseteq X$ sottoinsieme aperto e convesso di uno spazio normato allora indichiamo con

$$N_G(f) = \{x \in A : f \text{ non è G-differenziabile in } x\},$$
$$N_F(f) = \{x \in A : f \text{ non è F-differenziabile in } x\}.$$

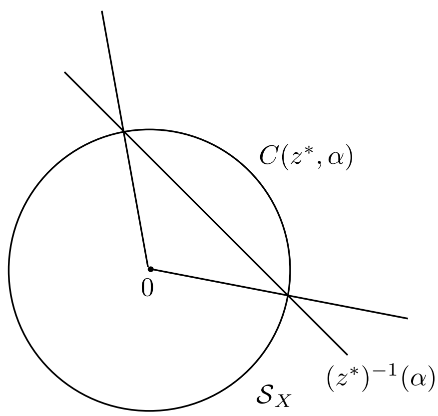
È semplice osservare che $N_G(f) \subseteq N_F(f)$. Parafrasando il teorema di Zajíček si ottiene che se X è normato separabile, allora $N_G(f)$ è di prima categoria secondo Baire. Se addirittura $X = \mathbb{R}^d$ allora $N_F(f) = N_G(f)$ è un insieme di prima categoria secondo Baire e di misura nulla secondo Lebesgue.

È naturale chiederesi quando $N_F(f)$ risulti essere di prima categoria secondo Baire. Vogliamo ottenere un risultato che ci fornisca delle condizioni sufficienti a tale scopo, ma per fare ciò abbiamo bisogno di una definizione.

Definizione 7.8 Sia X normato, $z^* \in \mathcal{S}_{X^*}$ e $\alpha \in (0,1)$. Indichieremo nel seguente modo il seguente insieme

$$C(z^*, \alpha) = \{x \in X : z^*(x) \geq \alpha\|x\|\},$$

spesso z^* e α vengono detti rispettivamente *direzione* e *ampiezza*.



Osservazione Enunciamo alcune propreità di base dell'insieme $C(z^*, \alpha)$.

- $C(z^*, \alpha)$ è non vuoto, infatti $0 \in C(z^*, \alpha)$;
- se $x \in C(z^*, \alpha)$ e $t > 0$ allora $tx \in C(z^*, \alpha)$, ovvero $C(z^*, \alpha)$ è un *cono*;
- $C(z^*, \alpha)$ è convesso;[8]
- se $x \in C(z^*, \alpha) \cap \mathcal{S}_X$ allora $z^*(x) \geq \alpha$.

Lasciamo come esercizio il seguente risultato, che tornerà utile per comprendere il corollario di fine sezione.

Esercizio 7.4.1 Siano Y, Z normati, $T : Y \to 2^Z$ e $y \in Y$. Sono tra loro equivalenti

1. $\mathrm{card}(T(y)) = 1$ e T è $(\|\cdot\|\text{-}\|\cdot\|)$-superiormente semicontinua in y;
2. $T(y) \neq \emptyset$ e $\lim_{\delta\to 0^+} \mathrm{diam}\, T(\mathcal{B}(y, \delta)) = 0$.[9]

[8] Ciò è dovuto al fatto che $\alpha\|\cdot\| - z^*(\cdot)$ è una funzione convessa.

[9] Si osservi che il limite $\lim_{\delta\to 0^+} \cdots$ è uguale a $\inf_{\delta>0} \cdots$ per monotonia in δ. Esso viene spesso chiamato *oscillazione di T in y*.

Teorema 7.11 (di Preiss-Zajíček) Siano X normato e $T : X \to 2^{X^*}$ operatore monotono. Se X^* è separabile, allora

$$NC(T) = \left\{ x \in X \, : \, T(x) \neq \emptyset \text{ e } \lim_{\delta \to 0^+} \operatorname{diam} T(\mathcal{B}(x, \delta)) > 0 \right\}$$

è di prima categoria secondo Baire.

Dimostrazione Anche in questo caso bisogna decomporre l'insieme in modo furbo. Fissiamo $\alpha \in (0,1)$ e $E \subseteq X^*$ numerabile e denso. Sappiamo che

- esiste $n_x \in \mathbb{N}$ tale che
$$\lim_{\delta \to 0^+} \operatorname{diam} T(\mathcal{B}(x, \delta)) > \frac{2}{n_x};$$
- esiste $e_x^* \in E$ tale che
$$\operatorname{dist}(e_x^*, T(x)) < \frac{\alpha}{n_x}.$$

Posto $A(n, e^*) = \{x \in NC(T) \, : \, n_x = n \text{ e } e_x^* = e^*\}$, allora

$$NC(T) = \bigcup_{\substack{n \in \mathbb{N} \\ e^* \in E}} A(n, e^*).$$

Vogliamo mostrare che gli insiemi $A(n, e^*)$ sono mai densi. Sia $x \in A(n, e^*)$, per ogni $\delta > 0$ esistono $x_1, x_2 \in \mathcal{B}(x, \delta)$ e $x_i^* \in T(x_i)$ tali che $\|x_1^* - x_2^*\| > \frac{2}{n}$, ed esiste $k \in \{1, 2\}$ tale che $\|x_k^* - e^*\| > \frac{1}{n}$.[10] Poniamo

$$z^* = \frac{x_k^* - e^*}{\|x_k^* - e^*\|}.$$

Sia ora $y \in A(n, e^*)$ qualsiasi, sappiamo che esiste $y^* \in T(y)$ tale che $\|y^* - e^*\| < \frac{\alpha}{n}$. Ora

$$\|x_k^* - e^*\| \cdot z^*(y - x_k) = (x_k^* - e^*)(y - x_k) = (x_k^* - y^*)(y - x_k) + (y^* - e^*)(y - x_k) \leq$$

per monotonia

$$\leq \|y^* - e^*\| \|y - x_k\|.$$

Quindi

$$z^*(y - x_k) \leq \frac{\|y^* - e^*\| \|y - x_k\|}{\|x_k^* - e^*\|} < \frac{\frac{\alpha}{n} \|y - x_k\|}{\frac{1}{n}} = \alpha \|y - x_k\|.$$

Abbiamo ottenuto che $y \notin x_k + C(z^*, \alpha)$, ovvero per ogni $\delta > 0$ esiste un $x_k \in \mathcal{B}(x, \delta)$ tale che per ogni $y \in A(n, e^*)$ vale $y \notin x_k + C(z^*, \alpha)$, cioè

$$A(n, e^*) \cap (x_k + C(z^*, \alpha)) = \emptyset.$$

In pratica ogni bolla ha un buco e quindi $A(n, e^*)$ è mai denso. □

Corollario Siano X normato, $A \subseteq X$ aperto convesso, $f : A \to \mathbb{R}$ convessa e continua. Se X^* è separabile, allora $N_F(f)$ è di prima categoria secondo Baire.

[10] Altrimenti, utilizzando la disuguaglianza triangolare, si contraddice $\|x_1^* - x_2^*\| > \frac{2}{n}$.

7.5 Mappa di dualità

L'idea di questa sezione è quella di ricavare delle proprietà geometriche, degli spazi di Banach, a partire dalla differenziabilità secondo Fréchet o Gâteaux della norma. Sia $C \subseteq X^*$ convesso, w^*-compatto e non vuoto, sappiamo già cosa intendiamo per funzione di supporto per C, ovvero

$$s_C(x) = \sup x(C),$$

osserviamo che $s_{\mathcal{B}_{X^*}}(x) = \|x\|$.

Osservazione Sia f convessa continua su A aperto convesso di X normato. Fissati $a \in A$ e $v \in X$ allora

$$\{x^*(v) \,:\, x^* \in \partial f(a)\} = [-f'_+(a,-v), f'_+(a,v)].$$

Dimostrazione

($\subseteq$) Segue dal teorema 7.5.

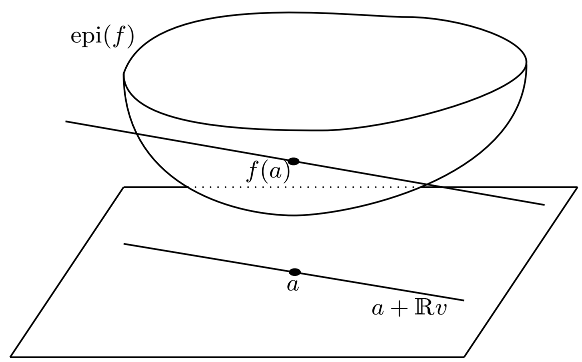
($\supseteq$) Posto $\varphi(t) = f(a+tv)$, fissiamo

$$m \in [-f'_+(a,-v), f'_+(a,v)] = [\varphi'_-(0), \varphi'_+(0)].$$

Sempre dal teorema 7.5 segue che $m \in \partial\varphi(0)$, ma allora

$$f(a+tv) \geq f(a) + mt \qquad \forall t \in \mathbb{R}.$$

Il diagramma permette di caire cosa abbiamo ottenuto



Quindi la retta in figura è tutta sotto l'epigrafico di f. Estendiamo, utilizzando il teorema di Hahn–Banach, ricordando che l'epigrafico ha interno non vuoto per continuità di f, ed otteniamo un funzionale lineare $x^* \in \partial f(a)$ tale che $x^*(v) = m$. □

Vogliamo arrivare ad un risultato che permetta di dare una corrispondenza tra gli insiemi convessi, w^*-compatti e non vuoti e le funzioni continue e sublineari. Per fare ciò abbiamo bisogno del seguente lemma.

Lemma Sia X normato e $p : X \to \mathbb{R}$ continuo e sublineare. Esiste $C \subseteq X^*$ convesso, w^*-compatto e non vuoto tale che

$$p(x) = s_C(x) \qquad \forall x \in X.$$

In realtà mostreremo di più, ovvero $C = \partial p(0)$.

Dimostrazione Osserviamo che

$$p'_+(0,v) = \lim_{t\to 0^+} \frac{p(tv)-p(0)}{t} = \lim_{t\to 0^+} \frac{tp(v)-0}{t} = p(v).$$

Dunque $x^* \in \partial p(0)$ se e solo se $x^*(v) \leq p'_+(0,v) = p(v)$. Per l'osservazione precedente si ottiene che

$$s_{\partial p(0)}(v) = \sup\{x^*(v) \,:\, x^* \in \partial p(0)\} = p'_+(0,v) = p(v). \qquad \square$$

Corollario Sia X normato. Se denotiamo

$$\mathcal{P} = \{p : X \to \mathbb{R} \,:\, p \text{ continua e sublineare}\},$$
$$\mathcal{C} = \{C \subseteq X^* \,:\, C \text{ convesso, } w^*\text{-compatto e non vuoto}\},$$

allora le applicazioni

$$\mathcal{P} \ni p \mapsto \partial p(0) \in \mathcal{C} \qquad \mathcal{C} \ni C \mapsto p = s_C \in \mathcal{P}$$

l'una inversa dell'altra, definiscono una corrispondenza biunivoca tra $\mathcal{P}$ e $\mathcal{C}$.

Ci interessa ora capire come è fatto il subdifferenziale della norma, per capirlo abbiamo bisogno di un lemma.

Lemma Siano X normato, $C \subseteq X^*$ w^*-compatto, convesso e non vuoto e $x \in X$. Valgono

1. $\partial s_C(x) \subseteq C$;
2. $\partial s_C(x) = \{c^* \in C \,:\, c^*(x) = s_C(x)\}$, ovvero è l'insieme dei punti di supporto a C che hanno come funzionale di supporto x.

Dimostrazione

1. Sia $x^* \in \partial s_C(x)$, dalla definizione di subdifferenziale e per sublinearità si ottiene
$$x^*(v) \leq s_C(x+v) - s_C(x) \leq s_C(x) + s_C(v) - s_C(x) = s_C(v).$$
Dunque $x^*(v) \leq s_C(v) - s_C(0)$, ovvero $x^* \in \partial s_C(0) = C$.
2. Facciamo vedere le due inclusioni.
   - ($\subseteq$) Se $x^* \in \partial s_C(x)$ allora per ogni $v \in X$ vale $x^*(v) \leq s_C(x+v) - s_C(x)$. Prendiamo $v = -\frac{1}{2}x$, si ottiene
   $$-\frac{1}{2}x^*(x) \leq -\frac{1}{2}s_C(x) \Longrightarrow s_C(x) \leq x^*(x).$$
   Ma per definizione, siccome $x^* \in C$, si ha $x^*(x) \leq s_C(x)$ e quindi $x^*(x) = s_C(x)$.
   - ($\supseteq$) Sia $c^* \in C$ tale che $c^*(x) = s_C(x)$. Si ha per ogni $y \in X$
   $$s_C(y) \geq c^*(y) = c^*(x) + c^*(y-x) = s_C(x) + c^*(y-x)$$
   che è la definizione di subdifferenziale, quindi $c^* \in \partial s_C(x)$. $\square$

Corollario Sia X normato. Dal fatto che $\|\cdot\| = s_{\mathcal{B}_{X^*}}(\cdot)$ segue

$$(\partial\|\cdot\|)(x) = \begin{cases} \mathcal{B}_{X^*} & x = 0 \\ \{x^* \in \mathcal{B}_{X^*} \,:\, x^*(x) = \|x\|\} & x \neq 0 \end{cases}$$

Osservazione Se $x \neq 0$ allora $(\partial\|\cdot\|)(x) = \{x^* \in \mathcal{S}_{X^*} \,:\, x^*(x) = \|x\|\}$.

Definiamo ora cosa intendiamo per mappa di dualità e vediamo alcune sue applicazioni riguardanti la geometria degli spazi normati.

Definizione 7.9 (Mappa di Dualità) Sia X normato. La *mappa di dualità di X* è la mappa multivoca $D_X : \mathcal{S}_X \to 2^{\mathcal{S}_{X^*}}$ definita come segue

$$D_X(x) = \{x^* \in \mathcal{S}_{X^*} \,:\, x^*(x) = 1\} = (\partial\|\cdot\|)(x).$$

Osservazione La mappa di dualità è "simmetrica" in X e X^*. Infatti se $x \in \mathcal{S}_X$ e $x^* \in \mathcal{S}_{X^*}$ si ha

$$x^* \in D_X(x) \Longleftrightarrow x \in D_{X^*}(x^*),$$

inoltre $D_X^{-1}(x^*) = D_{X^*}(x^*) \cap X$.

7.5.1 Proprietà geometriche degli spazi normati

Ovviamente sono imprescindibili alcune definizioni di carattere geometrico.

Definizione 7.10 (Spazio G-smooth/F-smooth/strettamente convesso) Sia X normato. Diciamo che

- X è *G-smooth* se $\|\cdot\|$ è G-differenziabile su $X \setminus \{0\}$;
- X è *F-smooth* se $\|\cdot\|$ è F-differenziabile su $X \setminus \{0\}$;
- X è *strettamente convesso* se $\mathcal{B}_X$ è strettamente convessa, ovvero se $x, y \in \mathcal{B}_X$ con $x \neq y$ allora $\left\|\frac{x+y}{2}\right\| < 1$.

Osservazione Grazie alla convessità della norma, utilizzando i teoremi 7.8 e 7.9, è possibile mostrare che:

- X è G-smooth se e solo se D_X è univoca (inoltre D_X è $(\|\cdot\|$-$w^*)$-continua);
- X è F-smooth se e solo se D_X è univoca e $(\|\cdot\|$-$\|\cdot\|)$-continua;
- X è strettamente convesso se e solo se $\mathcal{S}_X$ non contiene segmenti.

Valgono le seguenti proprietà.

Proposizione 7.4 Sia X normato. X è strettamente convesso se e solo se $\operatorname{ext}\mathcal{B}_X = \mathcal{S}_X$.

Dimostrazione Se $\operatorname{ext}\mathcal{B}_X = \mathcal{S}_X$ necessariamente $\mathcal{S}_X$ non contiene segmenti, altrimenti se $[y, z] \subseteq \mathcal{S}_X$, allora $\frac{y+z}{2} \in \mathcal{S}_X$ non è estermo per la bolla. Dimostriamo il viceversa. Se per assurdo $\operatorname{ext}\mathcal{B}_X \subsetneq \mathcal{S}_X$, allora esiste $x \in \mathcal{S}_X$ tale che[11]

$$x = \frac{x_1 + x_2}{2} \qquad x_i \in \mathcal{S}_X,$$

ma ciò è assurdo siccome $\left\|\frac{x_1+x_2}{2}\right\| < 1$. □

[11] Gli x_i devono appartenere a $\mathcal{S}_X$ altrimenti, applicando in modo semplice la disuguaglianza triangolare, si ottiene che $x \notin \mathcal{S}_X$.

Proposizione 7.5 Sia X normato. Se X^* è strettamente convesso, allora X è G-smooth.

Dimostrazione Vogliamo D_X univoca, ma $D_X(x) \subseteq \mathcal{S}_{X^*}$ è un convesso non vuoto. Essendo $\mathcal{S}_{X^*}$ strettamente convessa necessariamente $D_X(x)$ è ridotto a un solo punto. □

Proposizione 7.6 Sia X normato. Se X^* è G-smooth, allora X è strettamente convesso.

Dimostrazione Supponiamo per assurdo che X non sia strettamente convesso, sappiamo quindi che esistono $x_1, x_2 \in \mathcal{S}_X$ distinti tali che $[x_1, x_2] \subseteq \mathcal{S}_X$. Consideriamo

$$x = \frac{x_1 + x_2}{2} \in \mathcal{S}_X,$$

e $x^* \in D_X(x)$. x^* definisce un iperpiano di supporto che passa per x e quindi per x_1 e x_2. Da ciò segue $x^* \in D_X(x_1) \cap D_X(x_2)$, perciò $\{x_1, x_2\} \subseteq D_{X^*}(x^*)$. Ma questo implica che X^* non è G-smooth. □

Osservazione Per nessuna delle due proposizioni vale il viceversa, ma se lo spazio X è riflessivo allora le frecce si invertono.[12]

Si osservi che

$$D_X(\mathcal{S}_X) = \{x^* \in \mathcal{S}_{X^*} \,:\, x^* \text{ assume la norma}\}.$$

questo permette di dare una riscrittura ad alcuni noti teoremi:

- il teorema di James asserisce che se X è uno spazio di Banach allora
$$D_X(\mathcal{S}_X) = \mathcal{S}_{X^*} \iff X \text{ è riflessivo;}$$
- il teorema di Bishop–Phelps asserisce che se X è uno spazio di Banach allora
$$\overline{D_X(\mathcal{S}_X)} = \mathcal{S}_{X^*}\,.$$

Utilizzeremo questa versione di tali teoremi nel seguente interessate risultato.

Teorema 7.12 Sia X spazio di Banach. Se X^* è F-smooth allora X è riflessivo.

Dimostrazione Sia $x^{**} \in X^{**}$ qualsiasi e di norma 1. Per il teorema di Bishop–Phelps esiste $y^{**} \in D_{X^*}(\mathcal{S}_{X^*})$ tale che sia arbitrariamente vicino a x^{**}. Esiste dunque $y^* \in \mathcal{S}_{X^*}$ tale che $D_{X^*}(y^*) = \{y^{**}\}$ (è ridotto ad un solo punto per F-smoothness). Usiamo nuovamente il teorema di Bishop–Phelps e otteniamo che esiste $\{z_n^*\} \subseteq D_X(\mathcal{S}_X)$ tale che

$$z_n^* \xrightarrow{\|\cdot\|} y^*.$$

Esiste dunque qualche $z_n \in \mathcal{S}_X$ tale che $z_n^* \in D_X(z_n)$ e quindi $z_n \in D_{X^*}(z_n^*)$. Per F-smoothness (in particolare la continuità di D_{X^*})

$$z_n \xrightarrow{\|\cdot\|} y^{**}.$$

Siccome X è di Banach allora X è chiuso in X^{**}, quindi $y^{**} \in X$ e

$$\operatorname{dist}(x^{**}, X) = 0.$$

Per la chiusura di X segue $x^{**} \in X$. □

[12] È bene ricordare che se uno spazio è riflessivo è sempre possibile rinormarlo in maniera tale che la nuova norma risulti non essere strettamente convessa.

7.5.2 Applicazioni a spazi concreti

Vediamo come in alcuni spazi noti si applichino i risultati appena ottenuti.

- Sia $X = \ell^1$, sappiamo che $X^* = \ell^\infty$ con isometria
$$x^*(x) = \sum_{n=1}^{+\infty} x^*(n)x(n).$$
Sia $x^* \in \mathcal{S}_{\ell^\infty}$, si ha
$$x^* \in D_{\ell^1}(x) \Longleftrightarrow \sum_{n=1}^{+\infty} x^*(n)x(n) = 1 = \sum_{n=1}^{+\infty} |x(n)|,$$
dal fatto che $x^*(n)x(n) \leq |x(n)|$ segue
$$x^* \in D_{\ell^1}(x) \Longleftrightarrow x^*(n)x(n) = |x(n)| \quad \forall n$$
e dunque
$$x^* \in D_{\ell^1}(x) \Longleftrightarrow \left[\begin{array}{c} \text{se } x(n) \neq 0 \text{ allora } x^*(n) = \operatorname{sgn}(x(n)) \\ \text{se } x(n) = 0 \text{ allora } x^*(n) \in [-1,1] \text{ qualsiasi} \end{array} \right]$$
Quindi $\|\cdot\|_1$ è G-differenziabile in x se e solo se $x(n) \neq 0$ per ogni n. Vediamo in quali punti abbiamo F-differenziabilità, ovviamente dobbiamo richiedere $x(n) \neq 0$ per ogni n. Fissato $x \in \mathcal{S}_{\ell^1}$ esiste
$$x_n = (x(1), x(2), \ldots, x(n-1), -x(n), x(n+1), \ldots)$$
e $\|x - x_n\| = 2|x(n)| \to 0$ e $\|x_n\| = 1$. Si ha
$$D(x) = \{(x^*(1), x^*(2), \ldots)\} \qquad D(x_n) = \{(x^*(1), \ldots, -x^*(n), \ldots)\}$$
con $x^*(i) \in \{\pm 1\}$ e quindi $\|D(x) - D(x_n)\|_\infty = 2|x^*(n)| = 1$, perciò $D(x_n) \not\to D(x)$. Nessun punto di $\mathcal{S}_X$ è di F-differenziabilità.

- Se $X = c_0$ e $x \in \mathcal{S}_{c_0}$ allora
$$\begin{aligned} \|\cdot\|_\infty \text{ è G-differenziabile in } x &\Longleftrightarrow \operatorname{card}\{n : |x(n)| = 1\} = 1 \\ &\Longleftrightarrow \|\cdot\|_\infty \text{ è F-differenziabile in } x \end{aligned}$$

- Se $X = \ell^\infty$ e $x \in \mathcal{S}_{\ell^\infty}$ allora
$$\|\cdot\|_\infty \text{ è G-differenziabile in } x \Longleftrightarrow \operatorname{card}\{n : |x(n)| = 1\} = 1;$$
e per quanto riguarda la F-differenziabilità
$$\begin{aligned} \|\cdot\|_\infty \text{ è F-differenziabile in } x \Longleftrightarrow \operatorname{card}\{n : |x(n)| = 1\} = 1 \text{ e} \\ \sup_n \{|x(n)| : |x(n)| < 1\} < 1. \end{aligned}$$

- Se $X = \mathcal{C}^0([0,1])$ e $x \in \mathcal{S}_{\mathcal{C}^0([0,1])}$ allora $\|\cdot\|_\infty$ è G-differenziabile in x se e solo se
$$\operatorname{card}(\{r \in [0,1] : x(r) = 1\}) = 1,$$
mentre $\|\cdot\|_\infty$ non è mai F-differenziabile in x

- Se $X = \mathrm{L}^1([0,1])$ e $x \in \mathcal{S}_{\mathrm{L}^1([0,1])}$ allora $\|\cdot\|_1$ è G-differenziabile in x se e solo se
$$\mu(\{r \in [0,1] \,:\, x(r) = 0\}) = 0,$$
mentre $\|\cdot\|_1$ non è mai F-differenziabile in x

- Gli spazi $X = \mathrm{L}^p(\mu)$ con $(1 < p < +\infty)$ sono strettamente convessi (addirittura uniformemente convessi) e F-smooth (addirittura uniformemente F-smooth).

7.5.3 Spazi di Asplund

Vogliamo ora analizzare una classe particolare di spazi che, visti tutti i teoremi fin'ora enunciati, sembra naturale definire. Purtroppo per mancanza di tempo, e spesso per la complessità dei risultati, non dimostreremo nessuno dei teoremi che li riguardano.

Definizione 7.11 (Spazio di Asplund e debolmente di Asplund) Sia X spazio di Banach. Diciamo che

- X è uno *spazio di Asplund* se per ogni $A \subseteq X$ aperto, convesso e non vuoto e ogni funzione $f : A \to \mathbb{R}$ convessa e continua si ha che $N_F(f)$ è di prima categoria secondo Baire.

- X è uno *spazio debolmente di Asplund* o *w-Asplund* se per ogni $A \subseteq X$ aperto, convesso e non vuoto e ogni funzione $f : A \to \mathbb{R}$ convessa e continua si ha che $N_G(f)$ è di prima categoria secondo Baire.

I teoremi di Zajíček e di Preiss–Zajíček affermano rispettvamente che se X è uno spazio di Banach separabile allora è uno spazio debolmente di Asplund, se invece è uno spazio di Banach a duale separabile allora è uno spazio di Asplund. Enunciamo alcuni teoremi su tali spazi.

Teorema 7.13 Sia X uno spazio di Banach. Se

1. X è separabile, allora X è rinormabile in modo che X^* sia strettamente convesso;

2. X è rinormabile in modo che X^* sia strettamente convesso, allora X è rinormabile in modo G-smooth;

3. X è rinormabile in modo G-smooth, allora X è debolmente di Asplund.[13]

Osservazione Non vale nessuna delle implicazioni inverse. Addirittura non vale nemmeno "X è di Asplund, allora X è rinormabile in modo G-smooth". Un controesempio a tale fatto è dovuto ad Haydon.

Teorema 7.14 Sia X spazio di Banach separabile. Sono tra loro equivalenti:

1. X^* è separabile;

2. X è di Asplund;

3. X è rinormabile in modo F-smooth.

[13] Questo è il *teorema di Preiss–Phelps–Namioka*.

Teorema 7.15 Sia X spazio di Banach. Valgono

1. se X è rinormabile in modo F-smooth, allora ogni sottospazio separabile di X ha duale separabile;

2. ogni sottospazio separabile di X ha duale separabile se e solo se ogni sottospazio separabile di X è di Asplund;

3. ogni sottospazio separabile di X è di Asplund se e solo se X è di Asplund.

Osservazione Nella 1. non vale il viceversa, ancora una volta un controesempio a tale fatto è dovuto a Haydon.

Capitolo 8

Dualità

In questo capitolo cercheremo di dare una prima introduzione alla teoria della dualità di insiemi e funzioni convesse.

8.1 Dualità tra insiemi convessi

È, come sempre, indispensabile una definizione propedeutica.

Definizione 8.1 (Polare) Siano X normato, $A \subseteq X$ e $B \subseteq X^*$. Chiamiamo *polare di* A l'insieme

$$A^\circ = \{x^* \in X^* \,:\, x^*(a) \leq 1 \ \forall a \in A\} = \{x^* \in X^* \,:\, \sup x^*(A) = \sigma_A(x^*) \leq 1\}.$$

Chiamiamo *prepolare di* B l'insieme

$$^\circ B = \{x \in X \,:\, b^*(x) \leq 1 \ \forall b^* \in B\} = \{x \in X \,:\, \sup x(B) = s_B(x) \leq 1\}.$$

Osservazione È semplice osservare che

- A° è un convesso, w^*-chiuso a cui appartiene l'origine;
- $^\circ B$ è un convesso, chiuso a cui appartiene l'origine;
- $A^\circ = [\overline{\mathrm{conv}}(A)]^\circ = [\overline{\mathrm{conv}}(A \cup \{0\})]^\circ$;
- $^\circ B = {}^\circ[\overline{\mathrm{conv}}^{w^*}(B \cup \{0\})]$;
- $(\mathcal{B}_X)^\circ = \mathcal{B}_{X^*}$;
- fissato $r > 0$ vale $(r\,\mathcal{B}_X)^\circ = \frac{1}{r}\,\mathcal{B}_{X^*}$;
- se $C_1 \subseteq C_2$ allora $C_2^\circ \subseteq C_1^\circ$.

Enunciamo ora il risultato di dualità tra insiemi convessi.

Teorema 8.1 Sia X spazio vettoriale topologico localmente convesso.

1. Per ogni $A \subseteq X$ vale
$$^\circ(A^\circ) = \overline{\mathrm{conv}}(A \cup \{0\}) \qquad (A^\circ)^\circ = \overline{\mathrm{conv}}^{w^*}(A \cup \{0\}).$$
2. Per ogni $B \subseteq X^*$ vale
$$(^\circ B)^\circ = \overline{\mathrm{conv}}^{w^*}(B \cup \{0\}).$$

3. Definiti

$$\mathcal{F} = \{C \subseteq X \,:\, 0 \in C \text{ e } C \text{ convesso e chiuso}\}$$
$$\mathcal{F}_* = \{D \subseteq X^* \,:\, 0 \in D \text{ e } D \text{ convesso e } w^*\text{-chiuso}\}$$

Le mappe

$$\mathcal{F} \ni C \mapsto C^\circ \in \mathcal{F}_* \tag{a}$$
$$\mathcal{F}_* \ni D \mapsto {}^\circ D \in \mathcal{F} \tag{b}$$

sono l'una inversa dell'altra e definiscono una corrispondenza biunivoca tra $\mathcal{F}$ e $\mathcal{F}_*$.

Dimostrazione

1. Sia $A \subseteq X$, sicuramente $A \subseteq {}^\circ(A^\circ)$ e $0 \in {}^\circ(A^\circ)$. Da ciò segue che

$$\overline{\text{conv}}(A \cup \{0\}) \subseteq {}^\circ(A^\circ).$$

Supponiamo per assurdo che esista $x_0 \in {}^\circ(A^\circ) \smallsetminus \overline{\text{conv}}(A \cup \{0\})$ e applichiamo il teorema di Hahn–Banach. Esiste $y^* \in X^*$ tale che

$$y^*(x_0) > 1 > \sup y^*(\overline{\text{conv}}(A \cup \{0\})).$$

Abbiamo quindi

$$\sup y^*(A) \leq \sup y^*(\overline{\text{conv}}(A \cup \{0\})) < 1,$$

da ciò $y^* \in A^\circ$ e quindi $x_0 \notin {}^\circ(A^\circ)$, il che è assurdo. L'altra uguaglianza è simile.

2. Simile al punto 1.
3. Se $D \in \mathcal{F}_*$ e $C \in \mathcal{F}$ allora $D = ({}^\circ D)^\circ$ e $C = {}^\circ(C^\circ)$, questo dimostra che le due mappe sono suriettive. Se $C_1, C_2 \in \mathcal{F}$ sono tali che $C_1^\circ = C_2^\circ$ allora

$$C_1 = {}^\circ(C_1^\circ) = {}^\circ(C_2^\circ) = C_2,$$

e dunque la mappa (a) è iniettiva. In maniera analoga si dimostra l'iniettività della mappa (b). □

Osservazione Osserviamo che se X è normato e $C \in \mathcal{F}$ allora $0 \in \operatorname{int} C$ se e solo se C° è limitato. Infatti

$$0 \in \operatorname{int} C \Leftrightarrow \exists r > 0 \text{ tale che } r\,\mathcal{B}_X \subseteq C \Leftrightarrow \exists r > 0 \text{ tale che } (r\,\mathcal{B}_X)^\circ \supseteq C^\circ \Leftrightarrow$$
$$\Leftrightarrow \exists r > 0 \text{ tale che } C^\circ \subseteq \frac{1}{r}\mathcal{B}_{X^*} \Leftrightarrow C^\circ \text{ è limitato.}$$

Passiamo ora allo studio delle dualità non più di insiemi convessi, ma di sottospazi.

Definizione 8.2 (Ortogonale) Siano X normato, $A \subseteq X$ e $B \subseteq X^*$. Chiamiamo *ortogonale di* A l'insieme

$$A^\perp = \{x^* \in X^* \,:\, x^*(a) = 0 \;\forall a \in A\}.$$

Chiamiamo *preortogonale di* B l'insieme

$${}^\perp B = \{x \in X \,:\, b^*(x) = 0 \;\forall b^* \in B\}.$$

Osservazione Enunciamo alcune proprietà di base dell'ortogonale.

- Se $A \subseteq X$ allora $A^\perp = \bigcap_{a \in A} \ker(a)$ è un sottospazio w^*-chiuso.
- Se $B \subseteq X^*$ allora ${}^\perp B = \bigcap_{b^* \in B} \ker(b^*)$ è un sottospazio chiuso.
- Valgono $A^\perp = (\overline{\text{span}}(A))^\perp$ e ${}^\perp B = {}^\perp(\overline{\text{span}}^{w^*} B)$.
- Se A è un sottospazio di X e B è un sottospazio di X^* allora $A^\perp = A^\circ$ e ${}^\perp B = {}^\circ B$.

È possibile enunciare un teorema di dualità simile a quello dimostrato per i polari, in questo caso omettiamo la dimostrazione che ricalca la precedente.

Teorema 8.2 Sia X spazio vettoriale topologico localmente convesso.

1. Per ogni $A \subseteq X$ vale
$$^\perp(A^\perp) = \overline{\text{span}} A.$$

2. Per ogni $B \subseteq X^*$ vale
$$(^\perp B)^\perp = \overline{\text{span}}^{w^*} B.$$

3. Definiti
$$\mathcal{L} = \{M \subseteq X \;:\; M \text{ chiuso e lineare}\}$$
$$\mathcal{L}_* = \{N \subseteq X^* \;:\; N\ w^*\text{-chiuso e lineare}\}$$
Le mappe
$$\mathcal{L} \ni M \mapsto M^\perp \in \mathcal{L}_*$$
$$\mathcal{L}_* \ni N \mapsto {}^\perp N \in \mathcal{L}$$
sono l'una inversa dell'altra e definiscono una corrispondenza biunivoca tra $\mathcal{L}$ e $\mathcal{L}_*$.

Osservazione

- Se $M_1, M_2 \in \mathcal{L}$ si ha $M_1 \subseteq M_2$ se e solo se $M_2^\perp \subseteq M_1^\perp$.
- Se $M \in \mathcal{L}$ allora $M^* = \frac{X^*}{M^\perp}$ e $\left(\frac{X}{M}\right)^* = M^\perp$. Da ciò segue che se $\dim M$ e $\operatorname{codim} M$ sono finite allora
$$\dim M = \dim M^* = \dim\left(\frac{X^*}{M^\perp}\right) = \operatorname{codim} M^\perp;$$
$$\operatorname{codim} M = \dim\left(\frac{X}{M}\right) = \dim\left(\frac{X}{M}\right)^* = \dim M^\perp.$$

8.2 Dualità di Fenchel (o di Legendre)

In questa sezione vogliamo esplorare alcune proprietà della dualità tra funzioni convesse. Anzitutto diciamo che una funzione $g : X \to (-\infty, +\infty]$ è *propria* se
$$\text{dom}(g) = \{x \in X \;:\; g(x) < +\infty\}$$
è non vuoto.

Definizione 8.3 (Funzione coniugata) Siano X spazio vettoriale topologico localmente convesso, $g : X \to (-\infty, +\infty]$ e $\psi : X^* \to (-\infty, +\infty]$. Chiamiamo *funzione coniugata di* g la funzione $g^* : X^* \to (-\infty, +\infty]$, definita come

$$g^*(x^*) = \sup_{x \in X} \{x^*(x) - g(x)\}.$$

Chiamiamo *funzione preconiugata di* ψ la funzione ${}^*\psi(x) : X \to (-\infty, +\infty]$, definita come

$${}^*\psi(x) = \sup_{x^* \in X^*} \{x^*(x) - \psi(x^*)\}.$$

Osservazione Alcune osservazioni di base sono d'obbligo.

- g^* è convessa e w^*-inferiormente semicontinua, mentre ${}^*\psi$ è convessa e inferiormente semicontinua.
- g^* è propria se e solo se g è minorata da una funzione affine continua. Infatti g^* è propria se e solo se esiste $x^* \in X^*$ e $\alpha \in \mathbb{R}$ tali che $x^* - \alpha \leq g$ su X. In particolare se g è convessa e inferiormente semicontinua allora g^* è propria.
- ${}^*\psi$ è propria se e solo se ψ è minorata da una funzione affine w^*-continua. In particolare se ψ è convessa w^*-inferiormente semicontinua allora ${}^*\psi$ è propria.
- Vale

$$g^*(x^*) = \sup \{x^*(x) - t \,:\, x \in X \; g(x) \leq t\} =$$
$$= \sup \{\langle (x^*, -1), (x, t) \rangle \,:\, (x, t) \in \mathrm{epi}(g)\} = \sigma_{\mathrm{epi}(g)}((x^*, -1)).$$

Definizione 8.4 (Involucro convesso/chiuso/convesso chiuso) Siano X spazio vettoriale topologico localmente convesso e $g : X \to (-\infty, +\infty]$. Chiamiamo

- *involucro convesso di* g la funzione $\mathrm{co}(g)$ che soddisfa

$$\mathrm{epi}(\mathrm{co}(g)) = \mathrm{conv}(\mathrm{epi}(g));$$

- *involucro chiuso di* g la funzione $\mathrm{cl}(g)$ che soddisfa

$$\mathrm{epi}(\mathrm{cl}(g)) = \overline{\mathrm{epi}(g)};$$

- *involucro convesso chiuso di* g la funzione $\mathrm{clco}(g)$ che soddisfa

$$\mathrm{epi}(\mathrm{clco}(g)) = \overline{\mathrm{conv}}(\mathrm{epi}(g));$$

Osservazione È possibile dimostrare che

- $\mathrm{co}(g) = \sup \{h \,:\, h \leq g,\; h \text{ convessa}\}$;
- $\mathrm{co}(g)(x) = \inf \left\{\sum_{i=1}^n \lambda_i g(x_i) \,:\, x = \sum_{i=1}^n \lambda_i x_i \text{ combinazione convessa}\right\}$;
- $\mathrm{cl}(g) = \sup \{h \,:\, h \leq g,\; h \text{ inferiormente semicontinua}\}$;
- $(\mathrm{cl}(g))(x) = \liminf_{y \to x} g(y)$;
- $\mathrm{clco}(g) = \sup \{h \,:\, h \leq g,\; h \text{ convessa e inferiormente semicontinua}\}$.

Veniamo ora al nostro teorema di dualità.

Teorema 8.3 Siano X spazio vettoriale topologico localmente convesso, $g : X \to (-\infty, +\infty]$ e $\psi : X^* \to (-\infty, +\infty]$. Se ψ e g sono proprie allora

1. ${}^*(g^*) = \mathrm{clco}(g)$;
2. $({}^*\psi)^* = w^*\text{-}\mathrm{clco}(\psi)$;
3. $(g^*)^* = w^*\text{-}\mathrm{clco}(g)$.

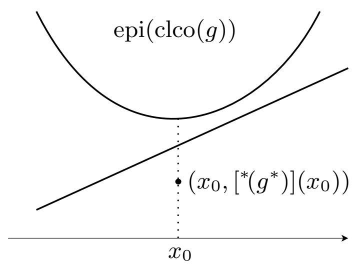
Dimostrazione Dimostreremo solo la 1. le altre sono simili. Una maggiorazione è semplice, infatti

$$\begin{aligned}[{}^*(g^*)](x) &= \sup_{x^*\in X^*}\{x^*(x) - g^*(x^*)\} = \sup_{x^*\in X^*}\left\{x^*(x) - \sup_{y\in X}\{x^*(y) - g(y)\}\right\} \leq \\ &\leq \sup_{x^*\in X^*}\{x^*(x) - x^*(x) + g(x)\} = \sup_{x^*\in X^*}\{g(x)\} = g(x).\end{aligned}$$

Da cui ${}^*(g^*) \leq \mathrm{clco}(g)$ siccome è convessa e inferiormente semicontinua per l'osservazione precedente. Vogliamo l'uguaglianza. Per assurdo supponiamo che esista $x_0 \in X$ tale che

$$[{}^*(g^*)](x_0) < (\mathrm{clco}(g))(x_0).$$

Il disegno aiuta a comprendere ciò che vogliamo fare.



Per il teorema di Hahn–Banach esiste $x_0^* \in X^*$ e $\alpha \in \mathbb{R}$ tali che

$$[{}^*(g^*)](x_0) < x_0^*(x_0) + \alpha \qquad x_0^*(x) + \alpha < g(x) \qquad \forall x \in X$$

Passando all'estremo superiore si ottiene

$$\sup_{x\in X}\{x_0^*(x) - g(x)\} \leq -\alpha \Longrightarrow g^*(x_0^*) \leq -\alpha.$$

Quindi

$$g^*(x_0^*) \leq -\alpha < x_0^*(x_0) - [{}^*(g^*)](x_0) \leq x_0^*(x_0) - x_0^*(x_0) + g^*(x_0^*).$$

Ma ciò è assurdo. □

Corollario Sia X spazio vettoriale topologico localmente convesso. Definite

$$\begin{aligned}\overline{\mathcal{C}}(X) &= \{f : X \to (-\infty, +\infty] \,:\, f \text{ convessa, lsc e propria}\},\\ \overline{\mathcal{C}}^*(X^*) &= \{\varphi : X^* \to (-\infty, +\infty] \,:\, \varphi \text{ convessa, } w^*\text{-lsc e propria}\},\end{aligned}$$

la *dualità di Fenchel* (o *di Legendre*), ovvero le mappe

$$\begin{aligned}\overline{\mathcal{C}}(X) \ni f \mapsto f^* \in \overline{\mathcal{C}}^*(X^*)\\ \overline{\mathcal{C}}^*(X^*) \ni \varphi \mapsto {}^*\varphi \in \overline{\mathcal{C}}(X),\end{aligned}$$

definisce una corrispondenza biunivoca tra $\overline{\mathcal{C}}(X)$ e $\overline{\mathcal{C}}(X^*)$.

Osservazione Vediamo alcune proprietà di base della dualità di Fenchel, siano $f, g \in \overline{\mathcal{C}}(X)$.

- Se X è riflessivo allora la dualità di Fenchel fornisce una corrispondenza biunivoca tra $\overline{\mathcal{C}}(X)$ e $\overline{\mathcal{C}}(X^*)$.
- Se $f \leq g$ allora $g^* \leq f^*$.
- Dalla definizione di funzione coninugata si ottiene la famosa *disuguaglianza di Young*, ovvero

$$x^*(x) \leq f(x) + f^*(x^*) \qquad \forall x \in X \ \forall x^* \in X^*.$$

- Se definiamo

$$\partial f(x) = \begin{cases} \emptyset & x \notin \mathrm{dom}(f) \\ \{x^* \in X^* \ : \ f(y) \geq f(x) + x^*(y-x) \ \forall y \in X\} & x \in \mathrm{dom}(f) \end{cases}$$

otteniamo che $f^*(x^*) + f(x) = x^*(x)$ se e solo se $x^* \in \partial f(x)$, e ciò accade se e solo se $x \in \partial f^*(x^*)$.

Dimostrazione L'uguaglianza vale se $x^* \in \mathrm{dom}(f^*)$ e $x \in \mathrm{dom}(f)$. Si ha

$$\begin{aligned} f^*(x^*) = x^*(x) - f(x) \Leftrightarrow \forall y \in X \quad x^*(y) - f(y) \leq x^*(x) - f(x) \Leftrightarrow \\ \Leftrightarrow \forall y \in X \quad f(x) + x^*(y-x) \leq f(y) \Leftrightarrow x^* \in \partial f(x). \end{aligned}$$

L'altra equivalenza discende dal fatto che $f(x) = [{}^*(f^*)](x)$. □

- $f^*(0) = -\inf_{x \in X} \{f(x)\}$.
- Esiste $x_0 \in X$ tale che $f(x_0) = \inf \{f(X)\}$ se e solo se $0 \in \partial f(x_0)$, il che è equivalente a $x_0 \in \partial f^*(0)$. Quindi f assume minimo su X se e solo se $0 \in \mathrm{dom}(f^*)$ e $\partial f^*(0) \cap X \neq \emptyset$.[1]
- Sia $D \subseteq X^*$ convesso e w^*-chiuso. Vogliamo calcolare $s_D^*(x^*)$. Ma

$$s_D^*(x^*) = \sup_{x \in X} \{x^* - s_D(x)\} = \delta_D(x^*),$$

in particolare $\|\cdot\|^* = \delta_{\mathcal{B}_{X^*}}$.
- Sia $C \subseteq X$ convesso chiuso. Si ha

$$\delta_C^*(x^*) = \sup_{x \in X} \{x^*(x) - \delta_C(x)\} = \sup_{x \in C} \{x^*(x)\} = \sigma_C(x^*).$$

- È possibile dimostrare che se $p > 1$ allora

$$\left(\frac{1}{p}\|\cdot\|^p\right)^*(x^*) = \frac{1}{q}\|x^*\|^q$$

con $\frac{1}{p} + \frac{1}{q} = 1$.

Esempio Se $p > 1$ e $x > 0$, vogliamo calcolare esplicitamente la coniugata della funzione $f(x) = \frac{1}{p}x^p$. Per definizione

$$f^*(y) = \sup_{x > 0} \left(xy - \frac{x^p}{p}\right).$$

Deriviamo la funzione tra parentesi e poniamola uguale a zero

$$y - \frac{px^{p-1}}{p} = 0 \Longrightarrow y = x^{p-1}.$$

[1] Se X è riflessivo è sufficiente richiedere che $0 \in \mathrm{dom}(f^*)$ e $\partial f^*(0) \neq \emptyset$.

Osservando che $x = y^{\frac{1}{p-1}}$ e sostituendo per $y > 0$

$$f^*(y) = y^{\frac{p}{p-1}} - \frac{y^{\frac{p}{p-1}}}{p} = \left(1 - \frac{1}{p}\right) y^{\frac{p}{p-1}} = \frac{y^q}{q},$$

con $\frac{1}{p} + \frac{1}{q} = 1$. E la disuguaglianza di Young risulta essere

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \qquad \forall x, y \geq 0.$$

8.3 Convoluzione infimale

L'idea di questa sezione è quella di capire che forma ha la coniugata di una somma di funzioni, per fare ciò abbiamo bisogno di una nuova operazione tra funzioni che permette di ottenere nuove funzioni convesse.

Definizione 8.5 (Convoluzione infimale) Siano X spazio vettoriale e $f, g : X \to (-\infty, +\infty]$ convesse. Chiamiamo *convoluzione infimale di f e g* la funzione $f \,\Box\, g : X \to [-\infty, +\infty]$ definita come segue

$$(f \,\Box\, g)(x) = \inf \{f(y) + g(x - y) \,:\, y \in X\}.$$

Osservazione È semplice osservare che

$$(f \,\Box\, g)(x) = \inf \{f(x_1) + g(x_2) \,:\, x_1 + x_2 = x \in X\} = (g \,\Box\, f)(x).$$

Vogliamo dimostrare che se f e g sono convesse allora anche $f \,\Box\, g$ è convessa. Chiamiamo *epigrafico stretto di f* l'insieme

$$\mathrm{epi_s}(f) = \{(x, t) \in X \times \mathbb{R} \,:\, f(x) < t\}.$$

È semplice osservare che f è convessa se e solo se $\mathrm{epi_s}(f)$ è convesso. Ricordiamo infine che se $t, a, b \in \mathbb{R}$ sono tali che $t > a + b$ allora esistono $t_1, t_2 \in \mathbb{R}$ tali che

$$t = t_1 + t_2 \qquad t_1 > a \qquad t_2 > b$$

Lemma Siano X spazio vettoriale e $f, g : X \to (-\infty, +\infty]$. Si ha

$$\mathrm{epi_s}(f \,\Box\, g) = \mathrm{epi_s}(f) + \mathrm{epi_s}(g).$$

In particolare se f e g sono convesse, allora $f \,\Box\, g$ è convessa.

Dimostrazione Sia $(x, t) \in \mathrm{epi_s}(f \,\Box\, g)$, ovvero

$$t > (f \,\Box\, g)(x) = \inf \{f(x_1) + g(x_2) \,:\, x_1 + x_2 = x \in X\},$$

ma allora esistono $x_1, x_2 \in X$ tali che $x = x_1 + x_2$ e $t > f(x_1) + g(x_2)$. Per quanto detto in precedenza esistono $t_1, t_2 \in \mathbb{R}$ tali che $t = t_1 + t_2$ e

$$t_1 > f(x_1) \qquad t_2 > g(x_2).$$

Rileggendo ciò che abbiamo fatto si ottiene che esistono $(x_1, t_1) \in \mathrm{epi_s}(f)$ e $(x_2, t_2) \in \mathrm{epi_s}(g)$ tali che

$$(x, t) = (x_1, t_1) + (x_2, t_2).$$

Ma ciò è la tesi. $\square$

Osservazione

- Si dimostra che $\delta_C \,\square\, \delta_D = \delta_{C+D}$.
- Se $f, g \in \overline{\mathcal{C}}(D)$ non è detto che $f \,\square\, g$ sia inferiormente semicontinua.
- La convoluzione infimale può essere identicamente uguale a $-\infty$. Per esempio se $X = \mathbb{R}$, $f(x) = x$ e $g(x) = -x$ allora

$$(f \,\square\, g)(x) = \inf\{y + y - x \,:\, y \in \mathbb{R}\} = -\infty \qquad \forall x \in \mathbb{R}.$$

- $(f \,\square\, g)^*(x^*) = f^*(x^*) + g^*(x^*)$.

  Dimostrazione

$$(f \,\square\, g)^*(x^*) = \sup_{x \in X} \{x^*(x) - \inf\{f(x_1) + g(x_2) \,:\, x_1 + x_2 = x \in X\}\} =$$
$$= \sup_{x \in X} \sup_{x_1 + x_2 = x} \{x^*(x_1) + x^*(x_2) - f(x_1) - g(x_2)\} = f^*(x^*) + g^*(x^*). \qquad \square$$

È naturale chiedersi sotto quali ipotesi è possibile riottenere che la convoluzione infimale sia inferiormente semicontinua e soprattutto se è vero che la coniugata di una somma è la convoluzione infimale delle coniugate. I seguenti teoremi ci forniscono una risposta.

Teorema 8.4 Sia X spazio di Banach riflessivo e $f, g \in \overline{\mathcal{C}}(X)$. Se f è coercitiva e g è limitata inferiormente, allora $f \,\square\, g \in \overline{\mathcal{C}}(X)$.

Teorema 8.5 (di Moreau–Rockafellar) Siano X spazio vettoriale topologico localmente convesso e $f, g \in \overline{\mathcal{C}}(X)$. Se esiste $x_0 \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$ tale che in x_0 una tra f e g sia continua, allora

$$(f + g)^*(x^*) = (f^* \,\square\, g^*)(x^*).$$

Il prossimo teorema ci dice come è fatta la coniugata dell'estremo superiore di una famiglia di funzioni convesse, inferiormente semicontinue e proprie.

Teorema 8.6 Siano X spazio vettoriale topologico localmente convesso. Consideriamo $\{f_j\}_{j \in J} \subseteq \overline{\mathcal{C}}(X)$, se poniamo

$$g = \sup_{j \in J} f_j \in \overline{\mathcal{C}}(X),$$

allora

$$g^*(x^*) = \left(w^* - \mathrm{clco}\left(\inf_{j \in J} f_j^*\right)\right)(x^*).$$

8.4 Calcolo del subdifferenziale

Ci interessa, in questa sezione, trovare una formula pratica per calcolare il subdifferenziale di una somma di funzioni convesse. Iniziamo con delle semplici osservazioni.

Osservazione

- È banale dimostrare che $\partial(f + x^*)(x) = \partial f(x) + x^*$.

- Se $f \in \overline{\mathcal{C}}(X)$ allora
$$\mathrm{dom}(f) = \bigcup_{n=1}^{+\infty} \{x \in X \,:\, f(x) \leq n\}$$
è un F_σ convesso. Quindi se X è uno spazio di Banach abbiamo che, per il teorema 1.7, $\mathrm{int}(\mathrm{dom}(f)) = \text{a-int}(\mathrm{dom}(f))$ e f è continua su tale insieme.
- Si può dimostrare che, detto *dominio effettivo* l'insieme
$$D(\partial f) = \{x \in \mathrm{dom}(f) \,:\, \partial f(x) \neq \emptyset\},$$
allora
$$D(\partial f) = \bigcup_{n\in\mathbb{N}} \{x \in \mathrm{dom}(f) \,:\, \partial f(x) \cap n\,\mathcal{B}_{X^*} \neq \emptyset\}$$
è un F_σ.[2]

Veniamo al teorema promesso.

Teorema 8.7 (di Moreau–Rockafellar "bis") Siano X spazio vettoriale topologico localmente convesso e $f, g \in \overline{\mathcal{C}}(X)$. Se esiste $x_0 \in \mathrm{dom}(f) \cap \mathrm{dom}(g)$ in cui una tra f e g sia continua, allora
$$\partial(f+g)(x) = \partial f(x) + \partial g(x) \qquad \forall x \in X.$$

Dimostrazione Dimostriamo le due inclusioni.

($\supseteq$) Sia $x^* \in \partial f(x)$ e $y^* \in \partial g(x)$, per definizione di subdifferenziale si ha per ogni $z \in X$
$$f(z) \geq x^*(z-x) + f(x) \qquad g(z) \geq y^*(z-x) + g(x)$$
Sommando si ottiene $(f+g)(z) \geq (x^*+y^*)(z-x) + (f+g)(x)$, ovvero $x^*+y^* \in \partial(f+g)(x)$.

($\subseteq$) Sia $u^* \in \partial(f+g)(x)$, per definizione
$$f(z) - f(x) \geq u^*(z-x) + g(x) - g(z) \qquad \forall z \in X.$$
Poste $\varphi(z) = f(z) - f(x)$ e $\psi(z) = u^*(z-x) + g(x) - g(z)$ si ha che la prima è convessa e inferiormente semicontinua, mentre la seconda è concava e superiormente semicontinua, inoltre $\varphi(x) = 0 = \psi(x)$. L'epigrafico di φ ha interno non vuoto, per continuità di f in x_0, e utilizziamo il teorema di Hahn–Banach per separare, ovvero esiste $x^* \in X^*$ tale che per ogni $z \in X$ si ha
$$f(z) - f(x) \geq x^*(z-x) \geq u^*(z-x) + g(x) - g(z),$$
ovvero $x^* \in \partial f(x)$ e $u^* - x^* \in \partial g(x)$. Combinando tutto ciò che abbiamo ottenuto si ha
$$u^* = x^* + u^* - x^* \in \partial f(x) + \partial g(x). \qquad \square$$

Osservazione Dalla dimostrazione si evince che in ogni caso (nel senso senza ipotesi aggiuntive) $\partial f(x) + \partial g(x) \subseteq \partial(f+g)(x)$.

Veniamo ora ad enunciare, senza dimostrazione, alcuni teoremi che permettono il calcolo del subdifferenziale di una funzione composta e di una funzione definita come massimo.

[2] Il fatto che gli insiemi dell'unione siano chiusi è un utile esercizio che fa uso delle net.

Teorema 8.8 (di Dubovitskii–Milyutin) Siano X spazio vettoriale topologico localmente convesso, $A \subseteq X$ aperto convesso, $n \in \mathbb{N}$ e

$$f_i : A \longrightarrow \mathbb{R} \qquad i = 1, \ldots, n$$

convesse e continue. Definita puntualmente $g = \max\{f_1, \ldots, f_n\}$, allora per ogni $x \in A$

$$\partial g(x) = \operatorname{conv}\left(\bigcup_{i \in I(x)} \partial f_i(x) \right),$$

dove $I(x) = \{i : f_i(x) = g(x)\}$ ovvero l'insieme degli indici attivi.

Teorema 8.9 (Generalizzazione del teorema di Dubovitskii–Milyutin) Siano X spazio vettoriale topologico localmente convesso, $A \subseteq X$ aperto convesso e

$$f_j : A \longrightarrow \mathbb{R} \qquad j \in J$$

convesse e continue. Se supponiamo inoltre che

- J sia uno spazio topologico compatto;
- per ogni $x \in X$ la funzione $j \mapsto f_j(x)$ è superiormente semicontinua;
- per ogni $x \in X$ si ha $g(x) = \sup_{j \in J} f_j(x) < +\infty$,

allora per ogni $x \in A$

$$\partial g(x) = \overline{\operatorname{conv}}^{w^*}\left(\bigcup_{j \in J(x)} \partial f_j(x) \right),$$

dove $J(x)$ è l'insieme degli indici attivi. Inoltre se $\dim X < +\infty$ o J è finito, allora

$$\partial g(x) = \operatorname{conv}\left(\bigcup_{j \in J(x)} \partial f_j(x) \right).$$

Teorema 8.10 (Subdifferenziale della mappa composta) Siano X spazio vettoriale topologico localmente convesso, $A \subseteq X$ aperto convesso, $I \subseteq \mathbb{R}$ intervallo aperto, $f : A \to I$ convessa e continua e $\varphi : I \to \mathbb{R}$ convessa e non decrescente. La funzione

$$g = \varphi \circ f$$

è convessa e continua su A e per ogni $x \in A$ vale

$$\partial g(x) = \partial\varphi(f(x)) \cdot \partial f(x).$$

Capitolo 9

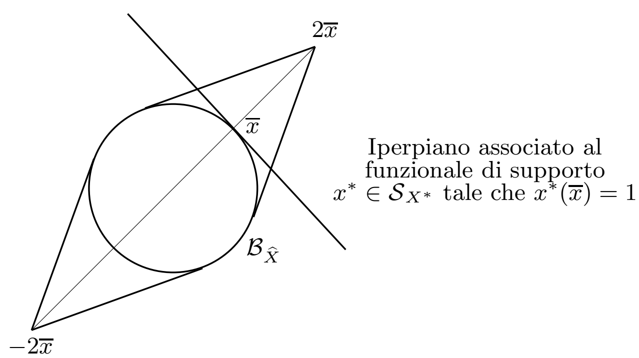
Teoremi di densità: Bishop–Phelps

Se X è uno spazio normato e $C \subseteq X$ convesso, chiuso e limitato, sappiamo che se C è w-compatto allora ogni $x^* \in X^*$ assume massimo su C. Sappiamo anche che il viceversa è vero se e solo se X è uno spazio di Banach riflessivo, ciò discende dal teorema di James. Ma se C non è w-compatto? Quanti funzionali lineari assumono il loro massimo su C? A questa domanda risponde il teorema di Bishop–Phelps che ci accingiamo a dimostrare. Ci sarà utile il seguente insieme

$$\Sigma(C) = \{x^* \in X^* \,:\, x^* \text{ assume il suo massimo su } C\},$$

il teorema di Bishop–Phelps afferma che se X è di Banach allora $\overline{\Sigma(C)} = X^*$.

Esempio Nel caso in cui lo spazio sia incompleto è possibile sempre esibire un insieme C chiuso, convesso e limitato tale che $\overline{\Sigma(C)} \neq X^*$. L'idea per costruire tale insieme è la seguente: sia X spazio normato incompleto e $\widehat{X}$ il suo completato.[1] Sia $\overline{x} \in \mathcal{S}_{\widehat{X}} \setminus \mathcal{S}_X$ e consideriamo gli oggetti in figura



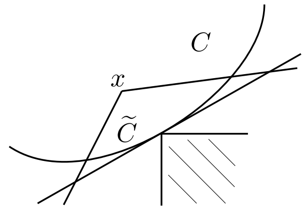
Consideriamo la bolla di una norma equivalente

$$C = \overline{\mathrm{conv}}_{\widehat{X}}(\mathcal{B}_{\widehat{X}} \cup \{\pm 2\overline{x}\}) \cap X.$$

Si può mostrare facilmente che x^* assume massimo su C solo in $2\overline{x}$, ma $2\overline{x} \notin X$. Quindi x^* non assume massimo su C. Lo stesso vale per ogni funzionale lineare y^* sufficientemente vicino a x^*.

[1] Ricordiamo che $X^* = (\widehat{X})^*$.

A differenza del teorema di James la dimostrazione del teorema di Bishop–Phelps parte da una idea geometrica semplice. Vediamo il seguente disegno



C è un chiuso, convesso e non vuoto. Vogliamo tanti funzionali che assumono la norma e per mostrarlo cercheremo di mostrare che ci sono tanti semispazi che "toccano" C. Osservando che "toccare C con un cosa dritta" è equivalente a "toccare C con una punta" (con interno non vuoto, per poi utilizzare il teorema di Hahn–Banach). La dimostrazione procederà poi nel seguente modo: fissiamo un funzionale limitato superiormente su C a distanza 0,[2] e x sufficientemente vicino all'iperpiano. Consideriamo ora un cono "molto aperto" con vertice x e la cui direzione è data dal funzionale fissato in partenza, trovando $C_1 = \widetilde{C} \cap C$. Ripetiamo il ragionamento costruendo gli insiemi C_n, usiamo la completezza dello spazio per ottenere un punto in cui il cono tocca da fuori. I coni che considereremo sono totalmente definiti da un funzionale lineare, e se il cono è sufficientemente ampio abbiamo appena ottenuto un funzionale lineare molto vicino a quello da cui eravamo partiti. Passiamo ora alla dimostrazione formale.

9.1 Le sue impure origini

Per dimostrare il teorema di Bishop–Phelps abbiamo bisogno di un congruo numero di lemmi preliminari. Iniziamo con un risultato classico.

Lemma 8 Sia (X, d) spazio metrico. Sono tra loro equivalenti:

1. X è completo;
2. siano $E_1 \supseteq E_2 \subseteq E_3 \supseteq \cdots$ una successione di chiusi non vuoti di X, se $\mathrm{diam}(E_n) \to 0$ allora
$$\bigcap_{n=1}^{+\infty} E_n \neq \emptyset.$$

Dimostrazione

(1. ⇒ 2.) Per ogni n scegliamo $x_n \in E_n$. Tale successione è di Cauchy e per completezza esiste $x_0 \in X$ tale che
$$x_n \longrightarrow x_0.$$
Per chiusura e inscatolamento $x_0 \in E_n$ per ogni n e dunque $x_0 \in \bigcap_{n=1}^{+\infty} E_n$.[3]

(2. ⇒ 1.) Sia $\{x_n\} \subseteq X$ una successione di Cauchy. Poniamo
$$E_n = \overline{\{x_i \,:\, i \geq n\}},$$
tali insiemi sono chiusi e inscatolati con diametro che tende a zero. Per ipotesi esiste $x_0 \in \bigcap_{n=1}^{+\infty} E_n$, è poi semplice verificare che $x_n \to x_0$. □

[2] Potrebbe anche non toccare C.

[3] Grazie alla richiesta fatta sul comportamento del diametro, necessariamente se l'intersezione è non vuota, allora contiene uno ed un solo punto.

D'ora in poi lavoreremo solo con funzionali di norma unitaria, infatti se un funzionale con tali caratteristiche assume massimo su un insieme lo stesso faranno tutti i suoi multipli positivi. Ricordiamo che per $E \subseteq X$, $x^* \in \mathcal{S}_{X^*}$ e $\alpha \in (0,1)$ denotiamo

$$\begin{aligned}
&\sigma_E(x^*) = \sup x^*(E);\\
&\Omega_1(E) = \{z^* \in \mathcal{S}_{X^*} : \sigma_E(z^*) < +\infty\};\\
&C(x^*, \alpha) = \{x \in X : x^*(x) \geq \alpha\|x\|\}.
\end{aligned}$$

È semplice verificare che $\operatorname{int} C(z^*, \alpha) = \{x \in X : x^*(x) > \alpha\|x\|\}$ e che se $x \in C(x^*, \alpha)$ allora $x + C(x^*, \alpha) \subseteq C(x^*, \alpha)$. Veniamo ora ad enunciare una proposizione di base che ci verrà utile per fare il gioco con i coni precedentemente annunciato.

Proposizione 9.1 Siano X spazio di Banach, $E \subseteq X$ chiuso e non vuoto, $x^* \in \Omega_1(E)$, $x \in E$ e $\varepsilon > 0$. Se

$$x^*(x) > \sigma_E(x^*) - \varepsilon,$$

allora per ogni $\alpha \in (0,1)$ esiste $z \in E$ tale che

- $E \cap (z + C(x^*, \alpha)) = \{z\}$;
- $\|z - x\| < \frac{\varepsilon}{\alpha}$.

Dimostrazione Fissiamo una successione $\{\varepsilon_n\}_{n=1}^{+\infty} \subseteq (0, +\infty)$ tendente a zero. Denotiamo $E_0 := E$, $x_0 := x$, $\varepsilon_0 := \varepsilon$ e definiamo induttivamente E_n e $x_n \in E_n$ come segue: poniamo

$$E_n = E_{n-1} \cap [x_{n-1} + C(x^*, \alpha)]$$

e scegliamo

$$x^*(x_n) > \sigma_{E_n}(x^*) - \varepsilon_n.$$

Osserviamo che $E_0 \supseteq E_1 \supseteq E_2 \supseteq \cdots$ sono insiemi chiusi, $x_n \in E_n$ e, per definizione, $x_{n-1} \in E_n$. Se $y \in E_n$ allora $y - x_{n-1} \in C(x^*, \alpha)$ e di conseguenza

$$\alpha\|y - x_{n-1}\| \leq x^*(y - x_{n-1}) = x^*(y) - x^*(x_{n-1}) \leq \sigma_{E_{n-1}}(x^*) - x^*(x_{n-1}) < \varepsilon_{n-1}.$$

Ma ciò implica che $\operatorname{diam}(E_n) \leq \frac{2\varepsilon_{n-1}}{\alpha} \to 0$. Per la completezza di X, utilizzando il lemma 8, si ottiene $\bigcap_{n=0}^{+\infty} E_n = \{z\}$. Elenchiamo ora le proprietà di z:

- $z \in E$;
- $z \in E_1$ e quindi $\|z - x\| = \|z - x_0\| < \frac{\varepsilon_0}{\alpha} = \frac{\varepsilon}{\alpha}$;
- $z \in x_{n-1} + C(x^*, \alpha)$, ovvero esiste $x_{z,x_{n-1}} \in C(x^*, \alpha)$ tale che $z = x_{n-1} + x_{z,x_{n-1}}$. Ma se $\widetilde{x} \in z + C(x^*, \alpha)$ allora esiste $\overline{x} \in C(x^*, \alpha)$ tale che

  $$\widetilde{x} = z + \overline{x} = x_{n-1} + x_{z,x_{n-1}} + \overline{x} \in x_{n-1} + C(x^*, \alpha),$$

  ove la conclusione è dovuta alla convessità del cono $C(x^*, \alpha)$. Dunque

  $$[z + C(x^*, \alpha)] \subseteq [x_{n-1} + C(x^*, \alpha)];$$

- vale $E \cap [z + C(x^*, \alpha)] \subseteq E_n$;[4]

[4] Questo è semplice, infatti sappiamo che $E_{n-1} \cap (z + C(x^*, \alpha)) \subseteq E_n$ e poi si ragiona a ristroso ricordandosi la definizione di E_{n-1} e il punto precedente.

- $\operatorname{diam}(E_n) \to 0$ da cui $\operatorname{diam}(E \cap [z + C(x^*, \alpha)]) = 0$ e quindi
$$E \cap [z + C(x^*, \alpha)] = \{z\}. \qquad \square$$

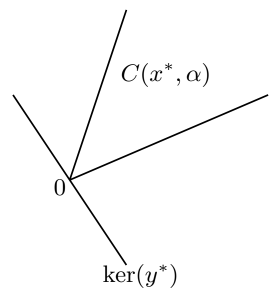
Abbiamo infine bisogno di un lemma tecnico per poter poi enunciare il teorema di Bishop–Phelps.

Lemma 9 Siano X normato, $x^*, y^* \in \mathcal{S}_{X^*}$ e $\alpha \in \left(0, \frac{1}{2}\right)$. Valgono le seguenti implicazioni

1. se $\|x^* - y^*\| \leq \alpha$ allora $\inf y^*(C(x^*, \alpha)) = 0$;
2. se $\inf y^*(C(x^*, \alpha)) = 0$ allora $\|x^* - y^*\| \leq 2\alpha$.

Dimostrazione

1. Questa implicazione vale addirittura per $\alpha \in (0,1)$ ed è equivalente a richiedere che $y^* \geq 0$ su $C(x^*, \alpha)$; possiamo rappresentare tale situazione con il seguente disegno.



   Sia $x \in C(x^*, \alpha)$ vale
$$y^*(x) = x^*(x) - (x^* - y^*)(x) \geq x^*(x) - \|x^* - y^*\|\|x\| \geq x^*(x) - \alpha\|x\|.$$
   Dal fatto che $x \in C(x^*, \alpha)$ allora $x^*(x) \geq \alpha\|x\|$ e quindi $y^*(x) \geq 0$.

2. Sia $Y = \ker(y^*)$, utilizzando l'ipotesi si ottiene
$$Y \cap \operatorname{int}(C(x^*, \alpha)) = \emptyset.$$
   Sia $y \in Y$, per quanto appena detto $x^*(y) \leq \alpha\|y\|$ e dunque $\|x^*\|_{|_Y} \leq \alpha$. Possiamo estendere x^* ad un funzionale $\widetilde{x}^*$ su tutto X tale che
$$\|\widetilde{x}^*\| \leq \alpha.$$
   Osservando che $x^* - \widetilde{x}^*$ si annulla su Y, ovvero $\ker(y^*) \subseteq \ker(x^* - \widetilde{x}^*)$ e, da un semplice fatto di algebra lineare, segue che esiste $c \in \mathbb{R}$ tale che[5]
$$cy^* = (x^* - \widetilde{x}^*).$$
   Proviamo a stimare la norma si $\widetilde{x}^*$,
$$\|\widetilde{x}^*\| = \|x^* - (x^* - \widetilde{x}^*)\| \geq \|x^*\| - \|x^* - \widetilde{x}^*\| = \|x^*\| - \|cy^*\| = 1 - |c|.$$
   Dobbiamo ora analizzare due casi
   - Se $c \geq 0$ allora $|1 - c| \leq \alpha$. La tesi segue da un semplice conto
$$\|x^* - y^*\| = \|\widetilde{x}^* + cy^* - y^*\| \leq \|\widetilde{x}^*\| + |1 - c|\|y^*\| \leq \alpha + |1 - c| \leq 2\alpha.$$

[5] Ovviamente dal fatto che $\ker(y^*)$ è un iperpiano, segue che $\ker(x^* - \widetilde{x}^*)$ è un iperpiano oppure tutto lo spazio.

- Se $c < 0$ allora $|1 + c| \leq \alpha$. Vogliamo mostrare che tale caso risulta impossibile. Con un conto simile al precedente

  $$\|x^* + y^*\| \leq 2\alpha.$$

  Ma dal fatto che $\alpha < \frac{1}{2}$ segue $2\alpha < 1 = \|x^*\|$. Esiste $x_1 \in \mathcal{S}_X$ tale che

  $$\|x^*(x_1)\| > 2\alpha > \alpha > \alpha\|x_1\|$$

  e dunque $x_1 \in C(x^*, \alpha)$. Ma allora

  $$2\alpha \geq \|x^* + y^*\| \geq (x^* + y^*)(x_1) > 2\alpha + \underbrace{y^*(x_1)}_{\geq 0} \geq 2\alpha$$

  il che è assurdo. Quindi il caso $c < 0$ non può accadere. $\square$

9.2 La sua scandalosa carriera

Finalmente possiamo enunciare il teorema cosí a lungo atteso.

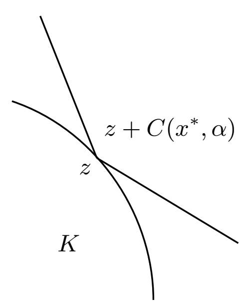
Teorema 9.1 (di Bishop–Phelps) Siano X spazio di Banach, $K \subseteq X$ convesso, chiuso e non vuoto, $\varepsilon > 0$ e $\alpha \in \left(0, \frac{1}{2}\right)$. Se $x^* \in \Omega_1(K)$ e $x \in K$ sono tali che

$$x^*(x) > \sigma_K(x^*) - \varepsilon,$$

allora esistono $z \in K$ e $z^* \in \Omega_1(K)$ tali che

- $z^*(z) = \sup z^*(K) = \sigma_K(z^*)$;
- $\|z - x\| < \frac{\varepsilon}{\alpha}$;
- $\|z^* - x^*\| \leq 2\alpha$.

Dimostrazione Applichiamo la proposizione 9.1 e otteniamo $z \in K$ come in figura tale che $\|z - x\| < \frac{\varepsilon}{\alpha}$.



Utilizziamo il teorema di Hahn–Banach e separiamo $\operatorname{int} C(x^*, \alpha)$ da K con un funzionale z^*. Necessariamente otteniamo che $z^*(z) = \sigma_K(z^*)$. Inoltre dal fatto che

$$z^*(z + C(x^*, \alpha)) \geq \sigma_K(z^*),$$

si ottiene che $z^*_{|C(x^*,\alpha)} \geq 0$. Applicando il lemma 9 abbiamo la tesi. $\square$

Enunciamo ora il seguente corollario che costuisce la prima e più naturale applicazione del teorema di Bishop–Phelps.

Corollario Siano X spazio di Banach e $K \subseteq X$ convesso, chiuso e non vuoto. Definiti

- $\operatorname{supp}(K) = \{x \in K : \exists x^* \in \Omega_1(K) \text{ tale che } x^*(x) = \sup x^*(K)\}$, i punti di supporto.
- $\Sigma_1(K) = \{z^* \in \Omega_1(K) : \exists z \in K \text{ tale che } z^*(z) = \sup z^*(K)\}$, i funzionali di supporto;

Segue

1. $\operatorname{supp}(K)$ è denso in ∂K;
2. $\Sigma_1(K)$ è denso in $\Omega_1(K)$.[6]

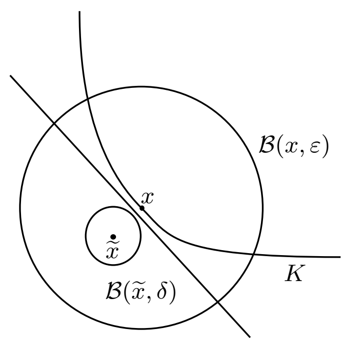
Dimostrazione Dimostreremo solo il primo punto siccome la dimostrazione del punto 2. è semplice. Fissiamo $\varepsilon > 0$ tale che $\sqrt{\varepsilon} < \frac{1}{2}$, e consideriamo $\widetilde{x} \in X \smallsetminus K$ tale che $x \in \mathcal{B}(\widetilde{x}, \varepsilon)$. Sappiamo che esiste $\delta > 0$ tale che $\mathcal{B}(\widetilde{x}, \delta) \subseteq \mathcal{B}(x, \varepsilon) \smallsetminus K$. Applichiamo il teorema di Hahn–Banach e separiamo $\mathcal{B}(\widetilde{x}, \delta)$ ottenendo il funzionale lineare x^*, possiamo supporre senza perdita di generalità che $\|x^*\| = 1$ e che valga

$$\sigma_K(x^*) = \sup x^*(K) < x^*(\widetilde{x}).$$

Ora sottraendo ambo i membri dell'ultima disuguaglianza per $x^*(x)$ otteniamo

$$\sigma_K(x^*) - x^*(x) < x^*(\widetilde{x} - x) \leq \|x^*\| \cdot \|\widetilde{x} - x\| \leq \varepsilon \Longrightarrow x^*(x) > \sigma_K(x^*) - \varepsilon.$$

Applichiamo ora il teorema di Bishop–Phelps con $\alpha = \sqrt{\varepsilon}$, quindi esiste $z \in \operatorname{supp}(K)$ tale che $\|z - x\| < \sqrt{\varepsilon}$. Il seguente disegno permette di capire la costruzione geometrica fatta.



□

Vediamo ora un teorema che è la versione per funzioni convesse del teorema di Bishop–Phelps, omettiamo la sua dimostrazione siccome risulta essere simile a quella già fatta nel corollario precedente.[7] Ricordiamo che se $f \in \overline{\mathcal{C}}(X)$ allora utiliziamo le seguenti notazioni

$$\begin{aligned} \operatorname{dom}(f) &= \{x \in X : f(x) < +\infty\}; \\ \operatorname{dom}(f^*) &= \{x^* \in X^* : f^*(x^*) < +\infty\}; \\ D(\partial f) &= \{x \in \operatorname{dom}(f) : \partial f(x) \neq \emptyset\}; \\ \partial f(X) &= \{x^* \in X^* : \exists x \in \operatorname{dom}(f) \text{ tale che } x^* \in \partial f(x)\}. \end{aligned}$$

È possibile mostrare che $\operatorname{dom}(f^*) = \{x^* \in X^* : f - x^* \text{ è limitata inferiomente}\}$.

[6] Si osservi che K è limitato se (e solo se per il teorema di Banach–Steinhaus) $\Omega_1(K) = \mathcal{S}_{X^*}$.
[7] In questo caso si utilizzano gli epigrafici.

Teorema 9.2 (di Brønsted–Rockafellar) Siano X spazio di Banach e $f \in \overline{\mathcal{C}}(X)$. Valgono

1. $D(\partial f)$ è denso in $\mathrm{dom}(f)$;
2. $\partial f(X)$ è denso in $\mathrm{dom}(f^*)$.

Capitolo 10

Miscellanea

Lo scopo di questo capitolo è quello di prendere coscienza dell'esistenza di alcuni teoremi di grande portata e di vedere alcune loro applicazioni.

10.1 Un'applicazione del teorema di James

Vogliamo arrivare ad una condizione sufficiente per la riflessività. Abbiamo bisogno di una definizione preliminare.

Definizione 10.1 (Insieme uniformemente convesso) Siano X normato e $C \subseteq X$ convesso e non vuoto. C si dice *uniformemente convesso* se

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ tale che } \forall x, y \in C \text{ con } \|x - y\| \geq \varepsilon \text{ allora } \operatorname{dist}\left(\frac{x+y}{2}, \partial C\right) \geq \delta.$$

Osservazione

- È equivalente a richiedere che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $x, y \in C$ se $\operatorname{dist}\left(\frac{x+y}{2}, \partial C\right) < \delta$ allora $\|x - y\| < \varepsilon$.
- Se C è uniformemente convesso, e non è ridotto ad un unico punto, allora C ha interno non vuoto. Infatti se per assurdo C ha interno vuoto, allora $C = \partial C$. Se $x, y \in C$ si ha
$$\frac{x+y}{2} \in C = \partial C,$$
ma $\operatorname{dist}\left(\frac{x+y}{2}, \partial C\right) > 0$.

Teorema 10.1 Sia X spazio di Banach. Se X contiene un insieme non vuoto, limitato, convesso e uniformemente convesso, allora X è riflessivo.

Dimostrazione Chiamato C l'insieme in questione, senza perdita di generalità possiamo supporre che C sia chiuso. è sufficiente mostrare che C è w-compatto. Sia $x^* \in \mathcal{S}_{X^*}$ e $s = \sup x^*(C)$. Consideriamo $\{x_n\} \subseteq C$ tale che $x^*(x_n) \to s$ e valutiamo

$$\operatorname{dist}\left(\frac{x_n + x_m}{2}, \partial C\right) \leq \operatorname{dist}\left(\frac{x_n + x_m}{2}, (x^*)^{-1}(s)\right) =$$

applicando la formula di Ascoli

$$= s - x^*\left(\frac{x_n + x_m}{2}\right) = \frac{1}{2}(s - x^*(x_n)) + \frac{1}{2}(s - x^*(x_m)) \xrightarrow{n,m \to +\infty} 0.$$

Dalla uniforme convessità si ottiene che $\|x_n - x_m\| \to 0$ e dunque $\{x_n\}$ è di Cauchy. Per completezza e per chiusura esiste $x_0 \in C$ tale che $x_n \to x_0$, e quindi

$$x^*(x_0) = \lim x^*(x_n) = s.$$

Applicando il teorema di James si ottiene che C è w-compatto, ma $\operatorname{int} C \neq \emptyset$ e dunque $\mathcal{B}_X$ è w-compatta. Da ciò la tesi. □

Osservazione Non vale il viceversa.

10.2 Il teorema di Josefson–Nissenzweig

Iniziamo questa sezione osservando che in generale esistono funzioni convesse e continue che sono illimitate sulla bolla unitaria. Sia X uno tra gli spazi c_0 e ℓ^p $(1 \leq p < +\infty)$ e definiamo

$$f(x) = \sum_{n=1}^{+\infty} n|x(n)|^n.$$

Possiamo osservare che

$$f(x) = \sup_{N\in\mathbb{N}} \sum_{n=1}^{N} n|x(n)|^n,$$

ma questo mostra che f è convessa e inferiormente semicontinua, siccome X è uno spazio di Banach allora f è continua. Ma f non è limitata sulla bolla unitaria, infatti $f(e_n) = n$.

Il seguente teorema (la cui dimostrazione è molto complicata) permette di costruire esempi simili in ogni spazio normato infinito-dimensionale.

Teorema 10.2 (di Josefson–Nissenzweig) Sia X normato infinito-dimensionale. Esiste $\{u_n^*\} \subseteq \mathcal{S}_{X^*}$ tale che

$$u_n^* \xrightarrow{w^*} 0.$$

Osservazione Questo teorema afferma che, per spazi infinito-dimensionali, la topologia debole* non coincide mai, dal punto di vista sequenziale, con quella forte. Sappiamo invece che ciò può succedere per la topologia debole, in tal caso gli spazi si dicono godere della proprietà di Schur.

Corollario Sia X normato. Sono tra loro equivalenti

1. X è finito-dimensionale;
2. ogni funzione convessa e continua su X è limitata sulla bolla chiusa;
3. ogni funzione convessa e continua su X è limitata superiormente sulla bolla unitaria.

Dimostrazione Osserviamo che 1. $\Rightarrow$ 2. segue dalla compattezza delle bolle e 2. $\Rightarrow$ 3. è ovvio. Dimostriamo $\neg$1. $\Rightarrow$ $\neg$3. Passiamo al completato $\widehat{X}$, e utilizzando il teorema di Josefson–Nissenzweig otteniamo una successione $\{u_n^*\} \subseteq \mathcal{S}_{X^*}$ tale che

$$u_n^* \xrightarrow{\sigma(X^*,\widehat{X})} 0.$$

Definiamo

$$f(\overline{x}) = \sum_{n=1}^{+\infty} n|u_n^*(\overline{x})|^n \qquad \forall \overline{x} \in \widehat{X},$$

essa, per ragionamenti analoghi a quelli fatti ad inizio sezione, risulta essere convessa e continua su $\widehat{X}$. Dimostriamone l'illimitatezza, sappiamo che per ogni n esiste $x_n \in \mathcal{B}_X$ tale che $u^*(x_n) > 1 - \frac{1}{n}$. Calcoliamo ora

$$f(x_n) \geq n|u^*(x_n)|^n > n\left(1 - \frac{1}{n}\right)^n \xrightarrow{n\to+\infty} +\infty.$$

Dunque la funzione $f_{|X}$ è quella cercata. □

10.3 Il teorema di Krein–Šmulyan

Sappiamo che in uno spazio normato se C è convesso, allora C è w-chiuso se e solo se $C \cap n\,\mathcal{B}_X$ è w-chiuso per ogni n. È naturale chiedersi se un teorema simile valga anche per la topologia debole*. Tale risultato prende proprio il nome di teorema di Krein–Šmulyan.

Teorema 10.3 (di Krein–Šmulyan) Siano X spazio di Banach e $C \subseteq X^*$ convesso. C è w^*-chiuso se e solo se per ogni $n \in \mathbb{N}$ si ha che $C \cap n\,\mathcal{B}_{X^*}$ è w^*-chiuso.

Corollario Siano X spazio di Banach separabile e $C \subseteq X^*$ convesso. C è w^*-chiuso se e solo se C è sequenzialmente w^*-chiuso.

Dimostrazione Daremo solo l'idea della dimostrazione. Si utilizza il teorema di Krein–Šmulyan e il fatto che se X è separabile allora $\mathcal{B}_{X^*}$ è metrizzabile. Infine si sfrutta il fatto che successioni w^*-convergenti sono limitate in norma. □

Corollario Siano X spazio di Banach separabile e $l : X^* \to \mathbb{R}$ lineare. Sono tra loro equivalenti

1. $l \in X$;
2. l è w^*-continuo per successioni;
3. $\ker(l)$ è sequenzialmente w^*-chiuso.

Dimostrazione Le implicazioni 1. ⇒ 2. e 2. ⇒ 3. sono banali. Dimostriamo 3. ⇒ 1. Sappiamo che $l \in X$ se è w^*-continuo, ovvero se $\ker(l)$ è w^*-chiuso. Ma sfruttando l'ipotesi e il corollario precedente si ha che $\ker(l)$ è w^*-chiuso, ovvero la tesi. □

10.4 Funzioni midconvesse

Vogliamo studiare una classe particolare di funzioni che hanno molto in comune con le funzioni convesse.

Definizione 10.2 (Funzione midconvessa) Siano X spazio vettoriale e $C \subseteq X$ convesso. Una funzione $f : C \to \mathbb{R}$ si dice *midconvessa*, o *Jensen-convessa*, o *jenseniana* se per ogni $x, y \in C$ si ha

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x) + f(y)).$$

Osservazione Esistono funzioni, anche su $\mathbb{R}$, che sono midconvesse ma non convesse.

Teorema 10.4 Siano X spazio vettoriale, $C \subseteq X$ convesso e $f : C \to \mathbb{R}$. Se f è midconvessa su C allora per ogni $x_1, \ldots, x_n \in C$ e per ogni $\lambda_1, \ldots, \lambda_n \in [0, +\infty) \cap \mathbb{Q}$ con $\sum_{i=1}^n \lambda_i = 1$ vale

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Dimostrazione Dividiamo la dimostrazione in due casi.

- Supponiamo che $\lambda_1 = \cdots = \lambda_n = \frac{1}{n}$. L'idea è quella di fare un'induzione strana.
  - Si lascia per esercizio di dimostrare il caso $n = 2^k$, che risulta essere facile facendo induzione.
  - Supponiamo che la tesi sia vera per n, vogliamo provarla per $n-1$. Poniamo

$$s = \sum_{i=1}^{n-1} x_i \qquad \sigma = \sum_{i=1}^{n-1} f(x_i).$$

Osserviamo che $\frac{1}{n-1} = \frac{1}{n}\frac{n}{n-1} = \frac{1}{n}\left(1 + \frac{1}{n-1}\right)$ e ora facendo un breve conto

$$f\left(\frac{s}{n-1}\right) = f\left(\frac{1}{n}\left(s + \frac{s}{n-1}\right)\right) =$$
$$= f\left(\frac{1}{n}\left(x_1 + \cdots + x_{n-1} + \frac{s}{n-1}\right)\right) \leq \frac{1}{n}\left(\sigma + f\left(\frac{s}{n-1}\right)\right).$$

Da cui $f\left(\frac{s}{n-1}\right)\left(1 - \frac{1}{n}\right) \leq \frac{\sigma}{n}$, ovvero $f\left(\frac{s}{n-1}\right) \leq \frac{\sigma}{n-1}$.

- Dimostriamo ora il caso generale. Possiamo scrivere

$$\lambda_i = \frac{p_i}{q} \qquad \begin{array}{c} \forall i = 1, \ldots, n \\ 1 \leq p_i \leq q,\ p_i \in \mathbb{Z},\ q \in \mathbb{N} \end{array}$$

con $\sum_{i=1}^n p_i = q$. È sufficiente uno sporco trucco per ottenere la tesi

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = f\left(\frac{1}{q}(p_1 x_1 + \ldots + p_n x_n)\right) =$$
$$= f\left(\frac{1}{q}(\underbrace{x_1 + \cdots + x_1}_{p_1} + \underbrace{x_2 + \cdots + x_2}_{p_2} + \cdots + \underbrace{x_n + \cdots + x_n}_{p_n})\right) \leq$$

utilizzando la prima parte della dimostrazione (ci troviamo di fronte ad una media aritmetica)

$$\leq \frac{1}{q}\sum_{i=1}^n p_i f(x_i) = \sum_{i=1}^n \lambda_i f(x_i).$$ □

Osservazione Questo teorema assicura che se una funzione midconvessa è continua allora necessariamente è convessa.

Teorema 10.5 Esiste una funzione $f : \mathbb{R} \to \mathbb{R}$ midconvessa, ma non convessa.

Dimostrazione Consideriamo $\mathbb{R}$ come spazio vettoriale su $\mathbb{Q}$ e sia $\mathcal{B}$ una sua base di Hamel, ovvero per ogni $x \in \mathbb{R}$ esistono unici $b_1, \ldots, b_n \in \mathcal{B}$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}$ tali che

$$x = \sum_{i=1}^{n} \alpha_i b_i.$$

Senza perdita di generalità possiamo supporre che $0 \in \overline{\mathcal{B}}$. Possiamo definire $f : \mathcal{B} \to \mathbb{R}$ qualsiasi di modo che f non sia limitata in nessun intorno dell'origine, ovvero

$$\sup f(\mathcal{B} \cap (-\varepsilon, \varepsilon)) = +\infty \qquad \forall \varepsilon > 0.$$

Estendiamo ora f (in modo unico) ad una funzione lineare. Quindi f è midconvessa ma non convessa (siccome non è continua nell'origine). □

Osservazione Modificando in maniera opportuna la dimostrazione, è possibile ottenere un esempio simile in ogni spazio vettoriale reale. Infatti se X è normato è possibile decomporlo come $X = H \oplus tv$ e definiamo la nostra funzione f del teorema precedente sulla retta tv, infine la estendiamo in modo costante su $H \oplus \overline{t}v$ per ogni $\overline{t}$ fissato.

Vediamo ora un importante teorema di caratterizzazione della convessità di una funzione midconvessa. La dimostrazione risulterebbe essere troppo pesante per i nostri scopi e preferiamo dunque ometterla.

Teorema 10.6 Siano $I \subseteq \mathbb{R}$ intervallo aperto e $f : I \to \mathbb{R}$ funzione midconvessa. Sono tra loro equivalenti

1. f è convessa;
2. f è continua;
3. f è localmente limitata su I;
4. f è misurabile secondo Lebesgue;
5. f è limitata superiormente su un insieme di misura positiva secondo Lebesgue;
6. f è limitata superiormente su un insieme di seconda categoria secondo Baire.