L. Pandolfi
Spazi di Banach

CAF in *Matematica per Le Scienze dell'Ingegneria*

Corso di *Elementi di Analisi Reale e Complessa*

19 gennaio 2006

Indice

1 Spazi di Banach **3**

1.1 Introduzione all'analisi funzionale 4
1.1.1 L'equazione $Ax = \phi$ 4
1.1.2 L'equazione $\lambda x - Ax = y$ 7
1.1.3 L'equazione di Fredholm a nucleo degenere 8
1.1.4 L'equazione di prima specie 10
1.1.5 Ricapitolazione 11
1.2 Spazi lineari normati 12
1.2.1 Dimostrazioni posposte 17
1.3 Spazi prodotto 22
1.4 Gli esempi principali di spazi di Banach 24
1.4.1 Gli esempi di spazi lineari normati 24
1.4.2 Le dimostrazioni della completezza 29
1.4.3 Teorema del doppio limite 35
1.5 Sottospazi di spazi lineari normati 38
1.5.1 Identità approssimate e dimostrazione del teorema di Weierstrass 39
1.6 La compattezza 45
1.6.1 Dimostrazioni posposte 46
1.7 Operatori lineari 51
1.7.1 Proprietà geometriche degli operatori lineari 51
1.7.2 La continuità degli operatori lineari 55
1.7.3 Funzionali lineari continui ed iperpiani 59
1.7.4 Lo spazio $\mathcal{L}(X, Y)$ 62
1.7.5 Inversi di un operatore 67
1.8 Il teorema di Baire e le sue conseguenze 72
1.8.1 Proiezioni 76
1.8.2 Appendice: Applicazioni 79
1.8.3 Dimostrazioni posposte 83

1.9 Lo spazio duale 87
1.9.1 Applicazioni: Insiemi convessi 90
1.9.2 Applicazioni: Funzioni convesse 94
1.9.3 Dimostrazioni posposte 97
1.10 Convergenza debole e debole stella 105
1.10.1 Dimostrazioni posposte 113
1.11 Esempi di spazi duali 115
1.11.1 Relazione tra le convergenze debole e debole stella . . . 125
1.12 Lo spettro di un operatore 126
1.12.1 Proiezioni spettrali 134
1.13 Trasformazioni non lineari 140
1.13.1 Teorema delle contrazioni e applicazioni 140
1.13.2 I differenziali 144

Capitolo 1

Spazi di Banach

1.1 Introduzione all'analisi funzionale

L'*Analisi funzionale* è nata quando, tra la fine del XIX secolo e il primo trentennio del XX, sono stati raccolti in un'unica teoria molti risultati provenienti da varie parti dell'analisi matematica; in particolare, dalla teoria delle *equazioni integrali* della forma

$$x(s) = \mu \int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi + y(s)\,, \qquad s \in [a,b] \tag{1.1}$$

ove μ è un parametro e $K(s,\xi)$, $y(s)$ sono funzioni note mentre la funzione $x(s)$ è incognita.

L'intervallo $[a,b]$ è limitato. Equazioni di questo tipo si chiamano *equazioni di Fredholm di seconda specie* e si incontrano per esempio nella soluzione di problemi di elasticità (problemi che hanno stimolato i primi studi di Fredholm).

La funzione $K(s,\xi)$ si chiama il *nucleo* dell'equazione di Fredholm.

Si chiama *equazione di Fredholm di prima specie* l'equazione

$$\int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi = y(s)\,, \qquad s \in [a,b] \tag{1.2}$$

che pure si incontra nelle applicazioni, ma che ha una teoria sostanzialmente più delicata, per una ragione che vedremo.

Come introduzione all'analisi funzionale, vediamo come sia possibile trattare l'equazione di Fredholm (di prima o di seconda specie) quando il nucleo è *degenere*, ossia quando

$$K(s,\xi) = \sum_{i=1}^{m} a_i(s)b_i(\xi)\,, \qquad s\,,\ \xi \in [a,b]\,. \tag{1.3}$$

Per semplicità supporremo che le funzioni $a(s)$ e $b(\xi)$ siano continue su $[a,b]$.

Nel caso del nucleo degenere la soluzione dell'equazione di Fredholm si riduce ad un problema di algebra lineare e quindi richiamiamo alcune nozioni che dovremo usare.

1.1.1 L'equazione $Ax = \phi$

Ricordiamo che $\mathbb{C}$ indica l'insieme dei numeri complessi e che $\mathbb{C}^n$ indica il prodotto cartesiano di n copie di $\mathbb{C}$. In $\mathbb{C}^n$ esistono infinite basi, ciascuna di n elementi. Si chiama *base canonica* la base $\{e_1\,,\ldots\,,e_n\}$, avendo indicato con e_i il vettore le cui componenti sono tutte nulle, salvo la i–ma, che vale 1.

Se x è un vettore di $\mathbb{C}^n$, scriveremo

$$x = \text{col} \left[\begin{array}{ccc} x_1 & \dots & x_n \end{array} \right] ,$$

intendendo con ciò che

$$x = \sum_{i=1}^{n} x_i e_i \, .$$

Affermazioni del tutto analoghe valgono per $\mathbb{C}^m$. Per evitare ambiguità, indicheremo con $\{\epsilon_1 \, , \dots \, , \epsilon_m\}$ la sua base canonica. Anzi, per chiarezza indicheremo con lettere romane i vettori di $\mathbb{C}^n$ e con lettere greche quelli di $\mathbb{C}^m$.

Sia A una matrice $m \times n$,

$$A = \left[\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right] .$$

L'espressione $Ax = \phi$ si usa per rappresentare in modo compatto il sistema di m equazioni in n incognite

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & \phi_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & \phi_2 \\ \vdots & & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & \phi_m \, . \end{array} \tag{1.4}$$

Vogliamo ricordare le condizioni, note dai corsi precedenti, perché:

1. L'equazione $Ax = \phi$ ammetta al più una soluzione;
2. L'equazione $Ax = \phi$ ammetta almeno una soluzione;
3. L'equazione $Ax = \phi$ ammetta esattamente una soluzione per ogni $\phi \in \mathbb{C}^m$.

Consideriamo prima di tutto il problema 1), unicità di soluzione. Si noti la formulazione del problema di unicità: non si richiede che una soluzione debba esistere *per ogni* ϕ; Si richiede che *se una soluzione esiste* allora questa sia *unica*.

Se per la medesima ϕ si hanno due soluzioni x' ed x'' tra loro diverse, allora $y = x' - x''$ verifica

$$Ay = A(x' - x'') = Ax' - Ax'' = \phi - \phi = 0 \, .$$

Dunque, se per una ϕ la soluzione non è unica, allora esiste $y \neq 0$ e tale che $Ay = 0$. E viceversa:

Teorema 1 *Se per* una ϕ *il problema (1.4) ammette soluzione e questa è unica, allora*

$$\ker A = \{x \mid Ax = 0\} = \{0\}\,.$$

Viceversa, se $\ker A = \{0\}$ *allora il problema (1.4) ammette soluzione unica* per ogni ϕ.

Osservazione 2 Si noti che il problema dell'unicità, come è stato enunciato, sembra dipendere dalla scelta di ϕ. Ossia, si potrebbe pensare che la soluzione di (1.4) sia unica per alcune ϕ ma non per altre. Invece, il teorema precedente mostra che si ha unicità di soluzione per ogni ϕ se e solo se si ha unicità di soluzione per una ϕ. ∎

L'insieme

$$\ker A = \{x \mid Ax = 0\}$$

si chiama il *nucleo* di A.

Passiamo ora a studiare il problema dell'esistenza di soluzioni. Chiamiamo *immagine* di A l'insieme

$$\operatorname{im} A = \{Ax \mid x \in \mathbb{C}^n\}\,.$$

Dunque, $\operatorname{im} A$ è esattamente l'insieme dei vettori ϕ per i quali l'equazione (1.4) è risolubile.

Si ricordi che $\operatorname{im} A$ è un s.spazio di $\mathbb{C}^m$. Ora un generico s.spazio V di $\mathbb{C}^m$ può caratterizzarsi descrivendone gli elementi; e ciò abbiamo fatto per definire $\operatorname{im} A$; ma può anche identificarsi per mezzo del suo s.spazio *ortogonale* $V^\perp$.

Definitione 3 Un vettore $\psi \in \mathbb{C}^m$ si dice ortogonale a V quando è ortogonale a tutti gli elementi di V; e $V^\perp$ è il s.spazio di tutti i vettori ψ ortogonali a V. ∎

Indicando col simbolo $\langle \phi, \psi \rangle$ il prodotto interno dei vettori ϕ e ψ,

$$V^\perp = \{\psi \mid \langle \phi, \psi \rangle = 0 \ \forall \phi \in V\}\,.$$

Quando $V = \operatorname{im} A$, è facile identificare $V^\perp$. Introduciamo per questo la matrice A^*, aggiunta di A. Si sa che A^* verifica

$$\langle Ax, \phi \rangle = \langle x, A^*\phi \rangle\,. \tag{1.5}$$

Teorema 4 *si ha*

$$(\operatorname{im} A)^\perp = \ker A^*\,, \qquad \operatorname{im} A = (\ker A^*)^\perp\,.$$

Dim. Se $\phi \perp \operatorname{im} A$ allora per ogni x vale

$$\langle Ax, \phi \rangle = 0 \qquad \text{ossia} \qquad \langle x, A^*\phi \rangle = 0 \quad \forall x \in \mathbb{C}^n ;$$

e quindi $A^*\phi = 0$. Viceversa, se $A^*\phi = 0$ allora per ogni $x \in \mathbb{C}^n$ vale

$$0 = \langle x, A^*\phi \rangle = \langle Ax, \phi \rangle \qquad \forall x \in \mathbb{C}^n$$

e quindi $\phi \perp \operatorname{im} A$. Ciò prova la prima uguaglianza. La seconda si può provare in modo analogo, ma può anche dedursi dalla prima, ricordando che $\left(V^\perp\right)^\perp = V$ per ogni s.spazio di $\mathbb{C}^m$. ∎

E quindi:

Teorema 5 *L'equazione (1.4) è risolubile se e solo se* $\phi \perp (\ker A^*)$.

Questo teorema assume un aspetto particolarmente semplice se vogliamo che l'equazione (1.4) sia risolubile per ogni ϕ:

Teorema 6 *L'equazione (1.4) è risolubile* per ogni ϕ *se e solo se* $\ker A^* = \{0\}$.

Si può dunque enunciare:

Teorema 7 (alternativa di Fredholm) *La matrice A identifica una trasformazione* iniettiva *se e solo se la matrice A^* identifica una trasformazione* suriettiva; *La matrice A identifica una trasformazione* suriettiva *se e solo se la matrice A^* identifica una trasformazione* iniettiva.

L'alternativa di Fredholm si incontra anche in situazioni più generali dello studio di sistemi di n equazioni in m incognite. Però, nel caso particolare del problema (1.4) si può essere anche più precisi e notare che se $\ker A = \{0\}$ e inoltre A è suriettiva, allora deve essere $n = m$.

Osservazione 8 L'alternativa di Fredholm per le matrici è utile perché in pratica è assai più semplice verificare l'unicità di soluzione piuttosto che l'esistenza di soluzione.

1.1.2 L'equazione $\lambda x - Ax = y$

Sia ora $X = Y$ e quindi A una matrice quadrata, e studiamo l'equazione

$$\lambda x - Ax = y\,, \qquad \text{ossia} \quad (\lambda I - A)x = y \tag{1.6}$$

ove I indica la matrice identità. Naturalmente quest'equazione è un caso particolare di (1.4) e quindi si tratta di adattare a questo caso particolare i risultati già trovati. In particolare:

Teorema 9 *L'equazione (1.6) ammette* una e una sola *una soluzione* per ogni y *se e solo se l'equazione*

$$(\bar{\lambda}I - A^*)\psi = \phi$$

ammette esattamente una soluzione. Ciò avviene se e solo se

$$\det(\lambda I - A) \neq 0\,. \tag{1.7}$$

Si noti che la condizione (1.7) è specifica del caso $n = m$ e quindi non ha analogo tra i risultati trovati prima.

L'insieme dei numeri λ per cui vale l'unicità di soluzione si chiama *l'insieme risolvente* della matrice A e si indica col simbolo $\rho(A)$; il suo complementare si chiama lo *spettro* della matrice e si indica col simbolo $\sigma(A)$. Gli elementi di $\sigma(A)$ si chiamano gli *autovalori* di A.

Si noti che

$$\det(\lambda I - A)$$

è un polinomio di grado $n \geq 1$; e quindi lo spettro di una matrice non è vuoto.

Gli autovalori sono in generale numeri complessi, anche se la matrice A è reale. Si ricordi che se la matrice A è reale allora $\lambda \in \sigma(A)$ se e solo se $\bar{\lambda} \in \sigma(A)$.

In generale si ha che se $\lambda \in \sigma(A)$ allora $\bar{\lambda} \in \sigma(A^*)$.

Si lascia per esercizio di enunciare risultati analoghi per l'equazione

$$x = \mu A x + y\,. \tag{1.8}$$

Notiamo però che quest'equazione è risolubile in modo banale quando $\mu = 0$ perché in tal caso essa si riduce a $x = y$.

1.1.3 L'equazione di Fredholm a nucleo degenere

Studiamo ora l'equazione (1.1) nel caso particolare in cui il nucleo ha forma (1.3), ossia nel caso del *nucleo degenere*.

Si noti che il parametro μ moltiplica l'integrale, così che l'equazione di Fredholm così scritta viene ad essere di seconda specie per ogni valore di μ. Avessimo invece scritto

$$\lambda x(s) = \int_a^b K(s,\xi)x(\xi)\,\mathrm{d}\xi + y(s)\,, \qquad s \in [a,b]$$

per $\lambda = 0$ avremmo trovato un'equazione di prima specie.

L'equazione (1.1) si scrive

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) \int_a^b b_r(\xi) x(\xi)\, \mathrm{d}\xi + y(s) \tag{1.9}$$

ossia

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)\,, \qquad x_r = \int_a^b b_r(\xi) x(\xi)\, \mathrm{d}\xi\,.$$

I numeri x_i dipendono dalla funzione incognita $x(s)$ e quindi non sono noti; però, moltiplicando i due membri dell'uguaglianza per $b_i(s)$ e integrando da a a b, si trova che essi risolvono

$$x_i = \mu \sum_{r=1}^{m} k_{ir} x_r + y_i\,, \qquad \left\{ \begin{array}{rcl} k_{ir} & = & \int_a^b b_i(s) a_r(s)\, \mathrm{d}s \\ y_i & = & \int_a^b b_i(s) y(s)\, \mathrm{d}s\,. \end{array} \right. \tag{1.10}$$

Dunque, i numeri x_i risolvono un sistema della forma (1.8). Introduciamo allora la matrice K i cui elementi sono k_{ir} e i vettori x e y, i cui elementi sono rispettivamente i numeri x_i e y_i e scriviamo il sistema (1.10) come

$$(I - \mu K)x = y\,. \tag{1.11}$$

Dunque ogni soluzione dell'equazione di Fredholm (1.9) identifica una soluzione x di (1.11) e si vede anche che vale il viceversa: se x risolve (1.11) allora

$$x(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)$$

risolve l'equazione (1.9). Mostriamo infatti che questa funzione, sostituita a destra ed a sinistra di (1.9), verifica l'uguaglianza. A sinistra si trova

$$\mu \sum_{r=1}^{m} a_r(s) x_r + y(s)$$

mentre sostituendo a destra e tenendo conto di (1.10) (e scambiando il nome degli indici) si trova

$$\begin{aligned} &\mu \sum_{r=1}^{m} a_r(s) \int_a^b b_r(\xi) \left\{ \mu \sum_{i=1}^{m} a_i(\xi) x_i + y(\xi) \right\} \mathrm{d}\xi + y(s) \\ &= \mu \sum_{r=1}^{m} a_r(s) \left\{ \mu \sum_{i=1}^{m} k_{ri} x_i + y_r \right\} + y(s) = \mu \sum_{r=1}^{m} a_r(s) x_r + y(s)\,, \end{aligned}$$

come si voleva.

Possiamo quindi trasferire all'equazione integrale di Fredholm i risultati che abbiamo enunciato per i sistemi di equazioni lineari:

Teorema 10 *L'equazione di Fredholm (1.1) con nucleo degenere ammette al più una soluzione se e solo se l'equazione*

$$\psi = \bar{\mu} K^* \psi + \phi \tag{1.12}$$

è risolubile per ogni ϕ; l'equazione di Fredholm è risolubile per ogni funzione $y(x)$ se e solo se l'equazione (1.12) ammette unicità di soluzione.

Il teorema precedente dà una specie di "alternativa di Fredholm", per l'equazione integrale, ma trascritta mediante le matrici. E' però facile vedere che l'equazione (1.12) corrisponde all'equazione integrale

$$\psi(\xi) = \bar{\mu} \int_a^b \overline{K(\xi, s)} \psi(\xi) \, \mathrm{d}\xi + \phi(\xi) \,, \tag{1.13}$$

equazione che si chiama *aggiunta* della (1.1). Possiamo quindi enunciare l'*alternativa di Fredholm* per le equazioni integrali di Fredholm a nucleo degenere:

Teorema 11 *Quando il nucleo è degenere, la (1.1) ammette unicità di soluzione se la sua aggiunta è risolubile per ogni $\phi(s)$; la (1.1) è risolubile per ogni $y(s)$ se la sua aggiunta ammette unicità di soluzione.*

Ancora, è assai più facile verificare l'unicità piuttosto che l'esistenza di soluzioni, e ciò spiega l'interesse di questo teorema.

1.1.4 L'equazione di prima specie

Le considerazioni precedenti possono tutte ripetersi nel caso dell'equazione di prima specie (1.2). In tal caso, invece della (1.10) si trova l'equazione

$$\sum_{r=1}^{m} k_{ir} x_r = y_i$$

ossia

$$Kx = y \,. \tag{1.14}$$

Per vedere la differenza tra questa e l'equazione (1.11), esaminiamo il caso particolare in cui

$$k_{ir} = \int_a^b b_i(\xi) a_r(\xi) \, \mathrm{d}\xi = \begin{cases} 0 & \text{se} \quad i \neq r \\ k_{ii} \neq 0 & \text{se} \quad i = r \,. \end{cases}$$

In tal caso la (1.11) e la (1.14) divengono rispettivamente

$$\begin{bmatrix} (1-\mu k_{11}) & & & & \\ & (1-\mu k_{22}) & & & \\ & & \ddots & & \\ & & & (1-\mu k_{mm}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

$$\begin{bmatrix} k_{11} & & & \\ & k_{22} & & \\ & & \ddots & \\ & & & k_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Ambedue queste equazioni sono, in principio, facilmente risolubili e le soluzioni sono, rispettivamente,

$$\begin{bmatrix} y_1/[1-\mu k_{11}] \\ y_2/[1-\mu k_{22}] \\ \vdots \\ y_m/[1-\mu k_{mm}] \end{bmatrix}, \qquad \begin{bmatrix} y_1/k_{11} \\ y_2/k_{22} \\ \vdots \\ y_m/k_{mm} \end{bmatrix}.$$

In generale, al crescere dell'indice i, i numeri k_{ii} divengono "velocemente" piccoli e quindi nel caso dell'equazione di prima specie, gli errori commessi nella misura di $y(s)$, e quindi nel calcolo degli y_i, si amplificano velocemente. Ciò non accade per il caso dell'equazione di seconda specie, salvo nel caso in cui $1/\mu$ è circa uguale ad uno dei numeri k_{ii}; e anche in questo caso la sola componente i–ma può provocare dei problemi numerici.

Queste considerazioni mostrano che, almeno nel caso del nucleo degenere, la soluzione delle equazioni integrali di prima specie è assai più delicata della soluzione di quelle di seconda specie.

1.1.5 Ricapitolazione

Ora poniamoci alcuni problemi, la cui soluzione guiderà la scelta degli argomenti di analisi funzionale che studieremo. Il primo è questo: tutti i nuclei espressi mediante polinomi sono degeneri. Ma per esempio

$$k(t,s) = e^{s\xi}$$

non è un nucleo degenere. Quindi possiamo chiedere se sia possibili approssimare nuclei abbastanza regolari mediante nuclei degeneri. A questo problema risponderà il *teorema di Weierstrass,* teorema 43.

Per applicare il teorema 11 è necessario ricondursi alla situazione matriciale del teorema 10. In tal caso l'unicità di soluzione si ha quando $1/\mu$ non è un'autovalore della matrice A. Se però il nucleo non è degenere, ciò certamente non può farsi; e allora ci si chiede se l'alternativa di Fredholm, opportunamente reinterpretata, continui a valere e, in caso affermativo, come sia possibile verificare l'unicità di soluzione.

Vorremo infine capire se anche nel caso dei nuclei non degeneri l'equazione di prima specie sia più delicata di quella di seconda, e chiarire la ragione di ciò.

Fatte queste premesse, passiamo ad uno studio sistematico dell'Analisi Funzionale.

1.2 Spazi lineari normati

Richiamiamo, dai corsi precedenti, le definizioni di spazio lineare e di norma. In questo corso gli spazi lineari avranno sempre, come campo scalare, il campo $\mathbb{R}$ o, più frequentemente, $\mathbb{C}$. Per indicare genericamente uno di questi due campi useremo il simbolo Φ, specificando, se necessario, quando si intende $\Phi = \mathbb{R}$ oppure $\Phi = \mathbb{C}$.

Definitione 12 (di spazio lineare) Si dice che X è uno spazio lineare su Φ quando è un gruppo commutativo (rispetto ad un'operazione usualmente indicata col segno $+$) ed inoltre esiste un'operazione (usualmente indicata con notazione moltiplicativa) che ad ogni coppia $(\alpha, x) \in \Phi \times X$ associa un elemento di X e che verifica le seguenti proprietà:

$$\begin{array}{lll} \alpha(x+y) = \alpha x + \alpha y & \forall \alpha \in \Phi\,, & \forall x\,,\ y \in X \\ (\alpha+\beta)x = \alpha x + \beta x & \forall \alpha\,,\ \beta \in \Phi\,, & \forall x \in X \\ 1x = x & & \forall x \in X\,. \ \blacksquare \end{array}$$

Regole di calcolo che discendono dalla precedente definizione sono:

$$\begin{array}{ll} 0x = 0 & \forall x \in X \\ (-\alpha x) = -(\alpha x) & \forall \alpha \in \Phi\,,\ \forall x \in X \\ \alpha 0 = 0 & \forall \alpha \in \Phi\,. \end{array}$$

Ricordiamo alcuni esempi di spazi lineari.

Esempio 13 Gli spazi $\mathbb{R}^n$ e $\mathbb{C}^n$ sono esempi di spazi lineari rispettivamente su $\mathbb{R}$ e su $\mathbb{C}$. (lo spazio $\mathbb{C}$ è anche uno spazio lineare su $\mathbb{R}$). Gli spazi dei

polinomi, delle funzioni continue, delle funzioni integrabili secondo Riemann; lo spazio delle successioni; o quello delle successioni convergenti; o quello delle successioni limitate; o quello delle successioni convergenti a zero sono sono spazi lineari su $\mathbb{R}$ oppure su $\mathbb{C}$ a seconda che sia $\mathbb{R}$ oppure $\mathbb{C}$ l'insieme dei valori. ∎

Sia X uno spazio lineare su Φ e sia $\mathcal{G}$ un suo s.insieme finito o meno. Si dice che $\mathcal{G}$ è un insieme di *generatori* se per ogni $x \in X$ esistono un numero naturale n; elementi $g_1, \ldots, g_n \in \mathcal{G}$; scalari $\alpha_1, \ldots, \alpha_n \in \Phi$, tali che

$$x = \sum_{i=1}^{n} \alpha_i g_i \,.$$

Se in X esiste un insieme $\mathcal{G}$ finito di generatori, si dice che X ha *dimensione finita*; si dice che ha *dimensione infinita* altrimenti.

La *dimensione* di X è il numero degli elementi di uno dei generatori di X che ha il minimo numero di elementi.

Gli spazi $\mathbb{R}^n$ hanno dimensione n; lo spazio $\mathbb{C}^n$ ha dimensione n su $\mathbb{C}$ e dimensione $2n$ su $\mathbb{R}$.

Ha dimensione finita lo spazio dei polinomi di grado al più n. Se come campo scalare si prende il campo cui appartengono i coefficienti, la dimensione è $n+1$. Nonostante questi esempi importanti, in questo corso noi studieremo solamente spazi lineari di dimensione infinita.

Ricordiamo ora la definizione di norma.

Definitione 14 Sia X uno spazio lineare su Φ. Si chiama *norma* su X una funzione, che si indica col simbolo $||x||$, da X in $\mathbb{R}$ (anche quando il campo scalare è $\mathbb{C}$), che soddisfa:

$$\begin{array}{ll} ||x|| \geq 0 & \forall x \in X \\ ||x|| = 0 \quad \text{se e solo se } x = 0 & \\ ||\alpha x|| = |\alpha| \cdot ||x|| & \forall \alpha \in \Phi\,,\ \forall x \in X \\ ||x+y|| \leq ||x|| + ||y|| & \forall x\,, y \in X\,. \end{array}$$ ∎

L'ultima proprietà della norma si chiama *disuguaglianza triangolare*.

Quando ci sia possibilità di confusione, si scrive $||\cdot||_X$ per la norma dello spazio X.

Uno spazio lineare dotato di norma si chiama uno *spazio lineare normato* e da ora in poi scriveremo semplicemente s.l.n..

Se X è uno s.l.n., la funzione

$$d(x,y) = ||x-y||$$

è una *distanza* per la quale

$$d(x+z, y+z) = d(x,y)\,;$$

ossia, è una distanza *invariante per traslazioni*.

L'introduzione di una norma su X permette di definire:

1. gli *intorni* $B(x_0,\epsilon) = \{x \mid ||x-x_0|| < \epsilon\}$, con $\epsilon > 0$;
2. gli *insiemi aperti* di X, e quindi una *topologia* su X: $A \subseteq X$ è aperto se esso contiene un intorno di ciascun suo punto;
3. gli insiemi *limitati*: un insieme si dice limitato quando è contenuto in un intorno di 0;
4. la *convergenza* di successioni: si dice che la successione (x_n) converge ad x_0,
$$\lim x_n = x_0\,,$$
se per ogni $\epsilon > 0$ esiste N_ϵ tale che da $n > N_\epsilon$ segue
$$||x_n - x_0|| < \epsilon \qquad \text{ossia } x_n \in B(x_0,\epsilon).$$

In modo del tutto analogo si definiscono limiti e continuità di funzioni che operano tra s.l.n-ti: per esempio $f(\cdot)$ da X in Y è *continua* nel punto $x_0 \in X$ se è ivi definita e se per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che

$$||x-x_0||_X < \delta\,,\ x \in \mathrm{dom} f(\cdot)\,, \quad \text{segue} \quad ||f(x)-f(x_0)||_Y < \epsilon\,.$$

Esattamente come nel caso dei valori assoluti e dei moduli dei numeri complessi, si prova:

Teorema 15 *Vale:*

$$\Big|\, ||x|| - ||y|| \,\Big| \le ||x-y|| \tag{1.15}$$

Dim. Va provato che:

$$-||x-y|| \le ||x|| - ||y|| \le ||x-y||\,. \tag{1.16}$$

Usando la disuguaglianza triangolare si vede che

$$||x|| \le ||x-y|| + ||y|| \qquad \text{ed anche} \quad ||y|| \le ||y-x|| + ||x||\,,$$

ossia la (1.16). ∎

E' conseguenza della (1.15):

Teorema 16 *La norma, come trasformazione dallo s.l.n. X a $\mathbb{R}$, è uniformemente continua.*

Inoltre:

Teorema 17 *Sia $f(x)$ una funzione da uno s.l.n. X ad uno s.l.n. Y. Si equivalgono le condizioni*

$$\lim_{x\to x_0} f(x) = 0 \qquad e \qquad \lim_{x\to x_0} ||f(x)||_Y = 0$$

Introduciamo ora la seguente definizione:

Definitione 18 Diciamo che una successione (x_n) di elementi di X è *fondamentale* quando per ogni $\epsilon > 0$ esiste un indice N_ϵ tale che, per ogni n, m maggiori di N_ϵ, vale

$$||x_n - x_m|| < \epsilon\,. \quad \blacksquare$$

Con la stesa dimostrazione che si conosce per le successioni di numeri reali, si prova che ogni successione convergente è fondamentale.

Lo s.l.n. X si dice *completo* se ogni successione fondamentale è convergente. $\blacksquare$

Si vede facilmente che tutte le regole di calcolo dei limiti che valgono in $\mathbb{R}^n$ o in $\mathbb{C}^n$ valgono in qualunque s.l.n. (ma ovviamente il teorema della funzione monotona, che dipende dalla relazione di ordine, non ha corrispondente). Importanti eccezioni sono le seguenti: si sa che sia $\mathbb{R}^n$ che $\mathbb{C}^n$ sono spazi completi. Invece:

Teorema 19 *Esistono s.l.n-ti non completi.*

Questo teorema suggerisce:

Definitione 20 Uno spazio lineare normato e completo si chiama *spazio di Banach*. $\blacksquare$

Vedremo in seguito esempi di spazi di Banach. Dovrebbe essere noto, come conseguenza del *teorema dei limiti,* che lo spazio lineare $C(a,b)$ è completo. Per una dimostrazione si veda il paragrafo 1.4.2.

Si sa che esistono successioni (q_n) la cui immagine è densa in $\mathbb{R}^n$ o in $\mathbb{C}^n$. Invece:

Teorema 21 *Esistono s.l.n-ti nei quali nessuna successione ha immagine densa.*

L'ultimo teorema suggerisce la seguente definizione:

Definitione 22 Sia X uno s.l.n. Se esiste una successione (x_n) a valori in X la cui immagine è densa in X, lo spazio X si dice $\boxed{\textit{separabile}}$. ∎

Il teorema 21 può quindi enunciarsi dicendo che esistono s.l.n-ti non separabili.

Per concludere quest'introduzione, ricordiamo che sul medesimo s.l.n. X possono introdursi più norme. Per esempio su $\mathbb{C}^n$ sono norme tra loro diverse le seguenti. Se $x = \text{col}\left[\begin{array}{ccc} x_1 & \dots & x_n \end{array}\right]$:

$$||x||_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p}, \quad (\text{se } p \geq 1) \qquad ||x||_\infty = \max\{|x_i|\,,\ i = 1\ \dots, n\}\,.$$

(E' facile provare che $||\cdot||_1$ e $||\cdot||_\infty$ sono norme. Gli altri casi sono più difficili. Si noti però che $||\cdot||_2$ è l'usuale norma euclidea).

Tuttavia, su $\mathbb{C}^n$, la proprietà

$$\lim x_n = x_0 \tag{1.17}$$

è indipendente dalla particolare norma che si usa per verificarla.

Invece:

Teorema 23 *Esistono s.l.n-ti X sui quali si possono definire due diverse norme, $||\cdot||_1$ e $||\cdot||_2$, tali che la (1.17) valga per una norma, ma non per l'altra.*

Questo teorema suggerisce di definire:

Definitione 24 Siano $||\cdot||_1$ e $||\cdot||_2$ due norme sul medesimo s.l.n. X. Si dice che le due norme sono $\boxed{\textit{equivalenti}}$ se la condizione

$$\lim ||x_n - x_0||_1 = 0 \qquad \text{per ogni } x_0$$

implica la condizione

$$\lim ||x_n - x_0||_2 = 0 \qquad \text{per ogni } x_0,$$

e viceversa. ∎

E' importante poter decidere se due norme sono equivalenti. Il teorema seguente dà un test utile:

Teorema 25 *Sia X uno s.l.n. e siano $||\cdot||_1$ e $||\cdot||_2$ due norme su X. Esse sono equivalenti se e solo se esistono due numeri m ed M tali che*

$$m > 0\,, \quad \textit{e inoltre} \quad m||x||_1 \leq ||x||_2 \leq M||x||_1\,. \tag{1.18}$$

Osservazione 26 Si ricordi che i chiusi nella topologia di uno spazio metrico sono tutti e soli gli insiemi sequenzialmente chiusi. Dunque, le due norme sono equivalenti quando subordinano la stessa topologia. ∎

Abbiamo notato che esistono spazi lineari normati e non completi. E' bene sapere:

Teorema 27 *Sia X uno s.l.n. Si costruisce uno s.l.n. $\hat{X}$ con queste proprietà:*

- *$\hat{X}$ è completo;*
- *un s.spazio X_0 di $\hat{X}$, denso in $\hat{X}$, è isometricamente isomorfo ad X.*

L'ultima affermazione vuol dire che esiste un isomorfismo J tra X ed X_0 tale che

$$x \in X\,, x = J\tilde{x}\,,\ \tilde{x} \in X_0 \Longrightarrow ||x||_X = ||\tilde{X}||_{\hat{X}}\,.$$

1.2.1 Dimostrazioni posposte

Passiamo ora alla dimostrazione dei teoremi che abbiamo enunciato.

Dimostrazione del TEOREMA 19. Un esempio di spazio lineare normato non completo si costruisce come segue: i suoi elementi sono le funzioni $x(\cdot)$ continue su $[-1, 1]$ con l'usuale struttura lineare. La norma è

$$||x(\cdot)|| = \int_{-1}^{1} |x(s)|\,\mathrm{d}s\,.$$

E' facile vedere che la successione (x_n):

$$x_n(t) = \begin{cases} -1 & \text{se} \quad -1 \leq t \leq -\frac{1}{n} \\ nt & \text{se} \quad -\frac{1}{n} \leq t \leq \frac{1}{n} \\ 1 & \text{se} \quad \frac{1}{n} \leq t \leq 1 \end{cases}$$

è fondamentale. Per ogni t, la successione numerica $(x_n(t))$ converge al numero $\mathrm{sgn}(t)$. Essendo la successione (x_n) limitata, segue che

$$\lim \int_{-1}^{1} |x_n(s) - \mathrm{sgn}(s)|\,\mathrm{d}s = 0\,.$$

Se fosse anche $\lim x_n(\cdot) = \phi(\cdot)$ nello spazio in cui stiamo lavorando, ossia con $\phi(\cdot)$ continua avremmo

$$\int_{-1}^{1} |\phi(s) - \mathrm{sgn}(s)|\,\mathrm{d}s \leq \lim ||\phi - x_n|| + \lim \int_{-1}^{1} |x_n(s) - \mathrm{sgn}(s)|\,\mathrm{d}s = 0\,;$$

ossia, $\phi(t) = \mathrm{sgn}(t)$ q.o. $t \in [-1, 1]$. Ciò non può aversi perché $\phi(\cdot)$ dovrebbe essere continua mentre la funzione segno ha un salto. ∎

Dimostrazione del TEOREMA 21. Un esempio di s.l.n. non separabile è il seguente, che si indica col simbolo l^∞: i suoi elementi sono le successioni (x_n) limitate e la norma è

$$||(x_n)||_\infty = \sup |x_n|\,.$$

Sia S una successione di elementi di l^∞, $S = (x^{(n)})$. Notiamo che per ogni n il simbolo $x^{(n)}$ indica un elemento di l^∞, ossia una successione $(x_k^{(n)})$ di indice k. Mostriamo che l'immagine di S non è densa in l^∞ costruendo un elemento $y = (y_k) \in l^\infty$ che dista almeno 1 da ciascun elemento $x^{(n)}$ della successione S. Costruiamo y specificandone gli elementi y_k. Per scegliere y_1, primo elemento della successione y, si guarda il primo elemento della successione $x^{(1)}$ e si pone:

$$y_1 = \begin{cases} 0 & \text{se} \quad |x_1^{(1)}| > 1 \\ 2 & \quad\ \text{altrimenti}\,. \end{cases}$$

In questo modo, qualunque siano i successivi elementi di y, si ha

$$||y - x^{(1)}|| \geq 1\,.$$

Per scegliere y_k si guarda la k-ma successione $x^{(k)}$ e si pone

$$y_k = \begin{cases} 0 & \text{se} \quad |x_k^{(k)}| > 1 \\ 2 & \quad\ \text{altrimenti}\,. \end{cases}$$

Indipendentemente dai valori degli y_r con $r \neq k$, $||y - x^{(k)}|| \geq 1$. Dunque, la successione $y \in l^\infty$ che abbiamo costruita verifica

$$||y - x^{(k)}|| \geq 1$$

per ogni k; e quindi la successione S non è densa in l^∞. ∎

Dimostrazione del TEOREMA 23. Si consideri lo spazio lineare che indichiamo con X, i cui elementi sono funzioni continue su $[-1, 1]$, con le due norme

$$||x||_1 = \int_{-1}^{1} |x(s)|\,\mathrm{d}s\,, \qquad ||x||_\infty = \max_{[-1,1]} |x(t)|\,.$$

La norma $||\cdot||_\infty$ corrisponde alla *convergenza uniforme*.

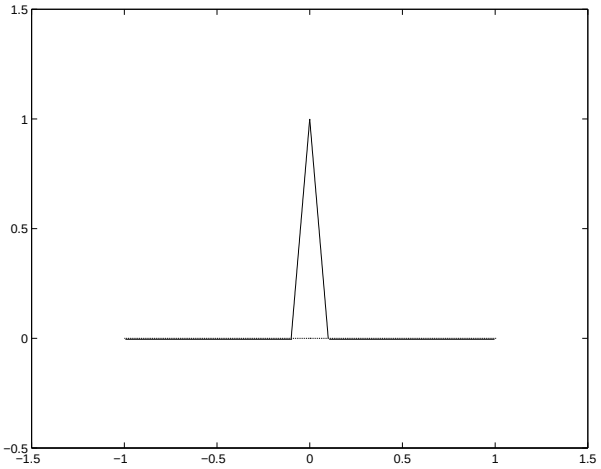
Sia (x_n) una successione in X. Se essa converge ad x_0 per $||\cdot||_\infty$, ossia se converge uniformemente, allora vale anche

$$\lim \int_{-1}^{1} |x_n(t) - x_0(t)| \, \mathrm{d}t = 0$$

e quindi anche $\lim ||x_n - x_0||_1 = 0$. Esitono però successioni convergenti in $||\cdot||_1$ ma non nella norma della convergenza uniforme. Sia $x_n(\cdot)$ la funzione il cui grafico è disegnato in figura 1.1:

$$x_n(t) = \begin{cases} 0 & t < -\frac{1}{n} \\ n\left(t + \frac{1}{n}\right) & -\frac{1}{n} \le t \le 0 \\ -n\left(t - \frac{1}{n}\right) & 0 \le t \le \frac{1}{n} \\ 0 & t > \frac{1}{n}\,. \end{cases}$$

Figura 1.1:



L'area del triangolo tende a zero, e quindi $\lim x_n = 0$ in $||\cdot||_1$; ma la successione (x_n) non è fondamentale in $||\cdot||_\infty$.

Dimostrazione del TEOREMA 25. Proviamo la condizione sufficiente. Si sappia che $x_n \to x_0$ in $||\cdot||_2$, ossia si sappia che

$$||x_n - x_0||_2 \to 0\,.$$

Dalla prima disuguaglianza in (1.18) segue

$$m||x_n - x_0||_1 \leq ||x_n - x_0||_2 \to 0\,,$$

ossia $x_n \to x_0$ in $||\cdot||_1$ perché $m > 0$.

Se si sa che $x_n \to x_0$ in $||\cdot||_1$, segue la convergenza in $||\cdot||_2$ dalla seconda delle (1.18).

Proviamo il viceversa. Si sappia che ogni successione convergente in $||\cdot||_2$ converge, ed ha il medesimo limite, anche in $||\cdot||_1$; ma supponiamo per assurdo che la prima disuguaglianza in (1.18) non valga per nessuna scelta di $m > 0$. In tal caso, per ogni n esiste x_n tale che

$$||x_n||_2 = 1 \qquad \text{e} \qquad ||x_n||_1 > n\,.$$

Definiamo $y_n = x_n/n$. E' ovvio che (y_n) tenda a zero in $||\cdot||_2$ mentre invece $||y_n||_1 \geq 1$ per ogni n; e quindi (y_n) non converge a zero rispetto a $||\cdot||_1$. Questo non si dà, per ipotesi, e quindi esiste $m > 0$ per cui vale la prima diseguaglianza in (1.18).

Analogamente si vede che se la convergenza in $||\cdot||_1$ implica la convergenza in $||\cdot||_2$ allora vale la seconda diseguaglianza in (1.18) per un certo M. ∎

Dimostrazione del TEOREMA 27. La costruzione di $\hat{X}$ è del tutto simile alla costruzione di Cantor dei numeri reali, e viene solamente accennata. Si considera l'insieme delle successioni fondamentali in X e in questo insieme si stabilisce la relazione di equivalenza

$$(x_n) \sim (y_n) \qquad \text{se} \qquad \lim(x_n - y_n) = 0\,.$$

Si vede facilmente che questa è una relazione di equivalenza. Indichiamo con $[(x_n)]$ la classe di equivalenza cui (x_n) appartiene. Si vede facilmente che l'insieme delle classi di equivalenza diviene uno spazio lineare se dotato delle operazioni

$$[(x_n)] + [(y_n)] = [(x_n + y_n)]\,, \qquad \alpha[(x_n)] = [(\alpha x_n)]\,.$$

In questo spazio, che indichiamo con $\hat{X}$, si introduce la norma

$$||\,[(x_n)]\,|| = \lim ||x_n||\,.$$

Questa definizione ha senso perché se (x_n) è fondamentale allora $(||x_n||)$ è una successione fondamentale di numeri, come si vede facilmente dal Teorema 15. Si vede inoltre che la norma così definita dipende dalla classe di equivalenza e non dal rappresentante, ed ha effettivamente le proprietà di una norma su $\hat{X}$.

Lo spazio X_0 è quelo delle classi di equivalenza che hanno un rappresentante costante. E' immediato verificare che X_0 ed X sono isometrici.

Notiamo ora che X_0 è denso in $\hat{X}$ perché ogni elemento $[(x_n)]$ di $\hat{X}$ si approssima mediante la successione $([(y_n)]_r)$ così costruita: $y_n = x_n$ se $n < r$; altrimenti $y_n = x_r$. ∎

1.3 Spazi prodotto

Ricordiamo che il *prodotto cartesiano* $X \times Y$ di due insiemi X ed Y è l'insieme delle *coppie ordinate* (x, y), il cui primo elemento è in X ed il secondo in Y.

Se X ed Y sono spazi lineari, si può rendere $X \times Y$ uno spazio lineare definendo

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)\,, \qquad \alpha(x, y) = (\alpha x, \alpha y)\,.$$

Se inoltre X ed Y sono dotati di norma, rispettivamente $||\cdot||_X$ e $||\cdot||_Y$, si possono definire norme su $X \times Y$. Per esempio

$$||(x, y)||_p = [||x||_X^p + ||y||_Y^p]^{1/p} \quad \forall p \geq 1\,, \qquad ||(x, y)||_\infty = \max\{||x||_X + ||y||_Y\}\,.$$

E' un fatto che non proviamo, il seguente: le norme precedenti su $X \times Y$ sono equivalenti.

Come casi particolari, si possono considerare i casi

$$Y = X \qquad \text{oppure} \qquad X = \Phi$$

perché, ricordiamo, il valore assoluto su $\mathbb{R}$ o il modulo su $\mathbb{C}$, sono norme. Tenendo presente questi due casi particolari, possiamo provare:

Teorema 28 *Le due trasformazioni*

$$\begin{aligned} &(x, y) \to x + y \quad &da \quad X \times X \to X \\ &(\alpha, x) \to \alpha x \quad &da \quad \Phi \times X \to X \end{aligned}$$

sono continue.

Dim. Per provare la prima affermazione, si fissa (x_0, y_0) e si nota che, lavorando con $||\cdot||_1$ su $X \times Y$, la disuguaglianza triangolare implica

$$\begin{aligned} &||(x + y) - (x_0 + y_0)||_X = ||(x - x_0) + (y - y_0)||_X \leq ||x - x_0||_X + ||y - y_0||_X \\ &= ||(x - x_0, y - y_0)||_1 = ||(x, y) - (x_0, y_0)||_1\,. \end{aligned} \tag{1.19}$$

Segue da qui la continuità della trasformazione $(x, y) \to x + y$.

Per provare la continuità del prodotto, fissiamo (α_0, x_0) con $\alpha_0 \neq 0$ e valutiamo:

$$\begin{aligned} &||\alpha x - \alpha_0 x_0||_X = ||(\alpha - \alpha_0)x + \alpha_0(x - x_0)||_X \\ &\leq |\alpha - \alpha_0| \cdot ||x||_X + |\alpha_0| \cdot ||x - x_0||_X\,. \end{aligned}$$

Si ha quindi

$$||\alpha x - \alpha_0 x_0||_X < \epsilon$$

se si sceglie

$$||x - x_0||_X \leq \epsilon/2\alpha_0$$

(e quindi anche $||x||_X \leq ||x_0||_X + \epsilon/2\alpha_0$) ed anche

$$|\alpha - \alpha_0| < \frac{\epsilon}{2[||x_0||_X + \epsilon/2\alpha_0]} .$$

In particolare quindi se si sceglie la coppia (α, x) in un intorno di (α_0, x_0) la cui esplicita determinazione si lascia al lettore.

Il caso $\alpha_0 = 0$ si lascia per esercizio. ∎

Osservazione 29 Si noti la stretta somiglianza della dimostrazione del teorema precedente con le usuali dimostrazioni sui limiti della somma e del prodotto di funzioni di una variabile; e si noti, da (1.19), che la somma è anche uniformemente continua; una proprietà che invece non vale per il prodotto. ∎

Il teorema precedente implica che in particolare sono continue le trasformazioni

$$x \longrightarrow x_0 + x\,, \qquad x \longrightarrow \alpha_0 x$$

(con x_0 ed α_0 fissati) dette rispettivamente *traslazioni* ed *omotetie*.

1.4 Gli esempi principali di spazi di Banach

Mostriamo ora gli esempi di spazi lineari normati che sono più importanti per le applicazioni. Essi sono tutti spazi completi, ossia spazi di Banach. La dimostrazione della completezza viene vista successivamente.

Sottolineiamo da subito che gli spazi che si incontrano nelle applicazioni hanno un simbolo standard, che indica sia lo spazio vettoriale che la norma su esso.

1.4.1 Gli esempi di spazi lineari normati

Introduciamo ora gli esempi più importanti di s.l.n. con i simboli comunemente usati per indicarli.

- Il simbolo $C([a,b])$ (più semplicemente $C(a,b)$).

Questo simbolo indica lo s.l.n. i cui elementi sono le funzioni continue sull'intervallo $[a,b]$ chiuso e limitato. La struttura lineare è quella usuale e la norma è quella della convergenza uniforme:

$$||x|| = \max_{[a,b]} |x(t)| .$$

In generale, se K è un compatto di Φ^n, con $C(K)$ si intende lo spazio delle funzioni continue su K, con norma $\max_K |x(t)|$.

I valori assunti dalla funzione sono numeri o vettori; talvolta sono matrici.

Osservazione 30 Sottolineiamo nuovamente che che ciascuno dei simboli seguenti indica uno spazio lineare con la norma che è ad esso associata nella definizione corrispondente. Quindi, per esempio, non useremo il simbolo $C(a,b)$ per indicare lo spazio delle funzioni continue con una norma integrale, per esempio quella introdotta nella dimostrazione del Teorema 23. Per questo, in quella dimostrazione abbiamo indicato genericamente con X tale s.l.n. ∎

- I simboli l^p, $1 \le p \le +\infty$.

Questi simboli indicano spazi di successioni. Gli elementi sono le successioni (x_n) tali che:

$$\begin{cases} \sum |x_n|^p < +\infty & \text{se } 1 \le p < +\infty, \quad ||(x_n)||_p = \left[\sum |x_n|^p\right]^{1/p} \\ \sup |x_n| < +\infty & \text{se } p = +\infty, \qquad ||(x_n)||_\infty = \sup |x_n|. \end{cases}$$

E' immediato verificare che gli spazi di successioni appena descritti sono s.l.n-ti rispetto alle usuali operazioni di somma elemento per elemento e di prodotto $\alpha(x_n) = (\alpha x_n)$. La verifica è diretta se $p = 1$ oppure $p = +\infty$ mentre fa uso della disuguaglianza di Minkovski per le serie se $1 < p < +\infty$.

- il simbolo c_0.

Si usa per per indicare lo s.l.n. delle successioni (x_n) tali che $\lim x_n = 0$. La norma in c_0 è

$$||(x_n)|| = \sup |x_n| .$$

Dunque c_0 è un s.spazio di l^∞.

- I simboli $\mathcal{L}^p(a,b)$ ed $L^p(a,b)$, $1 \le p \le +\infty$.

Il simbolo $\mathcal{L}^p(a,b)$ si usa per indicare gli s.l.n-ti i cui elementi sono le funzioni $f(\cdot)$ tali che, rispettivamente,

$$\int_a^b |f(x)|^p \, \mathrm{d}x \,, \qquad \operatorname{ess\,sup} |f(t)| < +\infty$$

(si ricordi il paragrafo **??**). Però le funzioni

$$f(\cdot) \to \left[\int_a^b |f(x)|^p \, \mathrm{d}x \right]^{1/p} , \qquad \operatorname{ess\,sup} |f(t)|$$

non sono norme su questi spazi: una funzione nulla in tutti i punti salvo uno ha nulli sia l'integrale che l'estremo superiore essenziale. Si rimedia a questo problema definendo una relazione di equivalenza, $f \sim g$, con

$$f \sim g \Longleftrightarrow f(t) = g(t) \quad \text{q.o. } t \in (a,b).$$

E' chiaro che due funzioni equivalenti hanno i medesimi integrali, finiti o meno, ed il medesimo estremo superiore essenziale. Si definisce una struttura lineare sull'insieme delle classi di equivalenza ponendo

$$[f] + [g] = [f+g] \,, \qquad \alpha[f] = [\alpha f]$$

(è facile vedere che questa definizione dipende solo dalle classi di equivalenza e non dai rappresentanti scelti per definire le operazioni). Quindi, per $1 \le p \le +\infty$, si definiscono i simboli $L^p(a,b)$ come gli spazi lineari delle classi di equivalenza di elementi di $\mathcal{L}^p(a,b)$ (stesso esponente p), dotati delle norme

$$|| \, [f] \, ||_p = \left[\int_a^b |f(x)|^p \, \mathrm{d}x \right]^{1/p} \qquad 1 \le p < +\infty \,, \qquad || \, [f] \, ||_\infty = \operatorname{ess\,sup} |f(t)| \,.$$

Si vede facilmente che queste funzioni dipendono solo dalla classe di equivalenza e non dai rappresentanti usati per calcolarle.

Osserviamo nuovamente che gli spazi sopra introdotti sono s.l.n-ti grazie alla disuguaglianza di Minkowski per gli integrali (si veda ancora il paragrafo **??**); e che per $p = 1$ è equivalente integrare $x(\cdot)$ o il suo valore assoluto.

Definizioni analoghe si danno per funzioni di più variabili. Se queste sono definite su un insieme K il simbolo che si usa è $L^p(K)$.

Osservazione 31 Mentre nella definizione di $C(K)$ l'insieme K deve essere compatto, tale condizione non è richiesta nella definizione di $L^p(K)$. ∎

Bisogna sapere:

Teorema 32 *Esistono successioni* $(x_n(\cdot))$ *di funzioni misurabili e limitate su un intervallo* $[a, b]$ *e tali che: 1) la successione di numeri* $(x_n(t))$ non converge *per* nessun *valore di* t*; ma 2) esiste una funzione integrabile* x_0 *tale che*

$$\lim_n \int_a^b |x_n(t) - x_0(t)|\, dt = 0\,.$$

Dim. Mostriamo un esempio di successione di funzioni con le proprietà del teorema, e con $x_0(t) \equiv 0$. Sia $[a, b] = [0, 1]$.

Costruiamo la successione in due passi; e quindi le funzioni verranno a dipendere da due indici n ed r. Ovviamente sarà possibile riscrivere la successione in modo da farla dipendere da un solo indice.

Al passo n dividiamo l'intervallo $[0, 1]$ in n intervalli uguali, I_0^n, I_1^n,..., I_n^{n-1}. definiamo quindi le funzioni $x_{n,r}$, $0 \le r \le (n-1)$ come segue:

$$x_{n,r}(t) = (-1)^n \begin{cases} (-1)^r & \text{se } t \in I_r^n \\ 0 & \text{altrimenti.} \end{cases}$$

L'integrale di $|x_{n,r}|$ vale $1/n$, e quindi vale la proprietà 2) con $x_0 = 0$; e si vede immediatamente che in ogni punto t infinite funzioni prendono il valore $+1$, infinite altre prendono il valore -1; e quindi vale anche la proprietà 1). ∎

Osservazione 33 Nonostante il teorema 32, nel seguito sempre confonderemo le funzioni con le loro classi di equivalenza, e quindi scriveremo f per indicare la classe di equivalenza $[f]$ di cui f è un rappresentante. ∎

- il simbolo $W^{1,p}(\Omega)$ (*spazi di Sobolev*).

Sia $\Omega \subseteq \mathbb{R}^n$ un aperto (limitato o meno). Si usa il simbolo $W^{1,p}(\Omega)$ per indicare lo spazio delle (classi di equivalenza di) funzioni $u(\cdot) \in L^p(\Omega)$ tali che esistono funzioni (ossia, classi di equivalenza) g_1, g_2,..., g_n in $L^p(\Omega)$, tali che

$$\int_\Omega u(x)\nabla\phi(x)\,\mathrm{d}x = \int_\Omega \left[\begin{array}{ccc} g_1(x) & \dots & g_n(x) \end{array}\right] \phi(x)\,\mathrm{d}x$$

per ogni funzione ϕ di classe C^∞ a supporto compatto in Ω.

La funzione g_i si chiama la i-ma *derivata parziale debole* di u.

Lo spazio $W^{1,p}(\Omega)$ si dota della norma

$$||u|| = \left[\int_\Omega |u(x)|^p\,\mathrm{d}x + \sum_{j=1}^n \int_\Omega |g_j(x)|^p\,\mathrm{d}x\right]^{1/p}$$

oppure della norma ad essa equivalente

$$||u||_{L^p(\Omega)} + \sum_{j=1}^n ||g_j||_{L^p(\Omega)}\,.$$

Se $n = 1$ allora esiste una sola derivata parziale debole e vale il seguente teorema, che non proviamo:

Teorema 34 *Se $n = 1$ ed $\Omega = (a, b)$, ogni funzione $u \in W^{1,p}(a, b)$ è assolutamente continua; e la sua derivata debole coincide con la derivata usuale, calcolata q.o.*

Se $n = 1$ una norma equivalente alle precedenti ed un po' più comoda da usare è

$$\left[|u(a)|^p + \int_a^b |u'(x)|^p\,\mathrm{d}x\right]^{1/p}\,.$$

Osservazione 35 Se $n > 1$ l'esistenza di tutte le derivate deboli non implica la continuità della funzione. ∎

Quando $p = 2$ si usa anche il simbolo $H^1(\Omega)$, invece di $W^{1,2}(\Omega)$. Questo simbolo non va confuso con quello che si usa per denotare gli spazi di Hardy, che ora definiamo.

- I simboli $H^p(D)$ (*spazi di Hardy*)

Col simbolo D indichiamo il disco $D = \{z \mid |z| < 1\}$ del piano complesso. Col simbolo $H^p(D)$, $1 \leq p < +\infty$, si indica lo spazio lineare i cui elementi sono le funzioni olomorfe $\phi(z)$ tali che

$$\sup_{r<1} \int_0^{2\pi} |\phi(re^{it})|^p \, \mathrm{d}t \qquad \text{con} \qquad ||\phi|| = \sup_{r<1} \left[\int_0^{2\pi} |\phi(re^{it})|^p \, \mathrm{d}t \right]^{1/p} .$$

Col simbolo $H^\infty(D)$ si indica lo spazio delle funzioni olomorfe limitate su D e norma $||\phi|| = \sup_D |\phi(z)|$.

Definizioni analoghe si danno sostituendo a D il semipiano Π_+,

$$\Pi_+ = \{z \mid \Re e\, z > 0\} .$$

In questo caso l'integrazione sulla circonferenza si sostituisce con l'integrazione sulle rette parallele all'asse immaginario (e il fattore $(1/2\pi)$ che compare nella definizione della norma si sostituisce col fattore $1/\pi$):

$$\sup_{x>0} \left[\frac{1}{\pi} \int_{-\infty}^{+\infty} |\phi(x+iy)|^p \, \mathrm{d}y \right]^{1/p} .$$

Invece, la norma di $H^\infty(\Pi_+)$ è

$$||\phi||_{H^\infty(\Pi_+)} = \sup_{x>0} |\phi(x+iy)| .$$

1.4.2 Le dimostrazioni della completezza

Ricordiamo che si chiama *Spazio di Banach* uno s.l.n. che è anche completo e che esistono s.l.n-ti che non sono completi, si veda il Teorema 19. Inoltre, al paragrafo 1.4.1 abbiamo presentato numerosi esempi di s.l.n-ti.

Vale:

Teorema 36 *Tutti gli s.l.n-ti presentati al paragrafo 1.4.1 sono completi.*

Per provare questo teorema dovremo esaminare separatamente i vari spazi del par. 1.4.1, fissare l'attenzione su una generica successione (x_n) fondamentale e associarle in qualche modo un elemento x_0 dello spazio, che intuiamo essere il limite della successione. Dobbiamo quindi provare che effettivamente $x_0 = \lim x_n$; ossia dobbiamo provare:

a) la funzione x_0 appartiene a X;

b) la funzione x_0 è limite di (x_n) nella norma di X.

Questo richiederà dimostrazioni diverse a seconda dello spazio che consideriamo.

Prima di studiare le singole dimostrazioni, ricordiamo:

Lemma 37 *Sia X uno s.l.n. e sia (x_n) una successione fondamentale in X. La successione (x_n) è limitata, ossia esiste M tale che*

$$||x_n|| < M \qquad \forall n\,.$$

Completezza degli spazi l^∞, c_0, $L^\infty(\Omega)$ e $C(K)$

Si ricordi che l'insieme K del simbolo $C(K)$ è un compatto di Φ^n mentre nessuna condizione si impone all'insieme Ω che compare nel simbolo $L^\infty(\Omega)$; e notiamo che sia l^∞ che c_0 sono spazi di funzioni sull'insieme $\Omega = \mathbb{N}$.

Le dimostrazioni della completezza di questi spazi sono tra loro simili, basate sul Teorema del doppio limite, provato in appendice.

Indichiamo con X uno degli spazi che stiamo considerando e sia (x_n) una successione fondamentale. Gli spazi che stiamo considerando sono accomunati da questo: sono spazi di funzioni su un certo insieme (indicheremo con t i suoi elementi) e la norma è definita in modo tale che se (x_n) è fondamentale allora ciascuna delle funzioni $(x_n(t))$ è una successione fondamentale di numeri; e quindi converge. Questa affermazione si interpreta se $X = L^\infty(\Omega)$ in questo modo: gli elementi della successione sono classi di equivalenza $[x_n]$ di funzioni.

Si fissa un rappresentante, che ancora indichiamo x_n, in ciascuna classe. Il limite della successione di numeri $(x_n(t))$ esiste q.o. su Ω.

Dunque, per ogni valore di t (o q.o. su Ω se $X = L^\infty(\Omega)$) si può definire

$$y(t) = \lim x_n(t) \ :$$

una funzione che si spera appartenga allo spazio che stiamo considerando e che sia limite di (x_n). Proviamo:

a) la funzione y appartiene ad X.

Questo è facile se X non è né $C(K)$ né c_0. Infatti in tal caso basta notare che y è una funzione limitata come limite puntuale di una successione (x_n) che, essendo fondamentale, è limitata: $|x_n(t)| < M$ per ogni t e per ogni n. Inoltre, se $X = L^\infty(\Omega)$, la funzione y può costruirsi a partire da un qualsiasi rappresentante delle classi di funzioni $[x_n]$, e ne è limite puntuale q.o.; e dunque è misurabile.

Se $X = C(K)$ allora la continuità di y seguirà dalla convergenza uniforme della successione di funzioni continue (x_n), che proveremo al punto **b),** grazie al Corollario 40.

Sia $X = c_0$. In questo caso ciascuna delle successioni $x_n = x_n(k)$ tende a zero per $k \to +\infty$, e, come proveremo al punto **b),** la successione stessa converge uniformemente ad $y = (y(k))$. Il *Teorema del doppio limite,* teorema 39, mostra che $\lim_k y(k) = 0$, ossia che $y \in c_0$.

Proviamo ora

b) la funzione y è limite della successione (x_n).

La funzione y è stata costruita come limite puntuale di $x_n(t)$. Mostriamo che in realtà il limite è nel senso della norma. Per questo fissiamo $\epsilon > 0$ ed un numero N_ϵ tale che se n, m sono maggiori di N_ϵ allora valga

$$||x_n - x_m|| < \epsilon \, .$$

Valutiamo ora $|y(t) - x_n(t)|$ per $n > N_\epsilon$ come segue:

$$|y(t)-x_n(t)| \le |y(t)-x_{n+r}(t)|+|x_{n+r}(t)-x_n(t)| \le |y(t)-x_{n+r}(t)|+\epsilon \, . \quad (1.20)$$

Questa disuguaglianza vale per ogni t e per ogni $r > 0$. Esiste r (dipendente da t) tale che $|y(t) - x_{n+r}(t)| < \epsilon$. Il numero r esiste perché $y(t)$ (per il valore fissato di t) è limite della successione di numeri $(x_{n+r}(t))$ (l'argomento precedente vale q.o. su Ω se $X = L^\infty(\Omega)$).

Dunque,

$$|y(t) - x_n(t)| \le \inf_r \{|y(t) - x_{n+r}(t)|\} + \epsilon < 2\epsilon \, .$$

Completezza dello spazio $H^\infty(D)$

Se (x_n) è una successione fondamentale in $H^\infty(D)$, essa è anche una successione fondamentale in $L^\infty(D)$ e quindi converge in $L^\infty(D)$ ad una funzione x_0, che è limitata. Per provare la completezza di $H^\infty(D)$ basta mostrare che x_0 è olomorfa. Ciò discende dal Teorema di Weierstrass, Teorema **??**.

Ciò prova che $x_0 \in H^\infty$, $x_0 = \lim x_n$, come si voleva.

Non proveremo la completezza degli spazi $H^p(D)$, proprietà che vale per ogni $p \in [1, +\infty]$.

Completezza degli spazi l^p con $1 \leq p < +\infty$

Indichiamo col simbolo $(x_n) = (x_n(k))$ una successione di elementi di l^p. Sia essa fondamentale. Da

$$|x_n(k) - x_m(k)| = \left[\sum_{r=0}^{+\infty} |x_n(r) - x_m(r)|^p\right]^{1/p}$$

segue che per ogni k la successione numerica $(x_n(k))$ è fondamentale e quindi convergente. Ciò induce a definire la successione x_0 con

$$x_0(k) = \lim_n x_n(k)\,.$$

Proviamo ora

a) la successione x_0 appartiene ad l^p.

Notiamo per questo che la successione (x_n) di l^p, essendo fondamentale, è limitata: esiste M indipendente da n e tale che

$$\left[\sum_{k=0}^{+\infty} |x_n(k)|^p\right]^{1/p} \leq M\,.$$

Segue che per ogni ν vale

$$\left[\sum_{k=0}^{\nu} |x_0(k)|^p\right]^{1/p} = \lim_n \left\{\left[\sum_{k=0}^{\nu} |x_n(k)|^p\right]^{1/p}\right\} \leq M\,.$$

Passando al limite rispetto a ν, si vede che $x_0 \in l^p$.

b) vale $x_0 = \lim x_n$ in l^p.

Si fissi $\epsilon > 0$ e sia $N = N(\epsilon)$ tale che se n, m superano N allora vale

$$||x_n - x_m|| < \epsilon \,.$$

Per n, m maggiori di N_ϵ e per ogni ν vale

$$\left[\sum_{r=0}^{\nu} |x_n(r) - x_m(r)|^p\right]^{1/p} \leq \left[\sum_{r=0}^{+\infty} |x_n(r) - x_m(r)|^p\right]^{1/p} < \epsilon \,.$$

Tenendo fermi ν ed n, si passi al limite per $m \to +\infty$. Si trova:

$$\left[\sum_{r=0}^{\nu} |x_n(r) - x_0(r)|^p\right]^{1/p} \leq \epsilon$$

e questa disuguaglianza vale per ogni ν. Prendendo l'estremo superiore rispetto a ν si vede che, quando $n > N(\epsilon)$, vale

$$||x_n - x_0||_{l^p} \leq \epsilon \,.$$

Questo volevamo provare.

Completezza degli spazi $L^p(\Omega)$, $p < +\infty$

In questa parte conviene distinguere tra gli elementi $[x]$ di $L^p(\Omega)$, ossia le classi di equivalenza, e i loro rappresentanti.

Sia $([x_n])$ una successione fondamentale di $L^p(\Omega)$. Il primo passo per mostrarne la convergenza è di costruire una funzione x_0, la cui classe di equivalenza è candidata ad essere limite di $([x_n])$. Gli esempi precedenti suggeriscono di costruire x_0 calcolando il limite puntuale dei valori di opportuni rappresentanti delle classi $[x_n]$. Questo metodo però non può applicarsi nel caso di $L^p(\Omega)$ perché si sa che una successione di funzioni che converge in media può non convergere in nessun punto, si veda il Teorema 32. Usiamo quindi una strategia diversa.

Ricordiamo una proprietà generale delle successioni fondamentali: una successione fondamentale che ha una sottosuccessione convergente è essa stessa convergente. Consideriamo una successione $([x_n])$ di elementi di $L^p(\Omega)$ e per ogni classe fissiamo un rappresentante che indichiamo x_n. In questo modo si trova una successione (x_n) di funzioni definite q.o. su Ω, e tali che per ogni $\epsilon > 0$ esiste $N = N(\epsilon)$ con questa proprietà: se n, è maggiore di N_ϵ allora per ogni m vale

$$\left[\int_\Omega |x_n(s) - x_{n+m}(s)| \, \mathrm{d}s\right]^{1/p} < \epsilon \,.$$

Facciamo vedere che una successione di funzioni (x_n) con tale proprietà ammette una sottosuccessione convergente q.o.

La sottosuccessione si costruisce con la regola seguente:

- si fissa $\epsilon = 1$ ed il numero $N(1)$. Si sceglie $n_1 = N(1) + 1$;
- si fissa $\epsilon = 1/2$ e si sceglie $n_2 = N(1/2) + 1$;
- in generale, con $\epsilon = 1/2^k$, si sceglie $n_k = N(1/2^k) + 1$.

Si considera quindi la sottosuccessione $(y_k) = (x_{N(k)+1})$.

La successione (y_k) gode della seguente proprietà:

$$||y_r - y_{r+1}||_{L^p(\Omega)} \leq \frac{1}{2^r}\,.$$

Proviamo che la successione di funzioni (y_k) converge q.o. E' strumentale a ciò introdurre la serie telescopica

$$\sum_{k=1}^{\infty} z_k\,, \qquad z_k = y_{k+1} - y_k\,.$$

Essendo

$$\sum_{k=1}^{m} z_k = y_{m+1} - y_1\,,$$

per provare la convergenza della successione, basta provare quella della serie.

La costruzione della successione (y_k) implica la convergenza assoluta della serie:

$$\sum_{k=1}^{\infty} \left[\int_\Omega |z_k(s)|^p \,\mathrm{d}s\right]^{1/p} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1\,. \tag{1.21}$$

Consideriamo la successione di funzioni

$$g_n(s) = \sum_{k=1}^{n} |z_k(s)|\,.$$

Questa successione è monotona crescente e quindi esiste

$$g(s) = \lim g_n(s) = \sum_{k=1}^{+\infty} |z_k(s)|$$

e inoltre, dalla disuguaglianza di Minkowski,

$$\left[\int_\Omega |g_n(s)|^p \,\mathrm{d}x\right]^{1/p} \leq \sum_{k=1}^{n} \left[\int_\Omega |z_k(s)|^p \,\mathrm{d}x\right]^{1/p} < \sum_{k=1}^{+\infty} \frac{1}{2^k} = 1\,.$$

Dunque, dal *teorema della convergenza monotona,* $|g(s)|^p$ è integrabile su Ω. Questo implica che la funzione $g(s)$ è finita q.o. su Ω (si veda il paragrafo **??**). Dunque, per q.o. $x \in \Omega$, la serie numerica

$$\sum_{k=1}^{+\infty} |z_k(s)|$$

converge; e dunque anche la serie numerica

$$\sum_{k=1}^{+\infty} z_k(s)$$

converge q.o. su Ω. Ciò mostra che la successione $(y_k(s))$ converge q.o. su Ω e permette di definire una funzione

$$x_0(s) = \lim y_k(s) = \sum_{k=1}^{+\infty} z_k(s) + y_1(s)\,.$$

La (1.21) mostra che $x_0 \in L^p(\Omega)$. Inoltre,

$$\left[\int_\Omega \left| x_0(s) - \sum_{k=1}^{n} z_k(s) - y_1(s) \right|^p \mathrm{d}s\right]^{1/p} \leq \sum_{k=n+1}^{+\infty} \left[\int_\Omega |z_k(s)|^p \,\mathrm{d}s\right]^{1/p} \leq \sum_{k=n+1}^{+\infty} \frac{1}{2^k} \to 0\,.$$

Ciò prova che la successione di classi di equivalenza $([y_k])$ converge alla classe di equivalenza $([x_0])$. Ciò è quanto volevamo provare.

Osserviamo che, in particolare, abbiamo anche provato un teorema sulla convergenza in media:

Teorema 38 *Sia $1 \leq p < +\infty$. Se una successione di funzioni (x_n) in $\mathcal{L}^p(\Omega)$ converge in media di ordine p ad x_0 allora esiste una sottosuccessione della (x_n) che converge ad x_0 q.o. su Ω.*

Completezza degli spazi di Sobolev

Ricordiamo: sia Ω è un aperto di $\mathbb{R}^n$ e siano $u \in L^p(\Omega)$ una funzione a valori scalari e $v \in L^p(\Omega)$ una funzione a valori vettori n–dimensionali. Si dice che

$$v = \nabla u$$

se per ogni ϕ di classe C^∞ ed a supporto compatto in Ω vale

$$\int_\Omega u(s) \nabla\phi(s)\,\mathrm{d}s = -\int_\Omega v(s)\phi(s)\,\mathrm{d}s\,.$$

In questo caso si dice che $u \in W^{1,p}(\Omega)$ e

$$||u||_{W^{1,p}(\Omega)} = \left[\int_\Omega |u(s)|^p \, \mathrm{d}s + \int_\Omega |v(s)|^p \, \mathrm{d}s\right]^{1/p} .$$

Dunque, se (u_n) è fondamentale in $W^{1,p}(\Omega)$, le due successioni (u_n) e $(v_n) = (\nabla u_n)$ sono fondamentali in norma $L^p(\Omega)$; e dunque convergenti,

$$u_n \to u_0 \,, \qquad v_n \to v_0 \,.$$

Passando al limite si trova quindi

$$\begin{aligned} -\int_\Omega v_0(s)\phi(s) \, \mathrm{d}s &= -\lim \int_\Omega \nabla u_n(s)\phi(s) \, \mathrm{d}s \\ &= \lim \int_\Omega u_n(s)\nabla\phi(s) \, \mathrm{d}s = \int_\Omega u_0(s)\nabla\phi(s) \, \mathrm{d}s \end{aligned}$$

per ogni funzione ϕ di classe $C^\infty(\Omega)$, a supporto compatto.

Dunque $u_0 \in W^{1,p}(\Omega)$ ha per gradiente v_0, ed (u_n) converge ad u_0 in $W^{1,p}(\Omega)$.

Ciò prova la completezza di $W^{1,p}(\Omega)$.

1.4.3 Teorema del doppio limite

Il *Teorema del doppio limite* riguarda una successione di funzioni (x_n), definite su un qualsiasi insieme Ω. Per esempio, per ogni n la x_n può essere una successione $(x_n(k))$, oppure può essere una funzione $x_n(t)$ definita su un intervallo $[a, b]$.

Indichiamo genericamente con t gli elementi di Ω.

Le funzioni prendono valori in uno spazio completo.

Supponiamo che per ogni n esista

$$\lim_t x_n(t) = L_n \,.$$

In questa scrittura può essere $t \to t_0$ oppure $t \to +\infty$, $|t| \to +\infty$ e simili. Per sottolineare questo, scriviamo genericamente

$$\lim_{t\to\alpha} x_n(t) = L_n \,. \tag{1.22}$$

Vale allora:

Teorema 39 (Teorema del doppio limite) *Per ogni n, esista* finito *il limite L_n definito in (1.22).*

Se (x_n) converge ad x_0 uniformemente *su Ω allora* esiste finito

$$\lim_{t\to\alpha} x_0(t) = L_0$$

e vale

$$L_0 = \lim_n L_n\,;$$

ossia vale

$$\lim_n \left[\lim_{t\to\alpha} x_n(t)\right] = \lim_{t\to\alpha}\left[\lim_n x_n(t)\right]\,.$$

Dim. Proviamo prima di tutto che $L_0 = \lim_{t\to\alpha} x_0(t)$ esiste finito. Dato che le funzioni prendono valori in uno spazio completo, basta provare che per ogni $\epsilon > 0$ esiste un intorno I_ϵ di α tale che:

$$t'\,, t''\in I_\epsilon \quad\Longrightarrow\quad ||x_0(t') - x_0(t'')|| < \epsilon\,.$$

Valutiamo $||x_0(t') - x_0(t'')||$ come segue:

$$||x_0(t') - x_0(t'')|| \le ||x_0(t') - x_n(t')|| + ||x_n(t') - x_n(t'')|| + ||x_n(t'') - x_m(t'')||\,.$$

Usando la convergenza uniforme, si scelgano n, m così grandi da avere

$$||x_0(t') - x_n(t')|| < \epsilon\,, \qquad ||x_n(t'') - x_m(t'')|| < \epsilon\,.$$

Con n ed m ormai fissati, si usi l'esistenza del limite finito

$$\lim_{t\to\alpha} x_n(t) = L_n\,.$$

Ciò implica che esiste un intorno I_ϵ di α tale che se $t'\in I_\epsilon$, $t''\in I_\epsilon$, allora

$$||x_n(t') - x_n(t'')|| < \epsilon\,.$$

Si noti che I_ϵ dipende anche da n, ma n è stato scelto, a sua volta, dipendente dal solo ϵ.

A questo punto sappiamo che esiste finito

$$L_0 = \lim_{t\to\alpha} x_0(t)\,.$$

Proviamo ora che

$$L_0 = \lim_n L_n\,.$$

Valutiamo $||L_0 - L_n||$ come segue:

$$||L_0 - L_n|| \leq ||L_0 - x_0(t)|| + ||x_0(t) - x_n(t)|| + ||x_n(t) - L_n|| \,.$$

Sia $\epsilon > 0$ fissato. Esiste un intorno I_ϵ di α tale che

$$t \in I_\epsilon \quad \Longrightarrow \quad ||L_0 - x_0(t)|| < \epsilon \,.$$

Scegliamo $t \in I_\epsilon$.
Fissiamo ora N_ϵ tale che

$$n > N_\epsilon \quad \Longrightarrow \quad ||x_0(t) - x_n(t)|| < \epsilon \,.$$

Il numero N_ϵ esiste, e non dipende da t, grazie alla convergenza uniforme.

Infine, da $L_n = \lim_{t\to\alpha} x_n(t)$, esiste un intorno $I'_\epsilon \subseteq I_\epsilon$ di α tale che, per $t \in I'_\epsilon$ vale

$$||x_n(t) - L_n|| < \epsilon \,.$$

L'intorno I'_ϵ dipende, oltre che da ϵ, anche da n. Comunque, da questa proprietà deduciamo che per ogni n esistono valori di $t \in I_\epsilon$ per cui

$$||x_n(t) - L_n|| < \epsilon$$

e quindi, per $n > N_\epsilon$, vale

$$||L_0 - L_n|| \leq 2\epsilon + \inf_{t\in\Omega} ||x_n(t) - L_n|| \leq 3\epsilon \,.$$

Ciò è quanto dovevamo provare. ∎

In particolare si ha:

Corollario 40 *Una successione (x_n) di funzioni continue definite su un qualunque insieme Ω ivi converga uniformemente ad x_0. Allora x_0 è continua su Ω.*

1.5 Sottospazi di spazi lineari normati

Sia X uno s.l.n., la cui norma indicheremo col simbolo $||\cdot||$. Sia Y un s.spazio di X. Ricordiamo che questo significa

$$\forall y_1\,,\ y_2 \in Y\,, \quad \forall \alpha\,,\ \beta \in \Phi \Longrightarrow \quad \alpha y_1 + \beta y_2 \in Y\,.$$

In particolare, Y stesso è uno spazio lineare e viene ad essere uno s.l.n. se ad Y si restringe la funzione norma definita su X. In tal caso diremo che Y è un s.spazio dello s.l.n. X e diremo che la norma su Y è quella *indotta* dalla norma di X. Notiamo che talvolta potrà essere necessario considerare su Y una norma diversa da quella indotta da X. Ciò va sempre esplicitamente specificato per evitare ambiguità. In caso contrario assumeremo sempre che la norma su Y sia quella indotta dalla norma di X.

Quando X ha dimensione finita, i sottospazi sono le controimmagini di $\{0\}$ sotto l'azione di trasformazioni lineari; e si sa che:

Teorema 41 *In dimensione finita, tutte le trasformazioni lineari sono continue; e quindi tutti i s.spazi sono chiusi.*

Invece, se X ha dimensione infinita, esso ammette sia s.spazi chiusi che non chiusi. Esempi di s.spazi chiusi sono ovviamente $\{0\}$ ed X stesso. Vediamo un esempio di s.spazio non chiuso.

Esempio 42 Si considera lo spazio $C(a,b)$ ed in esso il s.spazio Y dei polinomi. E' chiaro che Y non è chiuso perchè la restrizione ad $[a,b]$ della funzione esponenziale è limite uniforme di polinomi. Infatti, la serie di Taylor

$$\sum_{k=0}^{+\infty} \frac{t^k}{k!} = e^t$$

converge uniformemente sui compatti. ∎

E' importante sapere che non soltanto funzioni "regolari" possono approssimarsi uniformemente con polinomi. Anzi

Teorema 43 (Teorema di Weierstrass) *Sia f una funzione continua su un intervallo limitato e chiuso $[a,b]$,* a valori reali. *Esiste una successioni $\{p_n\}$ di* polinomi a coefficienti reali *che* converge *ad f,* uniformemente *su $[a,b]$.*

La dimostrazione di questo teorema è importantissima perché permette di introdurre il concetto di "identità approssimata". Ad essa è dedicato il paragrafo 1.5.1.

Una ulteriore proprietà puramente algebrica degli spazi lineari è la seguente: ogni loro s.spazio ammette *complementare*. Ricordiamo che un spazio lineare Z è un complementare di un spazio lineare Y (ambedue s.spazi di X) se:

$$Z \cap Y = \{0\}\,, \qquad Z + Y = X\,;$$

equivalentemente, se ogni elemento x di X si rappresenta in modo unico come somma di un elemento di Z e di uno di Y.

Quando si lavora con spazi normati, è naturale chiedere se tutti i s.spazi chiusi ammettano complementare, anch'esso chiuso. In dimensione finita la risposta è affermativa. Invece:

Teorema 44 *Esistono s.l.n-ti X, completi, dotati di s.spazi chiusi i quali sono privi di complementare chiuso.*

Quando un s.spazio Y è dotato di complementare Z, la dimensione (finita o meno) di Z si chiama la *codimensione* di Y.

Particolarmente importanti sono i sottospazi chiusi di codimensione 1, e anche i s.insiemi della forma

$$x_0 + H\,,$$

con H sottospazio chiuso di codimensione 1. Si chiamano tali s.insiemi gli *iperpiani* di X.

La dimostrazione del Teorema 44 consiste nella esplicita costruzione di un s.spazio chiuso privo di complementare, per esempio di l^p, con $1 \leq p < 2$ oppure con $p > 2$. La costruzione è alquanto macchinosa e viene omessa.

Osservazione 45 E' bene ricordare che gli spazi l^p, $1 \leq p \leq +\infty$, sono completi. Si veda il paragrafo 1.4.2 per la dimostrazione. E' anche bene sapere che ogni s.spazio chiuso di l^2 ammette complementare. Si veda il paragrafo **??** per la dimostrazione. ∎

1.5.1 Identità approssimate e dimostrazione del teorema di Weierstrass

Una famiglia di funzioni $\{h_\nu(s)\}$ che gode delle proprietà **0**–**3** seguenti si chiama *identità approssimata*. Le proprietà richieste sono:

0. per ciascun valore di ν, la funzione $s \to h_\nu(s)$ è integrabile su $\mathbb{R}$;

1. $h_\nu(s) \geq 0$ per ogni s;

2. $\int_{-\infty}^{+\infty} h_\nu(s)\,\mathrm{d}s = 1$ per ogni $\nu > 0$;

3. per ogni $\epsilon > 0$ si ha:

$$\lim_{\nu\to 0+} \int_{-\infty}^{-\epsilon} h_\nu(s)\,\mathrm{d}s = 0\,, \qquad \lim_{\nu\to 0+} \int_{+\epsilon}^{+\infty} h_\nu(s)\,\mathrm{d}s = 0\,.$$

La ragione del termine "identità approssimata" è espressa dal teorema seguente, che prova che la famiglia $\{h_\nu\}$ "approssima" l'identità rispetto alla convoluzione. Per una giustificazione più precisa si veda il paragrafo **??**.

Teorema 46 *Sia $\{h_\nu(s)\}$ un'identità approssimata e sia $f(x)$ una funzione uniformemente continua e limitata su $\mathbb{R}$. Sia $u(\nu, x)$ la funzione*

$$u(\nu, x) = \int_{-\infty}^{+\infty} h_\nu(x-s) f(s)\, ds = \int_{-\infty}^{+\infty} h_\nu(s) f(x-s)\, ds\,.$$

Vale:

$$\lim_{\nu\to 0+} u(\nu, x) = f(x)\,.$$

Il limite è uniforme su $\mathbb{R}$.

Dim. La proprietà **1.** dice che l'integrale di $h_\nu(s)$ vale 1. Dunque, si può scrivere

$$\begin{aligned} u(\nu,x) - f(x) &= \int_{-\infty}^{+\infty} h_\nu(s)[f(x-s) - f(x)]\,\mathrm{d}s \\ &= \int_{-\infty}^{-\epsilon} h_\nu(s)[f(x-s) - f(x)]\,\mathrm{d}s + \int_{\epsilon}^{+\infty} h_\nu(s)[f(x-s) - f(x)]\,\mathrm{d}s \\ &+ \int_{-\epsilon}^{+\epsilon} h_\nu(s)[f(x-s) - f(x)]\,\mathrm{d}s\,. \end{aligned}$$

Il numero ϵ deve ancora determinarsi.

Si fissa un numero σ e si usa l'uniforme continuità di f su $\mathbb{R}$ per scegliere $\epsilon = \epsilon_\sigma$ in modo tale che

$$|f(x-s) - f(x)| < \sigma \qquad s \in (-\epsilon, \epsilon)\,.$$

Le proprietà **1.** e **2.** implicano che l'integrale su $(-\epsilon, \epsilon)$ ha modulo minore di σ per ogni $\nu > 0$.

Fissato tale numero ϵ_σ si usino le proprietà **1.** e **3.** e la limitatezza di f su $\mathbb{R}$ per trovare ν_σ tale che, se $\nu \in (0, \nu_\sigma)$, ciascuno degli integrali rimanenti sia in modulo minore di σ. ∎

Le identità approssimate che si incontrano più spesso in pratica si costruiscono come segue: si assegna una funzione integrabile $h(s)$, non negativa. Dividendola per il suo integrale, si può assumere

$$\int_{-\infty}^{+\infty} h(s)\,\mathrm{d}s = 1\,.$$

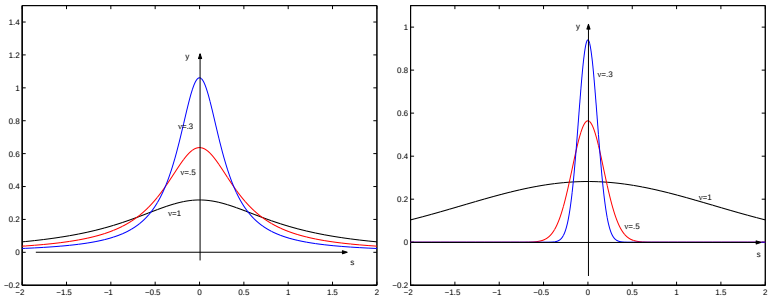
L'identità approssimata si ottiene ponendo

$$h_\nu(s) = \frac{1}{\nu} h(s/\nu)\,.$$

La figura 1.2 mostra i grafici di alcune delle funzioni $h_\nu(s)$ così ottenute a partire dalla funzione

$$\frac{1}{\pi(1+s^2)} \quad \text{(sinistra)} \qquad \frac{e^{-s^2}}{\sqrt{\pi}} \quad \text{(destra)}.$$

Figura 1.2: Esempi di identità approssimate



Osservazione 47 Implicitamente abbiamo supposto che ν sia un parametro continuo. Talvolta ν prende valori naturali, e l'identità approssimata è una successione di funzioni. In questo caso il limite per $\nu \to 0+$ si sostituisce col limite per $\nu \to +\infty$.

Diciamo infine che, così come si introducono le identità approssimate su $\mathbb{R}$, si possono anche introdurre le proprietà approssimate su $\mathbb{R}^n$. Le modifiche alla definizione (e alla costruzione a partire da una data funzione positiva) sono ovvie e vengono lasciate al lettore. ∎

Veniamo ora alla **dimostrazione del teorema 43.**

La dimostrazione del Teorema di Weierstrass è suggerita da certe considerazioni sull'equazione del calore, che conducono ad introdurre gli integrali

$$\frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{+\infty} e^{-s^2/4t} f(x-s)\,\mathrm{d}s = u(t,x)\,. \tag{1.23}$$

La famiglia delle funzioni

$$\frac{1}{\sqrt{4\pi t}} e^{-s^2/4t}$$

è un'identità approssimata (il cui parametro si indica con t perché nelle applicazioni all'equazione del calore indica il tempo). Essa è ottenuta a partire dalla funzione

$$h(s) = \frac{1}{\sqrt{\pi}} e^{-s^2} \;: \qquad h_t(s) = \frac{1}{\sqrt{4\pi t}} h\left(\frac{s^2}{4t}\right)\,.$$

Si veda la figura 1.2, a destra.

Sia ora f una funzione continua su $[a,b]$. Estendiamola in modo qualsiasi ad una funzione continua su $\mathbb{R}$, nulla per $x < a-1$ e per $x > b+1$. Indichiamo ancora con f la funzione così estesa.

E' chiaro che la funzione f è uniformemente continua su $\mathbb{R}$. Sia $u(t,x)$ la funzione definita in (1.23). Si vede facilmente che questa funzione è continua per $t > 0$ ed $x \in \mathbb{R}$.

Applichiamo il Teorema 46, ottenendo

$$f(x) = \lim_{t\to 0+} u(t,x) = \lim_{t\to 0+} \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{+\infty} e^{-(x-s)^2/4t} f(s)\,\mathrm{d}s\,.$$

Ora facciamo intervenire la condizione che f è nulla per $s < a-1$ e per $s > b+1$. Si ha così

$$u(t,x) = \frac{1}{\sqrt{4\pi t}} \int_{a-1}^{b+1} e^{-\frac{(x-s)^2}{4t}} f(s)\,\mathrm{d}s\,.$$

Si fissi $\sigma > 0$ e t_σ tale che

$$|u(t_\sigma, x) - f(x)| < \sigma/2\,. \tag{1.24}$$

Il numero t_σ esiste, grazie al Teorema 46.

La disuguaglianza (1.24) vale per ogni $x \in \mathbb{R}$. Limitiamoci però a considerare i valori di x in $[a-1, b+1]$. Si rappresenti

$$e^{-\frac{s^2}{4t_\sigma}} = \sum_{k=0}^{+\infty} \frac{1}{k!} \left(-\frac{s^2}{4t_\sigma}\right)^k$$

e la convergenza è uniforme sui compatti. Dunque, esiste N_σ tale che

$$\left| e^{-\frac{s^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{s^2}{4t_\sigma}\right)^k \right| < \frac{\sigma\sqrt{4\pi t_\sigma}}{4(b-a)M}\,, \qquad s \in [a-b-1, b-a+1]$$

con $M = \max|f|$.

Si ha quindi:

$$\begin{aligned}
&\left| u(t_\sigma, x) - \frac{1}{\sqrt{4\pi t_\sigma}} \int_{a-1}^{b+1} \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma}\right)^k f(s)\,\mathrm{d}s \right| \\
&= \frac{1}{\sqrt{4\pi t_\sigma}} \left| \int_{a-1}^{b+1} \left[e^{-\frac{(x-s)^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma}\right)^k \right] f(s)\,\mathrm{d}s \right| \\
&= \frac{1}{\sqrt{4\pi t_\sigma}} \left| \int_{x-b-1}^{x-a+1} \left[e^{-\frac{r^2}{4t_\sigma}} - \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{r^2}{4t_\sigma}\right)^k \right] f(x-r)\,\mathrm{d}r \right| < \sigma/2\,.
\end{aligned}$$

Combinando questa disuguaglianza con (1.24) si vede che la funzione $f(x)$ si approssima uniformemente su $[a, b]$ mediante i polinomi

$$\frac{1}{\sqrt{4\pi t_\sigma}} \int_{a-1}^{b+} \sum_{k=0}^{N_\sigma} \frac{1}{k!} \left(-\frac{(x-s)^2}{4t_\sigma}\right)^k f(s)\,\mathrm{d}s\,. \quad \blacksquare$$

Il teorema si estende a funzioni di n variabili, sostanzialmente con la medesima dimostrazione. Si ricorre per questo al seguente risultato:

Teorema 48 (Teorema di Tietze) *Ogni funzione continua su un compatti di $\mathbb{R}^n$ ammette estensione continua e limitata ad $\mathbb{R}^n$.*

Si fa quindi uso dell'identità approssimata

$$h_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{\frac{|x|^2}{4t}}. \tag{1.25}$$

Si noti che troncando la serie della funzione $e^{|x|^n}$ si trovano polinomi in $|x|$.

Osservazione 49 I punti $x \in \mathbb{R}^2$ si possono anche leggere come punti $x + iy$ del piano complesso e gli argomenti precedenti possono adattarsi al caso delle funzioni a valori complessi. Però, gli approssimanti che si ottengono troncando la serie di Taylor dell'identità approssimata (1.25) sono polinomi in $x^2 + y^2$ e non in $z = x + iy$. Non si trova quindi un'approssimazione mediante polinomi della variabile complessa z. ∎

1.6 La compattezza

Ricordiamo che un s.insieme K dello s.l.n. X si dice $\boxed{\textit{relativamente compatto}}$ quando ogni successione (x_n) a valori in K ammette s.successioni (x_{n_k}) convergenti. K si dice $\boxed{\textit{compatto}}$ quando è relativamente compatto e chiuso. Il teorema di Bolzano-Weierstrass può riformularsi dicendo che se $\Phi = \mathbb{R}$ oppure $\Phi = \mathbb{C}$ allora ogni s.insieme limitato e chiuso di Φ^n è compatto. Si ricordi che questa proprietà è cruciale per la dimostrazione del teorema di Weierstrass sull'esistenza di massimi e minimo.

Sfortunatamente, l'analogo del Teorema di Bolzano-Weierstrass non vale in spazi di Banach di dimensione infinita; anzi:

Teorema 50 *Sia X uno spazio normato. Se una sfera*

$$\{x \mid |x - x_0| = \epsilon\}$$

è compatta allora lo spazio ha dimensione finita.

Il Torema 50 implica in particolare:

Corollario 51 *Sia* $\dim X = +\infty$. *Se K è relativamente compatto, allora K non ha punti interni.*

Infatti, se K contiene punti interni esso contiene una palla chiusa, che deve essere compatta perché ogni s.insieme chiuso di un compatto è esso stesso compatto. Ciò non può darsi se $\dim X = +\infty$.

E' però vero che:

Teorema 52 *Se il s.insieme K di X è compatto, allora esso è limitato.*

Infatti, se K è illimitato esso contiene l'immagine di una successione (x_n) tale che $||x_n|| \to +\infty$; e si vede facilmente che (x_n) non ammette s.successioni convergenti.

Ricapitolando, in dimensione infinita gli insiemi compatti (rispetto alla topologia della norma) sono pochi (e ciò avrà conseguenze nefaste nei problemi di ottimizzazione) e difficili da caratterizzare. Di conseguenza i teoremi che caratterizzano gli insiemi compatti sono importanti. Il prototipo ed il più utile di essi è il *Teorema di Ascoli-Arzelà,* che caratterizza gli insiemi compatti di $C(a, b)$ (ricordiamo che con questo simbolo si intende in particolare che l'intervallo $[a, b]$ è limitato e chiuso).

Sia K compatto contenuto in $C(a, b)$. Abbiamo già notato, nel Teorema 52 che K deve essere limitato; ossia deve esistere $R > 0$ tale che

$$||x|| = \max_{[a,b]} |x(t)| < R \qquad \forall x \in K\,.$$

Trattandosi di limitatezza nella norma della convergenza uniforme, usa anche dire che K è *uniformemente* *limitato*.

Ricordiamo che ogni funzione $x \in C(a,b)$ è uniformemente continua perché $[a,b]$ è limitato e chiuso; ossia, per ogni $\epsilon > 0$ esiste un numero δ, che dipende da ϵ e dalla funzione x, tale che

$$|x(t') - x(t'')| < \epsilon\,, \qquad \forall t'\ t'' \in [a,b] \quad \text{per cui } |t' - t''| < \delta.$$

Si dice che l'insieme K è equicontinuo quando δ si può scegliere dipendente da ϵ ma non dall'elemento $x \in K$; ossia:

Definitione 53 *Si dice che l'insieme K è* equicontinuo *quando* per ogni

$\epsilon > 0$ esiste $\delta > 0$ tale che per ogni $x \in K$ e per ogni t', t'' in $[a,b]$ tali che $|t' - t''| < \delta$ si ha:

$$|x(t') - x(t'')| < \epsilon\,. \quad \blacksquare$$

Vale:

Teorema 54 (Teorema di Ascoli-Arzelà) *Gli insiemi relativamente compatti di $C(a,b)$ sono tutti e soli quelli uniformemente limitati ed equicontinui.*

1.6.1 Dimostrazioni posposte

Proviamo i teoremi enunciati.

Dimostrazione del TEOREMA 50. Basta mostrare una successione di elementi di norma 1 priva di sottosuccessioni fondamentali. Notiamo il seguente lemma (di immediata dimostrazione) che verrà utile anche in seguito:

Lemma 55 *Sia (x_n) una successione tale che $||x_n - x_m|| \geq \sigma > 0$ per ogni coppia n ed m. La successione (x_n) non ha s.successioni fondamentali.*

Per costruire una successione (x_n) con le proprietà richieste dal lemma, facciamo intervenire:

Lemma 56 (Lemma di Riesz) *Sia M un s.spazio chiuso di X, diverso da X. Per ogni $\epsilon > 0$, esiste un elemento $x \in X$ di norma 1, che dista da M* più *di $1 - \epsilon$.*

Dim. Ricordiamo che la distanza $d(x,M)$ di x da M è

$$d(x,M) = \inf\{||x - m|| \mid m \in M\}\,.$$

Essendo $M \neq X$, ed M chiuso, esiste $x_1 \neq M$ a distanza positiva da M:

$$d(x_1, M) = d > 0\,.$$

La definizione di distanza mostra che per ogni $\sigma > 0$ esiste $v_\sigma \in M$ tale che

$$||x_1 - v_\sigma|| < d(1+\sigma)\,.$$

Sia

$$y = x_1 - v_\sigma\,.$$

Ovviamente,

$$\delta = ||y|| = ||x_1 - v_\sigma|| \leq d(1+\sigma)$$

e inoltre $d(y, M) = d(x_1, M)$ perché $v_\sigma \in M$; ossia

$$d(y, M) = d(x_1, M) = d \geq \frac{1}{1+\sigma}||x_1 - v_\sigma|| = \frac{1}{1+\sigma}||y||\,.$$

In particolare,

$$||y|| \leq d(1+\sigma)\,. \tag{1.26}$$

Scegliamo ora

$$x_0 = \frac{y}{||y||}$$

e notiamo che

$$\begin{aligned} d(x_0, M) &= d\left(\frac{y}{||y||}, M\right) = \inf\{\left|\left|\frac{y}{||y||} - m\right|\right| \mid m \in M\} \\ &= \inf\{\frac{1}{||y||}\,||y - (m||y||)\,|| \mid m \in M\} = \inf\{\frac{1}{||y||}\,||x_1 - (v_\sigma + m||y||)\,|| \mid m \in M\} \\ &= \inf\{\frac{1}{||y||}\,||x_1 - m|| \mid m \in M\} = \frac{d}{||y||} > \frac{1}{1+\sigma}\,. \end{aligned}$$

L'ultima disuguaglianza discende da (1.26).

L'asserto segue se si è preventivamente scelto σ tale che $\frac{1}{1+\sigma} > 1 - \epsilon$. ∎

E' ora facile costruire una successione di elementi di norma 1, priva di s.successioni convergenti: si sceglie $x_1 \neq 0$ qualsiasi e si definisce

$$M_1 = \text{span}\{x_1\} = \{tx_1 \mid t \in \Phi\}\,.$$

Si sceglie quindi x_2, di norma 1, con

$$d(x_2, M_1) > \frac{1}{2}\,.$$

In particolare vale $||x_2 - x_1|| > 1/2$.

Definiti $x_1,\ldots,\ x_k$, si sceglie x_{k+1} di norma 1, distante almeno $1/2$ dallo spazio generato dai vettori $x_1,\ldots,\ x_k$.

Essendo $\dim X = +\infty$, $X \neq \text{span}\,\{x_1\,,\ \ldots\ \ x_k\}$ e quindi questa costruzione conduce ad una successione (x_n) i cui elementi distano due a due almeno $1/2$, e quindi priva di s.successioni convergenti. ∎

Dimostrazione del TEOREMA 52. La dimostrazione è analoga a quella che vale in dimensione finita: se A non è limitato, per ogni n esiste $a_n \in A$, con $||a_n|| > n$. La successione (a_n) è priva di s.successioni convergenti. ∎

Dimostrazione del TEOREMA 54. Proviamo la parte necessaria. Si è già detto che se K è compatto allora deve essere limitato. Proviamo che se è compatto in $C(a,b)$ allora esso è anche equicontinuo. Facciamo uso del lemma seguente:

Lemma 57 *Sia K compatto in uno spazio di Banach X e sia $\epsilon > 0$. Esiste un insieme finito di elementi $k_1,\ldots k_s$ di K, tali che*

$$K \subseteq \bigcup B(k_i,\epsilon)\,, \qquad B(k_i,\epsilon) = \{x \in X \mid ||x - k_i|| < \epsilon\}\,.$$

Dim. per assurdo sia K compatto e sia $\epsilon_0 > 0$ un numero tale che la proprietà non valga. Scelto un qualsiasi $x_1 \in K$, $B(x_1,\epsilon_0)$ non copre K; e quindi esiste $x_2 \in K$ che dista da x_1 più di ϵ_0. Ancora perchè ϵ_0 non soddisfa alla proprietà detta nel lemma, $B(x_1,\epsilon_0) \cup B(x_2,\epsilon_0)$ non copre K. Dunque esiste x_3 in K che dista più di ϵ_0 sia da x_1 che da x_2.

Iterando questo procedimento, si trova una successione (x_n) i cui punti distano l'uno dall'altro almeno ϵ_0, e quindi priva di s.successioni convergenti. Ciò contrasta con la compattezza di K. ∎

Proviamo ora che l'insieme K, compatto in $X = C(a,b)$, è equicontinuo. Si fissi $\epsilon > 0$ e si fissino $k_1,\ldots,k_r$ tali che

$$K \subseteq \bigcup B(k_i,\epsilon)\,.$$

Ciascuna delle funzioni k_i è una funzione uniformemente continua e quindi l'insieme delle k_i, che sono in numero finito, è equicontinuo: esiste $\delta > 0$ tale che se $|t' - t''| < \delta$ allora $|k_i(t') - k_i(t'')| < \epsilon$ per ogni i.

Sia ora $x \in K$ qualsiasi e k_{i_0} una funzione dell'insieme $\{k_1\,,\ldots\,,k_r\}$ che dista da x meno di ϵ. Valutiamo:

$$|x(t') - x(t'')| \leq |x(t') - k_{i_0}(t')| + |k_{i_0}(t') - k_{i_0}(t'')| + |k_{i_0}(t'') - x(t'')| < 3\epsilon\,.$$

Questa disuguaglianza vale per $|t' - t''| < \delta$ e per ogni $x \in K$. Notando che δ non dipende da $x \in K$ segue l'equicontinuità.

Proviamo ora la condizione sufficiente: proviamo che se $K \subseteq C(a,b)$ è sia limitato che equicontinuo allora K è relativamente compatto.

Fissiamo prima di tutto una successione iniettiva (t_n) la cui immagine è densa in $[a,b]$. Consideriamo quindi una qualunque successione (x_n) di elementi di K e consideriamo la successione di numeri $(x_n(t_1))$. Questa è una successione di numeri limitata perché K è limitato. Dunque ammette una s.successione convergente. Indichiamo col simbolo $x_n^{(1)}(t_1)$ questa successione di numeri e consideriamo la successione di funzioni $(x_n^{(1)})$. Valutiamo queste funzioni nel punto t_2 ottenendo la successione di numeri $(x_n^{(1)}(t_2))$. Estraiamo da questa una successione convergente, che indichiamo col simbolo $(x_n^{(2)}(t_2))$. Consideriamo quindi la successione di funzioni $(x_n^{(2)})$ e la successione di numeri $(x_n^{(2)}(t_3))$. Iteriamo il procedimento.

In questo modo si definiscono induttivamente le successioni di funzioni $(x_n^{(1)})$, $(x_n^{(2)})$,..., ciascuna delle quali è s.successione delle precedenti. Dunque, la successione $(x_n^{(k)}(t_r))$ è una successione di numeri che converge per ogni $r \leq k$. Inoltre, $(x_n^{(1)})$ è s.successione della (x_n).

Consideriamo ora la tabella seguente:

$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_4^{(1)}$	$x_5^{(1)}$	...
$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	$x_4^{(2)}$	$x_5^{(2)}$	...
$x_1^{(3)}$	$x_2^{(3)}$	$x_3^{(3)}$	$x_4^{(3)}$	$x_5^{(3)}$	...
$x_1^{(4)}$	$x_2^{(4)}$	$x_3^{(4)}$	$x_4^{(4)}$	$x_5^{(4)}$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Si ha:

- in ogni casella compare una delle funzioni della successione;
- gli elementi della prima riga costituiscono una s.successione della (x_n);
- ciascuna delle successive righe contiene gli elementi di una s.successione di quella che compare alla riga precedente;
- se si calcolano gli elementi della riga i–ma per $t = t_j$, con $j \leq i$, si trova una successione di numeri che converge.

Queste proprietà implicano che la successione diagonale $(x_n^{(n)})$ è una successione di funzioni con questa proprietà: le successioni di numeri $(x_n^{(n)}(t_k))$ convergono, per ogni k.

Notiamo che per ora abbiamo usato la sola limitatezza dell'insieme K. Usiamo ora l'equicontinuità per provare che la successione diagonale è fondamentale (e quindi convergente in $C(a,b)$ che, come si è detto, è uno spazio completo).

Sia $\epsilon > 0$. Si vuol provare l'esistenza di N_ϵ tale che se n, m sono maggiori di N_ϵ allora vale

$$||x_n - x_m|| < 3\epsilon \qquad \text{ossia} \quad |x_n(t) - x_m(t)| < 3\epsilon \quad \forall t \in [a,b].$$

Si fissi $\delta > 0$ tale che se $|t' - t''| < \delta$ allora ogni $x \in K$ verifica

$$|x(t') - x(t'')| < \epsilon\,.$$

Rappresentiamo l'intervallo $[a,b]$, che è limitato, come unione finita di intervalli di lunghezza minore di δ:

$$[a,b] = \bigcup_{s=1}^{\nu} [a_s, b_s]\,, \qquad b_s - a_s < \delta\,.$$

Per ciascun intervallo $[a_s, b_s]$, fissiamo uno dei punti della successione (t_n) che gli appartiene. Indichiamolo col simbolo $\bar{t}_s$.

Sia $t \in [a,b]$ qualsiasi e sia s tale che $t \in [a_s, b_s]$. Valutiamo, usando l'equicontinuità,

$$\begin{aligned} &|x_n^{(n)}(t) - x_m^{(m)}(t)| \le |x_n^{(n)}(t) - x_n^{(n)}(\bar{t}_s)| + |x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)| \\ &+|x_m^{(m)}(\bar{t}_s) - x_m^{(m)}(t)| \le 2\epsilon + |x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)|\,. \end{aligned}$$

Per ogni s, esiste N_s tale che, se n, m sono maggiori di N_s, vale

$$|x_n^{(n)}(\bar{t}_s) - x_m^{(m)}(\bar{t}_s)| < \epsilon$$

ed i punti $\bar{t}_s$ sono in numero finito e non dipendono dal punto t. Dunque, la dimostrazione si completa scegliendo $N_\epsilon = \max\{N_1\,, \ldots\,, N_\nu\}$. ∎

Il procedimento di estrarre la successione diagonale, dovuto a Cantor, va sotto il nome di *metodo* $\boxed{\textit{diagonale}}$ *di Cantor.*

1.7 Operatori lineari

Siano X ed Y due s.l.n-ti e sia f una trasformazione da X in Y. Non si richiede che il dominio di f sia tutto X. Per dire che f opera tra due s.l.n-ti, si dice che f è un *operatore*.

Si chiamano *funzionali* le trasformazioni che operano da X, s.l.n. sul campo scalare Φ, nel campo scalare Φ stesso.

Siano ora X ed Y due spazi lineari sul medesimo campo scalare. Si dice che f è un *operatore* *lineare* da X in Y quando

- dom$f$ è un sottospazio (non si richiede chiuso) di X;
- vale $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ per ogni x, y in X e per ogni α, β in Φ.

Quando si lavora con operatori lineari, invece della notazione $f(x)$ usa indicare l'operatore con una lettera maiuscola, per esempio F, A; e scrivere Fx, Ax invece di $F(x)$, $A(x)$; ossia si usa la "notazione moltiplicativa" nota dall'Algebra lineare.

Studiamo ora le proprietà degli operatori lineari.

1.7.1 Proprietà geometriche degli operatori lineari

Si ricordi che il *grafico* di una trasformazione $y = f(x)$ da un insieme X ad un insieme Y è l'insieme delle coppie

$$\{(x, y) \mid y = f(x)\}.$$

Invece, l'*immagine* ed il *nucleo*[1] sono rispettivamente s.spazi di Y e di X,

$$\text{im}f = \{y \mid y = f(x)\}, \qquad \ker f = \{x \mid f(x) = 0\}.$$

Si provi per esercizio:

Teorema 58 *Siano X ed Y due s.l.n-ti e sia f una trasformazione da X in Y. Vale:*

- *la trasformazione f è lineare se e solo se il suo grafico è un s.spazio di $X \times Y$;*
- *se f è lineare sia la sua immagine che il suo nucleo sono s.spazi.*

[1] il termine "nucleo" per indicare l'insieme degli zeri di una funzione si usa solamente se la funzione è lineare.

Notiamo ora:

Teorema 59 *Un operatore lineare da X in Y che ha immagine limitata è identicamente zero.*

Dim. Infatti, l'immagine di un operatore lineare è un sottospazio: questo è limitato se e solo se è il sottospazio 0. ∎

Col simbolo $B(x_0, r)$ indichiamo la *palla*

$$B(x_0, r) = \{x \mid ||x - x_0|| \leq r\}\,.$$

Lemma 60 *Sia $x_0 \in \mathrm{dom}A$. Vale*

$$B(x_0, r) \cap (\mathrm{dom}A) = x_0 + B(0, r) \cap (\mathrm{dom}A)\,.$$

Dim. Infatti, con $L = \mathrm{dom}A$,

$$B(x_0, r) \cap L = \{x \in L \mid ||x - x_0|| < r\}\,.$$

Sia $x \in B(x_0, r) \cap L$. Essendo x_0, x nel sottospazio L, $y = x - x_0$ è in L e verifica $||y|| < r$; ossia,

$$x = x_0 + y\,, \qquad y \in B(0, r) \cap L\,.$$

Dunque, $B(x_0, r) \cap L \subseteq x_0 + B(0, r) \cap L$. L'inclusione opposta si vede in modo analogo. ∎

Siano x_0 ed x_1 due punti di uno s.l.n. X. Il *segmento* di estremi x_0 ed x_1 è per definizione l'insieme dei punti

$$x = tx_0 + (1 - t)x_1\,, \qquad t \in [0, 1]\,.$$

Sia $K \subseteq X$ un insieme. Si dice che K è *convesso* quando il segmento che unisce due qualsiasi punti di K è contenuto in K; ossia quando

$$x_0\,,\ x_1 \in K\,, \quad t \in [0, 1] \Longrightarrow tx_0 + (1 - t)x_1 \in K\,.$$

Ovviamente, ogni palla è un insieme convesso. Invece, la *superficie* sferica

$$S(x_0, r) = \{x \mid ||x - x_0|| = r\}$$

non è convessa.

Gli insiemi convessi di $\mathbb{R}$ sono tutti e soli gli intervalli (limitati o meno).

Il risultato seguente è di ovvia dimostrazione:

Teorema 61 *Sia A lineare da X in Y. Se $K \subseteq \mathrm{dom}A$ è convesso in X, allora AK è convesso in Y.*

Le palle centrate in 0 sono insiemi che hanno in più una proprietà di simmetria: se $x \in B(0,r)$ allora $\alpha x \in B(0,r)$ per ogni α tale che $|\alpha| \leq 1$.

In generale, un insieme K si dice $\boxed{\textit{equilibrato}}$ se

$$|\alpha| \leq 1\,, \quad x \in K \quad \Longrightarrow \quad \alpha x \in K\,.$$

Si dice che $\tilde{K}$ è equilibrato rispetto ad un suo punto x_0 se

$$\tilde{K} = x_0 + K\,, \quad \text{con } K \text{ insieme equilibrato.}$$

Si vede immediatamente:

- un insieme di $\mathbb{C}$ equilibrato rispetto a z_0 e che contiene z' contiene anche il disco di centro z_0 e raggio $|z' - z_0|$. Affermazione analoga vale per gli insiemi equilibrati di $\mathbb{R}$, sostituendo i dischi con gli intervalli simmetrici rispetto a z_0. In particolare:

  Lemma 62 *Un insieme equilibrato di $\mathbb{R}$ oppure di $\mathbb{C}$ che è illimitato è uguale, rispettivamente, a $\mathbb{R}$ oppure a $\mathbb{C}$.*

- Se A è un operatore lineare da X in Y e se K è equilibrato in X, allora AK è equilibrato in Y; se K è equilibrato rispetto ad x_0 allora AK è equilibrato rispetto ad Ax_0.

Torniamo a considerare un operatore lineare A da uno s.l.n. X in un s.l.n. Y e consideriamo le palle $B(0,1)$, $B(0,r')$ e $B(x_0,r)$ di X. Supponiamo di sapere che $x_0 \in \mathrm{dom}\,A$. Si ha:

$$AB(0,r) = r(AB(0,1))\,, \qquad AB(x_0,r) = Ax_0 + r(AB(0,1))\,.$$

Di conseguenza,

Lemma 63 *Un operatore lineare A che è limitato su una palla è limitato su ogni altra palla.*

D'altra parte, con $S(0,1) = \{x \mid ||x|| = 1\}$,

$$\text{se } x \in B(0,r) \text{ allora } x = ||x|| \left(\tfrac{x}{||x||}\right) \text{ con } \tfrac{x}{||x||} \in S(0,1).$$

Dunque:

Lemma 64 *Vale:*

$$\sup_{||x||_X \leq 1} \{||Ax||_Y\} = \sup_{||x||_X = 1} \{||Ax||_Y\} .$$

In particolare, un operatore lineare A è limitato su una palla se e solo se è limitato sulla sfera $S(0,1)$.

Un operatore lineare che è limitato su una palla, non può "crescere troppo velocemente". Infatti:

Teorema 65 *Sia A lineare da X in Y e sia*

$$M_A = \sup_{||x||_X \leq 1} \{||Ax||_Y\} < +\infty .$$

Per ogni $x \in X$ vale

$$||Ax||_Y \leq M_A ||x||_X . \tag{1.27}$$

Viceversa, se esiste M tale che $||Ax||_Y \leq M||x||_X$ allora l'operatore lineare A è limitato su ogni palla.

Dim. Il viceversa è ovvio e quindi basta provare che se A è limitato su $B(0,1)$, allora vale la disuguaglianza (1.27). Basta considerare il caso $x \neq 0$. Se $x \neq 0$, $x/||x|| \in B(0,1)$ e quindi

$$\frac{1}{||x||} ||Ax|| = \left|\left| A \frac{x}{||x||} \right|\right| \leq M_A .$$

Questa è la disuguaglianza cercata. ∎

Dunque, se A è limitato su una palla, per esso vale

$$\lim_{||x||_X \to 0} ||Ax||_Y = 0$$

ed è quindi continuo in 0. Viceversa, il teorema della limitatezza locale (applicato alla funzione a valori reali $x \to ||ax||$) mostra che se A è continuo a 0 allora è limitato su una palla $B(0,r)$, e quindi su ogni altra palla. Vale dunque:

Teorema 66 *Un operatore lineare A da X in Y è continuo a zero se e solo se è limitato su una* qualsiasi *palla; equivalentemente, se e solo se esiste un numero M per cui vale*

$$||Ax||_Y \leq M||x||_X .$$

Ciò suggerisce un punto di partenza per lo studio della continuità degli operatori lineari. Prima di fare ciò, ricordiamo, dal Teorema 59 che l'unico operatore lineare e limitato da X in Y è quello identicamente zero; e quindi il termine "limitato" riferito ad operatori lineari rimane libero, e può essere usato con un significato diverso. Chiamiamo quindi operatore *lineare* $\boxed{\textit{limitato}}$ un operatore lineare che è limitato su una (qualsiasi) palla; ossia uno per il quale vale la disuguaglianza (1.27). Dunque:

Teorema 67 *Un operatore lineare A da X in Y è continuo a zero se e solo se è limitato.*

1.7.2 La continuità degli operatori lineari

E' noto:

Teorema 68 *Se* $\dim X < +\infty$ *e se la trasformazione A è lineare da X in Y allora A è continua.*

Gli esempi seguenti mostrano che, se X ha dimensione infinita, allora esistono sia operatori lineari continui che non continui. Un esempio banale di operatore lineare continuo è quello che ad ogni elemento di X associa lo 0 di Y, ossia il funzionale nullo. Un esempio meno banale è il seguente:

Esempio 69 Sia $X = C(a,b)$, $Y = \Phi$ e sia

$$\mathrm{dom}A = X\,, \qquad Ax = x(a)\,.$$

Essendo

$$||Ax - Ay||_\Phi = |x(a) - y(a)| \leq ||x - y||_X$$

si vede che A è addirittura uniformemente continuo. ∎

Mostriamo ora un esempio di funzionale lineare non continuo.

Esempio 70 Su $C(-1,1)$ si consideri l'operatore lineare Ψ definito da

$$\mathrm{dom}\Psi = \{x \text{ derivabili in } 1\}\,, \qquad \Psi x = x'(1)\,.$$

Quest'operatore, chiaramente lineare, non è continuo. Per mostrare ciò si consideri la successione delle funzioni x_n

$$x_n(t) = t^n/n\,.$$

Da $||x_n|| < 1/n$ segue che

$$\lim x_n = x_0 = 0$$

mentre per ogni n si ha:

$$\Psi x_n = 1\,;$$

ossia,

$$\lim \Psi x_n \neq \Psi x_0$$

e quindi Ψ non è continuo. ∎

Si noti che l'esempio precedente mostra che anche funzionali lineari importanti per le applicazioni possono essere discontinui; e, l'esempio specifico spiega perchè il problema della *derivazione numerica* è assai delicato.

Esempi di operatori lineari, rispettivamente continui e non continui, tra s.l.n-ti ambedue di dimensione infinita sono i seguenti:

Esempio 71 Sia $X = Y = C(a,b)$ e sia A con dominio uguale ad X,

$$(Ax)(t) = \int_a^t x(s)\,\mathrm{d}s\,.$$

E'

$$||Ax - Ay||_Y = \max_{[a,b]} \left| \int_a^t [x(s) - y(s)]\,\mathrm{d}s \right| \leq (b-a) \cdot ||x-y||_X\,;$$

e nuovamente si vede che l'operatore A è uniformemente continuo.

Un esempio di operatore lineare discontinuo è il seguente:

Sia $X = L^2(0,1)$, $Y = C([0,1])$. Il dominio di A sia lo spazio lineare delle classi di equivalenza dotate di rappresentante continuo. Sia

$$Ax = y\,, \qquad y(t) \equiv x(1)\,.$$

Le (classi di equivalenza delle) funzioni

$$x_n(t) = t^n$$

costituiscono una successione in $L^2(0,1)$, convergente a 0; mentre, per ogni x,

$$Ax_n \equiv 1\,.$$ ∎

Passiamo ora a studiare le proprietà degli operatori lineari che sono anche continui. Abbiamo notato che gli operatori lineari continui degli esempi precedenti sono anche uniformemente continui. Come ora vedremo, è questo un fatto generale.

Teorema 72 *siano X ed Y due s.l.n-ti e sia A lineare da X in Y. Vale:*

- *l'operatore A è continuo in ciascun punto del suo dominio se e solo se è continuo in un punto;*
- *l'operatore A è continuo se e solo se è uniformemente continuo;*
- *l'operatore A è continuo se e solo se è limitato.*

Dim. Proviamo che se A è continuo in un punto x_0 allora esso è continuo in qualunque altro punto x_1. Ovviamente, sia x_0 che x_1 devono appartenere al dominio di A. Sia $\epsilon > 0$ e sia $B(x_0, \delta)$ tale che

$$x \in B(x_0, \delta) \cap \mathrm{dom}A \Longrightarrow ||Ax - Ax_0|| < \epsilon \,.$$

Sia ora $x \in B(x_1, \delta) \cap \mathrm{dom}A$. Usando il Lemma 60 si vede che x può rappresentarsi come

$$x = x_1 + (x' - x_0)\,, \qquad \text{con} \quad x' \in \mathrm{dom}A \quad \text{e} \quad x' \in B(x_0, \delta)\,.$$

Dunque,

$$||Ax - Ax_1|| = ||A(x_1 - (x' - x_0)) - Ax_1|| = ||A(x' - x_0)|| = ||Ax' - Ax_0|| < \epsilon \,.$$

Ciò prova la continuità in x_1 e prova anche che il numero δ nel punto x_1 è il medesimo usato in x_0. Essendo x_1 arbitrario, si ha la continuità uniforme.

In particolare, la continuità in un qualsiasi punto x_0 equivale alla continuità in 0, e quindi alla limitatezza, si veda il Teorema 67. ∎

Frequentemente conviene verificare la continuità di un operatore verificando direttamente che è limitato.

Come si è visto, l'operatore lineare A è continuo se e solo se

$$M_A = \sup_{||x||_X \leq 1} ||Ax||_Y < +\infty \,.$$

Ci si può chiedere se l'estremo superiore sia in realtà un massimo. E' facile immaginare che l'estremo superiore non sarà un massimo se il dominio di A non è chiuso. Però:

Teorema 73 *Esistono s.l.n-ti* completi *X ed operatori lineari* continui *A con dominio* uguale *ad X e tali che*

$$\max\{||Ax||_Y \mid ||x||_X \leq 1\}$$

non *esiste.*

Dim. Si scelga $X = L^1(0,1)$ ed il funzionale

$$Lx = \int_0^1 sx(s)\,\mathrm{d}s\,.$$

E' immediato verificare che questo funzionale è continuo e che

$$||x|| \le 1 \quad \Longrightarrow \quad |Lx| \le 1\,;$$

anzi, si vede che

$$\sup\{||Lx||_Y\,,\ |\ ||x||_X \le 1\} = 1\,.$$

Infatti, se

$$x_n(t) = \begin{cases} 0 & \text{se} \quad 0 \le t \le 1 - 1/n \\ n & \text{se} \quad 1 - \frac{1}{n} \le t \le 1 \end{cases}$$

allora:

$$\lim Lx_n = 1\,.$$

Mostriamo che se $||x||_{L^1(0,1)} \le 1$ allora non vale $Lx = 1$. Sia infatti x tale che $Lx = 1$. In questo caso x non è zero q.o. e quindi esiste $\delta \in (0,1)$ tale che

$$\int_0^\delta |x(s)|\,\mathrm{d}s = \alpha > 0\,.$$

sottolineiamo che il numero δ si può scegliere *minore* di 1. Si scriva ora:

$$\begin{aligned} 1 &= \int_0^1 sx(s)\,\mathrm{d}s = \int_0^\delta sx(s)\,\mathrm{d}s + \int_\delta^1 sx(s)\,\mathrm{d}s \\ &\le \delta \int_0^\delta |x(s)|\,\mathrm{d}s + \int_\delta^1 |x(s)|\,\mathrm{d}s \\ &= \int_0^1 |x(s)|\,\mathrm{d}s - (1-\delta)\int_0^\delta |x(s)|\,\mathrm{d}s \end{aligned}$$

ossia

$$\int_0^1 |x(s)|\,\mathrm{d}s \ge 1 + (1-\delta)\alpha\,.$$

Dunque $||x|| > 1$ e quindi il massimo non viene raggiunto. ∎

Infine, introduciamo due operatori particolari, ed i loro simboli: col simbolo 0, riferito ad operatori che operano da X in Y, si intende l'*operatore* *nullo*, ossia quello che associa ad ogni $x \in X$ l'elemento 0 di Y. Col simbolo I, riferito ad operatori da X in X, si intende l'*operatore identità*, ossia quell'operatore che ad ogni x di X associa se stesso:

$$Ix = x\,.$$

1.7.3 Funzionali lineari continui ed iperpiani

Proviamo:

Teorema 74 *Sia Ψ un funzionale lineare definito su X non identicamente nullo, continuo o meno. Il suo nucleo ammette complementare di dimensione* 1.

Dim. Sia x_0 tale che $\Psi x_0 \neq 0$. Si noti che per ogni $x \in X$

$$n_x = x - x_0 \frac{\Psi x}{\Psi x_0} \in \ker \Psi \, .$$

Ogni $x \in X$ si rappresenta come

$$x = n_x + \alpha_x x_0 \, , \qquad \alpha_x = \frac{\Psi x}{\Psi x_0} \, .$$

Questa rappresentazione è unica perché se si ha anche

$$x = n'_x + \alpha'_x x_0$$

allora sottraendo si trova

$$0 = (n_x - n'_x) + (\alpha_x - \alpha'_x) x_0 \, .$$

Applicando il funzionale Ψ ai due membri si trova

$$0 = (\alpha_x - \alpha'_x) \Psi x_0$$

e quindi $\alpha_x = \alpha'_x$, perché $\Psi x_0 \neq 0$. E dunque si ha anche $n_x = n'_x$.

Ciò prova che $\{\beta x \, , \ \beta \in \mathbb{C} \}$ è uno spazio complementare di $\ker \Psi$. ∎

Osservazione 75 Il teorema precedente non richiede la completezza di X e nemmeno richiede la chiusura di $\ker \Psi$. Ciò nonostante, asserisce che $\ker \Psi$ ha un complementare di dimensione 1, e quindi chiuso. Mostriamo su un esempio la costruzione di tale complementare, nel caso in cui $\ker \Psi$ non sia chiuso. Sia X lo s.l.n. delle funzioni continue su $[-1, 1]$ e derivabili su $(-1, 1)$, dotato della norma del massimo. Questo spazio non è completo. Sia

$$\Psi x = x'(0) \, .$$

Il funzionale Ψ è ovunque definito, e non è continuo, come facilmente si vede riadattando gli argomenti presentati nell'esempio 70.

Il nucleo di Ψ è l'insieme delle funzioni di $C^1(-1,1)$ la cui derivata è nulla in 0. Non è difficile mostrare che questo spazio lineare è denso in $C(-1,1)$ e quindi in X. Ciò nonostante ammette complementare chiuso: ogni $x \in X$ si rappresenta in modo unico come

$$x(t) = [x(t) - x'(0)t] + x'(0)t$$

somma di un elemento di $\ker \Psi$ e di un multiplo di x_0, $x_0(t) = t \notin \ker \Psi$. ∎

Se Ψ è una qualsiasi trasformazione continua tra s.l.n-ti X ed Y, l'insieme degli zeri di Ψ è chiuso, come controimmagine continua di un chiuso. Il viceversa vale nel caso particolare dei funzionali lineari:

Teorema 76 *Sia Ψ un* funzionale *lineare su uno s.l.n. Esso è continuo se e solo se il suo nucleo è chiuso.*

Dim. Se il nucleo di Ψ è tutto X allora Ψ è costante e quindi continuo. Altrimenti, sia $x_0 \notin \ker \Psi$. Dato che $\ker \Psi$ è chiuso e diverso da X, esiste $\delta > 0$ tale che

$$\delta = \operatorname{dist}(x_0, \ker \Psi)\,.$$

Sia $B(x_0, \delta/2) = \{x \mid ||x - x_0|| < \delta/2\}$. L'immagine $\Psi B(x_0, \delta/2)$ di $B(x_0, \delta/2)$ è un insieme equilibrato di $\mathbb{R}$ oppure di $\mathbb{C}$, che non contiene 0, perché $B(x_0, \delta/2)$ non interseca $\ker \Psi$. Per il Lemma 62, $\Psi B(x_0, \delta/2)$ è limitato, e quindi Ψ è continuo. ∎

Si chiamano *iperpiani* i sottospazi chiusi di codimensione 1 e gli insiemi che si ottengono da essi per traslazione. Dunque:

Teorema 77 *Gli iperpiani sono tutti e soli gli insiemi della forma*

$$H = \{x \mid \Psi x = c\}$$

ove Ψ è un funzionale lineare e continuo.

Dim. Se ψ è un funzionale lineare continuo, il suo nucleo è chiuso, come contrimmagine dell'insieme chiuso $\{0\}$.

Viceversa, sia N un spazio lineare chiuso di codimensione 1. Costruiamo un funzionale lineare Ψ che ha N per nucleo, e che quindi è continuo per il Teorema 76.

Essendo N di codimensione 1, esiste $x_0 \notin N$ tale che ogni elemento di X si rappresenta in modo unico come

$$x = n + \alpha x_0\,, \qquad n \in N\,.$$

Il funzionale cercato è quello che ad x associa il numero α. ∎

Ossia, gli iperpiani sono gli insiemi di livello di funzionali lineari e continui.
Se Ψ è un funzionale lineare continuo, definiamo i due *semispazi*

$$H_+ = \{x \mid \Psi(x) > c\}\,, \qquad H_- = \{x \mid \Psi(x) < c\}\,.$$

I due semispazi H_+ ed H_- sono ovviamente disgiunti (perché le disuguaglianze sono strette). Le loro chiusure, che si chiamano anche semispazi chiusi, hanno in comune i punti dell'iperpiano $\{x \mid \Psi(x) = c\}$.

E' opportuno notare che le notazioni H_+ ed H_- non hanno significato intrinseco. Infatti, il funzionale Ψ che il teorema 77 associa ad H non è unico. Se, per esempio, $c = 0$, allora si identifica lo stesso iperpiano H sia col funzionale Ψ che col funzionale $-\Psi$; e lo scambio di Ψ con $-\Psi$ scambia tra di loro i due semispazi.

Notiamo infine che il teorema 76 vale per i funzionali. Non vale per generici operatori lineari, come mostra l'esempio seguente.

Esempio 78 Sia $X = Y = C(0,1)$ e sia

$$\operatorname{dom} A = C^1(0,1)\,, \qquad Ax = x'\,.$$

Argomenti analoghi a quelli visti all'esempio 70 mostrano che A non è continuo. Il suo nucleo è il s.spazio i cui elementi sono le funzioni costanti, e quindi è chiuso nonostante che l'operatore A non sia continuo. ∎

1.7.4 Lo spazio $\mathcal{L}(X, Y)$

Siano X ed Y due s.l.n-ti ed A, B due operatori lineari da X in Y. Definendo

$$\mathrm{dom}(A+B) = (\mathrm{dom}A) \cap (\mathrm{dom}B)\,, \qquad (A+B)x = Ax + Bx\,,$$

si ottiene chiaramente un operatore lineare $A+B$; ma in generale $\mathrm{dom}(A+B)$, $\mathrm{dom}A$ e $\mathrm{dom}B$ sono diversi e quindi non è possibile dare una struttura di spazio lineare all'insieme di tutti gli operatori lineari da X in Y. Per esempio, $B+(-B)$ non è in generale l'operatore 0, perchè l'operatore 0 è definito su X mentre $B+(-B)$ è solo definito su $\mathrm{dom}B$; e quindi $A+B+(-B)$ non è, in generale, l'operatore A. Se però ci si limita a considerare soltanto gli operatori lineari e continui si può ottenere di più. Vale infatti:

Teorema 79 *Sia A un operatore lineare e continuo da X in Y. Se Y è* completo *allora l'operatore A ammette un'unica estensione continua alla chiusura del suo dominio.*

Dim. Presentiamo i punti salienti della dimostrazione (del tutto analoga a quella che si usa per costruire l'estensione per continuità di funzioni reali), per mostrare il ruolo della completezza di Y.

Se x_0 è un punto della chiusura del dominio di A, esiste una successione (x_n) convergente ad x_0, $x_n \in \mathrm{dom}A$ (si noti che se $x_0 \in \mathrm{dom}A$ allora si può scegliere $x_n = x_0$ per ogni n).

Per la continuità di A si ottiene

$$||Ax_n - Ax_m||_Y = ||A(x_n - x_m)||_Y \leq M||x_n - x_m||_X\,. \tag{1.28}$$

La successione (x_n) è fondamentale in X, essendo per ipotesi convergente. Dunque, anche la successione (y_n), $y_n = Ax_n$ è fondamentale, però nello spazio Y. Essendo Y completo, si ha

$$\lim Ax_n = y_0\,.$$

Si definisce quindi

$$\tilde{A}x_0 = y_0.$$

Se (x'_n) è una seconda successione convergente ad x_0, vale

$$||Ax_n - Ax'_n||_Y \leq M||x_n - x'_n||_X \to 0$$

e quindi

$$\lim Ax_n = \lim Ax'_n\,;$$

ossia il valore y_0 dipende solo da x_0, e non dalla particolare successione (convergente ad x_0) scelta per calcolarlo. Dunque l'operatore $\tilde{A}$ che abbiamo costruito è un operatore univoco.

Ovviamente, $\tilde{A}$ estende A: se $x_0 \in \mathrm{dom}A$, scegliendo $x_n = x_0$ per ogni n si vede che

$$\tilde{A}x_0 = Ax_0 \,.$$

Si prova facilmente che l'operatore $\tilde{A}$ è lineare, ed è limitato.

Lasciamo per esercizio la dimostrazione della linearità e proviamo la limitatezza: se $x_n \to x_0$ ed $Ax_n \to y_0$,

$$||\tilde{A}x_0||_Y = \lim ||Ax_n||_Y \leq M \lim ||x_n||_X = M \cdot ||x_0||_X \tag{1.29}$$

(si ricordi che la norma è continua). ∎

Ricapitolando, $\tilde{A}$ è (l'unica) estensione continua di A alla chiusura del suo dominio. In particolare, se il dominio di A è denso in X, allora $\tilde{A}$ è definito su X.

Naturalmente, in pratica identificheremo A con $\tilde{A}$ (usando il simbolo più semplice A per ambedue gli operatori).

Da ora in poi, lavorando con operatori lineari e continui, assumeremo di averli estesi per continuità alla chiusura del dominio; e se non diversamente detto, assumeremo che il dominio sia X. Lavorando con operatori definiti su X, sia $A + B$ che αA (definito da $(\alpha A)x = \alpha(Ax)$ per ogni x) hanno dominio X e sono continui. Dunque, l'insieme degli operatori lineari continui su X è uno spazio lineare. Ciò che è più importante, esso può essere dotato di norma, come segue:

$$||A|| = \sup_{||x||_X \leq 1} ||Ax||_Y \,. \tag{1.30}$$

Si lascia al lettore la facile verifica che quella appena definita è una norma.

Conviene notare una conseguenza utile della definizione (1.30):

Corollario 80 *Per ogni $x \in X$ vale:*

$$||Ax||_Y \leq ||A|| \cdot ||x||_X \,. \tag{1.31}$$

Se anche Z è uno s.l.n. completo, e B è lineare e continuo da Y a Z, allora vale

$$||BA|| \leq ||B|| \cdot ||A|| \,. \tag{1.32}$$

Dim. Infatti, se $||\xi||_X \leq 1$, allora vale $||A\xi||_Y \leq ||A||$. Se $x \neq 0$ allora $\xi = x/||x||_X$ ha norma 1 e quindi

$$||A|| \geq ||A\xi||_Y = \frac{1}{||x||_X} \cdot ||Ax||_Y$$

ossia la (1.31).

La (1.31) mostra che:

$$||BAx|| \leq ||B|| \cdot ||Ax||_Y \leq ||B|| \cdot ||A|| \cdot ||x||_X \,.$$

Prendendo l'estremo superiore per $||x||_X \leq 1$, si trova la (1.32). ∎

La disuguaglianza (1.32) nel caso in cui $Z = Y = X$ mostra

$$||A^2|| \leq ||A||^2$$

e, più in generale,

$$||A^n|| \leq ||A||^n \,.$$

Conviene mostrare subito un modo equivalente per il calcolo di $||A||$:

Teorema 81 *Vale:*

$$||A|| = \min\{M \mid ||Ax||_Y \leq M||x||\} \,. \tag{1.33}$$

Dim. Indichiamo con M_0 il numero

$$M_0 = \inf\{M \mid ||Ax||_Y \leq M||x||\}$$

e proviamo che $M_0 = ||A||$, ossia che

$$M_0 = \sup_{||x||_X \leq 1} ||Ax||_Y \,.$$

Ciò in particolare mostra che l'estremo inferiore è un minimo.

La (1.31) implica che $M_0 \leq ||A||$. Per mostrare la disuguaglianza opposta, fissiamo $\delta > 0$ arbitrario. Vale, per ogni x,

$$||Ax||_Y \leq (M_0 + \delta) \cdot ||x||_X$$

e quindi, se $||x||_X \leq 1$,

$$||Ax||_Y \leq (M_0 + \delta) \cdot ||x||_X \leq M_0 + \delta \,.$$

Dunque, la disuguaglianza

$$||A|| \leq M_0 + \delta$$

vale per ogni $\delta > 0$. Passando all'estremo inferiore rispetto a δ si trova

$$||A|| \leq M_0 \tag{1.34}$$

e quindi l'uguaglianza (1.33). ∎

Possiamo ora tornare a considerare la diseguaglianza (1.29). Essa può ora scriversi

$$||\tilde{A}x_0|| \leq M||x_0|| \qquad M = \sup_{||x||\leq 1 \;\; x\in \mathrm{dom}A} ||Ax||$$

ossia

$$||\tilde{A}|| \leq M\,.$$

Però, $\tilde{A}$ estende A e quindi $||\tilde{A}|| \geq M$. dunque:

Corollario 82 *La norma dell'operatore $\tilde{A}$, estensione per continuità di A, è*

$$\sup_{||x||\leq 1 \;\; x\in \mathrm{dom}A} ||Ax||\,.$$

Ossia, il calcolo della norma di un operatore lineare continuo definito su X può effettuarsi a partire da una sua restrizione ad un sottospazio denso in X.

Si lascia per esercizio di provare la seguente ulteriore caratterizzazione di $||A||$:

Teorema 83 *Vale:*

$$||A|| = \sup_{||x||=1} ||Ax||_Y = \sup_{x\neq 0} \frac{||Ax||_Y}{||x||_X}\,.$$

Quando sia X che Y sono s.l.n-ti completi, ossia spazi di Banach, lo spazio degli operatori lineari e continui da X in Y, normato nel modo che abbiamo appena introdotto, si indica col simbolo $\boxed{\mathcal{L}(X,Y)}$ oppure $\boxed{\mathcal{B}(X,Y)}$. Due casi sono di uso particolarmente frequente e ad essi si riservano simboli speciali: il caso in cui $X = Y$ ed il caso, importantissimo, $X = \Phi$. Nel primo caso si usa il simbolo $\boxed{\mathcal{L}(X)}$ invece di $\mathcal{L}(X,X)$; nel secondo caso si usa il simbolo $\boxed{X^*}$ o $\boxed{X'}$ invece di $\mathcal{L}(X,\Phi)$. Lo spazio X^* si chiama lo $\boxed{\textit{spazio duale}}$ di X.

Infine, esaminiamo il problema della completezza dello spazio $\mathcal{L}(X,Y)$. La dimostrazione del teorema seguente usa la completezza dello spazio Y ma non quella dello spazio X. Per questa dimostrazione abbiamo bisogno di ricordare che una successione fondamentale è anche limitata; e che la successione (A_n) è limitata in $\mathcal{L}(X,Y)$ quando esiste un numero M, indipendente da n, tale che

$$||A_n|| < M\,.$$

Teorema 84 *Lo spazio $\mathcal{L}(X,Y)$ è completo.*

Dim. Dobbiamo mostrare che ogni successione (A_n) fondamentale in $\mathcal{L}(X,Y)$ è anche convergente. Sia allora (A_n) fondamentale. Per definizione di norma in $\mathcal{L}(X,Y)$, per ogni $\epsilon > 0$ esiste N_ϵ tale che per n, m maggiori di N_ϵ vale

$$||A_n - A_m|| \leq \epsilon \qquad \text{ossia} \qquad \sup_{||x||_X \leq 1} ||(A_n - A_m)x||_Y \leq \epsilon \,. \tag{1.35}$$

Segue che la successione $(A_n x)$ di elementi di Y è fondamentale per ogni x di norma minore o uguale ad 1; e quindi per ogni $x \in X$. Ciò permette di definire l'operatore B dato da

$$Bx = \lim A_n x \,.$$

Proviamo la linearità e la continuità di B e poi proviamo che $\lim A_n = B$. Da

$$B(\alpha x + \beta y) = \lim A_n(\alpha x + \beta y) = \lim \{\alpha A_n x + \beta A_n y\} = \alpha Bx + \beta By$$

segue la linearità. La continuità segue perché, se $||x|| \leq 1$,

$$||Bx|| = ||\lim A_n x|| = \lim ||A_n x|| \leq M||x||$$

con M indipendente da n perchè la successione (A_n), essendo fondamentale, è limitata.[2]

Mostriamo ora che $B = \lim A_n$, ossia che

$$\lim ||B - A_n|| = 0 \,. \tag{1.36}$$

E':

$$||B - A_n|| = \sup_{||x|| \leq 1} ||(B - A_n)x||_Y \,.$$

Sia $\epsilon > 0$ e sia N_ϵ tale che, per n, m maggiori di N_ϵ, valga (1.35). Fissato x con $||x|| < 1$, scriviamo

$$||Bx - A_n x||_Y = ||(B - A_m)x + (A_m - A_n)x||_Y \leq ||(B - A_m)x||_Y + \epsilon \,. \tag{1.37}$$

Notiamo che questa disuguaglianza vale per ogni x con $||x|| \leq 1$ e per tutti gli $m > N_\epsilon$.

Notiamo ora l'esistenza di un opportuno $m > N_\epsilon$ (dipendente sia da x che da ϵ) per cui vale anche

$$||(B - A_m)x|| < \epsilon \,.$$

Dunque, la (1.37) dà:

$$||Bx - A_n x||_Y \leq \inf_m \{||(B - A_m)x||_Y + \epsilon\} \leq 2\epsilon \qquad \forall n > N_\epsilon \,.$$

Ciò prova (1.36). ∎

[2] si noti che in questa dimostrazione si usa anche la continuità della norma.

1.7.5 Inversi di un operatore

In dimensione finita, l'equazione

$$Ax = y\,, \qquad x\,,\ y \in \mathbb{C}^n$$

con x, y vettori ed A trasformazione lineare, è risolubile per ogni y se e solo se

$$\ker A = \{0\}$$

ed in tal caso esiste l'operatore inverso A^{-1} di A che è lineare e che verifica ambedue le condizioni

$$\begin{cases} AA^{-1} &= I \\ A^{-1}A &= I\,; \end{cases}$$

anzi, se un operatore indicato con A^{-1} soddisfa una delle due uguaglianze precedenti esso soddisfa anche la seconda ed è l'operatore inverso di A, si veda il paragrafo (1.1.2)[3]. La situazione è più complessa in dimensione infinita. Infatti:

Esempio 85 Un operatore può avere nucleo nullo senza essere suriettivo. Per esempio, sia $X = Y = l^2$. Un operatore con tali proprietà è l'operatore S dato da:

$$Sx = S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, x_3, \ldots)\,.$$

Un operatore può avere nucleo non nullo ed essere suriettivo: per esempio,ancora con $X = Y = l^2$, Un operatore con tali proprietà è l'operatore T dato da:

$$Tx = T(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots)\,.$$

Dato che $\ker S = \{0\}$, si può definire un operatore A inverso di S, con $\mathrm{dom}\, A = \mathrm{im}\, S$ (si noti che, in quest'esempio particolare, A è la restrizione di T ad im S). Dunque, A verifica

$$ASx = x \qquad \forall x \in l^2$$

ma non verifica $SAx = x$ per ogni x di l^2.

Invece, si verifica facilmente l'esistenza di un operatore lineare B che verifica $TBx = x$ per ogni $x \in l^2$. Non vale però $BTx = x$ per ogni $x \in l^2$ nonostante che si stiano considerando trasformazioni da l^2 in sé. ∎

Queste considerazioni suggeriscono la seguente definizione:

[3]si ricordi che x ed y appartengono ambedue a $\mathbb{C}^n$. Se essi appartengono a spazi diversi allora l'ultima affermazione è falsa.

Definitione 86 Sia K un operatore lineare, limitato o meno, da X in Y. Se un operatore A, con dominio $\operatorname{im} K$, verifica

$$AKx = x \qquad \forall x \in X\,,$$

l'operatore A si chiama *inverso sinistro* di K. Se un operatore lineare B verifica

$$KBy = y \qquad \forall y \in Y$$

l'operatore B si chiama *inverso destro* di K.

Un operatore che è sia inverso destro che sinistro di K si chiama *inverso* di K e si indica col simbolo K^{-1}. ∎

Si noti che la definizione di linearità è stata esplicitamente richiesta nella definizione di inverso destro, ma non in quella di inverso sinistro. Ciò perché:

Teorema 87 *Un operatore lineare K da X in Y ammette inverso sinistro se e solo se* $\ker K = \{0\}$. *In tal caso l'inverso sinistro è unico, ed è lineare.*

Dim. Se esiste l'inverso sinistro A di K allora per ogni $x \in \mathrm{dom}K$ vale:

$$x = AKx$$

e quindi se $Kx = 0$ si ha anche $x = 0$. Dunque, l'esistenza dell'inverso sinistro implica $\ker K = \{0\}$, ossia che K è iniettivo.

Viceversa, sia $\ker K = \{0\}$. Allora K, essendo lineare, è (univoco e) iniettivo. Il suo inverso sinistro è l'operatore che a Kx associa x.

La definizione di inverso sinistro mostra che il suo grafico in $Y \times X$ è l'insieme $\{(Kx, x) \mid x \in X\}$. Questo è un s.spazio percé l'operatore K è lineare; e quindi anche l'inverso sinistro, avendo per grafico un s.spazio, è lineare, si veda il teorema 58. ∎

Per contrasto si noti che l'inverso destro non è unico e che possono anche esistere operatori B non lineari che verificano l'uguaglianza $KBy = y$ per ogni y:

Esempio 88 Sia $X = Y = l^2$. Sia T l'operatore introdotto nell'esempio 85. Per ogni numero naturale n e per ogni numero reale α, definiamo $B_{n,\alpha}$ ponendo

$$B_{n,\alpha}(y_1, y_2, y_3, \ldots) = (\alpha y_1^n, y_1, y_2, y_3, \ldots)\,.$$

L'operatore $B_{n,\alpha}$ è non lineare se $n > 1$ ed è lineare se $n = 1$. Per ogni scelta di n e di α si ha: $TB_{n,\alpha}y = y$ per ogni $y \in l^2$.

In particolare si vede la non unicità dell'inverso destro perfino con la condizione che esso debba essere lineare. ∎

Le considerazioni precedenti suggeriscono di privilegiare lo studio dell'inverso sinistro. Approfondendo tale studio, notiamo che non c'è relazione tra continuità di un operatore e continuità del suo inverso sinistro, come ora vediamo:

Esempio 89 Mostriamo l'esempio di un operatore continuo ed invertibile, con inverso sinistro non continuo.

Sia $X = Y = l^2$ e sia A l'operatore definito da

$$A(x_1, x_2, x_3, x_4, \ldots) = (x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \frac{1}{4}x_4, \ldots) .$$

E' facile vedere che l'operatore A è continuo e iniettivo, e che il suo inverso sinistro, definito su $\operatorname{im} A$, è l'operatore non continuo B:

$$B(y_1, y_2, y_3, \ldots) = (y_1, 2y_2, 3y_3, \ldots) .$$

Per avere un esempio di operatore illimitato il cui inverso è limitato, si scambino i ruoli degli operatori A e B appena introdotti. ∎

E' facile dare un test per la limitatezza dell'inverso sinistro: sia B inverso sinistro di A. Per definizione, l'operatore B è limitato se e solo se esiste $\rho > 0$ tale che

$$||By||_X \leq \rho ||y||_Y \qquad \forall y \in \operatorname{dom} B .$$

Un elemento y è in $\operatorname{dom} B$ se e solo se esiste x per cui $y = Ax$ e quindi la disuguaglianza precedente equivale a

$$m||x||_X \leq ||Ax||_Y \tag{1.38}$$

(con $m = 1/\rho > 0$). Questa condizione implica anche che $\ker A = \{0\}$. Dunque:

Teorema 90 *L'operatore A ammette inverso sinistro continuo se e solo se esiste $m > 0$* (disuguaglianza stretta!) *per cui vale (1.38).*

La condizione (1.38) non è di facile verifica ed in generale non è facile costruire l'espressione esplicita dell'inverso. Un caso semplice ed importante è il seguente:

Teorema 91 *Sia $A \in \mathcal{L}(X)$, con $||A|| < 1$* (disuguaglianza stretta!) *e si consideri l'operatore $I - A$. L'operatore $I - A$ è iniettivo e suriettivo, ossia invertibile, ed è*

$$(I - A)^{-1} = \sum_{k=0}^{+\infty} A^k . \tag{1.39}$$

Dim. Sia $q < 1$ tale che $||A|| < q$, ossia tale che $||Ax|| \leq q||x||$ per ogni x. Vale:

$$||(I-A)x|| = ||x - Ax|| \geq \Big|\, ||x|| - ||Ax|| \,\Big| = ||x|| - ||Ax|| \geq (1-q)||x||\,.$$

Dunque, l'operatore $I - A$ ammette l'inverso sinistro continuo, per il Teorema 90. Per trovare un'espressione per l'inverso sinistro consideriamo la serie in (1.39) (suggerita dalla serie geometrica!) Mostriamo prima di tutto che essa converge in $\mathcal{L}(X)$. Per questo consideriamo la successione delle somme parziali

$$S_n = \sum_{k=1}^{n} A^k\,.$$

Si ha:

$$||S_n - S_{n+m}|| = \left|\left| \sum_{k=n+1}^{n+m} A^k \right|\right| \leq \sum_{k=n+1}^{n+m} ||A^k|| \leq \sum_{k=n+1}^{n+m} q^k\,.$$

Essendo $q \in [0,1)$ si ha che la successione delle somme parziali è fondamentale e quindi convergente in $\mathcal{L}(X)$.

Notiamo che, in particolare,

$$||A^k|| \leq ||A||^k \leq q^k \qquad \text{così che} \quad \lim_k A^k = 0\,.$$

Per definizione

$$\left(\sum_{k=1}^{+\infty} A^k\right) x = \left(\lim_k \sum_{k=1}^{n} A^k\right) x = \lim_k \left(\sum_{k=1}^{n} A^k x\right) = \sum_{k=1}^{+\infty} A^k x\,.$$

Dunque:

$$\left(\sum_{k=0}^{+\infty} A^k\right)(I-A)x = \sum_{k=0}^{+\infty} A^k x - \sum_{k=0}^{+\infty} A^{k+1} x$$
$$= \lim \left\{ \sum_{k=0}^{n} A^k x - \sum_{k=0}^{n} A^{k+1} x \right\} = \lim\{x - A^{k+1}x\} = x$$

e quindi la serie in (1.39) rappresenta l'inverso sinistro di $(I - A)$. Con calcoli analoghi si vede che è anche inverso destro, e quindi inverso. In particolare segue che $I - A$ è suriettivo. ∎

La serie (1.39) si chiama *serie di* *von Neumann*.

Sottolineiamo ora che un operatore ammette inverso quando ammette sia inverso destro che sinistro; in particolare quando è sia iniettivo che suriettivo. Conviene indebolire un po' questa definizione.

Definitione 92 Sia A lineare da X in Y con dominio denso in X e con immagine densa in Y. Sia A iniettivo. L'inverso sinistro di A, definito su $\operatorname{im} A$ ed a valori in $\operatorname{dom} A$, si chiama $\boxed{\textit{inverso}}$ di A.

Il simbolo A^{-1} si usa anche per indicare l'inverso di A, nel senso generalizzato che abbiamo ora definito. ∎

Notiamo infine:

Teorema 93 *Siano A e B operatori lineari ambedue invertibili con*

$$\operatorname{im} B \subseteq \operatorname{dom} A\,.$$

Siano continui gli operatori inversi A^{-1} e B^{-1}. Allora $(AB)^{-1}$ esiste e vale

$$(AB)^{-1} = B^{-1}A^{-1}\,.$$

Dim. Immediato, notando che $\operatorname{dom} AB = \operatorname{dom} B$ e che

$$B^{-1}A^{-1}ABx = x \qquad \forall x \in \operatorname{dom} A\,.$$ ∎

1.8 Il teorema di Baire e le sue conseguenze

Una semplice osservazione che vale in $\mathbb{R}^2$ è la seguente: gli iperpiani per 0 in questo caso sono rette di equazione $y = mx$ oppure $x = 0$. Esse sono parametrizzate dal punto in cui intersecano la circonferenza $x^2 + y^2 = 1$. Dunque $\mathbb{R}^2$ non è unione di una famiglia numerabile di rette per 0; e questa osservazione si generalizza a rette qualsiasi, ed a dimensione $n > 2$. Vediamo come questo risultato si estende ad un generico spazio di Banach.

Proveremo il teorema seguente, non ovvio nemmeno in dimensione finita:

Teorema 94 (di Baire) *Sia X uno spazio di Banach e sia (A_n) una successione di s.insiemi di X, ciascuno dei quali è chiuso e privo di punti interni. Allora,*

$$\bigcup A_n \neq X\,.$$

Rimandando alla fine di questo paragrafo la dimostrazione, illustriamo varie conseguenze importanti di questo teorema.

Notiamo prima di tutto che un s.spazio di X, diverso da X stesso, non ha punti interni. Dunque vale in un generico spazio di Banach la proprietà che abbiamo notato sopra per $\mathbb{R}^2$, che conviene enunciare come segue:

Teorema 95 *Sia (X_n) una successione di s.spazi di uno spazio di Banach X. Se $X = \cup X_n$ allora esiste n_0 tale che $X = X_{n_0}$.*

Il Teorema di Baire è un potente strumento per lo studio delle proprietà degli operatori lineari tra due spazi di Banach X ed Y. Esso talvolta si usa direttamente; più spesso interviene grazie ai quattro teoremi seguenti. Il primo che presentiamo va sotto il nome di *Teorema di Banach-Steinhaus*. Esso concerne s.insiemi $\mathcal{A}$ di $\mathcal{L}(X, Y)$. Ricordiamo che $\mathcal{L}(X, Y)$ è uno spazio normato e quindi ha senso investigare quando $\mathcal{A}$ è un s.insieme limitato di $\mathcal{L}(X, Y)$. Ciò avviene se esiste M tale che $||A|| \leq M$ per ogni $A \in \mathcal{A}$.

Fissiamo ora un qualsiasi elemento $x \in X$ e consideriamo l'insieme dei "valori" Ax, $A \in \mathcal{A}$. Questo è un s.insieme di Y che è limitato se l'insieme $\mathcal{A}$ è limitato in $\mathcal{L}(X, Y)$. Infatti, $||Ax||_Y \leq ||A|| \cdot ||x|| \leq M \cdot ||x||$ per ogni $A \in \mathcal{A}$. Il teorema di Banach-Steinhaus permette di invertire questa proprietà:

Teorema 96 (di Banach-Steinhaus) *Sia $\mathcal{A}$ un s.insieme di $\mathcal{L}(X, Y)$. Supponiamo che per ogni $x \in X$ esista un numero M_x tale che*

$$||Ax|| \leq M_x \qquad \forall A \in \mathcal{A}\,. \tag{1.40}$$

(Sottolineiamo: M_x indipendente *da $A \in \mathcal{A}$). In questo caso $\mathcal{A}$ è un s.insieme limitato di $\mathcal{L}(X, Y)$.*

Dim. Indichiamo con $X_n \subseteq X$ l'insieme

$$X_n = \{x \mid ||Ax|| \leq n \quad \forall A \in \mathcal{A}\}\,.$$

La condizione (1.40) mostra che

$$\bigcup X_n = X\,.$$

Consideriamo ora un operatore $A \in \mathcal{A}$. Essendo A continuo, l'insieme $\{x \mid ||Ax|| \leq n\}$ è chiuso e quindi

$$X_n = \bigcap_{A\in\mathcal{A}} \{x \mid ||Ax|| \leq n\}$$

è esso stesso chiuso. Abbiamo quindi una famiglia di chiusi la cui unione è X. Per il *Teorema di Baire,* uno almeno deve avere punti interni. Sia esso X_N. Esiste $x_0 \in X_N$ ed esiste $\epsilon > 0$ per cui

$$\{x_0 + x \mid ||x|| \leq \epsilon\} \subseteq X_N\,.$$

Dunque, se $||x|| < \epsilon$ si ha

$$||Ax|| \leq ||A(x + x_0)|| + ||Ax_0|| \leq N + ||Ax_0|| = N + M_{x_0} = M\,,$$

con M *indipendente* da A. Ciò prova la limitatezza del s.insieme $\mathcal{A}$ di $\mathcal{L}(X, Y)$. ∎

Il Teorema di Banach-Steinhaus permette di passare da un'informazione *puntuale,* la limitatezza dell'insieme dei valori assunti in ciascun punto x, ad una limitatezza *uniforme* sulla sfera $\{x \mid ||x|| \leq 1\}$. Per questo esso va anche sotto il nome di *Teorema della limitatezza* *uniforme*.

In dimensione finita una trasformazione lineare invertibile non può "schiacciare" un aperto trasformandolo in un s.insieme di un s.spazio proprio. Si ricordi il ruolo importante di questa proprietà nella dimostrazione del teorema della funzione inversa e della funzione implicita.

Una proprietà analoga vale anche in spazi di Banach:

Teorema 97 (della mappa aperta) *Siano X ed Y spazi di Banach e sia $A \in \mathcal{L}(X, Y)$. Se A è suriettiva allora l'immagine di ogni aperto di X è un aperto di Y.*

Posponiamo la dimostrazione presentando invece due conseguenze del Teorema di Baire che si provano più facilmente mediante il teorema della Mappa aperta. Esse riguardano questo problema: abbiamo visto che gli operatori lineari tra X ed Y possono essere discontinui se X ha dimensione infinita. Gli esempi che abbiamo visto di operatori discontinui sono però esempi di operatori il cui dominio non è tutto X. Ci chiediamo se quando il dominio è tutto lo spazio allora l'operatore debba essere continuo. La risposta è negativa:

Teorema 98 *Siano X ed Y spazi di Banach. Esistono operatori lineari da X in Y, definiti su X e* non continui.

Si veda l'osservazione 133.
Però:

Teorema 99 (di Banach) *Siano X ed Y spazi di Banach e sia $A \in \mathcal{L}(X,Y)$ una trasformazione lineare* iniettiva *da X in Y. Se l'immagine di A è* chiusa *allora la trasformazione lineare A^{-1} (definita su* im$A$*) è* continua.

Dim. Ricordiamo che, per definizione di $\mathcal{L}(X,Y)$, un operatore $A \in \mathcal{L}(X,Y)$ ha dominio *uguale* ad X.

L'immagine di una trasformazione lineare è un s.spazio e in questo caso l'immagine è chiusa; dunque l'immagine di A è essa stessa uno spazio di Banach. Sostituendo Y con im$A$, possiamo supporre che A sia anche suriettiva.

La *trasformazione inversa* di A^{-1}, che è A, è suriettiva: per il teorema della mappa aperta, $A = (A^{-1})^{-1} \in \mathcal{L}(X,Y)$ trasforma aperti in aperti; e quindi A^{-1} è continua. ∎

Diamo infine un test importante per provare la continuità direttamente di un operatore (e non del suo operatore inverso).

Si prova facilmente che se $A \in \mathcal{L}(X,Y)$ (e quindi dom$A = X$) allora il *grafico* di A è chiuso in $X \times Y$.

Vale anche l'implicazione opposta:

Teorema 100 (del grafico chiuso) *Siano X ed Y spazi di Banach e sia A un operatore lineare da X in Y,* con dominio uguale ad X. *Se il grafico di A è chiuso allora A è continuo.*

Dim. Indichiamo con G il grafico di A. Per ipotesi, G è un s.spazio chiuso dello spazio di Banach $X \times Y$; e quindi è esso stesso uno spazio di Banach.

Introduciamo i due operatori, ovviamente lineari e continui:

$$P : G \to Y\,, \qquad P(x, Ax) = Ax$$
$$\Pi : G \to X\,, \qquad \Pi(x, Ax) = x\,.$$

Oltre che continuo, l'operatore Π è suriettivo, perché dom$A = X$ per ipotesi; ed è iniettivo perchè se $Ax_1 \neq Ax_2$ allora $x_1 \neq x_2$. Dunque esiste Π^{-1} e, per il Teorema di Banach, Π^{-1} è continuo. Dunque,

$$Ax = P(\Pi^{-1}x)$$

è continua. ∎

Bisogna notare che esistono anche operatori lineari il cui grafico è chiuso ma che non sono continui. Naturalmente, il loro dominio non sarà tutto lo spazio. Definiamo quindi:

Definitione 101 *Sia A uno operatore lineare tra due spazi di Banach X ed Y. L'operaore A si dice* chiuso *quando il suo grafico è chiuso in $X \times Y$.* ∎

L'esempio seguente mostra che operatore chiusi ma non continui non solo esistono ma sono anche importanti per le applicazioni:

Esempio 102 Sia $X = Y = C(0,1)$ e sia

$$\mathrm{dom}A = C^1(0,1)\,, \qquad Ax = x'\,.$$

Se $x \in \mathrm{dom}A$, vale

$$x(t) = x(0) + \int_0^t x'(s)\,\mathrm{d}s\,.$$

Come si verifica facilmente, l'operatore A non è continuo. Proviamo che il suo grafico è chiuso. Consideriamo quindi una successione nel grafico che è convergente e mostriamo che essa converge ad un punto del grafico. Sia quindi

$$(x_n, Ax_n) \to (x_0, y_0)\,.$$

Dobbiamo provare che $x_0 \in \mathrm{dom}A$ e che $Ax_0 = y_0$.

Si noti che y_0 è limite uniforme delle funzioni continue Ax_n e quindi y_0 è una funzione continua. Dunque dobbiamo provare che si può scrivere

$$x_0(t) = x_0(0) + \int_0^t y_0(s)\,\mathrm{d}s\,.$$

Questa uguaglianza segue da

$$x_n(t) = x_n(0) + \int_0^t x_n'(s)\,\mathrm{d}s$$

e da

$$\begin{cases} x_n \to x_0 & \text{uniformemente su } [0,1] \\ x_n' \to y_0 & \text{uniformemente su } [0,1]. \end{cases}$$ ∎

Osserviamo infine:

Corollario 103 *Siano X, Y e Z tre spazi di Banach. Sia $A \in \mathcal{L}(X,Y)$ e sia B un operatore lineare chiuso da Y in Z. Se*

$$\mathrm{im}\,A \subseteq \mathrm{dom}\,B$$

allora l'operatore composto BA è continuo.

Dim. Si vede facilmente che BA è definito su X, ed è chiuso. Dunque è continuo. ∎

1.8.1 Proiezioni

Un operatore $P \in \mathcal{L}(X)$ si dice una *proiezione* se

$$P^2 = P\,.$$

Si noti che le proiezioni vengono sempre a coppie. Infatti,

Teorema 104 *L'operatore P è una proiezione se e solo se l'operatore $I - P$ è una proiezione. Inoltre,* $\mathrm{im}P \cap \mathrm{im}(I-P) = \{0\}$.

Dim. Infatti,

$$(I-P)(I-P) = I - P - P + P^2 = I - P$$

se e solo se $P^2 = P$ ossia se e solo se P è una proiezione.

Se $x = Px' = (I-P)x''$ allora

$$Px' = x'' - Px'' \qquad \text{da cui} \quad Px' = Px'' - P^2x'' = Px'' - Px'' = 0$$

e quindi $x = Px' = 0$. ∎

Inoltre:

Teorema 105 *L'immagine di una proiezione è un s.spazio chiuso di X.*

Dim. Sia infatti (Px_n) una successione in $\mathrm{im}P$, $Px_n \to y_0$. Dobbiamo provare che $y_0 \in \mathrm{im}P$.

Poniamo $y_n = Px_n$ e notiamo che

$$Py_n = P^2x_n = Px_n \to y_0$$

e d'altra parte, essendo P continua, $Py_n = P^2y_n \to Py_0$. Dunque, $y_0 = Py_0 \in \mathrm{im}P$. ∎

Di conseguenza, ogni proiezione identifica sempre una coppia di s.spazi chiusi: l'immagine di P e quella di $(I-P)$. Questi s.spazi hanno in comune solo l'elemento 0. Inoltre, ogni x si rappresenta come

$$x = Px + (I-P)x\,.$$

Questa formula suggerisce un legame tra operatori di proiezione e complementare. Vale infatti:

Teorema 106 *L'immagine di una proiezione P è un s.spazio chiuso di X, dotato di complementare chiuso. Viceversa, sia X_1 un s.spazio di X chiuso e dotato di complementare chiuso X_2. Esiste una proiezione P la cui immagine è X_1.*

Dim. Sia P una proiezione ed $X_1 = \mathrm{im}P$. Definiamo

$$X_2 = \mathrm{im}(I - P)\,.$$

Si è già visto che $X_1 \cap X_2 = \{0\}$ e

$$X_1 + X_2 = \{Px + (I - P)y \mid x \in X\,,\ y \in Y\} = X\,.$$

Dunque, l'immagine di P è un s.spazio dotato di complementare.

Viceversa, sia

$$X = X_1 \oplus X_2\,,$$

somma diretta di due s.spazi chiusi. Questo vuol dire che per ogni x esistono x_1 ed x_2 unici e tali che

$$x = x_1 + x_2\,. \tag{1.41}$$

Si definisca $Px = x_1 = x_1 + 0$. Segue da qui che $P(Px) = x_1$.

L'operatore P è lineare perchè se $x' = x'_1 + x'_2$, $x'' = x''_1 + x''_2$, allora $\alpha x' + \beta x'' = \alpha(x'_1 + x'_2) + \beta(x''_1 + x''_2) = (\alpha x'_1 + \beta x'_2) + (\alpha x''_1 + \beta x''_2)$; e quindi $P(\alpha x' + \beta x'') = (\alpha x'_1 + \beta x'_2) = \alpha Px' + \beta Px''$. Se possiamo provare la continuità di P, abbiamo che P è una proiezione.

L'operatore P è definito su X e quindi, per provare che è continuo, basta provare che è chiuso. Sia quindi (x_n) una successione convergente ad x_0 e sia $y_n = Px_n$. Supponiamo che (y_n) converga ad y_0.

L'uguaglianza (1.41) mostra l'esistenza di un elemento $z_n \in X_2$ tale che

$$x_n = Px_n + z_n = y_n + z_n\,.$$

Di conseguenza, anche la successione (z_n) converge, a $z_0 = x_0 - y_0$.

I s.spazi essendo chiusi, vale $y_0 \in X_1$, $z_0 \in X_2$. Essendo inoltre

$$x_0 = y_0 + z_0\,, \qquad \text{si ha} \quad Px_0 = y_0\,.$$

Segue da qui che l'operatore P è chiuso e quindi continuo; dunque è una proiezione. ∎

Si consideri ora un esempio.

Esempio 107 Sia $X = \mathbb{R}^2$ normato dalla usuale norma

$$||(\xi, \eta)|| = \sqrt{\xi^2 + \eta^2}$$

e siano

$$X_1 = \{(\xi, 0) \mid \xi \in \mathbb{R}\}, \qquad X_2 = \{r(\cos\theta, \sin\theta) \mid r \in \mathbb{R}\}$$

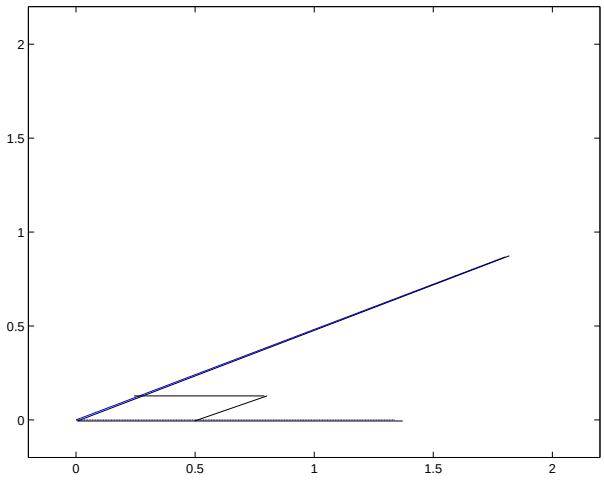
ove $\theta \in (0, \pi/2)$ è fissato. Dunque, X_2 è una retta per l'origine, non coincidente con X_1.

Ogni punto $x = (\xi, \eta)$ può rappresentarsi nella forma

$$x = (\xi - \frac{\eta}{\sin\theta}\cos\theta, 0) + \frac{\eta}{\sin\theta}(\cos\theta, \sin\theta),$$

si veda la figura seguente.

Figura 1.3:



L'operatore P:

$$P(\xi, \eta) = (\xi - \frac{\eta}{\sin\theta}\cos\theta, 0)$$

è una proiezione.

L'operatore P dipende dalla scelta di θ, $P = P_\theta$.

La norma di P_θ è

$$\max_{\xi^2+\eta^2=1} \frac{\left|\xi - \frac{\eta}{\sin\theta}\cos\theta\right|}{\sqrt{\xi^2 + \eta^2}} \geq |\mathrm{cotg}\,\theta|.$$

Dunque,

$$\lim_{\theta\to 0} ||P_\theta|| = +\infty. \quad \blacksquare$$

1.8.2 Appendice: Applicazioni

Il Teorema di Baire e le sue conseguenze sono strumenti potenti per provare l'esistenza di oggetti dalle proprietà "strane". Mostriamo due esempi.

Si costruiscono "esplicitamente", come somma di serie uniformemente convergenti di funzioni continue, delle funzioni che, pur essendo continue, non hanno derivata in nessun punto. Una dimostrazione, dovuta a Banach, dell'esistenza di tali funzioni si basa sul Teorema di Baire.

Teorema 108 *Esistono funzioni continue su un intervallo $[a, b]$, ovunque prive di derivata.*

Dim. Consideriamo in $C(0, 1)$ il s.insieme A_n i cui elementi sono funzioni f con questa proprietà: esiste $x \in [0, 1 - \frac{1}{n}]$ ed esiste $h \in (0, 1 - x)$ tale che

$$|f(x + h) - f(x)| \leq nh\,.$$

Si prova che:

- L'insieme A_n è chiuso e privo di punti interni.

Accettando queste proprietà che proveremo più avanti, il Teorema di Baire mostra che esiste una funzione continua $f(x)$ che non appartiene a $\cup A_n$.

Se $f \notin \cup A_n$ allora per ogni x e per ogni h vale

$$|f(x + h) - f(x)| > nh$$

e ciò per ogni n; ossia il rapporti incrementale è illimitato e quindi la derivata $f'(x)$ non esiste, e ciò per ogni x.

Per completare la dimostrazione, mostriamo che gli insiemi A_n sono chiusi e privi di punti interni.

Proviamo prima di tutto che A_n è chiuso. Sia per questo $f_k \to f$ (uniformemente su $[0, 1]$), con $f_k \in A_n$. Dunque, esiste $x_k \in [0, 1 - 1/n]$ tale che

$$|f_k(x_k + h) - f_k(x_k)| \leq nh\,.$$

Passando ad una s.successione, si può assumere $x_k \to x_0 \in [0, 1 - 1/n]$. Vale:

$$\begin{aligned}
&|f(x_0 + h) - f(x_0)| \\
&\leq |f(x_0 + h) - f_k(x_0 + h)| && (1.42)\\
&+|f_k(x_0 + h) - f_k(x_k + h)| && (1.43)\\
&+|f_k(x_k + h) - f_k(x_k)| && (1.44)\\
&+|f_k(x_k) - f_k(x_0)| && (1.45)\\
&+|f_k(x_0) - f(x_0)|\,. && (1.46)
\end{aligned}$$

Il termine (1.44) verifica

$$|f_k(x_k + h) - f_k(x_k)| \leq nh$$

Essendo f limite uniforme di f_k, per k sufficientemente grande i due addendi (1.42) ed (1.46) sono minori di un prefissato $\epsilon > 0$.

Usiamo ora il Teorema di Ascoli-Arzelà: essendo convergente, la successione (f_k) è equicontinua e: $||(x_0 + h) - (x_k + h)|| \to 0$. Dunque, per k grande, anche gli addendi (1.43) e (1.45) sono minori di ϵ.

Ricapitolando, la funzione f verifica, per ogni $\epsilon > 0$,

$$|f(x_0+h)-f(x_0)| \leq nh+4\epsilon \quad \text{e, essendo } \epsilon \text{ arbitrario,} \quad |f(x_0+h)-f(x_0)| \leq nh\,.$$

Ciò prova che ciascuno degli insiemi A_n è chiuso. Proviamo ora che ciascuno di essi è privo di punti interni. Fissato n ed $f \in A_n$, proviamo che per ogni $\epsilon > 0$ esiste $\zeta \notin A_n$ che dista da f meno di ϵ.

Si sa che esistono funzioni g continue e lineari a tratti tali che

$$||f - g|| < \epsilon/2\,.$$

Basta quindi provare che data una qualunque g continua e lineare a tratti si può costruire $\zeta \notin A_n$, che dista meno di $\epsilon/2$ da g. Sia per questo

$$\phi(x) = \text{distanza di } x \text{ dall'intero più vicino.}$$

Il grafico di $\phi(x)$ è in figura 1.4:

Fissiamo quindi una funzione g lineare a tratti. Essa è lipschitziana e quindi soddisfa $|g(x) - g(x')| < r|x - x'|$ per un r opportuno. Sia ζ la funzione

$$\zeta(x) = g(x) + \epsilon\phi(mx)\,.$$

Chiaramente, per ogni m, $||g - \zeta|| \leq \epsilon/2$. Vogliamo mostrare che, per un'opportuna scelta di m, $\zeta \notin A_n$. Notiamo

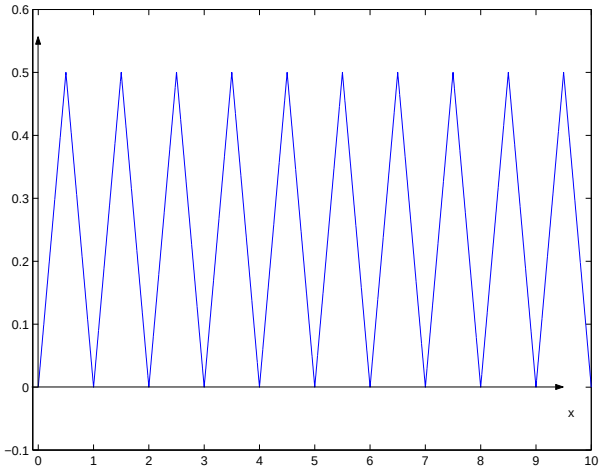
$$\begin{aligned} |\zeta(x) - \zeta(x')| &\geq \Big|\, \epsilon|\phi(mx) - \phi(mx')| - |g(x) - g(x')| \,\Big| \\ &= \Big|\, \epsilon m|x - x'| - |g(x) - g(x')| \,\Big| \geq (\epsilon m - r)|x - x'| \end{aligned}$$

se $m > r/\epsilon$, con x' tale che $|x - x'| < 1/2m$. Se ora m verifica anche $m > (n + r)/\epsilon$, allora $\zeta \notin A_n$. Ciò completa la dimostrazione. ∎

Sia ora $f(x)$ una funzione continua su $[-\pi, \pi]$. Si associ ad essa la serie

$$\sum_{n=-\infty}^{+\infty} f_n e^{inx}\,, \qquad f_n = \frac{1}{2\pi}\int_{-\pi}^{+\pi} f(s)e^{-ins}\,\mathrm{d}s$$

Figura 1.4:



che si chiama la *serie di* Fourier della funzione $f(x)$. Sotto ipotesi di regolarità, per esempio se la funzione $f(x)$ è di classe C^1 e inoltre $f(-\pi) = f(\pi)$, la serie converge ad $f(x)$ e questa condizione può indebolirsi, ma non fino alla sola continuità. Infatti:

Teorema 109 *Sia $x_0 \in [-\pi, \pi]$. Esiste una funzione f continua su $[-\pi, \pi]$ tale che $f(0) = f(2\pi)$ e tale che inoltre la serie di Fourier ad essa associata non converge in x_0.*

Dim. Indichiamo con $C_P(-\pi, \pi)$ il s.spazio di $C(-\pi, \pi)$ i cui elementi sono funzioni continue che verificano

$$f(-\pi) = f(\pi)\,.$$

Si vede facilmente che questo è un s.spazio chiuso di $C(-\pi, \pi)$, e quindi esso stesso uno spazio di Banach.

Studiamo le somme parziali della serie di Fourier. Indichiamo per questo con F_N l'operatore definito su $C_P(-\pi, \pi)$ da

$$(F_N f)(x) = \sum_{n=-N}^{N} f_n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \sum_{n=-N}^{N} e^{in(x-s)}\, \mathrm{d}s\,.$$

Si provi che

$$D_N(x) = \begin{cases} 2N+1 \quad \text{se } x = 0 \\ \sum_{n=-N}^{N} e^{inx} = \frac{\sin(N+1/2)x}{\sin x/2} \text{ se } x \neq 0\,. \end{cases}$$

La funzione $D_N(x)$ si chiama *nucleo di* Dirichlet.

Dunque,

$$(F_N f)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\frac{\sin[(N+1/2)(x-s)]}{\sin[(x-s)/2]}\,\mathrm{d}s\,.$$

Facciamo vedere che esiste una funzione $\tilde{f} \in C_P(-\pi,\pi)$ tale che

$$\sup_N |(F_N\tilde{f})(x_0)| = +\infty\,.$$

Ciò vuol dire che la serie di Fourier di questa funzione $\tilde{f}$ non converge in x_0.

Per completare la dimostrazione basta quindi provare l'esistenza di $\tilde{f}$. Supponiamo che tale funzione $\tilde{f}$ non esista. Allora, per ogni $f \in C_P(-\pi,\pi)$ si ha:

$$\sup_N |(F_N f)(x_0)| < +\infty\,.$$

In tal caso per ogni $f \in C_P(-\pi,\pi)$ esiste M_f per cui

$$|(F_N f)(x_0)| < M_f$$

per ogni N. E quindi, per il *teorema di Banach–Steinhaus,* esiste $M = M(x_0)$, indipendente da N, tale che

$$|(F_N f)(x_0)| < M(x_0)||f||_{C_P(-\pi,\pi)}\,.$$

Indicando con $F_{x_0,N}$ il funzionale che ad $f \in C_P(-\pi,\pi)$ associa $(F_N f)(x_0)$, la disuguaglianza precedente si scrive

$$||F_{x_0,N}|| < M(x_0)\,; \tag{1.47}$$

ossia, la famiglia dei funzionali lineari e continui $F_{x_0,N}$ è limitata. Calcoliamo esplicitamente la norma del funzionale $F_{x_0,N}$ e mostriamo che ciò non vale.

Per semplicità limitiamoci a fare il calcolo con $x_0 = 0$. In questo caso

$$F_{0,N} f = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t) f(t)\,\mathrm{d}t\,.$$

e

$$|F_{0,N} f| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t) f(t)|\,\mathrm{d}t \le \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\,\mathrm{d}t\right)||f||_{C_P(-\pi,\pi)}$$

così che

$$||F_{0,N}|| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\,\mathrm{d}t\,. \tag{1.48}$$

In realtà vedremo che vale l'uguaglianza. Accettando ciò,

$$
\begin{aligned}
||F_{0,N}|| &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_N(t)|\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin[(N+1/2)t]}{\sin t/2}\,\mathrm{d}t\right| \\
&> \frac{1}{\pi}\int_{-\pi}^{\pi}\left|\frac{\sin[(N+1/2)t]}{t}\right|\,\mathrm{d}t = \frac{2}{\pi}\int_{0}^{(N+1/2)\pi}\frac{|\sin t|}{t}\,\mathrm{d}t \\
&\geq \frac{2}{\pi}\sum_{k=0}^{N-1}\int_{k\pi}^{(k+1)\pi}\frac{|\sin t|}{t}\,\mathrm{d}t \geq \sum_{k=0}^{N-1}\frac{2}{(k+1)\pi^2}\int_{k\pi}^{(k+1)\pi}|\sin t|\,\mathrm{d}t \\
&= \frac{4}{\pi^2}\sum_{k=0}^{2n}\frac{1}{(k+1)} \longrightarrow +\infty\,.
\end{aligned}
$$

Ciò contrasta con la (1.47) e mostra che la funzione $\tilde{f}$ esiste.

Accenniamo ora alla dimostrazione del fatto che l'uguaglianza vale nella formula (1.48). Per mostrare ciò è sufficiente trovare una successione di funzioni (f_k) di norma al più uguale ad 1 e tale che

$$\lim_k |F_{0,N} f_k| = \frac{1}{2\pi}\int_{-\pi}^{\pi}|D_N(t)|\,\mathrm{d}t \tag{1.49}$$

Introduciamo per questo la funzione

$$y(x) = \operatorname{sign} D_N(x)$$

e una successione (f_k) di funzioni continue convergente puntualmente ad y e inoltre limitata da 1. L'uguaglianza (1.49) vale per questa successione di funzioni.

Ciò completa la dimostrazione. ∎

1.8.3 Dimostrazioni posposte

Dimostrazione del TEOREMA 94, Teorema di Baire.

Premettiamo un'osservazione: Sia B_n una successione di palle contenuta ciascuna nella precedente:

$$B_n = \{x \mid ||x - x_n|| < \epsilon_n\} \subseteq \{x \mid ||x - x_{n-1}|| < \epsilon_{n-1}\} = B_{n-1}$$

così che

$$||x_n - x_{n+m}|| < \epsilon_n\,.$$

Sia $\lim \epsilon_n = 0$. Da

$$||x_n - x_{n+m}|| < \epsilon_n$$

si vede che la successione (x_n) è fondamentale ossia convergente,

$$\lim x_n = \tilde{x}\,.$$

Il punto $\tilde{x}$ appartiene alla chiusura di B_n per ogni n e quindi si ha anche

$$\tilde{x} \in \bigcap_n \mathrm{cl}\, B_n\,.$$

Sia ora (A_n) una successione di insiemi chiusi e (B_n) una successione di palle con le proprietà appena dette e tali che, inoltre,

$$(\mathrm{cl}\, B_n) \cap A_n = \emptyset\,. \tag{1.50}$$

Allora, $\tilde{x}$ non appartiene a A_n per nessun n, grazie alla (1.50):

$$\tilde{x} \notin \bigcup A_n\,.$$

Infatti, se $\tilde{x} \in \cup A_n$ allora $\tilde{x} \in A_{n_0}$ per almeno un indice n_0. D'altra parte, si sa anche che $\tilde{x} \in \mathrm{cl}\, B_{n_0}$ e ciò non può essere perché gli insiemi A_{n_0} e $\mathrm{cl}\, B_{n_0}$ sono disgiunti.

Per provare il Teorema di Baire, costruiamo una successione di palle B_n che ha le proprietà dette sopra rispetto alla successione di insiemi (A_n), chiusi e privi di punti interni. Ciò porterà a trovare che $\tilde{x} \notin \cup A_n$ e quindi $\cup A_N \neq X$. Scegliamo $x_1 \notin A_1$ e una palla B_1 di centro x_1 e raggio minore di 1, tale che $(\mathrm{cl}\, B_1) \cap A_1 = \emptyset$. Sia $\epsilon_1 > 0$ il suo raggio. Non è restrittivo assumere $\epsilon_1 < 1$.

La palla B_1 esiste perché A_1 è chiuso e, essendo privo di punti interni, non è uguale ad X.

La palla B_1 non è contenuta in A_2 perché A_2 non ha punti interni. Dunque in B_1 esite un punto $x_2 \notin A_2$ e quindi interno al complementare dell'insieme chiuso A_2. Possiamo quindi scegliere una palla B_2 di centro x_2, contenuta in B_1 e di raggio minore di $\epsilon_1/2 < 1/2$, tale che $(\mathrm{cl}\, B_2) \cap A_2 = \emptyset$.

Sia $\epsilon_2 > 0$ il raggio di B_2.

Procedendo per induzione, scelti i punti $x_1, \ldots, x_k$ e le corrispondenti palle $B_1, \ldots, B_k$, scegliamo in B_k un punto $x_{k+1} \notin A_{k+1}$ e una sfera di centro x_{k+1} e raggio minore di $\epsilon_k/2 < 1/2^k$, tale che $(\mathrm{cl}\, B_{k+1}) \cap A_{k+1} = \emptyset$. Sia $\epsilon_{k+1} > 0$ il raggio di questa sfera.

La costruzione dei punti x_k e delle palle B_k può farsi perché gli A_k non hanno punti interni e sono chiusi.

Dato che $\epsilon_k \to 0$, esiste $\tilde{x} = \lim x_k$ e $\tilde{x} \in \cap_n \mathrm{cl}\, B_n$. Per quanto osservato sopra, $\tilde{x} \notin \cup A_n$, ossia

$$\bigcup A_n \neq X\,,$$

come si voleva provare. ∎

Dimostrazione del TEOREMA 97, Teorema della mappa aperta.

In questa dimostrazione interverrà la "differenza algebrica" di insiemi C, D di Y:

$$C - D = \{c - d \mid c \in C\,,\ d \in D\} = \bigcup_{d \in D} (C - d)\,.$$

Si noti che $C - C \neq \emptyset$ (e anche che $C - C \neq 0$, salvo nel caso in cui C ha un unico elemento).

Se C ha interno non vuoto anche

$$C - D = \bigcup_{d \in D} (C - d)$$

ha interno non vuoto grazie alla continuità delle traslazioni, per ogni insieme D (e quindi anche per $D = C$).

Inoltre, se C contiene punti interni, allora $(\mathrm{cl}\, C) - (\mathrm{cl}\, C)$ contiene un intorno di 0.

Proviamo ora il teorema 97. Ricordiamo che per definizione, un operatore $A \in \mathcal{L}(X, Y)$ ha dominio *uguale* ad X.

Per provare che l'operatore A, suriettivo, trasforma aperti in aperti, è sufficiente mostrare che l'immagine di una palla

$$B_{X,r} = \{x \in X \mid ||x|| < r\}$$

contiene una palla $B_{Y,\sigma}$,

$$B_{Y,\sigma} = \{y \in Y \mid ||y|| < \sigma\}\,.$$

Ciò prova che $A0$ è interno ad $\mathrm{im}\, A$ e, per traslazione, si trova che Ax_0 è interno ad $\mathrm{im}\, A$ per ogni x_0 (nuovamente, si usa la continuità delle traslazioni).

Precisamente, proveremo che esiste un intorno W di 0 in Y che è contenuto in $A(B_{X,2})$. Useremo per questo l'inclusione seguente, che vale per ogni $r > 0$ ed $r' > 0$:

$$B_{X,r} - B_{X,r'} \subseteq B_{X,r+r'}\,. \tag{1.51}$$

Consideriamo le palle

$$B_{X,2^n} = \{x \in X \mid ||x|| < 2^n\}\,.$$

Dato che

$$X = \bigcup_n B_{X,2^n}$$

e che A è suriettivo, si trova

$$Y = \bigcup_n \text{cl } (A(B_{X,2^n}))$$

e quindi, per il *teorema di Baire* almeno uno degli insiemi cl $(A(B_{X,2^n}))$ ha punti interni. Moltiplicando per numeri positivi si vede che ciascuno degli insiemi cl $(A(B_{X,r}))$ contiene punti interni, grazie alla continuità della moltiplicazione per scalari.

Da (1.51) si ha che

$$\text{cl}\, A(B_{X,r}) - \text{cl}\, A(B_{X,r}) \subseteq \text{cl } [A(B_{X,r}) - A(B_{X,r})] \subseteq \text{cl}\, A(B_{X,2r})$$

e quindi ciascun insieme cl $A(B_{X,2r})$ contiene un intorno di 0. Naturalmente, r è arbitrario: ogni insieme cl $A(B_{X,r})$ contiene un opportuno intorno di 0.

Rimane da provare che $A(B_{X,r})$ stesso contiene un intorno di 0. Sia W un intorno di 0 contenuto in cl $A(B_{X,1})$. Completiamo la dimostrazione mostrando che $W \subseteq A(B_{X,2})$. Per questo basta provare

$$\text{cl}\, A(B_{X,1}) \subseteq A(B_X, 2)\,.$$

Ciò mostriamo ora. Sia per questo $\tilde{y} \in \text{cl}\, A(B_{X,1})$. Mostriamo che $\tilde{y} \in A(B_{X,2})$. Si sa che cl $A(B_{X,1/2})$ contiene un intorno di 0. Dunque esiste $x_1 \in B_{X,1}$ tale che $||\tilde{y} - Ax_1||$ è così piccolo da aversi $\tilde{y} - Ax_1 \in \text{cl}\, A(B_{X,1/2})$.

In modo analogo, cl $AB_{X,1/4}$ contiene un intorno di 0 e quindi esiste $x_2 \in B_{X,1/2}$ per cui $(\tilde{y} - Ax_1) - Ax_2 \in \text{cl}\, A(B_{X,1/4})$. Iterando questo procedimento per ogni n si trova

$$\begin{array}{llllll}
x_1 \in B_{X,1} & \text{tale che} & \tilde{y} - Ax_1 \in \text{cl}\, A(B_{X,1/2}) & \text{cioè} & ||\tilde{y} - Ax_1|| \le 1/2 \\
x_2 \in B_{X,1/2} & \text{tale che} & (\tilde{y} - Ax_1) - Ax_2 \in \text{cl}\, A(B_{X,1/4}) & \text{cioè} & ||(\tilde{y} - Ax_1) - Ax_2|| \le 1/4 \\
\vdots & & & & \\
x_n \in B_{X,1/2^n} & \text{tale che} & \tilde{y} - \sum_{k=1}^n Ax_n \in \text{cl}\, A(B_{X,1/2^n}) & \text{cioè} & ||\tilde{y} - \sum_{k=1}^n Ax_n|| \le 1/2^n\,.
\end{array}$$

Sia ora

$$x = \sum_{i=1}^{+\infty} x_i \qquad \text{così che} \qquad ||x|| \le \sum_{i=1}^{+\infty} \frac{1}{2^n} < 2\,.$$

Per questo vettore x vale

$$||\tilde{y} - Ax|| = \lim ||\tilde{y} - Ax_n|| \le \lim \frac{1}{2^n} = 0\,, \quad \text{ossia} \quad \tilde{y} = Ax \quad \text{con} \quad ||x|| < 2\,.$$

Ciò mostra che ogni $\tilde{y} \in \text{cl}\, A(B_{X_1})$ è anche in $A(B_{X,2})$ e conclude la dimostrazione. ∎

1.9 Lo spazio duale

Abbiamo già visto la relazione tra i funzionali lineari e continui su X e la nozione geometrica di iperpiano. Ciò suggerisce di studiare più a fondo i funzionali lineari continui, sia singolarmente che nel loro insieme, studiando le proprietà dello spazio di Banach X^*.

E' ovvio che il funzionale 0, quello che ad ogni elemento di X associa l'elemento nullo del campo scalare, è in X^*. Non è affatto ovvio che esistano altri elementi di X^*. Infatti:

Teorema 110 *Esistono spazi lineari X, dotati di una metrica rispetto alla quale le operazioni di somma e moltiplicazioni per scalari sono continue e su cui nessun funzionale lineare diverso da 0 è continuo.*

Ossia, in tal caso, nello spazio X non si trovano iperpiani.

La dimostrazione è posposta.

E' quindi estremamente importante sapere che se X è uno spazio di Banach allora X^* ha "molti" elementi. Ciò è conseguenza del teorema seguente:

Teorema 111 (di Hahn-Banach) *Sia X uno s.l.n. e sia Y un suo s.spazio. Sia L_0 un funzionale* lineare continuo *su Y. Esiste un'estensione di L ad X tale che*

$$||L||_{X^*} = \sup\{|Ly| \mid y \in Y\,,\ ||y||_X = 1\}\,.$$

Ossia, ogni funzionale lineare e continuo su Y può estendersi ad X senza alterarne la norma.

Si noti che nell'enunciato precedente si può assumere che Y sia chiuso, perché l'operatore L_0 si può estendere per continuità alla chiusura del suo dominio.

Per certe applicazioni (allo studio degli insiemi convessi, si veda il paragrafo 1.9.1) è necessario provare una versione un po' più generale del Teorema 111; ossia è necessario provare i due teoremi seguente:

Teorema 112 *Sia X uno spazio lineare su $\mathbb{R}$. Esista una funzione p: $X \to \mathbb{R}$*

1. positivamente omogenea, *ossia tale che $p(tx) = tp(x)$ per ogni x e per ogni $t \geq 0$;*
2. subadditiva, *ossia tale che $p(x+y) \leq p(x) + p(y)$ per ogni x, y in X.*

Sia Y un s.spazio di X e sia L_0 un funzionale lineare definito su Y, tale che

$$L_0 x \leq p(x) \qquad \forall x \in Y\,.$$

Esiste un'estensione L di L_0 ad X che verifica

$$Lx \leq p(x) \qquad \forall x \in X\,.$$

Teorema 113 *Sia X uno spazio lineare su $\mathbb{C}$ e sia p da X in $\mathbb{R}$ una funzione tale che:*

1. *$p(tx) = |t| \cdot p(x)$ per ogni x in X e per ogni $t \in \mathbb{C}$;*
2. *$p(x+y) \leq p(x) + p(y)$.*

Sia Y un s.spazio di X e sia L_0 un funzionale lineare definito su Y, tale che

$$|L_0 x| \leq p(x) \qquad \forall x \in Y\,. \tag{1.52}$$

Esiste un'estensione L di L_0 ad X che verifica

$$|Lx| \leq p(x) \qquad \forall x \in X\,.$$

Posponiamo le dimostrazioni, notando che la dimostrazione del Teorema 113 si ridurrà, con un opportuno artificio, a quella del Teorema 112.

Dimostrazione del TEOREMA 111. La dimostrazione discende immediatamente dai Teoremi 112 e 113. Sia

$$M = \sup\{|L_0 x| \mid x \in Y\,,\ ||x|| = 1\}\,, \qquad p(x) = M||x||\,.$$

Nel caso reale, dal Teorema 112 si vede l'esistenza di L, funzionale lineare su X, che estende L_0 e tale che

$$Lx \leq M||x||$$

e quindi anche

$$-Lx = L(-x) \leq M||-x|| = M||x||$$

ossia

$$|Lx| \leq M||x||\,. \tag{1.53}$$

Nel caso complesso la disuguaglianza (1.53) figura direttamente nell'enunciato del Teorema 113. Dunque il Teorema 111 vale. ∎

Mostriamo ora alcune conseguenze importanti. La prima è che X^*, duale di X, ha "molti" elementi. Più precisamente,

Teorema 114 *Per ogni $x_0 \neq 0$ in X ed ogni $m \in \Phi$, esiste $L \in X^*$ tale che*

$$Lx_0 = m\,, \qquad ||L|| = \frac{m}{||x_0||}\,. \tag{1.54}$$

Dim. Si sceglie come spazio Y la retta per 0 ed x_0,

$$Y = \{\lambda x_0 \mid \lambda \in \Phi\}\,.$$

Il s.spazio ha dimensione 1 e su esso è facile definire

$$L_0(\lambda x_0) = \lambda m\,.$$

Il funzionale L_0 verifica (1.54) sui soli elementi di Y. Si usa quindi il Teorema di Hahn-Banach per estendere L ad X. ∎

In particolare si può segliere $m = 1$ oppure $m = ||x_0||$ oppure $m = ||x_0||^2$. In particolare, con quest'ultima scelta si trova

$$||L|| = ||x_0||\,.$$

Invece, scegliendo $m = ||x_0||$, si trova

Corollario 115 *Sia $x_0 \in X$. Vale:*

$$||x_0|| = \max_{||L||_{X^*}=1} |Lx_0|\,.$$

Dim. Si fissi $x_0 \neq 0$ in X. Per ogni $L \in X^*$ di norma 1 si ha:

$$|Lx_0| \leq ||x_0||\,, \qquad \text{ossia} \qquad \sup_{||L||_{X^*}=1} ||Lx_0|| \leq ||x_0||\,.$$

Il funzionale definito in (1.54) con $m = ||x_0||$ ha norma 1 e per esso

$$|Lx_0| = ||x_0||\,.$$ ∎

Osservazione 116 L'asserto del corollario precedente somiglia alla definizione della norma di un funzionale,

$$||L|| = \sup_{||x||=1} |Lx| \tag{1.55}$$

Ma abbiamo provato che In generale, l'estremo superiore in 1.55 *non* è un massimo, si veda il teorema 73.

Ciò mostra una prima differenza importante tra uno spazio ed il suo duale, che commenteremo in seguito nel contesto del Teorema di Banach-Alaoglu, Teorema 150. ∎

Il Teorema 114 in particolare afferma che se un elemento x_0 è diverso da 0 allora esiste un funzionale L di X^* che "lo vede non nullo"; e, traslando, se $x_1 \neq x_0$, esiste un $L \in X^*$ tale che

$$Lx_1 \neq Lx_0 \,.$$

Dunque, X^* ha così tanti elementi da distinguere quelli di X. Geometricamente, abbiamo provato l'esistenza di un iperpiano che non contiene ambedue gli elementi x_0 ed x_1. Questa osservazione può essere estesa fino a "separare" mediante iperpiani due insiemi convessi tra loro disgiunti. Prima di studiare gli insiemi convessi conviene però introdurre una notazione comoda per indicare l'azione degli elementi di X^* su X. Invece di scrivere Lx, Ax ecc, usa scrivere[4]

$$\langle\langle f, x \rangle\rangle$$

per indicare il valore che il funzionale $f \in X^*$ assume sull'elemento $x \in X$ (si noti: l'elemento di X^* è scritto prima di quello di X. In altri testi si trova scritto dopo). Inoltre, per distinguere immediatamente gli elementi di X da quelli di X^*, usa indicare questi ultimi con lettere greche, o con simboli del tipo x^*, y^* ecc. se i simboli x, y,... si riservano agli elementi di X.

1.9.1 Applicazioni: Insiemi convessi

Per semplicità supponiamo che il campo scalare sia $\mathbb{R}$.

Ricordiamo che un insieme A non vuoto si dice *convesso* quando ogni segmento di estremi in A è tutto contenuto in A; ossia quando

$$x\,,\ y \in A \Longrightarrow tx + (1-t)y \in A \qquad \forall t \in [0,1]\,.$$

Si verifica facilmente che gli iperpiani ed i semispazi sono insiemi convessi e che l'intersezione di una qualsiasi famiglia di insiemi convessi è un insieme che, se non è vuoto, è convesso.

Siano ora A e B due insiemi e sia $x^* \in X^*$. Si dice che l'iperpiano

$$\{x \mid \langle\langle x^*, x \rangle\rangle = \alpha\}$$

separa i due insiemi A e B se

$$A \subseteq \{x \mid \langle\langle x^*, x \rangle\rangle \leq \alpha\}\,, \qquad B \subseteq \{x \mid \langle\langle x^*, x \rangle\rangle \geq \alpha\}\,.$$

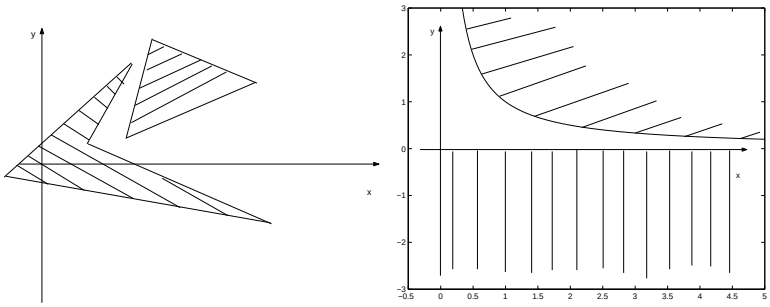
[4] In realtà la notazione comunemente usata è $\langle \cdot, \cdot \rangle$. Noi usiamo la notazione $\langle\langle \cdot, \cdot \rangle\rangle$ perché la notazione $\langle \cdot, \cdot \rangle$ si usa anche per indicare i "prodotti interni" nel contesto degli spazi di Hilbert. Dato che vedremo una relazione tra funzionali lineari e prodotti interni, è opportuno essere precisi nel distinguere gli uni dagli altri.

Si dice che la separazione è *stretta* se esiste $\epsilon > 0$ tale che

$$A \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \leq \alpha - \epsilon\}\,, \qquad B \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \geq \alpha\}\,.$$

E' ovvio che in generale due insiemi disgiunti non possono essere separati da iperpiani, si veda la figura 1.5 a sinistra. Per avere buoni risultati di separazione dovremo lavorare con insiemi convessi. La figura 1.5 a destra mostra che insiemi convessi e disgiunti non possono, in generale, separarsi strettamente.

Figura 1.5:



Valgono però i due teoremi seguenti, la cui dimostrazione viene posposta:

Teorema 117 *Sia $A \subseteq X$ un convesso aperto. Se $x_0 \notin A$ allora esiste un iperpiano che separa x_0 da A; ossia esiste $x^* \in X^*$ tale che*

$$\langle\langle x^*, x\rangle\rangle \leq \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in A\,.$$

e

Teorema 118 *Siano A, B convessi e disgiunti. Si ha:*

- *Se A è* aperto, *esiste un iperpiano che* separa *A e B;*
- *se A è* chiuso *e B è* compatto *allora esiste un iperpiano che* separa strettamente *A e B.*

Di conseguenza:

Teorema 119 *Sia H un s.spazio chiuso dello spazio di Banach X. Vale $H \neq X$ se e solo se esiste $x^* \in X^*$, $x^* \neq 0$, tale che*

$$H \subseteq \ker x^*\,.$$

Dim. Se $H \neq X$, esiste $x_0 \notin H$ e quindi per la seconda affermazione del Teorema 118, x_0 si separa strettamente da H: esistono $\epsilon > 0$ ed x^* tale che

$$\langle\langle x^*, x_0 \rangle\rangle \leq \alpha - \epsilon \,, \qquad \langle\langle x^*, h \rangle\rangle \geq \alpha \,, \ \forall h \in H \,.$$

Essendo H un s.spazio, segue

$$\alpha \leq \langle\langle x^*, -h \rangle\rangle = -\langle\langle x^*, h \rangle\rangle \,.$$

Dunque, $\langle\langle x^*, h \rangle\rangle = \alpha$ per ogni $h \in H$. Essendo H un s.spazio, 0 appartiene ad H e quindi $\langle\langle x^*, 0 \rangle\rangle = \alpha$, ossia $\alpha = 0$. Si ha dunque

$$\langle\langle x^*, h \rangle\rangle = 0 \qquad \forall h \in H \,,$$

come volevamo.

Il viceversa è ovvio. ∎

Inoltre,

Teorema 120 *Se X^* è separabile, anche X lo è.*

Dim. Sia $\{x_n^*\}$ un s.insieme denso in X^*. Si scelgano elementi $x_n \in X$ di norma 1 e tali che

$$\left| \langle\langle x_n^*, x_n \rangle\rangle \right| \geq \frac{1}{2} \lambda_n \,, \qquad \lambda_n = ||x_n^*|| \,.$$

Combinando linearmente gli x_n, con coefficienti razionali, si trova un s.insieme A numerabile di X. Procedendo per assurdo, mostriamo che A è denso in X, che pertanto è separabile.

Se A non fosse denso in X, la sua chiusura sarebbe un s.spazio H di X diverso da X. Dunque, per il Teorema 119, si potrebbe trovare $\phi^* \in X^*$ tale che

$$A \subseteq H \subseteq \ker \phi^* \,.$$

Si sa che $\phi^* = \lim x_{n_k}^*$ per una opportuna successione $(x_{n_k}^*)$ così che

$$\begin{aligned} 0 &= \lim ||\phi^* - x_{n_k}^*||_{X^*} \geq \lim \left| \langle\langle \phi^* - x_{n_k}^*, x_{n_k} \rangle\rangle \right| \\ &= \lim \left| \langle\langle x_{n_k}^*, x_{n_k} \rangle\rangle \right| \geq \lim \frac{1}{2} \lambda_n \end{aligned}$$

e quindi $\lim \lambda_n = 0$. D'altra parte, mentre

$$0 \neq ||\phi^*||_{X^*} = \lim ||x_{n_k}^*||_{X^*} = \lim \lambda_{n_k} \,.$$

La contraddizione trovata completa la dimostrazione. ∎

Osservazione 121 Invece, il duale di uno spazio separabile può non essere separabile. Mostreremo infatti che l^∞ è isometricamente isomorfo al duale di l^1. Si sa già che l^∞ non è separabile mentre è facile verificare che l^1 lo è. ∎

Un'ulteriore conseguenza importante del Teorema 118 è la seguente: ogni insieme A, anche non convesso, è contenuto nell'intersezione dei semispazi che lo contengono. Se A è convesso vale:

Teorema 122 *Un insieme A non vuoto è convesso e chiuso se e solo se è intersezione dei semispazi chiusi che lo contengono.*

Dim. L'intersezione di una famiglia di chiusi è un chiuso e l'intersezione di una famiglia di convessi, se non è vuota, è un convesso. Dunque, se l'insieme non vuoto A è intersezione dei semispazi chiusi che lo contengono, A è convesso e chiuso.

Viceversa, sia B l'intersezione dei semispazi chiusi che contengono l'insieme A. In generale, $B \supseteq A$. Dobbiamo provare che se A è convesso e chiuso allora vale l'uguaglianza. Per questo basta notare che se $x_0 \notin A$ allora, per il Teorema 118, esiste un iperpiano che separa strettamente A da x_0. E quindi x_0 non appartiene nemmeno a B. ∎

Si chiamano *iperpiani di* *supporto* all'insieme convesso e chiuso A quelli che godono della seguente proprietà: esiste $x_0 \in A$ tale che

$$A \subseteq \{x \mid \langle\langle x^*, x\rangle\rangle \leq \langle\langle x^*, x_0\rangle\rangle\}.$$

Più precisamente, in questo caso si dice che l'iperpiano è di supporto nel punto x_0.

Naturalmente, se esiste un iperpiano di supporto ad A in x_0 allora x_0 non è interno ad A. Più precisamente:

Teorema 123 *Sia A un convesso con interno non vuoto. Per ogni $x_0 \in \partial A$, esiste almeno un iperpiano di supporto ad A in x_0.*

Dim. Usiamo la proprietà seguente degli insiemi convessi, che non proviamo: se A è convesso ed il suo interno è non vuoto allora l'interno di A è convesso e denso in A.

Dunque $\operatorname{int} A$ è un convesso, ovviamente separato da $x_0 \in \partial A$: per il Teorema 117, esiste x^* tale che

$$\langle\langle x^*, x\rangle\rangle < \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in \operatorname{int} A$$

e la disuguaglianza, non stretta, si estende per continuità ad A.

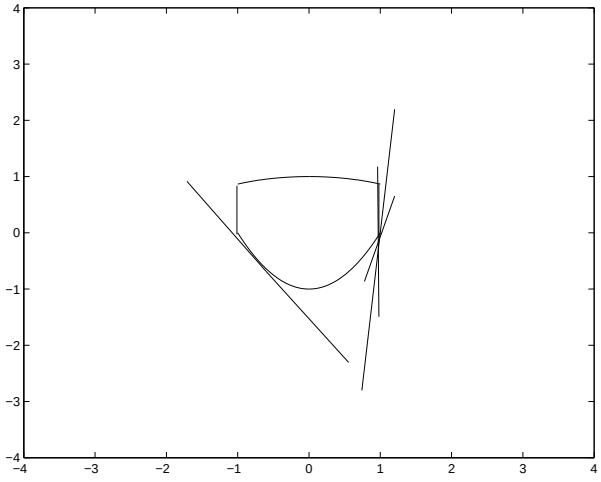
L'iperpiano di supporto è

$$\{x \mid \langle\langle x^*, x\rangle\rangle = \langle\langle x^*, x_0\rangle\rangle\}. \quad ∎$$

Osservazione 124 • Il teorema precedente può anche estendersi al caso in cui l'interno di A è vuoto. Essendo A convesso, si prova l'esistenza di un più piccolo s.spazio (traslato) contenente A. Il teorema si può enunciare in tale s.spazio. Non insistiamo su ciò.

• La figura 1.6 mostra che uno stesso iperpiano può essere di supporto in infiniti punti, e che in un punto si possono avere infiniti iperpiani di supporto. ∎

Figura 1.6:



Abbiamo dunque due modi di descrivere un insieme convesso e chiuso, tra loro equivalenti: il primo consiste nell'elencare i punti dell'insieme ed il secondo consiste nell'elencare gli iperpiani di supporto all'insieme. Questo secondo modo assume un aspetto particolare nel caso in cui l'insieme A è l'epigrafo di una funzione, caso che ora andiamo a studiare.

1.9.2 Applicazioni: Funzioni convesse

Ancora, supponiamo di lavorare in spazi lineari su $\mathbb{R}$ e studiamo certe proprietà delle funzioni convesse definite su uno spazio di *Banach* X. Per semplicità supporremo che esse siano definite su insiemi, ovviamente convessi, dotati di punti interni.

Ricordiamo che si chiama *epigrafo* di una funzione f definita su X, a valori reali, l'insieme

$$\mathrm{Epi}(f) = \{(x,t) \mid t \geq f(x)\} \subseteq X \times \mathbb{R}.$$

E' facile provare:

Teorema 125 *La funzione f è convessa se e solo se il suo epigrafo è un insieme convesso.*

Un iperpiano di supporto in (x_0, t_0) alla chiusura dell'epigrafo di f è rappresentato da una coppia $(x^*, \phi) \in X^* \times \mathbb{R}$ tale che, per ogni punto (x,t) dell'epigrafo, valga

$$\langle\langle x^*, x\rangle\rangle + \phi \cdot t \leq \langle\langle x^*, x_0\rangle\rangle + \phi \cdot t_0$$

ossia

$$\langle\langle x^*, (x - x_0)\rangle\rangle \leq (-\phi) \cdot (t - t_0)\,. \tag{1.56}$$

Notiamo che se $\phi = 0$ si ha

$$\langle\langle x^*, x\rangle\rangle \leq \langle\langle x^*, x_0\rangle\rangle \qquad \forall x \in \mathrm{dom} f\,.$$

In questo caso x_0 deve appartenere alla frontiera di $\mathrm{dom} f$. Supponendo di lavorare in punti interni al dominio, avremo $\phi \neq 0$ e si ha $-\phi > 0$. Infatti: se $x^* = 0$ allora deve essere

$$0 \leq (-\phi)(t - t_0) \qquad \forall t \geq t_0 \qquad \text{ossia } -\phi > 0.$$

Dunque non è restrittivo assumere $-\phi = 1$ perché dividendo i due membri per $-\phi$ si vede che la (1.56) può scriversi

$$\langle\langle x^*, x - x_0\rangle\rangle \leq t - t_0\,.$$

Supponiamo ora che l'epigrafo di f sia chiuso. In questo caso $t_0 = f(x_0)$ mentre $t \geq f(x)$. Dunque:

Teorema 126 *Se l'epigrafo di f è chiuso allora il funzionale $(x^*, 1) \in X^* \times I\!R$ è di supporto all'epigrafo di f in $(x_0, f(x_0))$ se e solo se*

$$\langle\langle x^*, x - x_0\rangle\rangle \leq f(x) - f(x_0) \qquad \forall x \in \mathrm{dom} f\,. \tag{1.57}$$

La (1.57) si scrive anche

$$f(x) \geq \langle\langle x^*, x - x_0\rangle\rangle + f(x_0)$$

e quindi, se $f(x)$ è convessa,

- il grafico è sopra a tutti gli iperpiani di supporto;

- il numero $f(x)$ è l'estremo superiore dei valori delle funzioni lineari affini il cui grafico è sotto quello di f. L'estremo superiore è un massimo, in ogni punto x, se l'epigrafo è chiuso.

La (1.57) e la pratica con i grafici delle funzioni da $\mathbb{R}$ in $\mathbb{R}$ suggerisce:

Definitione 127 Un funzionale x^* che verifica (1.57) si chiama **un** sottodifferenziale di f in x_0.

L'**insieme** dei sottodifferenziali di f in x_0 si chiama **il** *sottodifferenziale* di f in x_0.

Il sottodifferenziale di f in x_0 si indica col simbolo $\partial f(x_0)$. ∎

Osserviamo che il sottodifferenziale può essere l'insieme vuoto ma, come si è già notato, in tal caso il punto x_0 deve appartenere alla frontiera del dominio di f:

Esempio 128 Sia $X = \mathbb{R}$ e sia $\mathrm{dom} f = [0, +\infty)$, $f(x) = -\sqrt{x}$. Si vede facilmente che $\partial f(0) = \emptyset$. ∎

Invece, come si è notato il sottodifferenziale è non vuoto nei punti interni al dominio di f.

La proprietà di avere epigrafo chiuso è ovviamente importante nei problemi di ottimizzazione e merita una definizione specifica:

Definitione 129 Una funzione (anche non convessa) con epigrafo chiuso si chiama *semicontinua inferiormente*. Una funzione f si dice *semicontinua superiormente* se $-f$ è semicontinua inferiormente. ∎

E' facile immaginare che il sottodifferenziale sia uno strumento utile per il calcolo dei minimi. Infatti,

Teorema 130 *Sia f convessa e sia non vuoto il suo sottodifferenziale in x_0. Il punto x_0 è punto di minimo se e solo se $0 \in \partial f(x_0)$.*

Dim. Immediate, leggendo la (1.57) con $x^* = 0$. ∎

1.9.3 Dimostrazioni posposte

Dimostrazione del TEOREMA 110.

Indichiamo con X l'insieme delle (classi di equivalenza di) funzioni $f(x)$ definite su $(0,1)$ e tali che

$$\int_0^1 \sqrt{|f(x)|}\,\mathrm{d}x < +\infty\,.$$

La disuguaglianza

$$\sqrt{a+b} \le \sqrt{a} + \sqrt{b} \qquad \forall a \ge 0\,,\ b \ge 0 \tag{1.58}$$

permette di provare che X è uno spazio vettoriale e che

$$d(f,g) = \int_0^1 \sqrt{|f(x) - g(x)|}\,\mathrm{d}x$$

è una distanza invariante per traslazioni su X (per provare la disuguaglianza triangolare si usa la (1.58)).

Si può provare che lo spazio X è completo.

Lo spazio X ha una proprietà curiosa: l'unico convesso (non vuoto) di X che è anche aperto è X stesso. Accettando questa proprietà, che proveremo in seguito, si vede che l'unico funzionale L lineare e continuo su X è quello nullo. Infatti, per ogni $\epsilon > 0$, l'insieme $L^{-1}(-\epsilon, \epsilon)$ non è vuoto, perché contiene 0; è convesso perchè L è lineare e, se L è anche continuo, è aperto; e quindi è tutto lo spazio X. Dunque,

$$LX \subseteq \bigcap_{\epsilon>0} (-\epsilon, \epsilon) = \{0\}\,;$$

ossia, L è il funzionale nullo.

Per completare la dimostrazione, consideriamo un aperto V di X, non vuoto e convesso e proviamo che $V = X$. A meno di traslazioni, si può supporre $0 \in V$. Dunque può trovarsi $r > 0$ tale che

$$B(0,r) = \{x \mid d(x,0) < r\} \subseteq V\,.$$

Fissiamo una qualsiasi $f \in X$ e sia n tale che

$$\frac{1}{\sqrt{n}} d(f,0) < r\,.$$

Consideriamo ora la funzione integrale

$$F(x) = \int_0^x \sqrt{|f(s)|}\,\mathrm{d}s\,.$$

Vale $F(0) = 0$ ed $F(1) = d(f, 0)$. La continuità di F implica l'esistenza di un primo punto x_1 tale che

$$F(x_1) = \int_0^{x_1} \sqrt{|f(s)|}\, \mathrm{d}s = \frac{1}{n} d(f, 0)\,;$$

di un primo punto x_2 tale che

$$\int_{x_1}^{x_2} \sqrt{|f(s)|}\, \mathrm{d}s = \frac{1}{n} d(f, 0)\,.$$

Analogamente, si trovano esattamente n punti x_i tali che

$$\int_{x_{i-1}}^{x_i} \sqrt{|f(s)|}\, \mathrm{d}s = \frac{1}{n} d(f, 0) \qquad (\text{ove } x_0 = 0).$$

Per $1 \leq i \leq n$ definiamo le funzioni

$$g_i(s) = \begin{cases} nf(s) & \text{se } x_{i-1} < s < x_i \\ 0 & \text{altrimenti.} \end{cases}$$

In questo modo,

$$d(g_i, 0) = \frac{\sqrt{n}}{n} d(f, 0) = \frac{1}{\sqrt{n}} d(f, 0) < r\,,$$

ossia $g_i \in B(0, r) \subseteq V$. Inoltre,

$$nf = \sum_{i=1}^{n} g_i \qquad \text{ossia} \qquad f = \frac{1}{n} \sum_{i=1}^{n} g_i\,.$$

L'ultima uguaglianza mostra che anche f appartiene al convesso V.

L'arbitrarietà di f implica che $V = X$. ∎

Dimostrazione del TEOREMA 112.

La dimostrazione del Teorema 112 fa uso del *Lemma di Zorn*, che ora enunciamo. Sia $\mathcal{F}$ un insieme parzialmente ordinato. Chiamiamo *catena* ogni s.insieme di $\mathcal{F}$ che è totalmente ordinato.

Sia M un qualsiasi s.insieme di $\mathcal{F}$ e sia $z \in \mathcal{F}$. Diciamo che z è un *maggiorante* di M se per ogni $m \in M$ vale $m \leq z$; diciamo che z_0 è un *elemento massimale* di M se $z_0 \in M$ e se, inoltre, le condizioni $m \in M$ ed $m \geq z_0$ implicano $m = z_0$.

Un elemento massimale di M confrontabile *con tutti* gli elementi di M si chiama un *massimo*.

Si noti che un elemento massimale z_0 può non essere un massimo perché M può contenere elementi non confrontabili con z_0. E quindi un insieme M *parzialmente* ordinato può avere più elementi massimali, ma ha al più un massimo.

Invece, una catena ha al più un solo elemento massimale che è anche il suo massimo.

Il Lemma di Zorn si enuncia come segue:

Lemma 131 (di Zorn) *Sia $\mathcal{F}$ parzialmente ordinato. Se ogni catena di $\mathcal{F}$ ammette maggioranti, allora esiste in $\mathcal{F}$ un elemento massimale.*

Osservazione 132 Si noti che l'enunciato del Lemma di Zorn non richiede che ogni catena debba avere massimo; ossia, nell'applicare il Lemma di Zorn, basta mostrare l'esistenza in $\mathcal{F}$ di maggioranti delle singole catene. ∎

Possiamo ora provare il Teorema 112. In questo caso $\Phi = \mathbb{R}$.

Notiamo che l'enunciato del teorema non fa alcun riferimento alla presenza di norme su X; e quindi nessuna considerazione topologica può usarsi nella dimostrazione.

L'idea della dimostrazione del Teorema 112 è la seguente: si prova che se $Y \neq X$ allora esiste un'estensione lineare propria L_1 di L_0 e che L_1, definito sul sottospazio Y_1 propriamente contenente Y_0, soddisfa ancora alla disuguaglianza

$$L_1 x \leq p(x) \qquad \forall x \in Y_1 \,.$$

Per indicare che vale questa disuguaglianza, diremo che L_1 è *dominata* da p.

Questa è la parte tecnica della dimostrazione, che vedremo in seguito. Accettando ciò, si indichi con $\mathcal{F}$ la famiglia di tutte le estensioni di L_0 dominate da p; ossia, $L \in \mathcal{F}$ se L estende L_0 ed $Lx \leq p(x)$ per ogni x nel dominio di L.

In $\mathcal{F}$ si introduce una relazioni d'ordine parziale ponendo

$$L_1 \leq L_2 \qquad \text{se } L_2 \text{ estende } L_1.$$

Si vede immediatamente che ogni catena $\mathcal{C}$ in $\mathcal{F}$ ammette come maggiorante quell'elemento $\tilde{L} \in \mathcal{F}$ il cui dominio è l'unione dei domini di tutti gli elementi della catena $\mathcal{C}$ e così definito: se $x \in \mathrm{dom}\tilde{L}$ allora esiste $L' \in \mathcal{C}$ (non unico) il cui dominio contiene x. Si definisce

$$\tilde{L}x = L'x \,.$$

Si è detto che L' non è unico; ma l'essere $\mathcal{C}$ una catena prova che la trasformazione $\tilde{L}$ appena definita è univoca e lineare.

E' ovvio che $\tilde{L}$ è dominato da p.

Discende dal Lemma di Zorn che $\mathcal{F}$ ha un elemento massimale L: un'estensione L di L_0, dominata da p, non ulteriormente estendible, e quindi con dominio X.

Proviamo ora la parte tecnica della dimostrazione, consistente nella costruzione dell'estensione L_1. Prima però notiamo:

Osservazione 133 Sia X uno spazio di Banach di dimensione infinita e sia $\{e_n\}$ un sistema linearmente indipendente di elementi, tutti di norma 1. Definiamo

$$L_0 e_n = n$$

e quindi estendiamo L_0 a

$$\text{span}\,\{e_n\} = \{\sum_{\text{finita}} \alpha_i e_i\}\,.$$

L'argomento precedente basato sul *lemma di Zorn* può ripetersi anche senza far intervenire la funzione $p(x)$ e conduce a provare l'esistenza di un funzionale L definito su X e che estende L_0. Ciò mostra che esistono funzionali lineari non continui, il cui dominio è tutto lo spazio X. ∎

Veniamo ora a costruire una opportuna estensione L_1 di L_0, dominata da p.

Essendo $Y_0 \neq X$, esiste $x_0 \notin Y_0$. Scegliamo per Y_1 lo spazio lineare generato da Y_0 e dal x_0,

$$Y_1 = \{y + tx_0 \mid y \in Y_0\,,\ t \in \mathbb{R}\}\,.$$

Se vogliamo che L_1 sia un'estensione lineare di L_0, dovrà essere

$$L_1(y + tx_0) = L_1 y + tL_1 x_0 = L_0 y + tL_1 x_0$$

e il problema si riduce a trovare un valore ξ per $L_1 x_0$, in modo tale che L_1 sia dominato da p. Si vuole cioè che valga

$$L_0 y + t\xi \leq p(y + tx_0)\,, \qquad \forall y \in Y_0\,,\ t \in \mathbb{R}.$$

Distinguiamo i due casi $t > 0$ e $t < 0$. Se $t > 0$ allora si richiede

$$L_0 \frac{y}{t} + \xi \leq p(\frac{y}{t} + x_0) \qquad \text{ossia} \qquad \xi \leq p(\frac{y}{t} + x_0) - L_0 \frac{y}{t}\,; \tag{1.59}$$

Se $t < 0$ si richiede

$$L_0 \frac{y}{t} + \xi \geq -\frac{1}{-t} p(y + tx_0) = -p(-\frac{y}{t} - x_0)\,,$$

ossia

$$\xi \geq -L_0 \frac{y_0}{t} - p(-\frac{y}{t} - x_0)\,. \tag{1.60}$$

Dunque è da provare l'esistenza di un numero ξ che verifica la (1.59) per $t > 0$ e la (1.60) per $t < 0$.

Notiamo che, essendo y arbitrario in Y_0, anche y/t, $t > 0$ e $-y/t$, $t < 0$ sono arbitrari elementi di Y_0. Dunque, le (1.59), (1.60) equivalgono a

$$\sup_{y \in Y_0} \{-L_0 y - p(-y - x_0)\} \leq \xi \leq \inf_{y \in Y_0} \{p(y + x_0) - L_0 y\}\,.$$

Dunque. il numero ξ esiste se si può provare che

$$-L_0 \hat{y} - p(-\hat{y} - x_0) \leq -L_0 \tilde{y} + p(\tilde{y} + x_0) \qquad \forall \hat{y}\,,\ \tilde{y} \in Y_0\,.$$

Quest'eguaglianza equivale a

$$L_0(\tilde{y} - \hat{y}) \leq p(\tilde{y} + x_0) + p(-\hat{y} - x_0)$$

e questa vale certamente perché, per ipotesi, L_0 è dominato da p. Dunque si ha:

$$L_0(\tilde{y} - \hat{y}) \leq p(\tilde{y} - \hat{y}) = p(\tilde{y} + x_0 - x_0 - \hat{y}) \leq p(\tilde{y} + x_0) + p(-x_0 - \hat{y})\,.$$

Ciò completa la dimostrazione nel caso in cui $\Phi = \mathbb{R}$. ∎

Dimostrazione del TEOREMA 113.

Il caso $\Phi = \mathbb{C}$ si fa discendere dal precedente, procedendo come segue.

Si nota che se X è uno spazio lineare complesso, allora esso è anche uno spazio lineare reale. Indichiamo allora con i simboli $X_{(R)}$ ed $Y_{0,(R)}$ gli spazi lineari reali i cui elementi sono quelli di X e di Y_0. Sia inoltre $L_{0,(R)}$ il funzionale lineare definito su $Y_{0,(R)}$ da

$$L_{0,(R)} x = \Re e\, L_0 x\,.$$

La (1.52) mostra che $L_{0,(R)}$ è dominato da p e quindi che esiste un'estensione $L_{(R)}$ di $L_{0,(R)}$ ad $X_{(R)}$, ancora dominata da p. Inoltre, essendo $L_R x \leq p(x)$ per ogni $x \in X$, vale anche

$$L_{(R)}(-x) \leq p(-x) = p(x)$$

ossia,

$$-p(-x) \leq L_{(R)}(x) \leq p(x)\,. \tag{1.61}$$

Definiamo ora l'operatore

$$Lx = L_{(R)}x - iL_{(R)}(ix)\,.$$

Si verifica immediatamente che quest'operatore è lineare su X (con $\Phi = \mathbb{C}$). Inoltre, se $x \in Y_0$, vale

$$Lx = \Re e\, L_0x - i\Re e\,[iL_0x] = \Re e\, L_0x - i\Re e\,\{i\Re e\, L_0x - \mathcal{I}m\, L_0x\} = L_0x\,.$$

Dunque, L estende L_0 e inoltre vale $|Lx| \leq p(x)$ per ogni $x \in X$. Infatti se ciò non fosse potrebbe trovarsi x_0 tale che

$$|Lx_0| > p(x_0)\,.$$

Con $\theta = \arg Lx_0$ avremmo

$$L_Rx_0 = \Re e\, Lx_0 = |e^{i\theta}Lx_0| = |Lx_0| > p(x_0)\,.$$

Ciò contrasta con (1.61).

La dimostrazione è ora completa. ∎

Dimostrazione del TEOREMA 117.

Ricordiamo che si è supposto $\Phi = \mathbb{R}$.

Sia A un convesso aperto. A meno di una traslazione, si può supporre $0 \in A$. Definiamo il seguente funzionale p_A su X:

$$p_A(x) = \inf\{t > 0 \mid x \in tA\}$$

dove con tA si intende l'insieme $\{ty \mid y \in A\}$.

E' chiaro che il funzionale p_A è non negativo, positivamente omogeneo. Inoltre, esso è subadditivo. Infatti, sia $\epsilon > 0$ e siano $t_1 = p_A(x) + \epsilon$, $t_2 = p_A(y) + \epsilon$. Vale quindi

$$(x + y) \in [p_A(x) + \epsilon]A + [p_A(y) + \epsilon]A = [p_A(x) + p_A(y) + 2\epsilon]A\,.$$

Dunque, per ogni $\epsilon > 0$ vale $p_A(x+y) \leq p_A(x) + p_A(y) + 2\epsilon$. La subadditività segue dall'arbitrarietà di ϵ.

Consideriamo ora un punto $x_0 \notin A$ ed il s.spazio

$$X_0 = \{tx_0 \mid t \in \mathbb{R}\}\,.$$

La convessità di A implica la convessità di $A \cap X_0$ e quindi l'insieme $\{t \mid tx_0 \in A\}$ è un segmento a cui non appartiene 1. Si definisca il funzionale lineare L_0 su X_0,

$$L_0(tx) = t\,.$$

Questo funzionale verifica

$$L_0(x_0) = 1\,, \qquad L_0(tx_0) \le p_A(tx_0)\,.$$

La prima proprietà è immediata e la seconda è ovvia se $t < 0$. Che essa valga anche se $t > 0$ si vede notando che $p_A(x_0) \ge 1$ e quindi

$$L_0(tx_0) = t \le tp_A(x_0) = p_A(tx_0)\,.$$

Per il *Teorema di Hahn–Banach,* il funzionale L_0 si estende ad un funzionale L lineare e continuo su X, che ancora verifica $\langle\langle L, x\rangle\rangle \le p_A(x)$ e quindi

$$\langle\langle L, x\rangle\rangle \le p_A(x) \le 1 \quad \forall x \in A\,, \qquad \langle\langle L, x_0\rangle\rangle = 1\,. \ \blacksquare$$

Il funzionale p_A si chiama il *finzionale di* *Minkowski* del convesso A.

Dimostrazione del TEOREMA 118.

Il Teorema 118 è immediata conseguenza del Teorema 117. Studiamo prima il caso in cui A è aperto. Fissiamo $a_0 \in A$ e $b_0 \in B$ e sia $x_0 = a_0 - b_0$. Sia inoltre

$$C = A - B = \{a - b \mid a \in A\,,\ b \in B\} = \bigcup_{b\in B} (A - b)\,.$$

E' facile vedere che C è aperto (anche se B non lo è); e che $x_0 \in C$. Dunque, $0 \in D = C - x_0$ mentre $x_0 \notin D$. Infatti, se fosse $x_0 \in D$, esisterebbero $\tilde a \in A$, $\tilde b \in B$ per cui

$$x_0 = \tilde a - \tilde b - x_0 \quad \text{ossia} \quad \tilde a = \tilde b$$

mentre per ipotesi A e B sono disgiunti.

Dunque, per il Teorema 117, x_0 si separa da D con un funzionale x^*:

$$\langle\langle x^*, d\rangle\rangle \le \langle\langle x^*, x_0\rangle\rangle \qquad \forall d \in D\,.$$

Dunque,

$$\langle\langle x^*, (a - b + x_0)\rangle\rangle < \langle\langle x^*, x_0\rangle\rangle \qquad \forall a \in A\,, \quad b \in B\,;$$

Ossia,

$$\langle\langle x^*, a\rangle\rangle < \langle\langle x^*, b\rangle\rangle\,.$$

Ciò prova la prima parte del teorema.

La dimostrazione della seconda parte si riconduce a quella della prima, grazie al lemma seguente:

Lemma 134 *Siano K e C due s.insiemi di X, K compatto e C chiuso. Se $K \cap C = \emptyset$, allora esiste una sfera B di X tale che*

$$(K + B) \cap C = \emptyset \, .$$

Dim. Per ogni $k \in K$ esiste una opportuna sfera $B_{\epsilon(k)}$ tale che $k + 2B_{\epsilon(k)} \cap (C + B_{\epsilon(k)}) = \emptyset$. L'unione degli aperti $k_i + 2B_{\epsilon(k)}$ copre K e quindi esistono k_i, in numero finito, diciamo n, tali che

$$K \subseteq \bigcup_{i=1}^{n} \left(k_i + 2B_{\epsilon(k_i)} \right) .$$

Sia B la sfera di minimo raggio tra le sfere $B_{\epsilon(k_i)}$. Allora, $\bigcup \left(k_i + 2B_{\epsilon(k_i)} \right)$ non interseca $C + B$ e quindi vale anche:

$$K + B \subseteq \bigcup_{i=1}^{+\infty} \left(k_i + B_{\epsilon(k_i)} + B \right) \subseteq \bigcup_{i=1}^{n} \left(k_i + B_{\epsilon(k_i)} + B_{\epsilon(k_i)} \right)$$

e quest'insieme non interseca C. ∎

E' ora facile provare la seconda parte del Teorema 113. Sia B_ϵ una sfera tale che $A_1 = A + B_\epsilon$ non intersechi B. Gli insiemi A_1 e B verificano le ipotesi della prima parte del teorema e quindi esiste x^* tale che

$$\sup\{\langle\langle x^*, a \rangle\rangle \mid a \in A_1\} \leq \inf\{\langle\langle x^*, b \rangle\rangle \mid b \in B\} \, .$$

Ciò prova che x^* separa A e B. Proviamo ora che la separazione è stretta. Notiamo che

$$\langle\langle x^*, A_1 \rangle\rangle = \{\langle\langle x^*, x \rangle\rangle \mid x \in A_1\}$$

è un intervallo aperto di $\mathbb{R}$. Infatti, sia $\alpha = \langle\langle x^* a_0 \rangle\rangle \in \langle\langle x^*, A_1 \rangle\rangle$ e $B_\sigma = \{x \mid ||x|| < \sigma\}$ una sfera tale che $a_0 + B_\sigma \in A_1$. Allora, $\langle\langle x^*, (a_0 + B_\sigma) \rangle\rangle$ è un intervallo non ridotto ad un sol punto perché il funzionale x^* non è zero; e dunque α è interno ad $\langle\langle x^*, A_1 \rangle\rangle$. D'altra parte, $\langle\langle x^*, A \rangle\rangle$ è compatto e contenuto nell'aperto $\langle\langle x^*, A_1 \rangle\rangle$ e quindi ha distanza positiva dall'insieme $\langle\langle x^*, B \rangle\rangle$, che non interseca $\langle\langle x^*, A_1 \rangle\rangle$. Ciò prova che la separazione è stretta. ∎

Osservazione 135 Abbiamo enunciato i teoremi di separazione assumendo $\Phi = \mathbb{R}$ per semplicità. Tali teoremi possono estendersi al caso $\Phi = \mathbb{C}$. In questo caso le diseguaglianze valgono tra le parti reali. Si hanno cioè condizioni del tipo

$$\sup\{\Re e \, La \mid a \in A\} \leq \inf\{\Re e \, Lb \mid b \in B\} \, . \; ∎$$

1.10 Convergenza debole e debole stella

Si è visto che in uno spazio di Banach di dimensione infinita i compatti sono "rari". D'altra parte, si sa dai corsi precedenti che la proprietà di compattezza è cruciale nello studio dei problemi di minimo. Infatti:

Teorema 136 (di Weierstrass) *Una funzione* f continua *su un insieme* compatto K *ammette sia minimo che massimo.*

Accenniamo alla dimostrazione, che dovrebbe essere nota: per provare l'esistenza del minimo, si costruisce una *successione minimizzante* (k_n) in K, ossia una successione tale che

$$\lim f(k_n) = \inf\{f(k) \mid k \in K\}\,.$$

Si usa la compattezza di K per estrarre una sottosuccessione (k_{n_r}) convergente a $k_0 \in K$; e la continuità di f per concludere

$$f(k_0) = \lim f(k_{n_r}) = \inf\{f(k) \mid k \in K\}\,. \tag{1.62}$$

Ciò prova che k_0 è punto di minimo.

Osservazione 137 Alla medesima conclusione si giunge se la funzione f, invece di essere continua, ha soltanto epigrafo chiuso, ossia è *semicontinua inferiormente*. Infatti in tal caso da $k_{n_r} \to k_0$, $f(k_{n_r}) \to \inf\{f(k) \mid k \in K\}$. Se l'epigrafo è chiuso si ha

$$(k_0, \inf\{f(k) \mid k \in K\}) \in \mathrm{Epi} f\,;$$

ossia, invece della (1.62), vale

$$f(k_0) \leq \lim f(k_{n_r}) = \inf\{f(k) \mid k \in K\}\,.$$

Ovviamente la disuguaglianza stretta non può valere, e quindi, nella sola ipotesi di semicontinuità inferiore, segue la (1.62); ossia segue che k_0 è punto di minimo. ∎

E' facile vedere che:

- se un insieme è compatto in una topologia, tale rimane anche in topologie meno fini;
- se una funzione è continua oppure semicontinua inferiormente rispetto ad una topologia, tale rimane anche in topologie più fini.

Dunque in uno spazio di Banach conviene introdurre topologie meno fini di quella della norma, in modo da avere più insiemi compatti; ma sufficientemente fini in modo tale che almeno opportune classi di funzioni continue rimangano, se non continue, almeno semicontinue inferiormente.

Noi ci limiteremo ad introdurre concetti di convergenza per topologie meno fini di quella della norma. Non descriveremo invece la topologia.

Definitione 138 Sia (x_n) una successione in uno spazio di Banach X. Diciamo che (x_n) *converge* *debolmente* ad x_0 se

$$\lim \langle\langle x^*, x_n \rangle\rangle = \langle\langle x^*, x_0 \rangle\rangle \qquad \forall x^* \in X^* \,. \quad \blacksquare$$

Per indicare la convergenza debole si usa uno dei due simboli

$$\mathrm{w}-\lim x_n = x_0 \qquad \text{oppure} \qquad x_n \rightharpoonup x_0 \,.$$

La definizione stessa di convergenza debole mostra che:

Teorema 139 *Ogni $x^* \in X^*$ è continuo rispetto alla convergenza debole di successioni di X.*

Supponiamo ora di sapere che stiamo lavorando nello spazio di Banach X^*, duale dello spazio di Banach X. In X^* si può definire la convergenza debole, come sopra, facendo intervenire il suo duale; ma si può anche definire la *convergenza debole* *stella*, come segue:

Definitione 140 Sia (x_n^*) una successione in X^*. Diciamo che (x_n^*) *converge in senso* *debole stella* ad x_0^* se

$$\lim \langle\langle x_n^*, x \rangle\rangle = \langle\langle x_0^*, x \rangle\rangle \qquad \forall x \in X \,. \quad \blacksquare$$

Per indicare la convergenza debole stella si usa uno dei due simboli

$$\mathrm{w}^*-\lim x_n^* = x_0 \qquad \text{oppure} \qquad x_n^* \rightharpoonup x_0 \,.$$

(la notazione con la mezza freccia è la stessa per la convergenza debole e la convergenza debole stella. Il contesto evita ambiguità).

E' facile vedere che la convergenza in norma implica la convergenza debole, rispettivamente debole stella. Il viceversa, in dimensione infinita, non vale.

Esempio 141 Sia $X = L^2(0, 2\pi)$. Vedremo che $(L^2(0, 2\pi))^*$ è isometricamente isomorfo ad $L^2(0, 2\pi)$ stesso. In particolare ogni $g \in L^2(0, 2\pi)$ si interpreta come elemento di $(L^2(0, 2\pi))^*$, quell'elemento definito da

$$f \longrightarrow \int_0^{2\pi} \bar{g}(x) f(x)\, \mathrm{d}x\,.$$

Si sa, dalla teoria della serie di Fourier, che ogni $f \in L^2(0, 2\pi)$ si rappresenta

$$f(x) = \sum_{k=-\infty}^{k=+\infty} f_k e^{ikx}\,, \qquad f_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-ikx}\, \mathrm{d}x = \langle\langle e_k, f \rangle\rangle$$

con $e_k = e^{ikx}$. Si sa, dalla teoria della serie di Fourier, che

$$\lim f_k = 0\,.$$

Questo vuol dire che,

$$\mathrm{w}^*\!-\lim e_k = 0\,.$$

Ma,

$$||e_k||_2 = \sqrt{2\pi}$$

e quindi la successione (e_k) non tende a zero in norma. ∎

Vale però:

Teorema 142 *Il teorema di unicità del limite vale sia per la convergenza debole che per la convergenza debole stella.*

Dim. Se

$$\mathrm{w}\!-\lim x_n = x_0 \qquad \text{ed anche} \qquad \mathrm{w}\!-\lim x_n = y_0$$

allora si ha

$$\langle\langle x^*, (x_0 - y_0) \rangle\rangle = \lim\{\langle\langle x^*, x_n \rangle\rangle - \langle\langle x^*, x_n \rangle\rangle\} = 0$$

per ogni $x^* \in X^*$. Dunque x_0 ed y_0 sono indistinguibili da elementi di X^*. E' conseguenza del Teorema di Hahn-Banach che $x_0 = y_0$.

L'asserto relativo alla convergenza debole stella è invece elementare: sia

$$\mathrm{w}^*\!-\lim x_n^* = x_0^* \qquad \text{ed anche} \qquad \mathrm{w}^*\!-\lim x_n^* = y_0^*\,.$$

Si ha

$$\langle\langle [x_0^* - y_0^*], x \rangle\rangle = \lim\{\langle\langle x_n^*, x \rangle\rangle - \langle\langle x_n^*, x \rangle\rangle\} = 0$$

per ogni $x \in X$. E quindi $x^* = y^*$. ∎

Esistono alcune relazioni importanti tra la convergenza debole, oppure debole stella, e le proprietà che valgono in norma. Tra queste:

Teorema 143 *Ogni successione convergente in senso debole stella è limitata nella norma di X^*.*

Dim. Sia $x \in X$. Se $\mathrm{w}^* - \lim x_n^* = x^*$ allora

$$\lim \langle\langle x_n^*, x \rangle\rangle = \langle\langle x^*, x \rangle\rangle \, .$$

Dunque, per ogni $x \in X$ esiste un numero M_x, dipendente da x, tale che

$$|\langle\langle x_n^*, x \rangle\rangle| < M_x \, .$$

Il teorema di Banach-Steinhaus mostra che l'insieme dei funzionali x_n^* è limitato nella norma di X^*. ∎

Un asserto analogo vale anche per la convergenza debole, ed anzi il risultato relativo alla convergenza debole è un corollario del precedente. Per provarlo, abbiamo bisogno di una premessa. Ogni $x \in X$ si può vedere come funzionale lineare continuo su X^*. Basta associare ad esso il funzionale

$$x^* \to \langle\langle x^*, x \rangle\rangle \, . \tag{1.63}$$

Dunque, ogni elemento $x \in X$ si può porre in corrispondenza con un elemento $jx \in (X^*)^*$, duale di X^*.

Lo spazio $(X^*)^*$ si chiama il *biduale* di X e si indica col simbolo X^{**}.

Vale:

Teorema 144 *La trasformazione j da X in X^{**} definita in (1.63) è isometrica.*

Dim. Dobbiamo provare che

$$||x||_X = ||jx||_{X^{**}} \, ,$$

ossia:

$$||x||_X = \sup_{||x^*||_{X^*}} |\langle\langle x^*, x \rangle\rangle| \, .$$

Questo è l'asserto del Corollario (115). ∎

Osservazione 145 In particolare segue che j è una trasformazione iniettiva. Vedremo che, in generale, essa non è suriettiva. ∎

Possiamo ora provare:

Corollario 146 *Ogni successione debolmente convergente in X è limitata in norma.*

Dim. Sia (x_n) una successione in X, debolmente convergente ad x_0. Ciò vuol dire che

$$\lim \langle\langle x^*, x_n \rangle\rangle = \langle\langle x^*, x_0 \rangle\rangle \qquad \forall x^* \in X^* ,$$

ossia

$$\lim \langle\langle jx_n, x^* \rangle\rangle = \langle\langle jx_0, x^* \rangle\rangle \qquad \forall x^* \in X^* .$$

Dunque, la successione (jx_n) converge in X^{**} a jx_0, nel senso debole stella ed è quindi limitata in X^{**}. Il Teorema 144 implica che (x_n) è limitata in X. ∎

Osservando con attenzione la dimostrazione del Teorema 143 si vede che:

Teorema 147 *vale:*

- *se* $\mathrm{w}-\lim x_n = x_0$ *allora* $\liminf ||x_n||_X \geq ||x_0||_X$*;*
- *se* $\mathrm{w}^* - \lim x_n^* = x_0^*$ *allora* $\liminf ||x_n^*||_{X^*} \geq ||x_0^*||_{X^*}$.

Dim. Proviamo l'asserto relativo alla convergenza debole stella.

Per ogni x, con $||x||_X \leq 1$, vale

$$\langle\langle x_n^*, x \rangle\rangle \leq ||x_n^*||_{X^*}$$

e dunque

$$\langle\langle x_0^*, x \rangle\rangle = \lim \langle\langle x_n^*, x \rangle\rangle \leq \liminf ||x_n^*||_{X^*} .$$

Da qui:

$$||x_0^*||_{X^*} = \sup_{||x||_X = 1} \langle\langle x_0^*, x \rangle\rangle = \sup_{||x||_X = 1} \lim \langle\langle x_n^*, x \rangle\rangle \leq \liminf ||x_n^*||_{X^*} .$$

L'asserto relativo alla convergenza debole si deduce dall'asserto relativo alla convergenza debole stella mediante il Teorema 144:

$$||x_0||_X = \sup_{||x^*||_{X^*}=1} \langle\langle x^*, x_0 \rangle\rangle = \sup_{||x^*||_{X^*}=1} \lim \langle\langle x^*, x_n \rangle\rangle = \sup_{||x^*||_{X^*}=1} \lim \langle\langle jx_n, x^* \rangle\rangle \leq$$
$$\sup_{||x^*||_{X^*}=1} \{\liminf ||jx_n||_{X^{**}} ||x^*||_{X^*}\} = \liminf ||jx_n||_{X^{**}} = \liminf ||x_n||_X . \quad ∎$$

Il Teorema 147 può riformularsi dicendo che la norma di X è debolmente semicontinua inferiormente e la norma di X^* è debolmente stella semicontinua inferiormente.

Grazie al *Teorema di Weierstrass,* segue che la norma ammette minimo sugli insiemi che sono compatti rispetto alla convergenza debole oppure debole stella.

Proviamo ora:

Teorema 148 *Sia A un insieme sequenzialmente debolmente chiuso in X. Allora, A è anche chiuso in norma. Se A è sequenzialmente debolmente stella chiuso in X^*, esso è anche chiuso nella norma di X^*.*

Dim. Sia A chiuso rispetto alla convergenza debole. Sia (x_n) una successione in A, che converge in norma ad un x_0.

La convergenza in norma implica la convergenza debole, e quindi (x_n) converge debolmente ad x_0. Dato che A è debolmente chiuso, si ha $x_0 \in A$; e quindi A è chiuso in norma. L'assero relativo alla convergenza debole stella si prova in modo analogo. ∎

L'esempio 141 mostra che la superficie di una sfera di $L^2(0, 2\pi)$, pur essendo chiusa in norma, non è debolmente chiusa. Combinando questo col teorema precedente si vede che che la topologia della convergenza debole è effettivamente meno fine di quella della norma.

Vale:

Teorema 149 *Sia A un sottoinsieme di uno spazio di Banach X. Sia A* convesso *e* chiuso *rispetto alla norma. L'insieme A è anche* chiuso *rispetto alla convergenza debole.*

Dim. Sia (x_n) una successione a valori in A, debolmente convergente ad x_0. Dobbiamo provare che x_0 appartiene ad A. Ciò discende dai teoremi di separazione: se $x_0 \notin A$, allora esiste $x^* \in X^*$ che separa x_0 da A. Inoltre, la separazione è stretta perché l'insieme A è chiuso in norma e l'insieme costituito dal solo punto x_0 è compatto. Dunque, esiste $\epsilon > 0$ tale che

$$\langle\langle x^*, x_0\rangle\rangle + \epsilon < \inf\{\langle\langle x^*, x\rangle\rangle \mid x \in A\} \leq \inf\{\langle\langle x^*, x_n\rangle\rangle\}.$$

Ciò contrasta con l'ipotesi che (x_n) converge debolmente ad x_0. ∎

E' bene notare esplicitamente che l'asserto analogo per la topologia debole stella non vale. Però nella topologia debole stella successioni compatte si identificano facilmente:

Teorema 150 (di Alaoglu) *Ogni successione limitata in X^* ammette s.successioni convergenti in senso debole stella.*

La dimostrazione è posposta.

L'asserto analogo non vale in X e ciò suggerisce di dare un nome particolare agli spazi X tali che $jX = X^{**}$.

Definitione 151 Se uno spazio X è isometrico al suo biduale, lo spazio X si dice *riflessivo*.

In uno spazio riflessivo, la convergenza debole equivale alla debole stella e quindi:

Teorema 152 *Ogni successione limitata in uno spazio riflessivo ammette s.successioni debolmente convergenti.*

Il Teorema 149 ha un corollario che bene illustra la geometria della convergenza debole:

Corollario 153 (di Mazur) *Sia (x_n) una successione debolmente convergente ad x_0. Esiste una successione (s_n) in X, tale che:*

- *$s_n = \sum_{k=1}^{n} \lambda_k x_k$ per opportuni numeri $\lambda_k \in [0,1]$ tali che $\sum_{k=1}^{n} \lambda_k = 1$;*
- *la successione (s_n) converge ad x_0 in norma.*

Ossia, x_0 è limite in norma di una successione ottenuta da combinazioni convesse degli x_n.

Dim. Sia $A = \overline{\text{co}}\{x_n\}$, il più piccolo convesso chiuso contenente gli x_n. Per il Teorema 149, x_0 appartiene ad A. Da qui segue l'asserto. ∎

Applichiamo ora il Teorema 149 allo spazio $X \times \mathbb{R}$. Sia A l'epigrafo di una funzione convessa e continua su X. L'insieme A è convesso e chiuso e quindi sequenzialmente chiuso. Dunque vale:

Teorema 154 *Una funzione convessa e continua in norma è anche semicontinua inferiormente rispetto alla convergenza debole.*

Questo teorema ed il Teorema di Alaoglu mostrano che il Teorema di Weierstrass può applicarsi alla minimizzazione di funzionali convessi su insiemi limitati e chiusi su spazi riflessivi; e ciò è importantissimo nella teoria dell'ottimizzazione.

Per completare queste considerazioni, mostriamo che, in contrasto con l'asserto del Teorema di Alaoglu, in un generico spazio di Banach esistono successioni limitate che non ammettono s.successioni convergenti nella topologia debole; e quindi che esistono spazi di Banach che non sono riflessivi. Infatti:

Teorema 155 *Esistono successioni limitate in $L^1(0,1)$, prive di s.successioni debolmente convergenti; e quindi $L^1(0,1)$ non è riflessivo.*

Dim. Si è visto al Teorema 139 che ogni elemento $x^* \in X^*$ è continuo rispetto alla convergenza debole in X. Se ogni successione limitata avesse s.successioni

debolmente convergenti, allora, per il Teorema di Bolzano-Weierstrass, ogni elemento di X^* raggiungerebbe massimo su

$$\left\{ x \mid \int_0^1 |x(s)| \, \mathrm{d}s \leq 1 \right\} .$$

L'esempio riportato al Teorema 73 mostra che ciò non accade. ∎

Osservazione 156 Ciò mostra una profonda differenza tra lo spazio $L^1(a,b)$ e gli spazi $L^p(a,b)$ con $1 < p < +\infty$. Vedremo infatti che questi ultimi spazi sono riflessivi e quindi tutte le successioni limitate in $L^p(0,1)$, con $1 < p < +\infty$ ammettono s.successioni debolmente convergenti. ∎

Continuità e continuità debole di operatori lineari

Fa parte della definizione di convergenza debole che ogni elemento di X^*, ossia ogni funzionale lineare su X che è continuo in norma, è anche continuo rispetto alla convergenza debole. E' importante sapere che gli elementi di X^* sono tutti i funzionali lineari su X, continui rispetto alla convergenza debole:

Teorema 157 *Sia ϕ un funzionale lineare su X. Esso è continuo rispetto alla convergenza debole se e solo se è continuo in norma.*

Dim. Se ϕ è continuo rispetto alla convergenza debole lo è anche in norma perché ogni successione convergente in norma converge anche debolmente, al medesimo limite. Il viceversa è provato al Teorema 139. ∎

Un asserto analogo vale in realtà per generici operatori lineari, ma la dimostrazione è più profonda:

Teorema 158 *Siano X ed Y due spazi di Banach e sia A un operatore lineare da X in Y, con dominio uguale ad X e debolmente continuo, ossia tale che*

$$\textit{se} \qquad x_n \rightharpoonup x_0 \quad \textit{in } X \qquad \textit{allora} \qquad Ax_n \rightharpoonup Ax_0 \quad \textit{in } Y .$$

In tal caso, l'operatore A è continuo in norma.

Dim. Proviamo che nelle ipotesi del teorema, l'operatore lineare A, definito su X, è chiuso.

Sia $x_n \to x_0$ tale che $Ax_n \to y_0$ (rispettivamente nelle norme di X e di Y). Allora, $x_n \rightharpoonup x_0$ e quindi, per le ipotesi, $Ax_n \rightharpoonup Ax_0$; D'altra parte, $Ax_n \to y$ implica $Ax_n \rightharpoonup y$. Per l'unicità del limite debole, $y = Ax_0$ ossia la successione

di punti ((x_n, Ax_n)) se converge, converge ad un punto del grafico, che quindi è chiuso.

Ciò prova che l'operatore lineare A è chiuso e quindi, avendo dominio uguale ad X, continuo. ∎

Si osservi che la dimostrazione precedente fa uso sia del teorema di Hahn-Banach che del Teorema di Baire.

1.10.1 Dimostrazioni posposte

Dimostrazione del TEOREMA 150.

Il teorema vale in qualsiasi spazio duale. Però la dimostrazione che presentiamo usa un'ipotesi ulteriore, non richiesta dal teorema: presentiamo la dimostrazione nel caso in cui lo spazio X è separabile. Dunque supponiamo l'esistenza di una successione (x_n) la cui immagine è densa in X. Se vale quest'ipotesi, la dimostrazione si ottiene usando il *procedimento diagonale di Cantor*.

Sia (y_n^*) una successione limitata in X^*, $||y_n^*|| < \alpha$. Vogliamo estrarne una s.successione convergente.

Consideriamo la successione numerica $(\langle\langle y_n^*, x_1\rangle\rangle)$. Questa è limitata e quindi ammette una s.successione convergente. Indichiamola $(\langle\langle y_{n,1}^*, x_1\rangle\rangle)$. Consideriamo quindi $(\langle\langle y_{n,1}^*, x_2\rangle\rangle)$. Anche questa successione è limitata e quindi se ne estrae una s.successione convergente, che indichiamo $(\langle\langle y_{n,2}^*, x_2\rangle\rangle)$. Procedendo in questo modo si costruisce la tavola seguente:

$\langle\langle y_{1,1}^*, x_1\rangle\rangle$	$\langle\langle y_{2,1}^*, x_1\rangle\rangle$	$\langle\langle y_{3,1}^*, x_1\rangle\rangle$	$\langle\langle y_{4,1}^*, x_1\rangle\rangle$	$\dots$
$\langle\langle y_{1,2}^*, x_2\rangle\rangle$	$\langle\langle y_{2,2}^*, x_2\rangle\rangle$	$\langle\langle y_{3,2}^*, x_2\rangle\rangle$	$\langle\langle y_{4,2}^*, x_2\rangle\rangle$	$\dots$
$\vdots$				
$\langle\langle y_{1,k}^*, x_k\rangle\rangle$	$\langle\langle y_{2,k}^*, x_k\rangle\rangle$	$\langle\langle y_{3,k}^*, x_k\rangle\rangle$	$\langle\langle y_{4,k}^*, x_k\rangle\rangle$	$\dots$
$\vdots$				

la successione $(y_{n,k}^*)$ che figura su ciascuna riga è s.successione di tutte le successioni $y_{n,r}^*$, con $r < k$; e quindi $(\langle\langle y_{n,k}^*, x_r\rangle\rangle)$ converge, per ogni $r \leq k$.

Dunque, la successione diagonale $(y_{n,n}^*)$ ha la proprietà che $(\langle\langle y_{n,n}^*, x_r\rangle\rangle)$ converge per ogni r.

Proviamo che in realtà la successione $(\langle\langle y_{n,n}^*, x_0\rangle\rangle)$ converge per ogni $x_0 \in X$. Basta provare che essa è fondamentale, dato che essa prende valori in Φ.

Si fissi per questo $\epsilon > 0$ ed $x_0 \in X$, qualsiasi. Sia N tale che

$$||x_0 - x_N|| < \epsilon/(3\alpha)\,.$$

Si noti che N dipende da x_0. Si stimi quindi

$$\begin{aligned}&|\langle\langle y^*_{m,m}, x_0\rangle\rangle - \langle\langle y^*_{n,n}, x_0\rangle\rangle| \\ &\leq |\langle\langle y^*_{m,m}, x_0 - x_N\rangle\rangle| + |\langle\langle y^*_{m,m} - y^*_{n,n}, x_N\rangle\rangle + |\langle\langle y^*_{n,n}, x_N - x_0\rangle\rangle| \,.\end{aligned}$$

Il primo e l'ultimo addendo sono minori di $\epsilon/3$, per la scelta fatta di x_N. L'addendo intermedio è minore di $\epsilon/3$ se m, n superano un numero $\tilde{N}$, che dipende da ϵ e da N ossia da ϵ e da x_0. Ciò prova la convergenza della successione di numeri $(\langle\langle y^*_{n,n}, x\rangle\rangle)$, per ogni $x \in X$.

Si costruisce ora un funzionale y^*_0 su X ponendo

$$\langle\langle y^*_0, x\rangle\rangle = \lim \langle\langle y^*_{n,n}, x\rangle\rangle \,.$$

E' immediato verificare che y^*_0 è lineare. Inoltre, y^*_0 è continuo perché la succesione (y^*_n) è limitata, $||y^*_n||_{X^*} < \alpha$ per ogni n così che per $||x||_X < 1$,

$$\left|\langle\langle y^*_0, x\rangle\rangle\right| = \lim \left|\langle\langle y^*_{n,n}, x\rangle\rangle\right| \leq \alpha \,.$$

Dunque y^*_0 è continuo. E quindi, $y^*_0 = \mathrm{w}^* - \lim y^*_{n,n}$. Ciò è quanto volevamo provare. ∎

Osservazione 159 Nel caso in cui il duale X^* di X sia separabile, può sembrare possibile applicare il ragionamento precedente ad una successione (x_n) di X, arrivando a provare un analogo del teorema di Alaoglu in X. Ci si convinca che ciò è falso esaminando bene la dimostrazione e notando che, tentando di ripetere la dimostrazione, NON si proverebbe la convergenza debole di (x_n) in X, ma la convergenza debole stella di (jx_n) in X^{**}. ∎

1.11 Esempi di spazi duali

Sia X uno spazio di Banach. Per definizione, X^* è lo spazio (di Banach) dei funzionali lineari e continui su X. Il problema che vogliamo studiare ora è il seguente: se X è uno spazio "particolare", per esempio uno spazio di funzioni o di successioni, vogliamo vedere se esiste un altro spazio "particolare" Y che è *isometricamente isomorfo* ad X^*; ossia tale che esista una trasformazione L da Y in X^* che è 1) suriettiva; 2) isometrica; 3) lineare (se $\Phi = \mathbb{R}$) oppure antilineare (se $\Phi = \mathbb{C}$).[5] Ricordiamo che la trasformazione L è isometrica quando

$$||Ly||_{X^*} = ||y||_Y \,.$$

E quindi una trasformazione isometrica è necessariamente iniettiva.

In questo caso, Y viene ad avere tutte le proprietà topologiche di X^*. Inoltre, definendo $y_n \rightharpoonup y_0$ quando $\mathrm{w}^*-\lim Ly_n = Ly_0$ si trasferiscono ad Y le proprietà della convergenza debole stella di X^*. Si dice allora che Y è una *realizzazione* di X^* e, frequentemente, non si distingue tra Y ed X^*.

Un esempio particolare è ben noto: se $X = l^2(n)$, lo spazio euclideo n-dimensionale, allora una realizzazione del duale è lo spazio stesso.

Non sempre è possibile trovare delle realizzazioni concrete (e comode) di uno spazio duale; e d'altra parte esistono spazi di Banach che non sono isometricamente isomorfi a nessuno spazio duale. Per questo conviene elencare alcuni casi particolarmente importanti. Prima di presentare le dimostrazioni, raccogliamo i risultati nella tabella seguente:

spazio	duale
l^1	l^∞
$l^p\,, \quad p < +\infty$	$l^{p'}\,, \quad p' = p/(p-1)$
c_0	l^1
$L^1(\Omega)$	$L^\infty(\Omega)$
$L^p(\Omega)\,, \quad p < +\infty$	$L^{p'}(\Omega)\,, \quad p' = p/(p-1)$
$C(a,b)$	$NV(a,b)$

Lo spazio $NV(a,b)$ è definito in seguito.

Non abbiamo inserito nella tabella precedente gli spazi l^∞ ed $L^\infty(\Omega)$. Realizzazioni dei loro duali sono note, ma per descriverle avremmo bisogno di conoscenze di teoria della misura che non abbiamo presentato.

[5] Una trasformazione L si dice *antilineare* se vale $L(\alpha x + \beta y) = \bar{\alpha} Lx + \bar{\beta} Ly$.

Il duale di l^p, $1 \leq p < +\infty$

Per caratterizzare il duale di l^p ed anche di c_0 abbiamo bisogno di una particolare successione di elementi dello spazio l^p stesso, che indichiamo con $(e^{(n)})$. Dunque, ciascun $e^{(n)}$ è a sua volta una successione di numeri. Per definizione,

$$e_i^{(n)} = \begin{cases} 1 & \text{se } i = n \\ 0 & \text{altrimenti.} \end{cases} \tag{1.64}$$

Notiamo che $||e^{(n)}||_p = 1$ per ogni p, $1 \leq p \leq +\infty$ e che lo spazio lineare generato dagli elementi $e^{(n)}$ è denso in l^p per ogni p, $1 \leq p < +\infty$. Non è invece denso in l^∞.

Sia $X = l^p$ con $1 \leq p < +\infty$. In questo caso una realizzazione di X^* è

$$l^{p'} \qquad \text{con} \qquad p' = \begin{cases} +\infty & \text{se} \quad p = 1 \\ \frac{p}{p-1} & \text{se} \quad p > 1\,. \end{cases} \tag{1.65}$$

Proviamo ciò prima di tutto nel caso $p = 1$. Sia $(y_n) \in L^\infty$. Si vede immediatamente che il funzionale lineare x^*, dipendente da (y_n),

$$\langle\langle x^*, x \rangle\rangle = \sum_{n=0}^{+\infty} \bar{y}_n x_n \tag{1.66}$$

è lineare, ed è continuo perché

$$|\langle\langle x^*, x \rangle\rangle| = |\sum_{n=0}^{+\infty} \bar{y}_n x_n| \leq \sup_n\{|y_n|\} \cdot ||x||_1\,.$$

Inoltre, la trasformazione $y = (y_n) \to x^*$ è antilineare e si vede facilmente che è isometrica. Infatti, la disuguaglianza precedente mostra che

$$||x^*|| = \sup_{||x||_1 = 1} \langle\langle x^*, x \rangle\rangle \leq ||y||_\infty\,. \tag{1.67}$$

Per vedere che vale anche la disuguaglianza inversa, e quindi l'uguaglianza, si scelga la successione $x^{(N)}$ definita da

$$x_r^{(N)} = \begin{cases} \frac{y_r}{|y_r|} & \text{se } r = N \text{ e } y_r \neq 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Ovviamente, $||x^{(N)}||_1 \leq 1$ e

$$\sup_N \langle\langle x^*, x^{(N)} \rangle\rangle = \sup_N |y_N| = ||y||_\infty$$

e quindi in (1.67) vale l'uguaglianza.

Per completare la dimostrazione, dobbiamo far vedere che la trasformazione che ad $y \in l^\infty$ associa $x^* \in (l^1)^*$ data da (1.66) è suriettiva; ossia dobbiamo assegnare un qualsiasi $x^* \in (l^1)^*$ ed associargli un opportuno $y \in l^\infty$, in modo che valga (1.66). Per costruire y consideriamo la successione $e^{(n)}$ in (1.64). Definiamo y ponendo

$$y_n = \overline{\langle\langle x^*, e^{(n)} \rangle\rangle}\,.$$

Da

$$|y_n| \leq ||x^*||$$

segue che $y \in l^\infty$.

Sia ora $x \in l^1$. Associamogli la successione $x^{(N)}$ definita come segue:

$$x^{(N)} = \sum_{k=0}^{N} x_k e^{(k)} \qquad \text{ossia} \quad x_k^{(N)} = \begin{cases} x_k & \text{se } k \leq N \\ 0 & \text{altrimenti.} \end{cases}$$

Si ha

$$\langle\langle x^*, x \rangle\rangle = \lim_N \langle\langle x^*, x^{(N)} \rangle\rangle = \lim_N \langle\langle x^*, \sum_{k=0}^{N} x_k e^{(k)} \rangle\rangle$$
$$= \lim_N \sum_{k=0}^{N} x_k \langle\langle x^*, e^{(k)} \rangle\rangle = \lim_N \sum_{k=0}^{N} \bar{y}_k x_k = \sum_{k=0}^{+\infty} \bar{y}_k x_k\,.$$

Si noti che l'ultima uguaglianza si giustifica perchè già sappiamo che $y \in l^\infty$ e già sappiamo che, in tal caso,

$$x \to \sum_{k=0}^{+\infty} \bar{y}_k x_k$$

è continua su l^1.

Ciò completa l'analisi del caso $p = 1$.

In modo analogo **trattiamo il caso** $1 \leq p < +\infty$**.**

Siano $x = (x_n) \in l^p$ ed $y = (y_n) \in l^{p'}$. Dalla disuguaglianza di Hölder si vede che:

$$\sum_{k=0}^{+\infty} \bar{y}_n x_n \leq ||(y_n)||_{p'} \cdot ||(x_n)||_p = ||y||_{p'} \cdot ||x||_p\,.$$

Ciò mostra che la trasformazione lineare

$$x \to \sum_{k=0}^{+\infty} \bar{y}_k x_k$$

è continua su l^p e suggerisce di considerare la trasformazione L da $Y = l^{p'}$ in X^*:

$$(Ly)(x) = \sum_{k=0}^{+\infty} \bar{y}_n x_n \,,$$

che è chiaramente iniettiva e inoltre

$$||Ly||_{X^*} \leq ||y||_{p'} \,.$$

Si vede che vale l'uguaglianza. Infatti, sia

$$x_n = \left(|y_n|^{p'/p} \frac{y_n}{|y_n|} \right) \frac{1}{||y||_{p'}^{(p'/p)}} \quad \text{se } y_n \neq 0,\ x_n = 0 \text{ altrimenti.}$$

Si vede facilmente che $x = (x_n)$ è un elemento di $X = l^p$ di norma 1. Per esso vale

$$(Ly)(x) = \frac{1}{||y||_{p'}^{p'/p}} ||y||_{p'}^{p'} = ||y||_{p'} \,.$$

In questo modo si è trovata una trasformazione antilineare L che è isometrica da $l^{p'}$ in $(l^p)^*$.

Per concludere, basta mostrare che L è suriettiva, ossia che ogni elemento di $(l^p)^*$ si rappresenta come in (1.65). Sia allora $x^* \in (l^p)^*$. Dobbiamo prima di tutto trovare una successione da associare a x^*. Per questo usiamo ancora la successione $e^{(n)}$ definita in (1.64) e definiamo

$$y_i = \overline{\langle\langle x^*, e^{(i)} \rangle\rangle} \,.$$

In questo modo si costruisce una successione $y = (y_i)$.

Proviamo prima di tutto che $y \in l^{p'}$, $||y|| \leq ||x^*||$. Proveremo infine che

$$\langle\langle x^*, x \rangle\rangle = \sum_{i=1}^{+\infty} \bar{y}_i x_i \,. \tag{1.68}$$

Introduciamo la successione $x^{(n)} \in l^p$ definita da

$$x^{(n)} = x_i^{(n)} \qquad \text{con} \quad x_i^{(n)} = \begin{cases} |y_i|^{p'-1} \frac{y_i}{|y_i|} & \text{se} \quad i \leq n \text{ e } y_i \neq 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Ovviamente, $x^{(n)} \in l^p$ per ogni n e

$$\langle\langle x^*, x^{(n)} \rangle\rangle \leq ||x^*|| \cdot ||x^{(n)}||_p$$

e, d'altra parte,

$$\langle\langle x^*, x^{(n)} \rangle\rangle = \sum_{i=1}^{n} |y_i|^{p'} .$$

Passando al limite rispetto ad n si trova

$$||y||_{p'} \leq ||x^*|| .$$

Dunque, $y \in l^{p'}$ e $||y|| \leq ||x^*||$.

Proviamo ora che vale la (1.68). Fissato l'elemento $x \in l^p$, consideriamo la successione $x^{(n)}$ di elementi di l^p,

$$x^{(n)} = \sum_{i=0}^{n} x_i e^i , \quad \text{ossia} \quad x_r^{(n)} = \begin{cases} x_r & \text{se} \quad r \leq n \\ 0 & \text{se} \quad r > n . \end{cases}$$

Si vede che, se $p < +\infty$,

$$x = \lim_{r \to +\infty} x^{(n)} .$$

Questo limite si calcola nella norma di l^p. Dunque, essendo x^* continuo,

$$\lim_n \langle\langle x^*, x^{(n)} \rangle\rangle = \langle\langle x^*, x \rangle\rangle .$$

Inoltre,

$$\lim_n \langle\langle x^*, x^{(n)} \rangle\rangle = \lim \sum_{r=0}^{+\infty} \bar{y}_r x_n^{(r)} = \lim \sum_{r=0}^{n} \bar{y}_r x_n = \sum_{i=0}^{+\infty} \bar{y}_r x_r$$

perchè si è già provato che $y \in l^{p'}$ e quindi che $x \to \sum_{i=0}^{+\infty} \bar{y}_r x_r$ è un funzionale continuo. Si trova così che vale la rappresentazione

$$\langle\langle x^*, x \rangle\rangle = \sum_{i=0}^{+\infty} \bar{y}_r x_r .$$

Osservazione 160 Nel caso particolare $p = 2$, una realizzazione del duale di l^2 è lo spazio stesso. ∎

Il duale di c_0

Ricordiamo che il simbolo c_0 indica il s.spazio di l^∞ i cui elementi sono le successioni che convergono a zero. Proviamo che il duale di c_0 è realizzato da l^1. Per provare questo notiamo che se $(\xi_k) \in l^1$ allora la trasformazione

$$x \to \sum_{k=1}^{+\infty} \bar{\xi}_k x_k \tag{1.69}$$

è lineare e continua su l^∞ e che la trasformazione da $\xi \in l^1$ al funzionale definito da (1.69) è isometrica e antilineare. Queste proprietà si conservano sostituendo l^∞ con c_0.

Dobbiamo provare che ogni $x^* \in (c_0)^*$ ammette la rappresentazione (1.69).

Notiamo prima di tutto che ogni l^p, $p < +\infty$, è un sottospazio di c_0 e che l'immersione di l^p in c_0 è continua. Dunque, ogni funzionale lineare continuo su c_0 è anche un funzionale lineare continuo su l^p, per ogni $p < +\infty$. Ciò suggerisce di porre ancora

$$\xi_i = \overline{\langle\langle x^*, e^{(i)} \rangle\rangle}\,.$$

Si trova così un vettore $\xi = (\xi_i)$, candidato ad essere un rappresentante di x^*.

Come si è detto, il vettore ξ è nel duale $l^{p'}$ di l^p per ogni $p < +\infty$. In particolare quindi è in l^∞. Proviamo che inoltre tale vettore è anche in l^1. Scegliamo per questo la seguente successione $x^{(n)}$ in c_0:

$$x_r^{(n)} = \begin{cases} \frac{\xi_r}{|\xi_r|} & \text{se } r \le n \text{ e } \xi_r \neq 0 \\ 0 & \text{altrimenti} \end{cases} \qquad = \sum_{r \le n \xi_r \neq 0} \frac{\xi_r}{|\xi_r|} e^{(r)}\,.$$

Dato che $(\xi_n) \in l^\infty$, la successione $(x^{(n)})$ è una successione limitata in c_0. Esiste quindi un numero M tale che

$$\left|\langle\langle x^*, x^{(n)} \rangle\rangle\right| = \sum |\xi_r| < M$$

per ogni n. Ciò prova che $(\xi_r) \in l^1$.

Ora, per ogni $x \in c_0$, $x = (x_i)$, vale:

$$\langle\langle x^*, x \rangle\rangle = \lim_N \langle\langle x^*, \sum_{i=0}^N x_i e_i \rangle\rangle = \lim_N \sum_{i=0}^N \bar{\xi}_i x_i = \sum_{i=0}^{+\infty} \bar{\xi}_i x_i\,.$$

Ciò completa la dimostrazione.

Osservazione 161 Notiamo nuovamente che nella dimostrazione si usa la densità in c_0 della successione $e^{(i)}$. Questa successione è densa in l^p per $1 \le p < +\infty$ e anche in c_0; ma non in l^∞. ∎

Il duale di $L^p(\Omega)$, $1 \le p < +\infty$

In questo caso X^* è isometricamente isomorfo a $L^{p'}(\Omega)$, con

$$p' = \frac{p}{p-1} \quad \text{se } p > 1; \qquad L^\infty(\Omega) \quad \text{se } p = 1.$$

Ciò vale con $\Omega \subseteq \mathbb{R}^n$, limitato o meno.

Accenniamo alla dimostrazione nel caso $p = 1$ e $\Omega = (a, b)$.

E' ovvio che per ogni $\xi \in L^\infty(a, b)$, il funzionale su $L^1(a, b)$ definito da

$$x \to \int_a^b \bar{\xi}(s)x(s)\,\mathrm{d}s$$

è continuo, di norma minore o uguale a $||\xi||_\infty$ e in realtà si vede che vale l'uguaglianza.

Viceversa, sia x^* un funzionale lineare e continuo su $L^1(a, b)$. Dobbiamo associargli una funzione $\xi(s) \in L^\infty(a, b)$ tale che

$$\langle\langle x^*, x\rangle\rangle = \int_a^b \bar{\xi}(s)x(s)\,\mathrm{d}s\,.$$

Introduciamo la famiglia delle funzioni $\chi_t(s)$, una funzione per ogni $t \in (a, b)$,

$$\chi_t(s) = \begin{cases} 1 & \text{se} \quad a < t < s \\ 0 & \qquad \text{altrimenti} \end{cases} \tag{1.70}$$

e studiamo i valori che x^* assume su queste funzioni. La ragione di ciò è che[6] le funzioni a costanti a tratti sono dense in $L^1(a, b)$.

Associamo ad $x^* \in (L^1(a, b))^*$ la funzione

$$g(t) = \langle\langle x^*, \chi_t\rangle\rangle$$

E':

$$|g(t) - g(t')| = |\langle\langle x^*, \chi_t - \chi_{t'}\rangle\rangle \le ||x^*|| \cdot ||\chi_t - \chi_{t'}||_1 = ||x^*|| \cdot |t - t'|\,.$$

Dunque la funzione $g(t)$ è lipschitziana e quindi è assolutamente continua. Per essa vale

$$g(t) = \int_a^t \bar{\xi}(s)\,\mathrm{d}s$$

e inoltre $\bar{\xi}(s)$ è q.o. la derivata di $g(t)$. Dunque $\xi \in L^\infty(a, b)$ perché il rapporto incrementale di g è limitato.

Notiamo che

$$\langle\langle x^*, \chi_t\rangle\rangle = g(t) = \int_a^t \bar{\xi}(s)\,\mathrm{d}s = \int_a^b \bar{\xi}(s)\chi_t(s)\,\mathrm{d}s\,.$$

[6]questa proprietà non è stata provata nella parte relativa all'integrale di Lebesgue. E' stato provato però che sono dense le funzioni semplici, ossia costanti su insiemi misurabili. La densità delle funzioni costanti a tratti, ossia costanti su intervalli, discende dal fatto che ogni insieme misurabile si approssima mediante unioni di intervalli.

Una qualunque funzione a scala si rappresenta come combinazione lineare di funzioni χ_t:

$$\psi(t) = \sum_{i=1}^{n} \psi_i \chi_{t_i}(s)$$

e quindi

$$\langle\langle x^*, \psi \rangle\rangle = \sum_{i=1}^{n} \psi_i \int_a^b \bar{\xi}(s)\chi_{t_i}(s)\,\mathrm{d}s = \int_a^b \bar{\xi}(s)\psi(s)\,\mathrm{d}s\,.$$

Abbiamo già notato che il funzionale

$$x \to \int_a^b \bar{\xi}(s)x(s)\,\mathrm{d}s$$

è continuo.

Sia ora $x \in L^1(a,b)$. Esiste una successione di funzioni a scala ψ_n convergente ad x in $L^1(a,b)$. Allora,

$$\langle\langle x^*, x \rangle\rangle = \lim_N \langle\langle x^*, \psi_N \rangle\rangle = \lim_N \int_a^b \bar{\xi}(s)\psi_N(s)\,\mathrm{d}s = \int_a^b \bar{\xi}(s)x(s)\,\mathrm{d}s\,.$$

Ciò è quanto volevamo provare.

Osservazione 162 Una dimostrazione del tutto analoga porta ad identificare il duale di $L^p(\Omega)$, per ogni $p < +\infty$. Invece gli argomenti precedenti non si estendono al caso $L^\infty(\Omega)$ perché le funzioni costanti a tratti non sono dense in $L^\infty(\Omega)$. ∎

Il duale di $C(a,b)$

Introduciamo ora una realizzazione del duale di $C(a,b)$, lo spazio di Banach delle funzioni continue sull'intervallo limitato e chiuso $[a,b]$, con la norma dell'estremo superiore. Fissiamo un elemento x^* del duale. Per rappresentarlo, procediamo in tre passi:

PASSO 1) Introduciamo lo spazio lineare di tutte le funzioni limitate su $[a,b]$, continue o meno, dotato della norma dell'estremo superiore. Si trova uno spazio di Banach che indicheremo col generico simbolo B.

$C(a,b)$ essendo un s.spazio di B, il funzionale lineare e continuo x^*, definito su $C([a,b])$, si estende ad in funzionale lineare continuo su B, con la stessa norma, per il Teorema di Hahn-Banach. Tale estensione non è unica. Fissiamone una, che indichiamo col simbolo $\tilde{x}^*$.

Per ogni $t \in [a,b]$, introduciamo le funzioni definite come in (1.70) e la funzione

$$v(t) = \langle\langle \tilde{x}^*, \chi_t \rangle\rangle\,.$$

Sia ora $f \in C(a,b)$. Essendo $[a,b]$ compatto, la funzione f è uniformemente continua e quindi si approssima in modo uniforme con funzioni costanti a tratti. Queste possono costruirsi scegliendo un insieme finito $\{t_i\}_{i=1}^n$ di punti di $[a,b]$, abbastanza fitto, e quindi definendo

$$z_n(t) = f(t_{i-1}) \qquad \forall t \in [t_{i-1}, t_i)$$

ossia

$$z_n(t) = \sum_{i=1}^n f(t_{i-1})[\chi_{t_i}(s) - \chi_{t_{i-1}}(s)]\,.$$

Si ha quindi

$$\langle\langle x^*, f\rangle\rangle = \langle\langle \tilde{x}^*, f\rangle\rangle = \lim\langle\langle \tilde{x}^*, z_n\rangle\rangle = \lim\sum_{i=1}^n f(t_{i-1})[\langle\langle \tilde{x}^*, \chi_{t_i}(s)\rangle\rangle - \langle\langle \tilde{x}^*, \chi_{t_{i-1}}(s)\rangle\rangle]$$

$$= \lim\sum_{i=1}^n f(t_{i-1})[v_{t_i}(s) - v_{t_{i-1}}(s)]\,.$$

Si noti che nel caso particolare in cui $v(t) = t$, tale limite è $\int_a^b f(s)\,\mathrm{d}s$.

PASSO 2) Introduciamo un simbolo per indicare il limite precedente,

$$\int_a^b f\,\mathrm{d}v = \lim\sum_{i=1}^n f(t_{i-1})[v_{t_i}(s) - v_{t_{i-1}}(s)] \tag{1.71}$$

(ovviamente, il limite non dipende dalla partizione scelta per definirlo, dato che esso deve essere $\langle\langle x^*, f\rangle\rangle$).

Il particolare integrale definito da (1.71) si chiama *integrale di* $\boxed{\textit{Stiltjes}}$.

Si osservi che la rappresentazione di x^* come integrale di Stiltjes usa la continuità uniforme di f; e quindi in generale $\tilde{x}^*$ non avrà tale rappresentazione.

PASSO 3) Ricapitolando, abbiamo rappresentato ogni elemento del duale di $C(a,b)$ come un integrale di Stiltjes. Dobbiamo ora studiare le proprietà di tale integrale, per trovare uno spazio di Banach che realizzi $[C(a,b)]^*$.

Si ha

$$\sum_{i=1}^n |v(t_i) - v(t_{i-1})| = \sum_{i=1}^n \{\mathrm{sgn}\,[v(t_i) - v(t_{i-1})]\}[v(t_i) - v(t_{i-1})]$$

$$\sum_{i=1}^n \langle \tilde{x}^*, \Big\{\mathrm{sgn}\,[v(t_i) - v(t_{i-1})][\chi_{t_i} - \chi_{t_{i-1}}]\Big\}\rangle$$

$$= \langle \tilde{x}^*, \sum_{i=1}^n \Big\{\mathrm{sgn}\,[v(t_i) - v(t_{i-1})][\chi_{t_i} - \chi_{t_{i-1}}]\Big\}\rangle$$

$$\leq ||\tilde{x}^*|| \sup_s \sum_{i=1}^n |\chi_{t_i}(s) - \chi_{t_{i-1}}(s)| = ||\tilde{x}^*|| = ||x^*||$$

perché la differenza $|\chi_{t_i}(s) - \chi_{t_{i-1}}(s)|$ vale 1 oppure 0.

Questa disuguaglianza vale per ogni suddivisione dell'intervallo $[a, b]$ in un numero finito di punti e quindi esiste un numero M, $M = ||x^*||$, tale che

$$V_a^b v = \sup_{\{t_i\}} \sum |v(t_i) - v(t_{i-1})| < M \, .$$

Funzioni v con questa proprietà si dicono *a variazione* *limitata*.

La struttura delle funzioni a variazione limitata è stata studiata con estrema precisione. Si prova in particolare che ogni funzione a variazione limitata è differenza di due funzioni monotone e che, quindi, i suoi punti di discontinuità sono salti. Si prova inoltre che

$$\sup_{||f||<1} \int_a^b f \, \mathrm{d}v = V_a^b v$$

e questo suggerisce di scegliere come spazio per rappresentare $[C(a, b)]^*$ uno spazio di funzioni a variazione limitata. Bisogna però notare che può aversi

$$\int_a^b f \, \mathrm{d}v = \int_a^b f \, \mathrm{d}v' \qquad \forall f \in C(a, b)$$

anche con $v \neq v'$. E quindi la rappresentazione che abbiamo trovato per x^* non è unica. Si prova però che l'uguaglianza può aversi, per ogni f, solo se v e v' differiscono per il valore che assumono in un punto di salto oppure nell'estremo sinistro a dell'intervallo. Ciò suggerisce di definire

$$NV(a, b)$$

lo spazio delle *funzioni a variazione limitata* *normalizzate* su $[a, b]$, continue a sinistra e nulle in a, dotato della norma

$$V_a^b(v) \, .$$

Si prova che questo spazio è di Banach e che vale:

Teorema 163 (Teorema di Riesz) *Lo spazio $NV(a, b)$ è una realizzazione del duale di $[C(a, b)]^*$ e ogni $x^* \in [C(a, b)]^*$ si rappresenta (in modo unico) come*

$$\langle\langle x^*, f \rangle\rangle = \int_a^b f \, dv \, , \qquad v \in NV(a, b) \, .$$

Il duale di $C(K)$

Ricordiamo che simbolo $C(K)$ l'insieme K è compatto.

Non abbiamo gli strumenti per studiare il duale di $C(K)$. Possiamo però descrivere come si rappresenta l'azione su $C(K)$ di un elemento x^* del suo duale. Per ogni $x^* \in (C(K))'$ si trovano una misura di Borel m ed una funzione $\psi(s)$ misurabile secondo Borel su K e tale che

$$|\psi(s)| = 1 \qquad \text{q.o. } s \in K$$

e per la quale vale

$$\langle\langle x^*, x \rangle\rangle = \int_K \psi(s) x(s) \, \mathrm{d}m$$

1.11.1 Relazione tra le convergenze debole e debole stella

Avendo a disposizione gli esempi precedenti, possiamo chiarire meglio le relazioni tra le convergenze debole e debole stella, quando queste si possano definire sul medesimo spazio, che in tal caso è lo spazio X^* duale di uno spazio di Banach X.

Queste due nozioni di convergenza non sono indipendenti. Infatti:

Teorema 164 *Sia x_n^* una successione in X^*. Se essa converge debolmente ad x_0^*, allora essa converge anche debole stella al medesimo x_0.*

Dim. Ciò discende dal fatto che si è già notato che X è isometrico ad un s.spazio di X^{**}. ∎

L'esempio seguente mostra che non vale l'implicazione inversa.

Esempio 165 Sia $X = c_0$, $X^* = l^1$ ed $X^{**} = l^\infty$.

In $X^* = l^1$ si consideri la successione $\{e^{(n)}\}$ definiti da (1.64).

Se $x \in c_0$, $x = (x_n)$ allora

$$\langle\langle x, e^{(n)} \rangle\rangle = x_n \to 0 \qquad \text{e quindi} \qquad \mathrm{w}^* - \lim e^{(n)} = 0 \, .$$

Invece, se $x \in l^\infty = X^{**}$ è la successione ogni cui elemento vale 1,

$\langle\langle x, e^{(n)} \rangle\rangle = 1$ e quindi la successione $(e^{(n)})$ non converge debolmente a 0. ∎

1.12 Lo spettro di un operatore

Se X ha dimensione finita è noto che molte informazioni si ottengono studiando gli autovalori della trasformazione, i quali hanno spesso iterpretazioni fisiche importanti. Vogliamo ora estendere questo tipo di studio a generici spazi di Banach. In tal caso la situazione è resa complessa dall'esistenza di operatori lineari non continui e di sottospazi non chiusi, due fatti che non si verificano in dimensione finita.

Si sa che, anche in dimensione finita, autovalori ed autovettori possono solo trovarsi se il campo scalare è quello dei numeri complessi. Per questo supporremo $\Phi = \mathbb{C}$.

In dimensione finita, un numero complesso z_0 si dice un autovalore di A se

$$(z_0 I - A)x = y$$

non è risolubile per ogni y; equivalentemente, se la soluzione x, quando esiste, non è unica. L'esempio seguente mostra che la non risolubilità per ogni y in dimensione infinita non equivale alla non unicità.

Esempio 166 Sia $X = l^p$, per un qualsiasi $p \in [1, +\infty]$ e siano T ed S definiti da

$$\begin{aligned} S\left[\begin{array}{cccc} x_0 & x_1 & x_2 & \dots \end{array}\right] &= \left[\begin{array}{cccc} 0 & x_0 & x_1 & \dots \end{array}\right] \\ T\left[\begin{array}{cccc} x_0 & x_1 & x_2 & \dots \end{array}\right] &= \left[\begin{array}{cccc} x_1 & x_2 & x_3 & \dots \end{array}\right]. \end{aligned} \tag{1.72}$$

E' chiaro che

$$\begin{aligned} &\ker S = \{0\} \qquad \text{e} \quad \operatorname{im} S \neq X\,, \\ &\ker T \neq \{0\} \qquad \text{e} \quad \operatorname{im} T = X\,. \ \blacksquare \end{aligned}$$

I due operatori precedenti sono particolarmente importanti. In particolare, l'operatore S si chiama *operatore di* $\boxed{\textit{traslazione}}$ (sottinteso, verso destra).

Ricordiamo inoltre che se l'equazione

$$(zI - A)x = y \tag{1.73}$$

è risolubile in modo unico allora $(zI - A)$ ammette inverso, definito sulla sua immagine e questo è un operatore lineare. Ma, in generale, l'inverso non è continuo.

Queste considerazioni suggeriscono la definizione seguente, che si applica ad ogni operatore lineare A da X in X, anche non continuo ma con dominio denso:

Definitione 167 Sia A lineare da X in X, con dominio denso. Si chiama *insieme* risolvente di A l'insieme dei numeri complessi z tali che la (1.73) ammette un'unica soluzione x per ogni y in un s.insieme denso di X e tali che, inoltre, l'inverso $(zI - A)^{-1}$ sia continuo.

Se z appartiene all'insieme risolvente di A, l'operatore $(zI-A)^{-1}$ si chiama l'*operatore* risolvente di A.

L'insieme risolvente si indica col simbolo $\rho(A)$. Il suo complementare si indica col simbolo $\sigma(A)$ e si chiama lo *spettro* dell'operatore A. ∎

Da un punto di vista logico, $z \in \sigma(A)$ se si verifica uno dei casi seguenti, mutuamente incompatibili:

i) $\ker(zI - A) \neq \{0\}$;

ii) $\ker(zI - A) = \{0\}$ ma $\mathrm{im}\,(zI - A)$ non denso in X;

iii) $\ker(zI - A) = \{0\}$, $\mathrm{im}\,(zI - A)$ denso in X ma $(zI - A)^{-1}$ non continuo.

Definiamo quindi:

- *spettro di* punti l'insieme dei numeri z per i quali si verifica il caso i);
- *spettro* residuo l'insieme dei punti per i quali si verifica il caso ii);
- *spettro* continuo l'insieme dei punti per i quali si verifica il caso iii).

Gli elementi dello spettro di punti si chiamano *autovalori* dell'operatore A.

Abbiamo definito una partizione dello spettro di A in tre s.insiemi. Essi si indicano rispettivamente con i simboli

$$\sigma_p(A)\,, \qquad \sigma_r(A)\,, \qquad \sigma_c(A)\,.$$

In dimensione finita solo il caso i) può verificarsi. Mostriamo che, invece, in dimensione infinita anche gli altri casi possono verificarsi.

Esempio 168 Sia S l'operatore definito in (1.72). Si vede facilmente che $0 \in \sigma_r(S)$ (invece, $0 \in \sigma_p(T)$).

Mostriamo un operatore con spettro continuo non vuoto. Sia $X = l^2$ e sia A definito come segue:

$$A\left[\begin{array}{ccccccc} x_0 & x_1 & x_2 & \dots & x_n & \dots \end{array}\right] = \left[\begin{array}{ccccccc} x_0 & \frac{1}{2}x_1 & \frac{1}{3}x_2 & \dots & \frac{1}{n+1}x_n & \dots \end{array}\right].$$

L'equazione $Ax = y$ è risolubile per ogni successione (y_n) definitivamente nulla, ossia per ogni y in un s.spazio denso di $X = l^2$, ed è

$$\left[\begin{array}{ccccccc} x_0 & x_1 & x_2 & \dots & x_n & \dots \end{array} \right] = \left[\begin{array}{ccccccc} y_0 & 2y_1 & 3y_2 & \dots & (n+1)y_n & \dots \end{array} \right] .$$

Dunque, l'inverso non è continuo. ∎

Grazie al teorema fondamentale dell'algebra, si sa che in dimensione finita lo spettro non è mai vuoto ed è un insieme finito. Mostriamo invece che esistono operatori lineari su spazi di Banach, con spettro vuoto ed operatori con risolvente vuoto.

Esempio 169 Sia $X = L^2(0,1)$ e siano A e B definiti come segue:

$$\begin{aligned} &\mathrm{dom}A = \{x \in C(0,1) \mid x' \in L^2(0,1)\}\,, \qquad Ax = x'\,, \\ &\mathrm{dom}B = \{x \in C(0,1) \mid x' \in L^2(0,1)\,,\ x(0) = 0\}\,, \qquad Bx = x'\,. \end{aligned}$$

Allora $\sigma(A) = \sigma_p(A) = \mathbb{C}$ perché per ogni z vale $(z-A)\chi_z = 0$, con $\chi_z(t) = e^{zt}$. Invece, $\sigma(B) = \emptyset$ perché

$$(zI - B)x = y \Longleftrightarrow \left\{ \begin{array}{l} x' = zx - y \\ x(0) = 0\,. \end{array} \right.$$

Dunque x è dato da

$$x(t) = -\int_0^t e^{z(t-s)} y(s)\,\mathrm{d}s$$

così che l'operatore $(zI - B)^{-1}$ è continuo. ∎

Nell'esempio precedente intervengono, vedremo non per caso, operatori che non sono continui; ma anche lo spettro di operatori continui può avere una struttura piuttosto complessa:

Esempio 170 Sia $X = l^2$ e sia T l'operatore definito in (1.72). Risolvendo

$$(zI - T)x = 0$$

si trova come soluzione:

$$x = q \left[\begin{array}{ccccc} 1 & z & z^2 & z^3 & \dots \end{array} \right]$$

con $q \in \mathbb{C}$ qualsiasi. Questa successione appartiene ad l^2 per ogni z di modulo minore di 1. Dunque, $\sigma_p(T) \supseteq \{z \mid |z| < 1\}$. ∎

Nonostante questi esempi, spettro e risolvente non possono essere insiemi qualsiasi. Infatti:

Teorema 171 *Se A è continuo, $\sigma(A) \subseteq \{z \mid |z| \leq ||A||\}$.*

Dim. Supponiamo che A sia continuo e che sia $|z| > ||A||$. Vogliamo provare che in tal caso $z \in \rho(A)$. Scriviamo per questo

$$(zI - A) = z(I - K)\,, \qquad K = \frac{1}{z}A \quad \text{così che} \quad ||K|| < 1\,.$$

Dunque,

$$(zI - A)^{-1} = \frac{1}{z}\sum_{n=0}^{+\infty} K^n = \frac{1}{z}\sum_{n=0}^{+\infty} \frac{A^n}{z^n}\,. \tag{1.74}$$

Quest'operatore è continuo, dato che $||K|| = ||A/z|| < 1$; e quindi $z \in \rho(A)$. ∎

Lo spettro non può essere un insieme qualsiasi nemmeno se l'operatore A non è continuo. Infatti:

Teorema 172 *Il risolvente è sempre un insieme aperto e quindi lo spettro è chiuso.*

Dim. Sia A qualsiasi, anche non continuo. Proviamo che il suo risolvente è aperto. Se esso è vuoto niente va provato. Dunque supponiamo che esista un numero $z_0 \in \rho(A)$ e mostriamo che esso è interno al risolvente; ossia proviamo l'esistenza di $\epsilon > 0$ (che dipende sia da z_0 che da A) tale che se $|z| < \epsilon$ allora $z + z_0 \in \rho(A)$. Per questo scriviamo

$$(z + z_0)I - A = zI + (z_0 I - A) = (z_0 I - A)\left[I + z(z_0 I - A)^{-1}\right]\,. \tag{1.75}$$

Per il Teorema 171, l'operatore

$$\left[I + z(z_0 I - A)^{-1}\right]$$

è invertibile se

$$|z| < \epsilon = \frac{1}{||(z_0 I - A)^{-1}||}$$

e in tal caso $(z + z_0)I - A$ è invertibile con inverso limitato, essendo composizione di operatori invertibili ciascuno con inverso limitato. ∎

La (1.75) permette anche di trovare un'espressione per $[(z + z_0)I - A]^{-1}$:

$$\begin{aligned}[(z + z_0)I - A]^{-1} &= \left\{\sum_{k=0}^{+\infty} z^n \left[(z_0 I - A)^{-1}\right]^n\right\}(z_0 I - A)^{-1} \\ &= \sum_{k=0}^{+\infty} z^n \left[(z_0 I - A)^{-1}\right]^{n+1}\,.\end{aligned} \tag{1.76}$$

Dunque, fissato $z_0 \in \rho(A)$, la funzione $[(z+z_0)I - A]^{-1}$ si esprime come serie di potenze di z, a coefficienti operatori limitati.

Chiameremo *funzioni* *olomorfe* a valori operatori quelle funzioni di $z \in \mathbb{C}$ che si esprimono localmente, in un opportuno intorno di ogni punto z_0 del loro dominio, mediante serie

$$\sum_{n=0}^{+\infty} K^n (z - z_0)^n$$

convergenti (nella norma di $\mathcal{L}(X)$). Dunque:

Corollario 173 *Se l'operatore A ha risolvente non vuoto, la funzione $z \to (zI - A)^{-1}$ è olomorfa su $\rho(A)$.*

Osservazione 174 Combinando il calcolo dell'esempio 170 con i teoremi 171 e 172, si vede che

$$\sigma(T) = \{z \mid |z| \leq 1\}\,. \quad \blacksquare$$

Torniamo ora a considerare un operatore continuo A. Si è detto che il suo risolvente non è vuoto, e anzi contiene l'esterno del disco $\{z \mid |z| \leq ||A||\}$. Naturalmente, esso può anche estendersi all'interno di tale disco; ma non può riempirlo. Infatti:

Teorema 175 *Sia $A \in \mathcal{L}(X)$. Lo spettro di A non è vuoto.*

Dim. La dimostrazione procede per assurdo, e va confrontata con la dimostrazione del Teorema fondamentale dell'algebra.

Dal Corollario 173 si sa che, se $\sigma(A) = \emptyset$, la funzione $(zI - A)^{-1}$ è olomorfa su $\mathbb{C}$. Dunque, per ogni $x \in X$, $y^* \in X^*$, la funzione

$$z \to f(z) = \langle\langle y^*, (zI - A)^{-1} x \rangle\rangle$$

è una funzione intera.

Dalla (1.76) si vede che per $|z| \to +\infty$, la funzione $f(z)$ tende a zero; e quindi $f(z)$ è una funzione intera e limitata, e quindi costante. Dato che un suo limite è nullo, essa deve essere identicamente zero.

Dunque, abbiamo provato che

$$\langle\langle y^*, (zI - A)^{-1} x \rangle\rangle \equiv 0 \qquad \forall x \in X\,, \quad \forall y^* \in X^*\,.$$

Dato che X^* distingue punti di X, deve essere

$$(zI - A)^{-1} x \equiv 0$$

per ogni $x \in X$. Ciò non può darsi perché $(zI - A)(zI - A)^{-1}x = x$ per ogni $x \in X$.

La contraddizione trovata prova il teorema. ∎

Nonostante che lo spettro di un operatore continuo non possa essere vuoto, può essere che esso sia un insieme molto più piccolo del disco di raggio $||A||$. Per esempio, in dimensione 2, la trasformazione lineare la cui matrice è

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

ha norma 1 ed il solo autovalore 0. E' un utile esercizio vedere che un caso analogo può darsi anche in dimensione infinita.

Esempio 176 Sia $X = L^2(0,1)$ e sia

$$Ax = \int_0^t x(s)\,\mathrm{d}s\,.$$

Si sa che A è continuo e si vede immediatamente che $0 \in \sigma(A)$, dato che A^{-1} è l'operatore di derivazione, $A^{-1}y = y'$, che non è continuo.

Mostriamo che ogni altro numero z appartiene al risolvente. Per questo risolviamo

$$(zI - A)x = y \quad \text{ossia} \quad zx(t) - \int_0^t x(s)\,\mathrm{d}s = y(t)\,.$$

Dividendo per z si trova

$$x(t) - \frac{1}{z}y(t) = \frac{1}{z}\int_0^t x(s)\,\mathrm{d}s = \frac{1}{z}\int_0^t \left[x(s) - \frac{1}{z}y(s)\right]\,\mathrm{d}s + \int_0^t \frac{1}{z^2}y(s)\,\mathrm{d}s\,.$$

Quest'uguaglianza mostra che la funzione

$$\xi(t) = x(t) - \frac{1}{z}y(t)$$

è derivabile quasi ovunque, che $\xi(0) = 0$ e che

$$\xi'(t) = \frac{1}{z}\xi(t) - \frac{1}{z^2}y(t) \qquad \text{ossia} \qquad \xi(t) = -\frac{1}{z^2}\int_0^t e^{\frac{1}{z}(t-s)}y(s)\,\mathrm{d}s\,.$$

Da qui,

$$x(t) = \frac{1}{z}y(t) - \frac{1}{z^2}\int_0^t e^{\frac{1}{z}(t-s)}y(s)\,\mathrm{d}s\,.$$

La trasformazione da y ad x è, per ogni fissato $z \neq 0$, lineare e continua. Dunque, $\sigma(A) = \{0\}$. ∎

Questi esempi suggeriscono di chiamare *raggio* *spettrale* $r(A)$ il numero

$$r(A) = \max\{|z| \mid z \in \sigma(A)\}\,.$$

Il raggio spettrale di un operatore continuo si esprime in modo che richiama la formula per il raggio di convergenza di una serie di potenze:

Teorema 177 *Sia $A \in \mathcal{L}(X)$. Vale:*

$$r(A) = \lim \sqrt[n]{||A^n||}\,.$$

Dim. Si prova, esattamente come nel caso scalare, che una serie di potenze a valori operatori converge in un disco di centro z_0, che si chiama ancora *disco di convergenza*,. Questo disco coincide col disco di convergenza della serie di potenze

$$\sum_{n=0}^{+\infty} ||A^n||(z - z_0)^n\,.$$

Applicando questo alla serie (1.74) si vede che

$$r(A) = \limsup \sqrt[n]{||A^n||}\,.$$

Si deve ora provare che in realtà esiste

$$\lim \sqrt[n]{||A^n||}\,.$$

Risulta più semplice provare l'esistenza del limite

$$\lim \frac{1}{n} \log ||A^n||\,.$$

Notiamo che

$$\log ||A^{n+m}|| \leq \log ||A^n|| \cdot ||A^m|| \leq \log ||A^n|| + \log ||A^m||\,.$$

Una successione di numeri (a_n), tutti positivi, che verificano

$$a_{n+m} \leq a_n + a_m$$

si dice *subadditiva*. La dimostrazione del Teorema 177 si completa usando il lemma seguente:

Lemma 178 *Se la successione (a_n) è subadditiva allora esiste*

$$\lim \frac{1}{n} a_n\,.$$

Dim. Si fissa un numero naturale m e si studiano i quozienti a_n/n con $n > m$. Notiamo che si può scrivere

$$n = md + r\,, \qquad 0 \le r < m$$

e quindi

$$a_n = a_{md+r} \le da_m + a_r\,.$$

Dividiamo per n e passiamo al limite per $n \to +\infty$.

Il numero a_r è funzione di n limitata al variare di r perché prende valori nell'insieme finito $a_1, \ldots, a_{m-1}$. Dunque $\lim a_r/n = 0$.

Ancora perché r prende un numero finito di valori,

$$\frac{d}{n} = \frac{n-r}{nm} = \frac{1}{m} - \frac{r}{nm} \to \frac{1}{m}\,.$$

Dunque,

$$\limsup \frac{a_n}{n} \le \frac{a_m}{m} \qquad \forall m\,.$$

E quindi

$$\limsup \frac{a_n}{n} \le \liminf \frac{a_n}{n}\,.$$

Ciò prova l'esistenza del limite. ∎

Il teorema è così provato. ∎

Esempio 179 Su $\mathbb{R}^2$ (riferito alla base canonica) consideriamo la trasformazione lineare descritta mediante la matrice

$$A = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}\,.$$

Per calcolare il raggio spettrale mediante la formula (177) bisogna prima di tutto calcolare le potenze di A:

$$A^{2n} = 2^n I\,, \qquad A^{2n+1} = 2^n A\,.$$

E' immediatamente evidente che $||A|| = 2$ e quindi

$$||A^{2n}||^{1/2n} = \sqrt{2}\,, \qquad ||A^{2n+1}||^{1/(2n+1)} = [\sqrt{2}]^{(2n+2)/(2n+1)}\,.$$

Dunque, il raggio spettrale è $\sqrt{2}$. Si noti che la successione $(||A^n||^{1/n})$ non è monotona. ∎

1.12.1 Proiezioni spettrali

Sia X uno spazio di Banach e sia A un operatore, anche non continuo, da X in sé, con dominio denso.

Abbiamo notato che il risolvente è una funzione analitica e ciò suggerisce di studiare l'analogo, scritto per gli operatori, della formula integrale di Cauchy:

$$\frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}\,\mathrm{d}z \tag{1.77}$$

ove γ è una curva semplice e chiusa[7] il cui sostegno è contenuto in $\rho(A)$. La funzione $f(z)$ è olomorfa.

Naturalmente, dovremo dare un senso all'integrale. Dato che la funzione a valori in $\mathcal{L}(X)$

$$z \to f(z)(zI-A)^{-1}$$

è uniformemente continua sul sostegno di γ, l'integrale si definisce coll'usuale metodo di Riemann, come "limite" delle somme di Riemann. Lasciamo al lettore i semplici dettagli.

Nonostante che la (1.77) abbia senso per ogni operatore lineare A da X in sé, purché il sostegno di γ sia contenuto in $\rho(A)$, noi ci limiteremo a considerare il cao degli operatori A continui.

Per interpretare la (1.77), consideriamo la funzione

$$f(z)=\sum_{n=0}^{+\infty} f_n z^n$$

e la serie corrispondente

$$\sum_{n=0}^{+\infty} f_n A^n\,. \tag{1.78}$$

Nel caso particolare in cui $f(z)=p(z)$ è un polinomio, la serie (1.78) è una somma finita e definisce un operatore che, ovviamente, si indica col simbolo $p(A)$. Per esempio, se $p(z)=z^2$, allora $p(A)=A^2$. Se $f(z)$ è una generica funzione analitica la cui serie converge in un intorno di 0, la serie (1.78) in generale non converge, ma certamente converge in $\mathcal{L}(X)$ se

$$||A|| < R$$

con R raggio di convergenza della serie di potenze di $f(z)$. Infatti in tal caso

$$\left|\left|\sum_{n=k}^{m} f_n A^n\right|\right| \le \sum_{n=k}^{m} f_n ||A||^n \le \sum_{n=k}^{m} f_n r^n\,, \qquad r<R\,,$$

[7] come al solito, orientata in verso positivo, ossia antiorario.

e la convergenza si vede dal test di Cauchy per la convergenza delle serie.

Ricapitolando, se $||A|| < R$, la serie (1.78) definisce un operatore di $\mathcal{L}(X)$, che indicheremo col simbolo

$$f(A)\,.$$

Ricordiamo ora che i coefficienti f_n si rappresentano come

$$f_n = \frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)}{\zeta^{n+1}}\,\mathrm{d}\zeta$$

con γ curva semplice e chiusa orientata positivamente, il cui sostegno è contenuto nel disco di convergenza di $f(z)$.

Supponiamo che la curva γ racchiuda il disco $\{z \mid |z| < r\}$, con

$$||A|| < r < R\,. \tag{1.79}$$

In tal caso si trova

$$\begin{aligned} f(A) &= \sum_{n=0}^{+\infty} f_n A^n = \frac{1}{2\pi i}\sum_{n=0}^{+\infty}\int_\gamma f(\zeta)\frac{A^n}{\zeta^{n+1}}\,\mathrm{d}\zeta \\ &= \frac{1}{2\pi i}\int_\gamma f(\zeta)\left(\sum_{n=0}^{+\infty}\frac{A^n}{\zeta^{n+1}}\right)\mathrm{d}\zeta = \frac{1}{2\pi i}\int_\gamma f(\zeta)(\zeta I - A)^{-1}\,\mathrm{d}\zeta\,. \end{aligned}$$

Si noti che questo calcolo vale grazie alla condizione (1.79) e, se vale (1.79), allora si ha anche

$$\sigma(A) \subseteq \{z \mid\ |z| < R\}\,.$$

Osservazione 180 Si noti che l'integrale (1.77) ha senso anche se γ, di sostegno in $\rho(A)$, racchiude solo una parte dello spettro di A. Però in tal caso non useremo la notazione $f(A)$ per indicarlo. ∎

In un caso particolare è facile calcolare l'integrale (1.77): supponiamo che $\rho(A)$ contenga una regione di Jordan Ω e supponiamo che il sostegno di γ appartenga a Ω. In questo caso un argomento analogo a quello usato nella dimostrazione del teorema 175, basato sul teorema di Hahn–Banach, prova che l'integrale è nullo. Ossia, il Teorema di Cauchy **??** vale anche per integrali della forma (1.77). Dunque, i casi interessanti saranno quelli nei quali γ "gira" intorno a punti di $\sigma(A)$. Per intuire cosa dobbiamo attenderci, consideriamo l'esempio seguente:

Esempio 181 Sia $X = \mathbb{C}^3$ a sia

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \text{così che} \qquad (zI - A)^{-1} = \begin{bmatrix} 1/z & 1/z^2 & 0 \\ 0 & 1/z & 0 \\ 0 & 0 & 1/(z-2) \end{bmatrix}.$$

Sia γ una curva semplice e chiusa che racchiude 0 e che lascia fuori 2. Si calcola immediatamente che

$$\frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\,\mathrm{d}z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Si trova così la proiezione sull'autospazio generalizzato dell'autovalore 0.

Operando in modo analogo con una curva che racchiude 2 e lascia fuori 0 si trova la proiezione sull'altro autospazio.

Abbiamo così calcolato l'integrale nel caso della funzione $f(z) = 1$. Se $f(z) = z$ un calcolo analogo dà

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

rispettivamente, a seconda della scelta della curva. Queste sono le restrizioni di A ai due autospazi. Si trova così una "diagonalizzazione a blocchi" della matrice A. ∎

Senza trattare l'integrale (1.77) in generale, vogliamo limitarci a considerare i due casi $f(z) = 1$ ed $f(z) = z$, che sono particolarmente importanti per le applicazioni, e che verranno usati nel paragrafo **??**.

Generalizzando l'esempio (181), supponiamo che $\sigma(A) = \sigma_1(A) \cup \sigma_2(A)$ e che una regione di Jordan Ω contenga $\sigma_1(A)$ e lasci fuori $\sigma_2(A)$. Sia γ una curva semplice e chiusa col sostegno in Ω, che gira intorno a $\sigma_1(A)$, come nella figura 1.7, a sinistra:

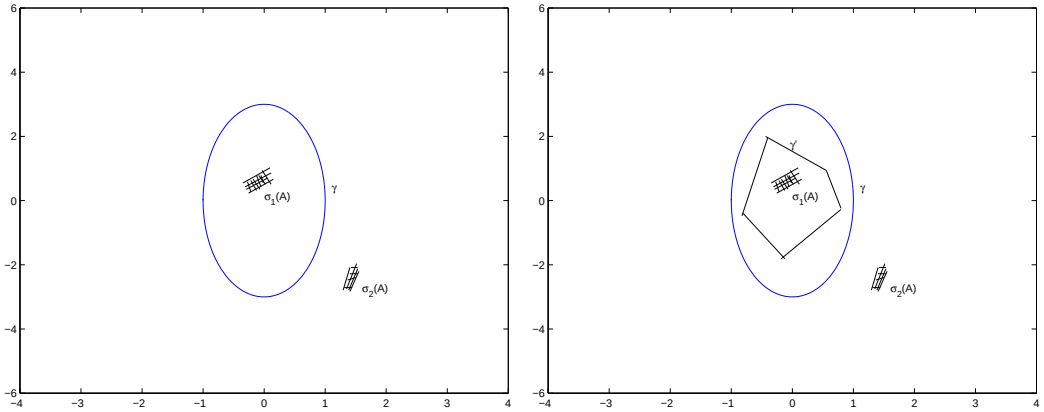
In tal caso:

Teorema 182 *Valgano le condizioni appena dette. L'operatore*

$$P = \frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\,dz$$

è una proiezione.

Figura 1.7:



Dim. Ricordiamo che vale il teorema di Cauchy. Usando ciò, una dimostrazione analoga a quella del teorema **??**, porta a concludere che due curve γ e γ' semplici e chiuse in Ω, che ambedue racchiudono $\sigma_1(A)$ e lasciano fuori $\sigma_2(A)$ (come in figura 1.7, a destra) definiscono il medesimo operatore P:

$$P = \frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\,\mathrm{d}z = \frac{1}{2\pi i}\int_{\gamma'} (\zeta I - A)^{-1}\,\mathrm{d}\zeta\,.$$

Dunque,

$$\begin{aligned} P^2 &= \left\{\frac{1}{2\pi i}\int_\gamma (zI - A)^{-1}\,\mathrm{d}z\right\}\left\{\frac{1}{2\pi i}\int_{\gamma'} (\zeta I - A)^{-1}\,\mathrm{d}\zeta\right\} \\ &= \frac{1}{2\pi i}\cdot\frac{1}{2\pi i}\int_\gamma \left[\int_{\gamma'} (zI - A)^{-1}(\zeta I - A)^{-1}\right]\,\mathrm{d}\zeta\;\mathrm{d}z\,. \end{aligned}$$

Non è restrittivo supporre che la curva γ racchiuda la curva γ', come nella figura 1.7, a destra. A questo punto usiamo una formula[8] che si chiama *prima formula del* *risolvente* e che è di verifica immediata:

$$(zI - A)^{-1}(\zeta I - A)^{-1} = \frac{1}{\zeta - z}\left[(zI - A)^{-1} - (\zeta I - A)^{-1}\right]\,.$$

[8] si noti che questa formula estende l'uguaglianza, valida tra numeri,

$$\frac{1}{(z-a)(\zeta - a)} = \frac{1}{\zeta - z}\left[\frac{1}{z-a} - \frac{1}{\zeta - a}\right]\,.$$

Usando questa formula, si trova

$$P^2 = \frac{1}{2\pi i}\int_\gamma \left[\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(zI - A)^{-1}\,\mathrm{d}\zeta - \frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(\zeta I - A)^{-1}\,\mathrm{d}\zeta\right]\,\mathrm{d}z\,.$$

Ora, dal Teorema di Cauchy,

$$\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(zI - A)^{-1}\,\mathrm{d}\zeta = 0$$

perché il punto z, che è sulla curva γ, è nella regione esterna a γ'.

Dunque rimane

$$\begin{aligned} P^2 &= -\frac{1}{2\pi i}\int_\gamma \left[\frac{1}{2\pi i}\int_{\gamma'} \frac{1}{\zeta - z}(\zeta I - A)^{-1}\,\mathrm{d}\zeta\right]\,\mathrm{d}z \\ &= \frac{1}{2\pi i}\int_{\gamma'} \left[\frac{-1}{2\pi i}\int_\gamma \frac{1}{\zeta - z}\,\mathrm{d}z\right](\zeta I - A)^{-1}\,\mathrm{d}\zeta\,. \end{aligned}$$

L'integrando

$$z \to \frac{1}{z - \zeta}$$

ha ζ per polo semplice, perché la curva γ' è racchiusa dalla curva γ. Dunque

$$\frac{-1}{2\pi i}\int_\gamma \frac{1}{\zeta - z}\,\mathrm{d}z = \frac{1}{2\pi i}\int_\gamma \frac{1}{z - \zeta}\,\mathrm{d}z = 1$$

e quindi

$$P^2 = \frac{1}{2\pi i}\int_{\gamma'}(\zeta I - A)^{-1}\,\mathrm{d}\zeta = P\,. \quad \blacksquare$$

Supponiamo ora che σ_1 e σ_2 sino due s.insiemi limitati di $\sigma(A)$, appartenenti alla regione interna rispettivamente di γ_1 e di γ_2, curve di Jordan di sostegno in $\rho(A)$ ed esterne l'una all'altra come in figura 1.8.

Poniamo

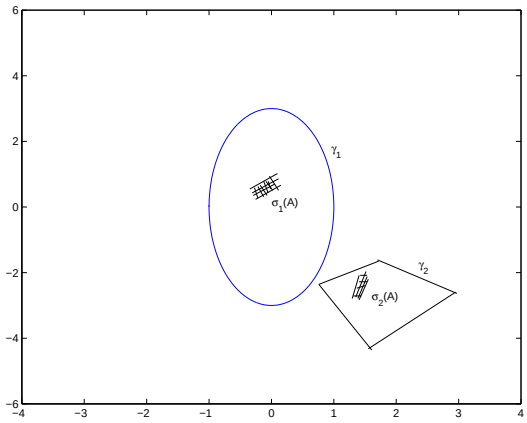
$$P_1 = \frac{1}{2\pi i}\int_{\gamma_1}(zI - A)^{-1}\,\mathrm{d}z\,, \qquad P_2 = \frac{1}{2\pi i}\int_{\gamma_2}(zI - A)^{-1}\,\mathrm{d}z\,.$$

Una dimostrazione analoga a quella del teorema precedente porta a:

Teorema 183 *Nelle ipotesi dette, si ha:* $P_1P_2 = P_2P_1 = 0$. *Inoltre* $X_1 = \operatorname{im} P_1$ *ed* $X_2 = \operatorname{im} P_2$ *sono spazi lineari invarianti per* A *e lo spettro della restrizione di* A *ad* $\operatorname{im} P_i$ *è l'insieme* σ_i. *Tale restrizione è data da*

$$Ax = \frac{1}{2\pi i}\int_{\gamma_i} z(zI - A)^{-1}x\,dz\,, \qquad \forall x \in X_i\,. \quad \blacksquare$$

Figura 1.8:



Omettiamo i dettagli della dimostrazione di questo teorema, che è analoga a quella del teorema 182.

Si noti che il Teorema 183 mostra una diagonalizzazione a blocchi dell'operatore A, analoga a quella vista nell'Esempio 181.

1.13 Trasformazioni non lineari

Fino ad ora abbiamo trattato soltanto di operatori lineari. Vogliamo ora presentare alcune considerazioni riguardanti i funzionali non lineari. Proviamo prima di tutto un teorema di punto fisso, ossia diamo una condizione per l'esistenza di soluzioni di un'equazione del tipo

$$f(x) = x\,.$$

In seguito, mostreremo come sia possibile estendere la prima formula degli incrementi finiti e la formula di Taylor.

1.13.1 Teorema delle contrazioni e applicazioni

Supponiamo che f sia una trasformazione da uno spazio di Banach X in sé stesso, non necessariamente lineare. Si dice che $f(x)$ è una *contrazione* se esiste un numero $\alpha \in [0, 1)$ tale che

$$||f(x) - f(x')|| \leq \alpha ||x - x'||\,.$$

Una contrazione è lipschitziana e quindi continua.

Se f è una qualsiasi trasformazione da X in sé, un punto $x_0 \in X$ si chiama *punto fisso* di f se

$$f(x) = x\,.$$

Vale:

Teorema 184 (teorema delle contrazioni) *Sia K un insieme chiuso dello spazio di Banach X che è invariante per la contrazione $f(x)$. Esiste uno ed un solo punto fisso di $f(x)$ che appartiene a K.*

Dim. Proviamo prima di tutto che il punto fisso, se esiste, è unico. Siano per questo x ed y due punti fissi. Vale per essi

$$||x - y|| = ||f(x) - f(y)|| \leq \alpha ||x - y||$$

ove α è strettamente minore di 1, per definizione di contrazione; e quindi la disuguaglianza precedente può solo valere se $x = y$.

Proviamo ora l'esistenza del punto fisso.

Fissiamo $k_0 \in K$ e costruiamo la successione

$$k_1 = f(k_0)\,, \ldots\,, k_n = f(k_{n-1})\,.$$

Si noti che $k_n \in K$ per ogni n, perché $f(K) \subseteq K$.

Proveremo che (k_n) è una successione fondamentale e quindi convergente dato che lo spazio X è completo. Essendo K chiuso, il limite x_0 di (k_n) è in K. Passando al limite nei due membri dell'uguaglianza

$$k_n = f(k_{n-1})$$

si trova

$$x_0 = f(x_0)$$

e quindi x_0 è punto fisso di f.

Per completare la dimostrazione, basta mostrare che (k_n) è una successione fondamentale.

Stimiamo prima di tutto

$$||k_n - k_{n-1}|| = ||f(k_{n-1}) - f(k_{n-2})|| \leq \alpha ||k_{n-1} - k_{n-2}|| \,.$$

Iterando si vede che

$$||k_n - k_{n-1}|| \leq \alpha^{n-1} ||k_1 - k_0|| \,.$$

Valutiamo ora

$$\begin{aligned} &||k_{n+m} - k_n|| \leq ||k_{n+m} - k_{n+m-1}|| + ||k_{n+m-1} - k_{n+m-2}|| + \cdots + ||k_{n+1} - k_n|| \\ &\leq \{\alpha^{n+m-1} + \alpha^{n+m-2} + \cdots + \alpha^n\} ||k_1 - k_0|| \leq \frac{\alpha^n}{1-\alpha} ||k_1 - k_0|| \,. \end{aligned}$$

Ciò prova che la successione (k_n) è fondamentale e completa la dimostrazione. ∎

Osservazione 185 Sottolineiamo che la successione (k_n) costruita nella dimostrazione del teorema converge al punto fisso per ogni scelta del valore iniziale k_0. ∎

Presentiamo ora una semplice modifica del teorema 184 che è spesso utile. Indichiamo con $f^{(n)}$ la funzione su X ottenuta componendo f con sé stessa n–volte:

$$f^{(1)}(x) = f(x) \,, \qquad f^{(k)}(x) = f\left(f^{(k-1)}(x)\right) \,.$$

Può accadere che f non sia una contrazione, ma che esista un numero ν per cui $f^{(\nu)}$ è una contrazione. Vale:

Corollario 186 *Se f è continua e se $f^{(\nu)}$ è una contrazione su un s.insieme K chiuso di X tale che $f(K) \subseteq K$, allora $f(x)$ ammette un punto fisso in K e questo è unico.*

Dim. Notiamo che se $f(x_0) = x_0$ allora vale anche

$$f(f(x_0)) = f(x_0) = x_0$$

e quindi x_0 è anche punto fisso di $f^{(2)}$. Iterando questo procedimento, si vede che x_0 è anche punto fisso della contrazione $f^{(\nu)}$. Ciò mostra l'unicità del punto fisso. Proviamone ora l'esistenza.

Si sa che esiste il punto fissa x_0 di $f^{(\nu)}$:

$$x_0 = f^{(\nu)}(x_0)\,.$$

Applicando f ai due membri dell'uguaglianza si vede che

$$f(x_0) = f(f^{(\nu)}(x_0)) = f^{(\nu)}(f(x_0))$$

ossia, anche $f(x_0)$ è punto fisso della contrazione $f^{(\nu)}$. L'unicità del punto fisso implica che

$$f(x_0) = x_0\,. \quad \blacksquare$$

Osservazione 187 Osserviamo che la ricerca dei punti fissi conduce anche alla ricerca di zeri di funzioni: il funto x_0 verifica $f(x_0) = 0$ se e solo se x_0 è punto fisso di $F(x) = x - f(x)$. $\blacksquare$

Applicazioni: il metodo delle tangenti

E' noto il metodo delle tangenti per la determinazioni di zeri di funzioni convesse di variabile reale. Mostriamo come tale metodo si ritrovi mediante il teorema delle contrazioni. Sia per questo $f(x)$ convessa su $\mathbb{R}$ e di classe C^2. Supponiamo che la derivata prima non si annulli e supponiamo che sia

$$\left|\frac{f(x)f''(x)}{f'(x)^2}\right| \leq \alpha < 1\,.$$

La funzione

$$F(x) = x - \frac{f(x)}{f'(x)}$$

ha un punto fisso x_0 se e solo se $f(x_0) = 0$ e viceversa (si ricordi che la derivata non si annulla).

Si calcola immediatamente che

$$F'(x) = \frac{f(x)f''(x)}{f'(x)^2}$$

e quindi, nelle ipotesi fatte, F è una contrazione. Ha quindi un punto fisso che si costruisce come segue: fissato un qualsiasi x_0, il punto x_1 è

$$x_1 = F(x_0) = x_0 - \frac{f(x_0)}{f'(x_0)}\,,$$

punto nel quale la tangente in $(x_0, f(x_0))$ al grafico di f,

$$y = f(x_0) + f'(x_0)(x - x_0)\,,$$

taglia l'asse delle ascisse.

Ciò è l'interpretazione geometrica del punto x_1 e quindi anche dei successivi punti x_n che approssimano lo zero di $f(x)$.

Equazioni integrali di Fredholm ed equazioni differenziali ordinarie

Sia $K(t, s, x)$ una funzione a valori reali continua su $[a, b] \times [a, b] \times \mathbb{R}$, lipschitziana nella terza variabile, uniformemente rispetto alla prima e alla seconda:

$$|K(t, s, x) - K(t, s, x')| \leq M|x - x'|$$

con M indipendente da t e da s. Consideriamo la trasformazione T_μ da $C(a, b)$ in sé definita da

$$(T_\mu x)(t) = \mu \int_a^b K(t, s, x(s))\,\mathrm{d}s + f(t)$$

con $f(t)$ funzione continua fissata. E' chiaro che

$$|(T_\mu x)(t) - (T_\mu x')(t)| \leq \mu \int_a^b M|x(s) - x'(s)|\,\mathrm{d}s \leq \mu M(b - a)||x - x'||$$

e quindi la trasformazione T è una contrazione se

$$\mu M(b - a) < 1\,. \tag{1.80}$$

Dunque:

Teorema 188 *Se $\mu M(b - a) < 1$, l'equazione di Fredholm*

$$x(t) = \mu \int_a^b K(t, s, x(s))\,ds + f(t) \tag{1.81}$$

ammette soluzione e questa è unica.

Si noti che la condizione (1.80) puó realizzarsi o con $[a, b]$ fissato, prendendo μ piccolo, o con μ fissato, spesso $\mu = 1$, prendendo $b - a$ piccolo.

Le ipotesi di questo teorema possono indebolirsi e in particolare si vede che anche l'operatore

$$(T_\mu x)(t) = \mu \int_a^t K(t, s, x(s))\, \mathrm{d}s + f(t)$$

è una contrazione se $K(t, s, x)$ è continua per $a \leq s \leq t \leq b$ ed uniformemente lipschitziana in $x \in \mathbb{R}$. Nel caso particolare in cui $f(t) = x_0$, costante, e $\mu K(t, s, x) = K(s, x)$, l'equazione (1.81) equivale a

$$x'(t) = K(t, x(t))\,, \qquad x(a) = x_0\,. \tag{1.82}$$

Dunque,

Teorema 189 *Sia $K(t, x)$ continua in t, x ed uniformemente lipschitziana in x. Il problema di Cauchy (1.82) ammette soluzione su (a, b), con b abbastanza piccolo, e tale soluzione è unica.*

1.13.2 I differenziali

Sia $f(x)$ una trasformazione da uno spazio di Banach X in uno spazio di Banach Y. Supponiamo che x_0 sia un punto interno al suo dominio.

Nel caso in cui $X = \mathbb{R}^n$ si sa che si possono definire le "derivate direzionali" in x_0 e la "matrice jacobiana", che rappresenta il "differenziale". Vogliamo estendere queste definizioni al caso in cui X è un generico spazio di Banach.

Sia v un qualsiasi elemento di X. Consideriamo il limite

$$\lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t}\,.$$

Questo limite può esistere o meno. Se esiste si indica col simbolo

$$D_v f(x_0)$$

e si chiama la *derivata* *direzionale* nella direzione v.

La derivata direzionale può esistere in una direzione e non esistere in altre direzioni; e, se anche esiste in ogni direzione, la trasformazione

$$v \longrightarrow D_v f(x_0) \tag{1.83}$$

è in generale non lineare, come prova l'esempio seguente.

Esempio 190 Si definisce una funzione $f(x,y)$ su $\mathbb{R}^2$ come segue: prima di tutto si fissa una successione di punti (x_k, y_k) due a due distinti, tutti di norma 1, ossia tutti appartenenti alla circonferenza

$$C = \{(x,y) \mid x^2 + y^2 = 1\}\,.$$

Fissato un qualsiasi punto (x,y) di $\mathbb{R}^2$ si considera il punto

$$\frac{(x,y)}{||(x,y)||}\,.$$

Questo può essere uno dei punti (x_k, y_k) o meno. Se non è uno di tali punti, si pone $f(x,y) = 0$. Se invece esiste un indice k per cui

$$\frac{(x,y)}{||(x,y)||} = (x_k, y_k)$$

allora si definisce

$$f(x,y) = k||(x,y)|| = k\sqrt{x^2+y^2}\,.$$

In particolare, $f(0,0) = 0$.

Fissata una qualsiasi direzione $v = (x,y)$, consideriamo i rapporti incrementali

$$\frac{f(tv)}{t} = \frac{f(tx,ty)}{t}\,.$$

Se $v/||v||$ non è uno dei punti (x_k, y_k), il valore del rapporto incrementale è zero per ogni t; e quindi

$$\lim_{t\to 0} \frac{f(tv)}{t} = 0\,.$$

Altrimenti, se esiste un indice k_v per cui

$$\frac{v}{||v||} = k_v(x_{k_v}, y_{k_v})$$

allora

$$\lim_{t\to 0} \frac{f(tv)}{t} = \lim_{t\to 0} \frac{tk_v||v||}{t} = k_v||v||\,.$$

Dunque, $df(x_0, v)$ esiste per ogni direzione v, ma non è funzione lineare di v. ∎

Quando invece l'operatore

$$v \longrightarrow df(x_0, v)\,,$$

è lineare, questo si chiama il *differenziale di* *Gâteaux* di f in x_0.

Se esiste il differenziale di Gâteaux di f in x_0 allora, per ogni v fissato, si ha

$$\lim_{t\to 0}\left|\left|\frac{f(x_0+tv)-f(x_0)}{t}-df(x_0,v)\right|\right|=0 \tag{1.84}$$

e quindi

$$\frac{f(x_0+tv)-f(x_0)}{t}=df(x_0,v)+\mathrm{o}(t;x_0,v)$$

ove $\mathrm{o}(t;x_0,v)$ indica una funzione della variabile reale t a valori in X e tale che

$$\lim_{t\to 0}\frac{\mathrm{o}(t;x_0,v)}{t}=0\,.$$

Si noti però che il limite non è generalmente uniforme rispetto a v. Si consideri infatti l'esempio seguente:

Esempio 191 Sia $X=\mathbb{R}^2$ e sia

$$f(x,y)=\begin{cases}1 & \text{se } x^2<x<x^4\\ 0 & \text{altrimenti.}\end{cases}$$

Si vede facilmente che il differenziale di Gâteaux di questa funzione in $(0,0)$ esiste e vale 0. Però, il limite (1.84) non è uniforme rispetto alla direzione. Infatti, sulla retta

$$x=t\,,\qquad y=mt$$

la disuguaglianza

$$\left|\frac{f(t,mt)}{t}\right|<\epsilon$$

vale, per $t>0$, quando $0\le t\le \sqrt[3]{m}\,$.

Si osservi che la funzione $f(x,y)$, pur essendo differenziabile secondo Gâteaux in $(0,0)$, non è continua. ∎

Si dice che una funzione $f(x)$ è *differenziabile secondo* *Fréchet* nel punto x_0 quando esiste un funzionale lineare L per cui

$$f(x_0+v)-f(x_0)=Lv+\mathrm{o}(v;x_0)\,.$$

Col simbolo $\mathrm{o}(v;x_0)$ si intende una funzione, questa volta da X in sé, tale che

$$\lim_{||v||\to 0}\frac{\mathrm{o}(v;x_0)}{||v||}=0\,.$$

Si richiede cioè che L verifichi

$$\lim_{||v||\to 0} \frac{||f(x_0+v)-f(x_0)-Lv||}{||v||} = 0\,.$$

Il funzionale lineare L si chiama il *differenziale di Fréchet* della funzione f in x_0, e si indica col simbolo

$$df(x_0)v\,.$$

E' facile provare:

- Se esiste il differenziale di Fréchet in un punto x_0 allora esiste anche quello di Gâteaux, e questi coincidono;
- se esiste il differenziale di Fréchet nel punto x_0 allora la funzione è continua in x_0.

La formula

$$f(x_0+v)-f(x_0) = df(x_0)v + \mathrm{o}(v;x_0)$$

generalizza la prima formula degli incrementi finiti.

Se esiste, $df(x_0)$ è un elemento di $\mathcal{L}(X,Y)$.

Quando il differenziale di Fréchet esiste in ogni punto di un intorno di x_0, la funzione

$$x \to df(x)$$

si indica col simbolo $f'(x)$ e si chiama la *funzione* $\boxed{\textit{derivata}}$ *secondo Fréchet* di $f(x)$. Questa funzione è generalmente non lineare, da X a $\mathcal{L}(X,Y)$. Può ben essere che questa sia a sua volta differenziabile secondo Fréchet nei punti di un intorno di x_0. Si può quindi definire la derivata seconda di f in x_0.

Procedendo analogamente, si definiscono anche le derivate successive.