

Analisi Funzionale

Alberto Tibaldi

19 gennaio 2013

Indice

1	Richiami su spazi metrici e spazi normati	6
1.1	Continuità, successioni di Cauchy	7
1.2	Introduzione alla compattezza	9
1.3	Teorema di estensione	10
1.3.1	Dimostrazione del teorema di estensione	11
1.4	Richiami sugli spazi vettoriali	14
1.4.1	Esempi di spazi vettoriali	15
1.5	Spazio vettoriale quoziente	19
1.5.1	Esempio 1: spazio \mathbb{R}^3	21
1.5.2	Esempio 2: spazi di successioni	23
1.6	Richiami sulle applicazioni lineari tra spazi vettoriali	24
2	Spazi normati	26
2.1	Dimostrazione della diseguaglianza triangolare per spazi l^p	29
2.1.1	Diseguaglianza di Young	29
2.1.2	Diseguaglianza di Hölder	30
2.1.3	Diseguaglianza di Minkowski	32
2.2	Palle negli spazi normati - convessità	34
2.2.1	Sintesi di norme a partire da palle	35
2.3	Spazi quoziente normati	36
2.3.1	Seminorme e quozienti	37
2.4	Spazi completi, completamento di spazi normati	39
2.4.1	Completezza su spazi normati	42
2.4.2	Teorema di completamento	44
2.5	Teorema di approssimazione di Weierstrass	48
2.5.1	Lemma dell'identità approssimata	48
2.5.2	Dimostrazione del lemma	49
2.5.3	Dimostrazione del teorema di Weierstrass	51

2.5.4	Cenni a variante del teorema di Weierstrass	55
3	Spazi con prodotto scalare	56
3.1	Definizione di prodotto scalare	56
3.1.1	Diseguaglianza di Cauchy-Schwarz	58
3.2	Ortogonalità	61
3.2.1	Continuità del prodotto scalare	61
3.2.2	Identità del parallelogramma	61
3.2.3	Teorema di Pitagora e sistemi ortonormali	63
3.3	Teorema di proiezione ortogonale su spazi di dimensione finita	63
3.4	Procedimento di Gram-Schmidt: ortogonalizzazione di sistemi di vettori	69
3.5	Rappresentazione di un vettore in termini di un sistema di vettori	70
3.6	Caratterizzazione dei sistemi completi	71
3.6.1	Esempio	74
3.7	Spazi separabili	75
3.7.1	Dimostrazione: separabilità di l^p	76
3.7.2	Osservazioni aggiuntive sugli spazi separabili	77
3.7.3	Teorema di esistenza di sistemi ortonormali completi in spazi separabili	79
3.8	Spazi di Hilbert	80
3.8.1	Definizione e osservazioni	80
3.8.2	Teorema di Riesz-Fischer	81
3.8.3	Teorema di proiezione su sottospazi	85
3.8.4	Conseguenze dei risultati proposti	90
3.8.5	Teorema di decomposizione ortogonale	91
4	Funzionali lineari	97
4.1	Definizione di funzionale lineare su spazi vettoriali	97
4.2	Funzionali continui	100
4.2.1	Norma di un funzionale limitato	102
4.2.2	Interpretazione geometrica della norma di un funzio- nale lineare	104
4.2.3	Esempi di calcolo di norme	108
4.3	Spazi duali	116
4.3.1	Problema dell'estensione di funzionali lineari continui .	117
4.4	Teorema di Hahn-Banach	119

4.4.1	Conseguenze del teorema di Hahn-Banach	119
4.5	Teoremi di rappresentazione	125
4.5.1	Annullatore di un sottospazio	125
4.5.2	Spazi biduali	128
4.5.3	Spazi riflessivi	130
4.6	Teoremi di rappresentazione	131
4.6.1	Teorema di rappresentazione di Riesz-Frechet	131
4.6.2	Teorema di Riesz di rappresentazione del duale di l^p , $1 \leq p < \infty$	135
4.6.3	Enunciato del teorema di rappresentazione del duale di $L^p(a, b)$	141
4.6.4	Rappresentazione del duale di $C([a, b])$	142
4.7	Convergenza debole	147
4.7.1	Lemma di Baire	150
4.7.2	Lemma di Baire e dimostrazione	150
4.7.3	Teorema di Banach-Steinhaus	152
4.7.4	Criterio di la convergenza debole	156
4.8	Convergenza debole* (convergenza debole stella)	161
4.8.1	Teorema di Alaoglu (versione sequenziale)	165

5 Operatori lineari 173

5.1	Introduzione ed esempi	173
5.1.1	Esempi di operatori lineari	175
5.1.2	Proprietà aggiuntive dello spazio degli operatori continui	183
5.2	Operatori compatti	185
5.2.1	Teorema di Ascoli-Arzelà	188
5.3	Spazio degli operatori compatti	196
5.3.1	Operatori di rango finito - Compattezza degli operatori di rango finito	199
5.4	Operatori aggiunti, operatori trasposti	203
5.4.1	Operatori aggiunti	203
5.4.2	Operatori trasposti	206
5.4.3	Teorema di Schauder	209
5.5	Operatori invertibili	212
5.5.1	Teorema della mappa aperta (teorema di Banach) . . .	213
5.6	Serie di Neumann	216
5.7	Lemma di Lax-Milgram	219

6	Teoria spettrale	227
6.1	Definizioni e proprietà fondamentali	227
6.2	Classificazione dello spettro	229
6.3	Teoria di Riesz-Fredholm	238
6.3.1	Lemma di Riesz	238
6.3.2	Corollario 1 del lemma di Riesz	240
6.3.3	Corollario 2 del lemma di Riesz	241
6.3.4	Proposizione: molteplicità autovalori di un operatore compatto	243
6.4	Spettro di un operatore compatto	244
6.4.1	Teorema: proprietà dello spettro di un operatore com- patto	244
6.4.2	Teorema di Fredholm	245
6.5	Operatori compatti e autoaggiunti	251
6.5.1	Operatori autoaggiunti	251
6.5.2	Lemma: esistenza di un autovalore	255
6.5.3	Teorema di Hilbert-Schmidt	258
6.5.4	Corollario del teorema di Hilbert-Schmidt	265
7	Teoria delle distribuzioni	268
7.1	Definizioni fondamentali	269
7.1.1	Definizione di distribuzione	273
7.1.2	Ordine di una distribuzione	279
7.1.3	Derivata di una distribuzione	280
7.1.4	Prodotto di una distribuzione per una funzione infini- tamente derivabile	282
7.1.5	Convergenza di distribuzioni	285
7.1.6	Supporto di una distribuzione	287
7.1.7	Distribuzioni a supporto compatto	289
7.2	Convoluzione	291
7.2.1	Convoluzione di funzioni	291
7.2.2	Proprietà del supporto	292
7.2.3	Proprietà di regolarità	293
7.2.4	Convoluzione tra una distribuzione e una funzione . . .	293
7.2.5	Convoluzione tra due distribuzioni	295
7.2.6	Proprietà della convoluzione	297
7.2.7	Applicazione: ruolo della convoluzione nelle PDE . . .	298
7.3	Distribuzioni temperate	300

7.4	Trasformata di Fourier	305
7.4.1	Trasformata di Fourier di funzioni: teoria L^1	305
7.4.2	Formula di inversione	309
7.4.3	Teoria L^2 della trasformata di Fourier	310
7.4.4	Trasformata di Fourier di distribuzioni	318
7.4.5	Antitrasformata di Fourier di distribuzioni	320
7.4.6	Proprietà aggiuntive della trasformata di Fourier di distribuzioni	326
7.4.7	Esempio pratico - Cenni ai teoremi di struttura	327
7.5	Spazi di Sobolev	328
7.5.1	Casi particolari di spazi di Sobolev	330

Capitolo 1

Richiami su spazi metrici e spazi normati

In questo capitolo prima verranno riprese le nozioni fondamentali riguardanti gli spazi metrici; verrà quindi proposto un teorema utile nel seguito della trattazione. Uno spazio metrico è un insieme E , unito a una funzione $d : E \times E \rightarrow \mathbb{R}$ che soddisfa le seguenti proprietà:

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(z, y) \forall x, y, z \in E$ (diseguaglianza triangolare)

Il fatto di definire una distanza permette di introdurre una topologia, ossia di definire le proprietà dei sottoinsiemi aperti di E . Lo strumento che viene utilizzato nell'ambito degli spazi metrici al fine di caratterizzare gli insiemi aperti è la **palla aperta**, di centro x_0 e raggio R :

$$B(x_0, R) = \{x \in E : d(x, x_0) < R\}$$

Un esempio può essere la palla in \mathbb{R}^2 , con la metrica euclidea.

Si è parlato di caratterizzare gli insiemi aperti; a tal fine, merita riprendere la definizione stessa di insieme aperto. Dato¹ $A \subset E$ è detto **aperto** se:

¹si osservi che nella trattazione, a meno che non si specifichi diversamente, \subset non esclude l'eguaglianza: \subseteq

$$\forall x \in A \exists R > 0 : B(x, R) \subset A$$

ossia, per ogni punto x appartenente all'insieme di interesse, deve esistere un certo R , tale per cui esista una palla aperta di raggio R contenuta interamente in A . Una volta definiti gli aperti è naturale definire i chiusi: $A \subset E$ è chiuso se e solo se il suo complementare è aperto: $E \setminus A$ deve essere aperto.

Si può dimostrare che tutti gli oggetti introdotti hanno le proprietà di una topologia: l'insieme stesso e l'insieme vuoto appartengono ad A , A si può scrivere come unione di un numero finito di aperti, A si può scrivere come intersezione di un numero arbitrario di aperti.

1.1 Continuità, successioni di Cauchy

Nell'ambito degli spazi metrici (come anche in quello degli spazi topologici, ma qui ci concentreremo sui metrici) è possibile definire il concetto di convergenza: data una successione $\{x_n\}$, essa è detta **convergente** a un punto x se:

$$\{x_n\} \longrightarrow x \iff d(x_n, x) \longrightarrow 0$$

Riprenderemo tra breve questa definizione, al fine di affiancarla alla definizione di continuità. Nell'ambito degli spazi metrici, una funzione si dice **continua**, se essa è continua **in ogni punto**. Questo significa che, considerando un'applicazione $A : E_1 \rightarrow E_2$, dove E_1 e E_2 sono spazi normati, si ha che A è continua se:

$$\forall x_0 \in E_1, \forall \varepsilon > 0 \exists \delta > 0 : x \in E_1, d(x, x_0) < \delta \implies d(A(x), A(x_0)) < \varepsilon$$

Questa è la classica definizione $\delta - \varepsilon$ tipica dei primi corsi di Analisi Matematica, dove però al fine di valutare i confronti tra le funzioni valutate nei vari punti si è utilizzato il concetto di metrica.

Si può dimostrare che il concetto di continuità (in ogni punto) coincide, nell'ambito degli spazi metrici, con la **continuità sequenziale**: considerando la successione $\{x_n\}$ a valori in E_1 , convergente a un punto x di E_1 , allora la successione $\{A(x_n)\}$ a valori in E_2 (spazio di arrivo dell'applicazione) converge a $A(x)$ nel medesimo:

$$x_n \longrightarrow x \iff A(x_n) \longrightarrow A(x)$$

questa è una versione *estesa* del teorema di relazione.

Nell'ambito dello studio dell'Analisi Funzionale, di solito si lavora quasi solo sulle successioni; per questo motivo, si cerca di ricondurre tutte le ipotesi a ipotesi sulla convergenza di successioni, al fine di caratterizzare il comportamento delle funzioni che intendiamo studiare.

Precedentemente si è parlato del concetto di insiemi aperti e di insiemi chiusi; quando si mettono in relazione questi insiemi, si introduce il concetto di **chiusura** di un insieme. Anche in questo caso è possibile esprimere questo concetto in termini di successioni: dato $A \subset E$, la sua chiusura \bar{A} è il più piccolo insieme chiuso che contiene A ; questo si può vedere come:

$$\bar{A} = \bigcap_{A \in K_\alpha} K_\alpha$$

dove K_α sono vari insiemi chiusi caratterizzati ciascuno da un indice α . Questo è un risultato valido anche nell'ambito degli spazi topologici. Nell'ambito degli spazi metrici, esiste un risultato più semplice da vedere: dato un generico $x \in E$, allora $x \in \bar{A}$ se e solo se esiste una successione di punti di A che converge a x . Questo risultato interessa solo gli spazi metrici, dal momento che esso **non è valido** nell'ambito degli spazi topologici. Il concetto è: **il limite della successione deve appartenere all'insieme stesso**, affinché esso sia un chiuso.

Si può a questo punto riprendere un'ulteriore definizione, valida solo nell'ambito degli spazi metrici, ossia quella di **successione di Cauchy**. Una successione $\{x_n\}$ a valori in E viene detta di Cauchy (o **fondamentale**) se:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : \forall n, m > N, d(x_n, x_m) < \varepsilon$$

La definizione è abbastanza simile a quella di convergenza a un certo punto limite, ma in realtà essa non fa intervenire alcun punto dello spazio E : questa definizione infatti riguarda esclusivamente il fatto che, per indici della successione sufficientemente grandi, sia possibile dire che i due punti siano arbitrariamente vicini tra loro (metrica arbitrariamente piccola).

Questa definizione è particolarmente utile dal momento che, come si vedrà, è molto facile da utilizzare. Le successioni di Cauchy inoltre hanno alcune proprietà: una successione convergente è sicuramente una successione

di Cauchy; infatti, sicuramente, se gli x_n devono avvicinarsi a un punto fissato, essi, appropinquandosi a esso, diverranno sempre più vicini tra loro; non è tuttavia detto che una successione di Cauchy converga a un punto: questo è il caso in cui si ha una successione definita in (a, b) , e il punto limite è b : la successione tende a convergere a un punto che non appartiene allo spazio metrico, dunque **non è convergente**, pur essendo di Cauchy.

1.2 Introduzione alla compattezza

Un'altra proprietà che un insieme può avere è la **limitatezza**: dato uno spazio metrico E , un sottoinsieme $A \subset E$ è detto **limitato** se:

$$\exists x_0 \in \mathbb{R} \ R > 0 : A \subset B(x_0, R)$$

ossia, se è possibile trovare una palla di un certo raggio R e di un certo centro x_0 tale per cui questo insieme vi sia chiuso. Allo stesso modo, a partire dall'idea di insieme limitato, una successione si dice limitata se l'insieme dei suoi valori è limitato. Un esempio di successioni limitate è l'insieme delle successioni di Cauchy; di conseguenza, essendo tutte le successioni convergenti anche successioni di Cauchy, tutte le successioni convergenti saranno limitate.

L'idea di limitatezza è utile per formulare il concetto di **completezza**: la completezza è alla base del **principio delle contrazioni**, classicamente usato nell'Analisi Matematica al fine di dimostrare il teorema di esistenza e unicità delle ODE, o per dimostrare il teorema della funzione implicita, nel caso di funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}^k$ (cosa non facile da fare senza di esso). Prima di procedere, si vuole dunque riprendere il concetto di contrazione ed enunciare il relativo teorema.

Dati E_1, E_2 due spazi metrici, una applicazione $A : E_1 \rightarrow E_2$ è detta **Lipschitz** se:

$$\exists c > 0 : d(A(x), A(y)) \leq c d(x, y) \quad \forall x, y \in E_1$$

ossia, l'applicazione è detta Lipschitz se essa non può **amplificare** la distanza tra le immagini, rispetto a quella tra i punti dello spazio di partenza, al di sopra di una certa costante c . Se poi $c < 1$, l'applicazione A in questione viene detta **contrazione**, dal momento che la distanza tra le immagini sarà sempre **minore** di una costante a sua volta minore di 1, e dunque sarà

una riduzione, anziché un'amplificazione, della distanza rispetto a quella di partenza.

Il teorema delle contrazioni afferma che, dato E uno spazio metrico completo e non vuoto, data $A : E \rightarrow E$ una contrazione, allora A ha un (uno solo) punto fisso:

$$\exists x \in E : A(x) = x$$

Questo è un primo risultato che sfrutta la completezza di uno spazio al fine di ottenere risultati. Verrà ora proposto e dimostrato un altro risultato, utilizzando ancora una volta il concetto di completezza.

1.3 Teorema di estensione

Si supponga di avere un'applicazione A_0 , definita su di un dominio $D \subset E_1$; spesso può essere utile **estendere** questa applicazione su un insieme più vasto, ossia su E_1 ; per **estensione** si intende dunque una funzione definita su un insieme più grande di quello di partenza; nell'insieme di partenza, tuttavia, il comportamento della funzione deve essere coincidente a quello della applicazione da cui si parte.

Nel dettaglio, il tipo di problema di estensione che si intende affrontare a questo punto è l'**estensione per continuità**: si richiede che l'applicazione finale sia **continua**. Vedremo che, tra le varie ipotesi che dovranno essere introdotte, vi è la **completezza** dello spazio di destinazione dell'applicazione.

Si considerino dunque E_1, E_2 due spazi metrici, e $D \subset E_1$; sia $A_0 : D \rightarrow E_2$, dove A_0 è Lipschitz². L'obiettivo è trovare una funzione $A : E_1 \rightarrow E_2$ continua, che estenda A_0 , ossia tale per cui:

$$A(x) = A_0(x), x \in D$$

in altre parole, A_0 è la **restrizione** di A in D ; si può pensare che quello che si stia facendo dunque sia il problema inverso rispetto alla determinazione della restrizione di una funzione su un sottospazio; in realtà, tuttavia, questo problema *inverso* è molto più complicato.

²Dovendo essere l'estensione continua, A_0 dovrà essere almeno continua; ciò non basta. In verità non è però necessario che sia Lipschitz, poiché è sufficiente che essa sia uniformemente continua; tuttavia, utilizzeremo questa ipotesi.

Vi sono alcune ipotesi aggiuntive: è naturale chiedere che D **sia denso** in E_1 : ogni punto di E_1 deve essere approssimabile arbitrariamente bene con un punto di D . Dunque:

$$D \subset E_1 \text{ denso} \implies \overline{D} = E_1$$

Come già detto, tra le varie ipotesi, si richiede la **completezza** di E_2 . La funzione estesa, A , sarà Lipschitz con la stessa costante c di A_0 .

Prima di procedere con la dimostrazione, si vuole proporre un'ulteriore nota: si è detto che la funzione A_0 deve essere **uniformemente continua**, poiché continua non è sufficiente. Il motivo può essere visto nel seguente esempio: dato $E_1 = [a, b]$, $D = (a, b)$, $A_0 : (a, b) \rightarrow \mathbb{R}$; considerando una funzione come in Figura ??, in cui la funzione diverge agli estremi, questa **non può essere estesa** a una funzione continua nell'intervallo chiuso E_1 : non è possibile estenderla per continuità! Per questo motivo, la continuità è una condizione non sufficiente, per la funzione.

1.3.1 Dimostrazione del teorema di estensione

Prima di tutto, si consideri l'ipotesi di continuità (che comunque si ha): dato $x \in E_1$, $E_1 = \overline{D}$, esiste una successione $\{x_n\}$ a valori in D tale per cui essa converga a x . Ogni successione convergente è di Cauchy, di conseguenza si può dire che $\{x_n\}$ sia una successione di Cauchy.

Sfruttiamo a questo punto l'ipotesi per cui A_0 è Lipschitz: questo ci garantisce che:

$$d(A_0(x_n), A_0(x_m)) \leq c d(x_n, x_m)$$

se dunque la successione $\{x_n\}$ è di Cauchy, anche la successione $\{A_0(x_n)\}$ è di Cauchy, nello spazio E_2 . Questo significa in altre parole che, fissata una certa soglia N appropriata, usando la Lipschitzianità appena introdotta, si ha che:

$$d(A_0(x_n), A_0(x_m)) \leq c\varepsilon, \quad n, m > N$$

Grazie alla Lipschitzianità si è dunque appena visto che l'applicazione A_0 manda successioni di Cauchy nello spazio di partenza in successioni di Cauchy nello spazio di arrivo.

Lo spazio di arrivo, E_2 , è però uno spazio completo; questo significa che ogni successione di Cauchy in esso è convergente. Di conseguenza, **esiste sempre il limite della successione in E_2** ; si può in questa maniera **definire** l'estensione $A(x)$ come:

$$\lim_{n \rightarrow \infty} A_0(x_n) := A(x)$$

Questo risultato è molto importante: a priori, infatti, non sapevamo che il limite esistesse.

Una volta ottenuta questa funzione $A(x)$, vorremmo capire se essa è ben definita; il valore di $A(x)$ deve sicuramente dipendere da x , ossia dal punto in cui la valutiamo, ma **non** dalla successione $\{x_n\}$ che utilizziamo per definirla. Si consideri dunque una successione $\{x'_n\}$ a valori in D , diversa dalla $\{x_n\}$ di partenza, ma sempre convergente a x . Si è già detto che A_0 è Lipschitz; quindi, applichiamo la definizione:

$$d(A_0(x_n), A_0(x'_n)) \leq c d(x_n, x'_n)$$

Riprendiamo ora una proprietà delle metriche: dati $x_n \rightarrow x$, $y_n \rightarrow y$, si ha:

$$d(x_n, y_n) \rightarrow d(x, y)$$

questa proprietà segue dal fatto che, applicando due volte la proprietà triangolare, si ha che:

$$d(x_n, y_n) \leq d(x, y) + d(x_n, x) + d(y_n, y)$$

da cui:

$$d(x_n, y_n) - d(x, y) \leq d(x_n, x) + d(y_n, y)$$

ma anche:

$$d(x, y) - d(x_n, y_n) \geq -d(x_n, x) - d(y_n, y)$$

che si scrive:

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y)$$

ma, per $n \rightarrow \infty$,

$$d(x_n, x) \rightarrow 0 \quad d(y_n, y) \rightarrow 0$$

e dunque

$$|d(x_n, y_n) - d(x, y)|$$

Tornando al caso di interesse, si ha:

$$d\left(\lim_{n \rightarrow \infty} A_0(x_n), \lim_{n \rightarrow \infty} A_0(x'_n)\right) \leq c d(x, x) = 0$$

dunque, la funzione A definita come limite delle due diverse successioni di Cauchy è ben definita.

Proviamo ora a determinare alcune proprietà di questa applicazione estesa, $A(x)$: è Lipschitz? È continua? Al fine di verificare entrambe le cose, si cerchi di verificarne la Lipschitzianità. Si considerino $\{x_n\}, \{y_n\}$ convergenti a due punti x e y . Dunque, per ipotesi, si ha che:

$$d(A_0(x_n), A_0(y_n)) \leq c d(x_n, y_n)$$

questo, per le ipotesi fatte su A_0 . Passando al limite:

$$A_0(x_n) \rightarrow A(x) \quad A_0(y_n) \rightarrow A(y)$$

ma dunque:

$$d(A(x), A(y)) \leq c d(x, y)$$

dunque, $A(x)$ è Lipschitz, ma dunque è pure continua.

Si può ora discutere l'**unicità** della funzione: data un'estensione continua di A_0 , essa è univocamente determinata dai valori di D , che è **denso in** E_1 ; dunque, dato $x \in E_1$, dato $\{x_n\} \subset D$, tale per cui $x_n \rightarrow x$, si ha che:

$$A(x) = \lim_{n \rightarrow \infty} A(x_n)$$

ma, essendo un'estensione, questo coincide con dire:

$$\lim_{n \rightarrow \infty} A(x_n) = \lim_{n \rightarrow \infty} A_0(x_n)$$

dal momento che, per $x \in D$, $A(x) = A_0(x)$. Un esempio di successione che si può utilizzare per calcolare $A(x)$ dentro D , ossia per $x \in D$, è la **costante**: infatti se $x \in D$, allora $x_n \in D$, e si può usare $x_n = x, \forall n$.

Questa dimostrazione mette in evidenza come la completezza possa essere utile per ottenere risultati nell'ambito dell'Analisi Funzionale.

1.4 Richiami sugli spazi vettoriali

Si consideri E un insieme, e lo si consideri chiuso rispetto a qualsiasi operazione di somma e prodotto, come si scriverà ora.

- operazione di somma: dati $x \in E$, $y \in E$, si ha che la somma è $E \times E \rightarrow E$, ed è tale per cui:

$$(x, y) \rightarrow x + y$$

- operazione di prodotto per scalare: dato $\alpha \in \mathbb{R}$, o $\alpha \in \mathbb{C}$, il prodotto per scalare è $\mathbb{R} \times E \rightarrow E$, o $\mathbb{C} \times E \rightarrow E$ (a seconda di dove sia definito lo scalare α), ed è tale per cui:

$$(\alpha, x) \rightarrow \alpha x$$

Il gruppo $(E, +)$ deve essere commutativo, ossia deve osservare le proprietà ora riportate.

- proprietà associativa:

$$(x + y) + z = x + (y + z)$$

- deve esistere l'elemento neutro:

$$x + 0 + x$$

- deve esistere l'elemento inverso:

$$\forall x \exists -x : x + (-x) = 0$$

- proprietà distributiva:

$$\alpha(x + y) = \alpha x + \alpha y$$

e

$$(\alpha + \beta)x = \alpha x + \beta x$$

dove α, β sono scalari.

- deve valere la proprietà associativa:

$$\alpha(\beta + x) = (\alpha\beta)x$$

deve esistere l'elemento neutro per la moltiplicazione a scalare:

$$1x = x$$

1.4.1 Esempi di spazi vettoriali

Verrà a questo punto riportato un elenco di esempi di spazi vettoriali, al fine di comprendere quanto il concetto di spazio vettoriale sia vasto. Per verificare che ciascuno di questi spazi è vettoriale, bisognerebbe prima verificare le due proprietà fondamentali, ossia la chiusura a somma e prodotto per scalare, quindi si dovrebbe verificare che $(E, +)$ sia commutativo.

- Il primo esempio di spazio vettoriale è \mathbb{R}^n : esso è uno spazio vettoriale su \mathbb{R} (nel senso che con $\alpha \in \mathbb{R}$, scalare, il prodotto è chiuso in \mathbb{R}^n . Allo stesso modo, \mathbb{C}^n è uno spazio vettoriale su \mathbb{C} , ma pure su \mathbb{R} (moltiplicare una n -pla di numeri complessi per un numero reale li mantiene complessi!).
- Lo spazio delle funzioni continue nell'intervallo $[a, b]$: $C([a, b])$ è uno spazio vettoriale su \mathbb{R} . Infatti, la somma di funzioni continue è ancora continua:

$$(f + g)(x) = f(x) + g(x)$$

allo stesso modo, moltiplicando una funzione continua per uno scalare si ha ancora una volta una funzione continua:

$$(\alpha f)(x) = \alpha f(x)$$

- Un altro esempio di spazio vettoriale è lo spazio dei polinomi con coefficienti in \mathbb{R} (o in \mathbb{C}).
- Un altro spazio vettoriale è lo spazio delle successioni limitate: l^∞ , definito come:

$$l^\infty = \{x = (x_n) : x_n \in \mathbb{R}, \exists c > 0 : |x_n| < c \forall n\}$$

infatti, date due successioni limitate, sommandole esse rimangono limitate; moltiplicando per uno scalare una successione, essa continua a essere limitata.

- Spazio delle successioni che danno luogo a serie assolutamente convergenti:

$$l^1 = \left\{ x = (x_n) : \sum_{n=1}^{\infty} |x_n| < \infty \right\}$$

con la notazione $< \infty$ si intende che la serie deve convergere. Questa proprietà si può dimostrare utilizzando la disuguaglianza:

$$|x_n + y_n| \leq |x_n| + |y_n|$$

e se dunque si suppone di avere due serie convergenti, la loro somma sarà per confronto convergente.

- Spazio delle successioni convergenti, C :

$$c = \left\{ x = (x_n) : \exists \lim_{n \rightarrow \infty} x_n, \text{ finito} \right\}$$

- Spazio delle successioni convergenti a 0:

$$c_0 = \left\{ x = (x_n) : \lim_{n \rightarrow \infty} x_n = 0 \right\}$$

- Spazio delle successioni assolutamente nulle: si tratta di successioni tali per cui esiste una certa soglia N al di sopra della quale ogni termine della successione è nullo:

$$c_{00} = \{x = (x_n) : \exists N : x_n = 0 \forall n > N\}$$

si osservi che N dipende dalla successione che si sta considerando.

Sono stati introdotti vari spazi di successioni; è possibile dire che:

$$c_{00} \subset l^1 \subset c_0 \subset c$$

Sicuramente, se al di sopra di un certo N la successione appartenente a c_{00} ha solo più elementi nulli, essa ha solo un numero finito di addendi, di conseguenza la serie sarà sicuramente convergente, dal momento che da un certo punto in poi si aggiungeranno solo zeri; questo significa che tutte le successioni appartenenti a c_{00} sono convergenti. Non è tuttavia detto che una serie, per convergere, debba avere un comportamento di questo genere; per questo motivo, è solo un sottospazio. Lo spazio delle successioni che generano serie assolutamente convergenti è un sottospazio di c_0 , dal momento che non tutte le successioni che convergono a 0 hanno una serie assolutamente convergente ad essa associata. Infine, ovviamente, non tutte le successioni convergenti convergono a 0.

Questi esempi di spazi vettoriali sono assolutamente generali: aldilà di ciò che si ha ben presente, come lo spazio euclideo, anche spazi di successioni possono essere vettoriali; questo significa che una successione è un elemento dello spazio, dunque un **vettore**. Questo permette di intuire quanto generale sia il concetto di spazio vettoriale.

Al fine di concludere i richiami sugli spazi vettoriali, si introduce dunque una rassegna di definizioni.

- **Sottospazio vettoriale:** dato E spazio vettoriale (chiamandolo **spazio ambiente**), un insieme E_1 è un sottospazio se, dati $x, y \in E_1$, α, β scalari, si ha che:

$$\alpha x + \beta y \in E_1$$

dunque, un sottospazio è anche uno spazio vettoriale, a sua volta.

- **Combinazione lineare:** dati x_1, x_2, \dots, x_n vettori appartenenti a uno spazio vettoriale, per combinazione lineare si intende un vettore della forma:

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

ossia, una **somma finita** di vettori.

- Sottospazio generato da una famiglia di vettori (**span**): dato S un sottoinsieme di E , si definisce

$$\text{span}(S)$$

l'insieme di tutte le combinazioni lineari degli elementi di S . S è un **sottospazio**, generato da S ; nel dettaglio, esso è il più piccolo sottospazio di E contenente interamente S , e, in altre parole, è dato dall'intersezione di tutti i sottospazi di E che contengono S .

- Un sistema di vettori $\{x_1, x_2, \dots, x_n\}$ si dice:
 1. **linearmente dipendente** se esistono scalari $\alpha_1, \dots, \alpha_n$ **non tutti nulli** e tali per cui la combinazione lineare

$$\alpha_1 x_1 + \dots + \alpha_n x_n$$

sia il vettore nullo;

2. **linearmente indipendente** se **non è** linearmente dipendente;
3. che genera E , se

$$E = \text{span} \{x_1, x_2, \dots, x_n\}$$

4. **base**, se è **linearmente indipendente** e **genera** E .

- **Dimensione**: se per lo spazio vettoriale si riescono a trovare N vettori linearmente indipendenti ma non più di N , la dimensione dello spazio vettoriale è N .

A questo punto, un'osservazione: se si riescono a trovare continuamente nuovi vettori linearmente indipendenti, vuol dire che esiste una successione di vettori:

$$\{x_1, x_2, \dots, x_n, \dots\} \subset E, \quad \forall n$$

tale per cui questo sistema di vettori sia sempre linearmente indipendente. In questo caso, se dunque non si riesce a trovare un *ultimo* vettore linearmente indipendente dagli altri, lo spazio è di dimensione infinita. Un esempio di spazio dotato di questa proprietà è lo **spazio dei polinomi in una variabile**:

$$\text{span} \{1, t, t^2, t^3 \dots\}$$

questo è di dimensione infinita poiché i polinomi possono avere grado infinito.

1.5 Spazio vettoriale quoziente

Si riprenderà a questo punto il concetto di spazio quoziente, e verrà introdotto nel dettaglio il concetto di spazio vettoriale quoziente. Si consideri prima di tutto $E_1 \subset E$, dove dunque E è lo **spazio ambiente**, mentre E_1 è un suo sottospazio. Si consideri quindi la seguente relazione di equivalenza definita su E :

$$x \sim y \iff x - y \in E_1$$

ossia, due vettori $x, y \in E$ sono equivalenti se la loro differenza appartiene al sottospazio E_1 . Prima di tutto, bisognerebbe dimostrare che questa è effettivamente una relazione di equivalenza, ossia verificare che soddisfi le tre proprietà delle relazioni di equivalenza:

- proprietà riflessiva:

$$x - x = 0$$

dove $0 \in E_1$, ovviamente.

- proprietà simmetrica: se $x - y \in E_1$, allora $y - x$ è l'opposto di $x - y$, e dunque anche esso appartiene a E_1 .
- proprietà transitiva.

Una volta definita la relazione di equivalenza appena proposta, è possibile definire l'**insieme quoziente** come l'insieme delle classi di equivalenza: E/E_1 . Ossia:

$$E/E_1 = \{[x] : x \in E\}$$

rivediamo a questo punto il concetto di classe di equivalenza: la classe di equivalenza $[x]$ si definisce come:

$$[x] = \{z \in E : z - x \in E_1\}$$

ossia: la classe di equivalenza rispetto alla relazione \sim appena definita è l'insieme dei punti z appartenenti a E tali per cui $z - x$ sia un punto appartenente al sottospazio E_1 . Questa definizione, al fine di essere compresa meglio, può essere associata a un significato geometrico: prima di tutto, si definisca $y = z - x$; il vettore y apparterrà a E_1 . La definizione si può scrivere:

$$[x] = \{z \in E : z - x \in E_1\} = \{x + y \in E : y \in E_1\}$$

a questo punto, si provi a introdurre un significato geometrico, come in Figura ???: si immagini, per fare un esempio, che E sia il piano, e che E_1 sia una certa retta, passante per l'origine del sistema di riferimento (la cosa è necessaria, dal momento che E_1 è un sottospazio, dunque uno spazio vettoriale, dunque deve passare per l'origine!); $x \in E$ è un generico punto del piano. L'insieme delle classi di equivalenza $[x]$ è una **retta parallela a E_1** : infatti, se si considera un qualsiasi $y \in E_1$, ossia un certo punto della retta E_1 , e se si calcola $x + y$, è possibile vedere che $x + y$ apparterrà sempre a una retta, parallela a E_1 : $[x]$. Sommando infatti y si aggiungerà sempre un vettore parallelo alla retta E_1 , al punto x di partenza considerato. Un altro modo di chiamare $[x]$ è :

$$[x] = x + E_1$$

ossia, a un certo punto x si somma un elemento appartenente al sottospazio E_1 : questa è la classe di equivalenza $[x]$. Si noti tuttavia che $[x]$ non sarà di sicuro uno spazio/sottospazio vettoriale: esso infatti non passa per l'origine, dunque non rispetta le proprietà!

Finora si è parlato di insiemi quoziente; ciò che si vuole ottenere ora è lo **spazio vettoriale quoziente**, ossia si vuole fornire a E/E_1 una struttura di spazio vettoriale. Per fare ciò, è necessario verificare che, date due classi di equivalenza $[x_1]$ e $[x_2]$, e uno scalare α , lo spazio quoziente sia chiuso rispetto alle operazioni di somma e prodotto per scalare. Queste operazioni possono essere definite naturalmente come segue:

$$[x_1] + [x_2] = [x_1 + x_2]$$

$$\alpha[x] = [\alpha x]$$

A questo punto, bisognerebbe verificare che queste operazioni non dipendano dalle rappresentanti che si considerano. Si considerino dunque:

$$[x'_1] = [x_1]$$

e

$$[x'_2] = [x_2]$$

Si può dire che:

$$[x'_1 + x'_2] = [x_1 + x_2]$$

oppure no? Se per ipotesi si ha che:

$$[x'_1] = [x_1]$$

allora si ha che, per la relazione di equivalenza precedentemente definita, si ha:

$$x'_1 - x_1 \in E_1 \quad x'_2 - x_2 \in E_1$$

ma dunque:

$$x'_1 + x'_2 - x_1 - x_2 = (x'_1 - x_1) + (x'_2 - x_2) \in E_1$$

infatti, si ha che entrambi gli elementi sotto parentesi tonda sono in E_1 , dunque sommandoli si rimane in E_1 . Allo stesso modo si può dimostrare la chiusura per prodotto con scalare.

Una volta visto che lo spazio quoziente è uno spazio vettoriale, esso può avere una dimensione. La dimensione dello spazio quoziente è detta **codimensione**:

$$\text{codim}(E_1) = \dim(E/E_1)$$

1.5.1 Esempio 1: spazio \mathbb{R}^3

Al fine di comprendere ciò, si propone il seguente esempio: $E = \mathbb{R}^3$, E_1 suo sottospazio, generato da due vettori $\{v_1, v_2\}$, linearmente indipendenti. È chiaro che v_1 e v_2 generano un piano nello spazio; è possibile trovare dunque v_3 , che non appartiene a questo piano, e dunque che non appartiene a E_1 .

A partire da ciò, si può capire cosa significhi **quoziante**: significa **ignorare le dimensioni appartenenti al sottospazio rispetto al quale si quoziante**; infatti, riprendendo ??, si può vedere che se a un vettore di E_1 sommo x , il risultato finale **non dipende dalle direzioni al di fuori di quelle appartenenti alla classe di equivalenza**. Vediamolo meglio: dato un generico $x \in \mathbb{R}^3$, esso può essere scritto mediante una certa combinazione lineare dei tre vettori v_1, v_2, v_3 per mezzo di alcuni specifici coefficienti:

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$$

in termini di classi di equivalenza, si ha che:

$$\alpha_1 v_1 + \alpha_2 v_2 \in E_1$$

ossia, i due vettori appartengono a E_1 ; dunque, per la classe di equivalenza, si ha che:

$$[x] = [\alpha_3 v_3]$$

come mai? Si riprenda Figura ??: le dimensioni appartenenti al sottospazio vengono ignorate, nel senso che se a un vettore appartenente a E_1 sommo un generico punto di E , l'unica variazione sarà nella direzione non compresa in E_1 . Riprendendo le precedenti definizioni, si ha:

$$[x] = \{x + y \in E : y \in E_1\}$$

se $x + y$ deve appartenere in generale a E , ma $y \in E_1$, allora l'unica parte di x che può variare è quella che **non appartiene a E_1** ; se si variasse anche l'altra parte, si potrebbe andare fuori da E_1 , perdendo le ipotesi; dunque, ciò che rimane è il vettore che non appartiene a E_1 :

$$[x] = [\alpha_3 v_3]$$

ossia, la classe di equivalenza di x è un multiplo di una classe fissata. Questo significa che lo spazio quoziente ha dimensione 1, dal momento che:

$$E/E_1 = \text{span}([v_3])$$

Si ha dunque che $[v_3]$ è una base per lo spazio quoziente. E_1 dunque ha codimensione pari a 1, mentre ha dimensione 2: questo significa che la codimensione è quanto manca per generare l'intero spazio ambiente E (nel nostro esempio, \mathbb{R}^3).

1.5.2 Esempio 2: spazi di successioni

Si consideri a questo punto un secondo esempio per fissare i concetti: dati

$$c = \{x = (x_n) \text{ convergenti} \}$$

$$c = \{x = (x_n) \text{ convergenti a } 0\}$$

si vuole calcolare una base per lo spazio quoziente. Al fine di fare ciò, si proceda in maniera simile a prima: si provi a scrivere un generico elemento dello spazio ambiente, ossia c , come qualcosa che sta nel sottospazio più qualcos'altro; facendo il quoziente, rimarrà solo "l'altro". Si consideri il seguente esempio di modo di procedere:

$$x = y + \alpha_n$$

si supponga che $y = (y_n) \in c_0$, ossia si immagini che sia una successione che converge a 0; questo significa che dunque α_n sarà una successione con limite convergente allo stesso limite di x (essendo y_n convergente a 0!). Si definisca dunque:

$$\alpha = \lim_{n \rightarrow \infty} x_n$$

al fine di scrivere un esempio, si consideri la seguente successione convergente (non a 0!):

$$\alpha_n = (\alpha, \alpha, \alpha, \dots, \alpha) = \alpha(1, 1, 1, \dots, 1) = \alpha\bar{1}$$

dove dunque $\bar{1}$ è la successione costante e con ogni elementi pari a 1. Ora, se si fa:

$$x - \alpha\bar{1}$$

questa differenza converge a 0! Questo significa che:

$$[x] = [\alpha\bar{1}] = \alpha[\bar{1}]$$

ossia, la successione costante è una base per lo spazio quoziente di c rispetto a c_0 : essa è una funzione che non converge a 0, dunque appartiene "all'altra parte", alla "parte rimanente" rispetto allo spazio ambiente. Dunque:

$$c/c_0 = \text{span}([\bar{1}])$$

e dunque, la codimensione di c_0 è 1.

1.6 Richiami sulle applicazioni lineari tra spazi vettoriali

Si considerino E_1, E_2 due spazi vettoriali (sullo stesso campo, che può essere \mathbb{R} o \mathbb{C}); si consideri A un'applicazione:

$$A : E_1 \rightarrow E_2$$

l'applicazione A viene detta **lineare** se:

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay, \quad \forall x, y \in E_1, \quad \alpha, \beta \text{ scalari}$$

si noti che l'applicazione A su x viene indicata come Ax e non come $A(x)$: questa è una tradizione quando si parla di applicazioni lineari.

Data un'applicazione lineare, è possibile definire il suo **nucleo** come l'insieme dei suoi zeri:

$$\ker(A) = \{x \in E_1 : Ax = 0\}$$

si osservi che:

$$\ker(A) \subset E_1$$

dunque, esso è un **sottospazio di** E_1 . Allo stesso modo, si può definire l'**immagine** di A :

$$A(E_1) = \{Ax : x \in E_1\} \subset E_2$$

in questo caso, l'immagine è un sottospazio di E_2 .

Per concludere, date le definizioni appena introdotte, vengono proposte alcune definizioni finali.

- A è **iniettiva** se:

$$\ker(A) = \{\emptyset\}$$

questo è vero solo dal momento che l'applicazione è **lineare**.

- A è **suriettiva** se:

$$A(E_1) \equiv E_2$$

- A è un **isomorfismo** se è sia iniettiva sia suriettiva. In questo caso, l'applicazione inversa A^{-1} esiste, ed è **lineare**

Capitolo 2

Spazi normati

Sia X uno spazio vettoriale (su \mathbb{R} o su \mathbb{C}); una **norma** è una funzione $\|\cdot\| : X \rightarrow \mathbb{R}$ che soddisfa le seguenti proprietà:

1. $\|x\| \geq 0$, dove $\|x\| = 0 \iff x = 0$. Il fatto che questa proprietà sia un “se e solo se” verrà ripreso in seguito.
2. $\|\alpha x\| = |\alpha| \|x\|$, questo $\forall x \in X$, e $\forall \alpha$, dove quest’ultimo è uno scalare;
3. $\|x + y\| \leq \|x\| + \|y\|$

Per **spazio normato** si intende uno spazio vettoriale sul quale è definita una norma; esso dunque si indica come $(X, \|\cdot\|)$. Uno spazio normato è sempre uno spazio metrico: è infatti possibile definire una distanza per X come segue:

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

Tutto ciò ha senso dal momento che **siamo in uno spazio vettoriale**; questo significa che gli spazi che stiamo trattando hanno sia una **struttura vettoriale**, sia una **struttura metrica** e le due sono compatibili tra loro: la norma induce una metrica, ma dunque è necessario che questa soddisfi tutte le proprietà del caso.

Si consideri a questo punto un esempio di una proprietà delle norme: data una successione $x_n \rightarrow x$, si ha che:

$$\|x_n - x\| \rightarrow 0$$

questo è un esempio di proprietà che si può considerare per le norme; le varie proprietà delle metriche (come la Lipschitzianità, e così via), possono essere estese anche alle norme. Altra relazione che esiste è la seguente disuguaglianza:

$$|||x| - |y|| \leq |x - y|$$

questa è una disuguaglianza nota con i numeri reali: dati a, b reali, si ha infatti che:

$$||a| - |b|| \leq |a - b|$$

questa relazione si può dimostrare dimostrando che:

$$\begin{cases} |a| - |b| \leq |a - b| \\ |b| - |a| \leq |a - b| \end{cases}$$

Si propone un accenno di dimostrazione: bisogna lavorare su $|a|$, sommando e sottraendo b , e quindi applicando la disuguaglianza triangolare:

$$|a| = |a - b + b| \leq |a - b| + |b| \implies |a| - |b| \leq |a - b|$$

si fa la stessa cosa con $|b|$:

$$|b| = |b - a + a| \leq |b - a| + |a| = |a - b| + |a| \implies |b| - |a| \leq |a - b|$$

dunque, è soddisfatta.

Detto ciò, si può a questo punto affermare come già anticipato che la norma è una funzione Lipschitz:

$$|||x| - |y|| \leq c|x - y|$$

essendo Lipschitz, è anche continua!

Si propongono a questo punto alcuni esempi di spazi normati, con la norma normalmente utilizzata per essi.

1. l^∞ , c , c_0 sono tutti spazi vettoriali che possono essere dotati di una norma. Di solito, per questi spazi, la norma utilizzata è:

$$||x||_\infty = \sup_k |x_k|$$

questa norma solitamente viene chiamata **norma infinito**. Si osservi che, per tutti questi spazi, questa norma è ben definita: gli spazi appena proposti sono sempre di successioni limitate, dunque il sup sarà un elemento non infinito.

2. Lo spazio delle funzioni continue, $C([a, b])$, dove, data una certa funzione $x(t)$ a esso appartenente, $t \in [a, b]$, si ha, come norma:

$$\|x\|_\infty = \sup_{t \in [a, b]} |x(t)|$$

anche in questo caso, è ben definita: una funzione continua su un chiuso $[a, b]$ è limitata (e presenta un massimo, per il teorema di Weierstrass); la norma è ben definita.

3. Lo spazio l^1 delle successioni che generano serie assolutamente convergenti:

$$l^1 = \left\{ x = (x_k) : \sum_{k=1}^{\infty} |x_k| < \infty \right\}$$

alcune parole in più su questo spazio: un esempio di successione che appartiene a esso è $x_k = \frac{1}{k^2}$, poiché a essa è associata una serie assolutamente convergente; la stessa cosa non si può dire per $x_k = \frac{1}{k}$: in questo caso, ossia per la serie geometrica, non si ha assoluta convergenza. In questo caso, quando la serie converge, come norma si può usare la **norma 1**:

$$\|x\|_1 = \sum_{k=1}^{\infty} |x_k|$$

4. Introduciamo a questo punto spazi che rappresentano una sorta di via di mezzo tra l^1 e l^∞ : se infatti l^1 è uno spazio di funzioni abbastanza “ristretto”, se si pensa che esso è composto solo da successioni che generano serie assolutamente convergenti, d'altra parte l^∞ è lo spazio delle successioni convergenti, dunque molto più ampio. Esistono delle vie di mezzo, ossia gli spazi l^p , per $1 < p < \infty$ (se $p = 1$ si ha l^1):

$$l^p = \left\{ x = (x_k) : \sum_{k=1}^{\infty} |x_k|^p < \infty \right\}$$

al crescere di p , gli spazi diventano sempre più “ampi”, ossia racchiudono più successioni. Si consideri per esempio $p = 2$, dunque l^2 : in esso la serie associata, $x_k = \frac{1}{k}$ converge, dal momento che si ha l’elevazione a quadrato. La norma per l^p di solito usata è:

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}}$$

la radice p -esima su tutto serve per soddisfare la proprietà 2, ossia quella di moltiplicazione per scalare; se non si avesse, infatti, si avrebbe un α^p al momento di portare lo scalare “fuori” dalla norma.

2.1 Dimostrazione della disuguaglianza triangolare per spazi l^p

In questa sezione si vuole dimostrare la disuguaglianza triangolare per quella che è stata definita come la norma di l^p ; attualmente non è possibile dire che sia effettivamente una norma dal momento che questa disuguaglianza non è verificata. A breve, si potrà dire. Questa dimostrazione richiede un certo numero di risultati intermedi, che verranno dimostrati passo-passo.

2.1.1 Disuguaglianza di Young

Si considerino $a, b > 0$, e $p, q > 1$ e tali per cui:

$$\frac{1}{p} + \frac{1}{q} = 1$$

quando p e q soddisfano questa relazione, si suol dire che essi sono **coniugati**. Dunque, date queste condizioni, vale la disuguaglianza di Young:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Questa è una generalizzazione del caso banale per $p = q = 2$, per cui:

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2}$$

questa da dimostrare è molto semplice: basta portare ab al secondo membro e riconoscere che è semplicemente il quadrato di un binomio, che è sempre maggiore o uguale a zero.

Per dimostrare la disuguaglianza nel caso più generale, prima di tutto si osservi che, se p e q sono coniugati, allora si hanno le seguenti relazioni:

$$\frac{1}{p-1} = q-1 \quad (p-1)q = p$$

A questo punto, si vuole proporre una dimostrazione geometrica della disuguaglianza: si disegnino sul piano xy le funzioni $x = x^{p-1}$, e poi $x = y^{\frac{1}{p-1}}$. Per quanto riguarda la prima, essa è semplicemente una parabola di $p-1$ -esimo grado dove la variabile indipendente è x ; per la seconda, sfruttando la relazione tra gli esponenti coniugati appena proposta, è possibile ottenere:

$$x = y^{\frac{1}{p-1}} = y^{q-1}$$

a questo punto, si calcolino le aree sottese da queste due curve; nel primo caso, l'area è $\frac{a^p}{p}$; nel secondo caso, l'area è $\frac{b^q}{q}$. Tutto ciò è rappresentato geometricamente in Figura ??.

Come è evidente dalla suddetta figura, la somma delle due aree sottese dalla curva è sicuramente maggiore dell'area del rettangolo ab , poiché si ha un "pezzo di area in più". Di conseguenza, abbiamo dimostrato geometricamente che:

$$ab \leq \frac{b^q}{q} + \frac{a^p}{p}$$

2.1.2 Disuguaglianza di Hölder

La disuguaglianza di Hölder afferma che, dati $x \in l^p$, $y \in l^q$, $p, q > 1$, e tali da essere coniugati, allora:

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_p \|y\|_q$$

dove¹ (per esempio):

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}}$$

Questo risultato è molto importante: esso ci dice che la successione:

$$(x_1 y_1, x_2 y_2, \dots) \in l^1$$

ossia, è convergente.

Si vuole proporre una dimostrazione per questo fatto. Supponiamo prima di tutto che $\|x\|_p$ e $\|y\|_q$ siano diversi da zero (altrimenti, la dimostrazione diventa banale). Si ha, dividendo il membro destro dell'equazione appena scritta:

$$\sum_{k=1}^{\infty} \frac{|x_k|}{\|x\|_p} \frac{|y_k|}{\|y\|_q} \leq 1$$

ossia, dividendo per il membro destro, a destra rimane solo "1". A questo punto, per compattezza, si chiamino:

$$a_k = \frac{|x_k|}{\|x\|_p} \quad b_k = \frac{|y_k|}{\|y\|_q}$$

in questo modo, in maniera compatta, si ha che:

$$\sum_{k=1}^{\infty} a_k b_k \leq 1$$

Ora, si veda che a_k e b_k sono due successioni; indicando:

$$a = (a_k) = (a_1, a_2, \dots) \quad b = (b_k) = (b_1, b_2, \dots)$$

si ha che:

$$\|a_k\|_p = 1 \quad \|b_k\|_q = 1$$

infatti (per esempio su a_k):

¹si vuole rimarcare il fatto che non si è ancora dimostrato che tutte queste sono effettivamente delle norme: stiamo dimostrando la disuguaglianza triangolare!

$$\|a_k\|_p = \left(\sum_{k=1}^{\infty} |a_k|^p \right)^{\frac{1}{p}} = \left(\sum_{k=1}^{\infty} \left| \frac{x_k}{\|x\|_p} \right|^p \right)^{\frac{1}{p}} = \frac{1}{\|x\|_p} \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}} = 1$$

e idem per b_k . A questo punto, è possibile applicare la disuguaglianza di Young, ricavata nella sezione precedente, per ottenere:

$$a_k b_k \leq \frac{a_k^p}{p} + \frac{b_k^q}{q}$$

da cui:

$$\sum_{k=1}^{\infty} a_k b_k \leq \sum_{k=1}^{\infty} \frac{a_k^p}{p} + \sum_{k=1}^{\infty} \frac{b_k^q}{q} = \frac{1}{p} \sum_{k=1}^{\infty} a_k + \frac{1}{q} \sum_{k=1}^{\infty} b_k = \frac{1}{p} + \frac{1}{q} = 1$$

2.1.3 Disuguaglianza di Minkowski

La disuguaglianza di Minkowski è l'equivalente della disuguaglianza triangolare nell'ambito della norma l^p . Dati $x, y \in l^p$, $1 \leq p < \infty$, si ha che:

$$x + y \in l^p \quad \|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Questa disuguaglianza è stata introdotta di Minkowski nei primi del 1900, nell'ambito di \mathbb{R}^n ; essa tuttavia può essere applicata in ambiti molto più generici, tanto che Riesz e altri, a partire da queste idee, introdussero l'idea di spazi normati.

Proponiamo, per questa disuguaglianza, la seguente dimostrazione: si ha che

$$\|x + y\|_p^p = \sum_{k=1}^{\infty} |x_k + y_k|^p$$

dove si è portata a sinistra l'elevazione a $\frac{1}{p}$. Questa espressione si può riscrivere come:

$$= \sum_{k=1}^{\infty} |x_k + y_k| |x_k + y_k|^{p-1}$$

per quanto riguarda il valore assoluto, è possibile applicare la disuguaglianza triangolare:

$$|x_k + y_k| \leq |x_k| + |y_k|$$

dunque,

$$\sum_{k=1}^{\infty} |x_k + y_k| |x_k + y_k|^{p-1} \leq \sum_{k=1}^{\infty} |x_k| |x_k + y_k|^{p-1} + \sum_{k=1}^{\infty} |y_k| |x_k + y_k|^{p-1}$$

Su ciascuno di questi termini è dunque possibile applicare la disuguaglianza di Hölder; se ne applica una in un esempio, poi si procede con i calcoli:

$$\sum_{k=1}^{\infty} |y_k| |x_k + y_k|^{p-1} \leq \|x\|_p \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q} \right)^{\frac{1}{q}}$$

infatti, si è semplicemente usato:

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \|x\|_p \|y\|_q$$

applicando la definizione di $\|y\|_q$. Dunque, si ottiene:

$$\dots \leq \|x\|_p \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q} \right)^{\frac{1}{q}} + \|y\|_q \left(\sum_{k=1}^{\infty} |x_k + y_k|^{(p-1)q} \right)^{\frac{1}{q}}$$

ma:

$$(p-1)q = p$$

da cui, riprendendo il termine iniziale:

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_q) \|x + y\|_q^{\frac{p}{q}}$$

a questo punto, si divida tutto per la norma della somma a membro destro:

$$\|x + y\|_p^{p-\frac{p}{q}} \leq \|x\|_p + \|y\|_q$$

ma, essendo:

$$p - \frac{p}{q} = 1$$

si ha:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Questa è la disuguaglianza triangolare! Dunque, effettivamente, quella che finora era stata chiamata norma- p senza aver mai dimostrato che era una norma, lo era!

2.2 Palle negli spazi normati - convessità

Si vuole a questo punto proporre alcune considerazioni riguardo le palle, focalizzandosi all'inizio sulle palle unitarie, negli spazi normati. La definizione di palla unitaria è:

$$B = \{x \in X : \|x\| \leq 1\}$$

La “forma” di questa palla, nello spazio, dipende dalla norma che si sceglie. Si consideri per esempio lo spazio \mathbb{R}^2 : se si utilizza la norma 1, $\|\cdot\|_1$, la palla unitaria sarà un rettangolo con le diagonali sugli assi principali:

$$|x| + |y| < 1$$

in questo caso dunque la palla non è “tonda”. Un secondo esempio, questa volta con le palle tonde, è per lo stesso spazio con la norma 2: $\|\cdot\|_2$: in questo caso si ha un cerchio:

$$\sqrt{x^2 + y^2} \leq 1$$

Terzo esempio può essere sempre lo stesso spazio, questa volta con la norma ∞ :

$$\sup\{|x|, |y|\}$$

in questo caso, si ha un quadrato con i lati paralleli agli assi. Le tre rappresentazioni si possono vedere in Figura ??.

Da questa prima osservazione dunque non sembra di poter dire nulla sulla forma delle palle. In verità, esiste una particolare proprietà: ogni palla, in uno

spazio normato, è **convessa**. Al fine di richiamare il concetto di convessità, si considerino le due rappresentazioni in Figura ??: un insieme E è convesso se, $\forall x_1, x_2 \in B$, l'intero segmento che collega i due punti appartiene all'insieme stesso; se è possibile scegliere due punti tali per cui il segmento che li collega non è interamente appartenente al dominio, allora l'insieme non è convesso.

Al fine di descrivere formalmente questa proprietà, si considerino due punti $x_1, x_2 \in E$; si scriva dunque la parametrizzazione della curva che collega x_1 a x_2 : dato $\alpha \in [0, 1]$ il parametro della curva, si ha:

$$(1 - \alpha)x_1 + \alpha x_2 \forall \alpha$$

questo è il “segmento” che collega i due punti. Calcoliamo la norma di ciò:

$$\|(1 - \alpha)x_1 + \alpha x_2\| \leq \|(1 - \alpha)x_1\| + \|\alpha x_2\| = (1 - \alpha)\|x_1\| + \alpha\|x_2\|$$

poiché sia x_1 sia x_2 appartengono alla palla, si ha che:

$$\|x_2\| \leq 1 \quad \|x_1\| \leq 1$$

da cui, finalmente:

$$\leq (1 - \alpha) + \alpha = 1$$

se invece di 1 si mette un generico raggio r , la dimostrazione funziona egualmente.

Questa dimostrazione mostra come una palla in uno spazio metrico sia **sempre** convessa; ovviamente, questo è vero quando la palla è “piena”: se si ha solamente la frontiera, non è possibile dire la stessa cosa. Tutto ciò vale tanto per palle aperte quanto per palle chiuse.

2.2.1 Sintesi di norme a partire da palle

Si vuole proporre un cenno su un problema opposto a quello appena affrontato: utilizzando le proprietà delle norme si è studiata la palla associata a una certa norma. Il fatto che però si disponga di un certo insieme convesso con una forma arbitraria, permette di **generare** una norma; anche questo problema è stato affrontato da Minkowski, mediante l'introduzione del cosiddetto **funzionale di Minkowski**:

$$\mathcal{P}(x) = \inf \left\{ t > 0 : \frac{x}{t} \in K \right\}$$

dove, dato X lo spazio ambiente, $K \subset X$, e K è convesso.

Per fissare l'idea del fatto che è possibile ottenere una norma a partire da un convesso, si consideri il seguente esempio: dato un'ellisse in cui un raggio è il doppio dell'altro, ricavare la norma la cui palla unitaria produca il suddetto.

Questo esercizio si può risolvere scrivendo semplicemente l'equazione dell'ellisse appena proposto:

$$\frac{x^2}{4} + \frac{y^2}{1} \leq 1$$

questa è *quasi* la norma: non rispetta la proprietà di moltiplicazione per scalare! Per terminare, basta mettere una radice quadra, che aggiusti tutto:

$$\sqrt{\frac{x^2}{4} + y^2}$$

2.3 Spazi quoziente normati

Precedentemente è stato introdotto il concetto di spazio quoziente: dato un certo spazio ambiente, quozientandolo rispetto a un suo sottospazio, si ottiene uno spazio che non dipende da tutte le dimensioni del sottospazio rispetto a cui è stata fatta l'operazione.

La domanda alla quale si vuole rispondere ora è: dato un sottospazio quoziente, questo è normato? La risposta è sì, ma solo a patto che il sottospazio rispetto a cui si quozienta sia un **chiuso**.

Formalizziamo meglio tutto ciò: dato X uno spazio normato, X_0 un suo sottospazio chiuso, allora X/X_0 è uno spazio normato con norma:

$$\|[x]\| = \inf_{z \in [x]} \|z\|$$

Si consideri Figura ???: il sottospazio X_0 si può, come già fatto, considerare X_0 come una retta nel piano, e $[x]$ sarà una retta parallela a essa (considerando un'analogia simile a quella di prima).

La norma, essendo stata definita come mostrato, rappresenta la distanza tra la retta X_0 , ossia il sottospazio rispetto a cui si quozienta, e la classe di

equivalenza $[x]$; in verità, sfruttando il trucco precedentemente usato, si può vedere che questa norma è indipendente da x :

$$\|[x]\| = \inf_{y \in X_0} \|x + y\|$$

dunque, questa fa variare solo $y \in X_0$: la distanza va identificata, una volta fissato un x arbitrario, variando y .

Si ha il solito punto da toccare: stiamo continuamente dicendo che questa è una norma, ma lo è veramente? Una proprietà non banale da dimostrare è quella per cui, se $[x] = 0$, allora si ha $x = 0$. Cosa significa ciò in altre parole? Se $[x] = 0$, significa sostanzialmente avere $x \in X_0$: quando si considera la classe con $x = 0$, allora significa che si passa per l'origine, e questa non è altri che la retta X_0 . Questo è il "valore nullo" in X/X_0 : quando si fanno variare le dimensioni in X_0 , ossia nel sottospazio! In altre parole ancora, tutti i vettori sono quelli per cui la differenza con 0 è in X_0 !

Se la norma è 0, il vettore è nullo; qui interviene l'ipotesi di chiusura: se non avessimo la chiusura, non potremmo muoverci su X_0 , poiché x potrebbe non essere in X_0 ! Infatti, se $\|[x]\| = 0$, significa che esiste una successione y_n tale per cui:

$$\|x + y_n\| \rightarrow 0$$

x è fissato, dunque $y_n \rightarrow -x$. Dal momento che però gli y_n appartengono a X_0 , che è chiuso, x , che è un limite, deve appartenere a X_0 ; se non avessimo la chiusura, il limite potrebbe non appartenere a X_0 , dunque potremmo avere rette a distanza nulla ma non coincidenti con X_0 .

Allo stesso modo sarebbe necessario dimostrare la validità della disuguaglianza triangolare, che tuttavia non è banale.

2.3.1 Seminorme e quozienti

Si supponga di avere una funzione $p(x)$ che soddisfi tutte le proprietà delle norme, tranne quella secondo cui il fatto di avere norma nulla implichi che il vettore che si consideri sia il vettore nullo:

$$\|p(x)\| = 0 \not\Rightarrow x = 0$$

In questo caso, la funzione $p(x)$ chiaramente non è una norma, ma è una **seminorma**. Un esempio di seminorma può essere, in \mathbb{R}^2 :

$$p(x, y) = |x|$$

questa norma non dipende da y ; questo vuol dire che, se $|x| = 0$, si deve certamente avere che $x = 0$, ma non che $y = 0$: y può assumere qualsiasi valore.

Il motivo per cui l'esistenza di queste seminorme è rilevante trova radici ancora una volta nella definizione degli spazi quoziente: al fine di liberarsi di questo problema, infatti, la soluzione è quozientare rispetto a un qualche sottospazio; quozientare significa **ignorare** alcune dimensioni; in altre parole, questo ci permette di ignorare, intuitivamente, le dimensioni che non “modificano” la seminorma, ottenendo che, sullo spazio quoziente, la seminorma diviene una norma a tutti gli effetti. Vediamo formalmente cosa significa tutto ciò: data p una seminorma su X , si consideri N la seguente porzione dell'immagine della seminorma:

$$N = \{x \in X : p(x) = 0\}$$

ossia, l'insieme dei valori per cui p è nulla (se p fosse un'applicazione lineare, N sarebbe il kernel di p). Questo sarà lo spazio rispetto a cui quozientare: si quozienta rispetto allo spazio nullo della norma, al fine di ottenere uno spazio in cui si varia “in tutte le altre dimensioni”! Proponiamo alcune osservazioni:

1. si ha che:

$$0 \leq p(\alpha x + \beta y) \leq p(\alpha x) + p(\beta y) = |\alpha| p(x) + |\beta| p(y)$$

2. si calcoli $p(x, y)$, dove x è fissato in X , mentre $y \in N$: questo coincide con l'operazione precedentemente fatta studiano i quozienti. Si ha che $p(x + y)$ è indipendente dalla scelta di $y \in N$: si considerino due punti $y_1, y_2 \in N$, e $x \in X$; si ha che, applicando la disuguaglianza triangolare:

$$p(x + y_1) = p(x + y_1 - y_2 + y_2) \leq p(x + y_2) + p(y_2 - y_1)$$

il secondo termine è 0: $y_1 - y_2 \in N$, dunque va a 0; quindi, si ha che, indipendentemente dal y_1 o y_2 scelti, il risultato è sempre lo stesso.

Questo significa che la seminorma p induce sullo spazio quoziente X/N una norma $\| [x] \| = p(x)$. Questa è ben definita (come si è visto prima, su X ,

poiché ha quasi tutte le proprietà della norma in esso), e in più la proprietà mancante.

Al fine di fissare le idee sulle seminorme, si vuole proporre il seguente esempio: dato lo spazio E definito come:

$$E = \{x : [a, b] \rightarrow \mathbb{R} \text{ funzione integrabile secondo Riemann} \}$$

è possibile definire su di esso una seminorma come segue:

$$p(x) = \int_a^b |x(t)| dt$$

Questa **non** è una norma, dal momento che esistono funzioni (non per forza continue: integrabili!) in cui l'integrale del modulo è nullo, ma la funzione no. Si consideri dunque il sottospazio degli zeri della seminorma, N :

$$N = \{x : p(x) = 0\} = \left\{ x : \int_a^b |x(t)| dt = 0 \right\}$$

Tuttavia, se si quozienta lo spazio E rispetto a N , si ottiene che $||[x]|| = p(x)$ è una **norma**; lo spazio quoziente, come già detto, è uno spazio in cui ciascun elemento è una classe di equivalenza, ossia quell'insieme di x tali per cui la relazione di equivalenza è soddisfatta; mentalmente, in verità, ciò che spesso si fa è considerare non un'intera classe di equivalenza, ma solamente una delle sue rappresentanti.

Per introdurre la prossima sezione, anticipiamo il fatto che questo spazio **non è uno spazio completo**.

2.4 Spazi completi, completamento di spazi normati

È appena stato prodotto un primo esempio di spazio normato non completo; al fine di comprendere esattamente di cosa si sta parlando, si ricorda la definizione di completezza: uno spazio si dice **completo** quando ogni successione di Cauchy è convergente. Uno spazio normato completo è detto **spazio di Banach**, in onore del matematico polacco che nei primi anni del 1900 introdusse la sua definizione. Questi spazi vengono anche detti B-spazi. Consideriamo alcuni esempi di spazi di Banach.

- Lo spazio delle funzioni continue su un intervallo chiuso, con la norma infinito:

$$C([a, b]), \text{ con la norma } \|x\|_\infty = \sup_{t \in [a, b]} \{|x(t)|\}$$

al fine di verificare la completezza di questo spazio, è possibile applicare il teorema delle contrazioni.

- $l^p (1 \leq p \leq \infty)$ sono spazi completi.
- c_0 è completo, con la norma di l^∞ .

Si propone ora un esempio di spazio non completo:

$$C([a, b]), \text{ con la norma } \|x\|_1 = \int_a^b |x(t)| dt$$

prima di tutto, un dettaglio: questa è veramente una norma, dal momento che in questo caso l'ingresso è una funzione **continua**; nel caso precedentemente citato, si consideravano funzioni solamente integrabili. Al fine di verificare questo fatto, è necessario trovare una successione di Cauchy che non sia convergente. Un esempio può essere la successione definita in Figura ??: una successione nulla per $t \leq \frac{1}{2}$, linearmente crescente da $\frac{1}{2}$ a $\frac{1}{2} + \frac{1}{n}$, e pari a 1 oltre questo valore.

Si può dimostrare agevolmente che questa successione è di Cauchy sullo spazio che stiamo considerando: al crescere di n infatti il segmento si impenna sempre di più; se dunque consideriamo:

$$\|x_n - x_m\|_{L^1} < \varepsilon$$

questa norma, ossia la misura della distanza tra le due funzioni, è semplicemente il segmento di area chiuso tra due di queste funzioni nella zona di transizione; se si sceglie dunque una coppia di valori n, m sufficientemente grandi, è possibile ridurre arbitrariamente questa area, riuscendo dunque a maggiorarla a un qualche ε .

Ciò che si deve dimostrare tuttavia è che non esiste una funzione $x(t)$ continua su $[0, 1]$ (per esempio) tale per cui $\|x_n - x\|_{L^1} \rightarrow 0$. In altre parole, essendo la norma definita come un integrale, si deve avere:

$$\int_0^1 |x_n(t) - x(t)| dt \rightarrow 0$$

e non deve esistere una $x(t)$ che soddisfi ciò. Al fine di effettuare questa dimostrazione, si proceda per assurdo: si supponga che questa $x(t)$ esista, se ne studino le caratteristiche e si cerchi di arrivare a una contraddizione. Al fine di fare ciò, si spezzi l'integrale in due pezzi:

$$\int_0^1 |x_n(t) - x(t)| dt \rightarrow \int_0^{\frac{1}{2}} |x_n(t) - x(t)| dt + \int_{\frac{1}{2}}^2 |x_n(t) - x(t)| dt$$

per quanto riguarda il primo pezzo di integrale, quando si ha che

$$\int_0^{\frac{1}{2}} |x_n(t) - x(t)| dt \rightarrow 0?$$

La risposta è: per $x \leq \frac{1}{2}$, si ha che $x_n(t) = 0$; quindi:

$$\int_0^{\frac{1}{2}} |x_n(t) - x(t)| dt = \int_0^{\frac{1}{2}} |x(t)| dt =$$

e dunque, è $x(t)$ che deve essere 0; $x(t)$ è una funzione non dipendente da n , dal momento che appartiene allo spazio, ma è una funzione, non una successione; essendo inoltre essa appartenente allo spazio delle funzioni continue, essa deve essere una funzione continua. L'unico esempio di funzione continua nulla è la funzione costante e uguale a 0:

$$x(t) = 0, x \in \left[0, \frac{1}{2}\right)$$

Consideriamo ora la seconda parte dell'integrale: $\forall \delta > 0$, si deve avere:

$$\int_{\frac{1}{2}+\delta}^2 |x_n(t) - x(t)| dt \leq \int_0^1 |x_n(t) - x(t)| dt \rightarrow 0$$

se n è sufficientemente grande, è possibile considerare di essere sulla parte successiva alla parte lineare di $x_n(t)$, dunque di avere $x_n(t) = 1$; questo significa che si ha:

$$= \int_{\frac{1}{2}+\delta}^1 |1 - x(t)| dt = 0$$

ma $x(t)$ è una funzione continua, indipendente da n e, dunque, per avere questo integrale uguale a 0, $x(t)$ deve valere 1 su tutto $(\frac{1}{2}, 1]$. Dunque, $x(t)$ deve essere continua, ma valere identicamente 0 in metà intervallo, e identicamente 1 nella seconda metà: questo non è possibile!

Abbiamo appena dimostrato che ci sono successioni di Cauchy che non convergono, dunque lo spazio **non è completo**. È possibile dimostrare che un altro esempio di spazio non completo è lo spazio delle successioni definitivamente nulle, con la norma di l^∞ ; anche in questo caso, al fine di effettuare la dimostrazione, è possibile cercare una successione di Cauchy che non converga.

2.4.1 Completezza su spazi normati

Si vuole a questo punto proporre un risultato riguardante la completezza che valga su spazi normati, ma non in generico su spazi metrici, in cui non si ha una struttura vettoriale. Si ha il seguente teorema: dato X uno spazio normato, allora X è di Banach se e solo se ogni serie assolutamente convergente in X è convergente in X .

Rivediamo alcune definizioni, per comprendere meglio questo risultato: una serie

$$\sum_{k=1}^{\infty} x_k \in X$$

è **assolutamente convergente** se è convergente la **serie delle norme**:

$$\sum_{k=1}^{\infty} \|x_k\| \in \mathbb{R}$$

questa cosa ha senso esclusivamente nell'ambito degli spazi normati, ossia a struttura vettoriale: in spazi metrici non vettoriali, infatti, non è possibile calcolare la somma di due elementi, dunque queste considerazioni non hanno senso. In \mathbb{R} questo risultato si applica sui moduli, invece che sulle norme. Questo risultato è la generalizzazione del fatto che, in \mathbb{R} , una serie assolutamente convergente è anche convergente; in questo caso si è semplicemente detto che ciò vale per uno spazio di Banach.

Vogliamo dimostrare una delle implicazioni: sia

$$\sum_{k=1}^{\infty} x_k$$

una serie assolutamente convergente:

$$\sum_{k=1}^{\infty} \|x_k\| \text{ converge}$$

indichiamo a questo punto con S_n la ridotta n -esima della serie dei vettori, e con s_n la ridotta delle norme:

$$S_n = \sum_{k=1}^n x_k$$

$$s_n = \sum_{k=1}^n \|x_k\|$$

essendo lo spazio completo, ogni successione di Cauchy è convergente; di conseguenza, è sufficiente dimostrare che la ridotta S_n sia una successione di Cauchy, al fine di poter dire che converge. Si consideri, con $n > m$, che:

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^n x_k \right\| \leq \sum_{k=m+1}^n \|x_k\| = s_n - s_m$$

Considerando la norma della differenza delle ridotte, si può applicare la disuguaglianza triangolare, portando il segno di norma fuori dalla sommatoria; quindi, si ottiene la somma delle norme, che è di fatto la differenza delle due sommatorie, $s_n - s_m$ (infatti, $S_n - S_m$ era data dalla somma di tutti gli elementi da 0 a n , meno la somma di tutti gli elementi da 0 a m , dunque si ottengono solo gli elementi da m a n , e così via anche per il secondo passaggio). Essendo

$$\sum_{k=1}^{\infty} \|x_k\|$$

convergente per ipotesi, dunque s_n è di Cauchy, ma dunque per n, m abbastanza grandi, si ha che:

$$\forall \varepsilon \exists N : \forall n > m > N, \|S_n - S_m\| \leq s_n - s_m < \varepsilon$$

Essendo dunque maggiorata $\|S_n - S_m\|$ da una differenza di elementi di una successione di Cauchy, allora S_n è una successione di Cauchy, dunque, per completezza, convergente!

Si consideri dunque la seguente osservazione: se X è uno spazio di Banach, allora la convergenza assoluta implica la convergenza! In verità, per essere precisi, quella che è implicata è la **convergenza incondizionata**:

$$\sum_{k=1}^{\infty} x_{\sigma(k)}$$

ossia, questa serie, indipendentemente dal fatto che si abbia una permutazione $\sigma(k)$ degli indici, converge!

$$\sigma : \mathbb{N} \setminus \{\emptyset\} \rightarrow \mathbb{N} \setminus \{\emptyset\}$$

è una biiezione. Tutto ciò è un modo complicato di dire che vale la proprietà commutativa.

Si noti che, per quanto riguarda la relazione tra convergenza incondizionata e convergenza assoluta, vale anche il viceversa, ossia il fatto che la convergenza incondizionata implica la convergenza assoluta, se la dimensione dello spazio normato X che si sta considerando è finita.

Riguardo a ciò, esiste un teorema di Riemann che afferma che, fissato un numero reale, è possibile riordinare i termini della serie in maniera tale da farla convergere a questo valore; allo stesso modo, è possibile riordinarli in maniera tale da farla divergere, se prima convergeva. Questo si basa sul fatto che, se si considerano una serie a segni alterni, considerando separatamente la parte della serie composta da soli valori positivi, e quella composta da soli valori negativi, esse non convergono; sommando i positivi e i negativi, si ottiene un certo equilibrio, che poi permette la convergenza. Si noti che in questo caso si parla di **convergenza semplice**.

2.4.2 Teorema di completamento

Si è parlato di spazi completi, e di spazi che possono essere non completi. Uno spazio non completo, tuttavia, può essere **completato**, ossia è possibile **immergerlo** in uno spazio completo, che *copra i buchi rimasti*. Se si ha un'isometria densa con uno spazio completo, è possibile appunto ottenere un completamento per uno spazio.

Si consideri dunque uno spazio normato X ; esiste uno spazio \hat{X} di Banach, e un'applicazione lineare $T : X \rightarrow \hat{X}$ che è una **isometria** (ossia, che conserva la metrica):

$$\|Tx\| = \|x\| \quad \forall x \in X$$

e l'immagine Tx è densa in \hat{X} . A queste condizioni, \hat{X} è detto **completamento** di X .

Non si propone una dimostrazione completa, sapendo che comunque essa ricorda la dimostrazione che si deve applicare per costruire l'insieme dei numeri reali a partire dai razionali. Si proporrà solo un cenno di dimostrazione. Si consideri uno spazio \mathcal{E} , definito come:

$$\mathcal{E} = \{x = (x_k) \text{ spazio delle successioni di Cauchy in } X\}$$

su questo, dunque, si definisce una seminorma p :

$$p : \mathcal{E} \rightarrow \mathbb{R}$$

tale per cui, per $x \in \mathcal{E}$, si ha:

$$p(x) = \lim_{n \rightarrow \infty} \|x_n\|$$

A questo punto, si considera un punto chiave della dimostrazione: x è una successione di Cauchy, della quale l'operazione di seminorma calcola il limite. Il punto fondamentale è il fatto che questo limite **sicuramente esiste**: stiamo infatti considerando una successione di Cauchy in \mathbb{R} , dal momento che stiamo lavorando sulle norme; di conseguenza, si ha che:

$$\left| \|x_n\| - \|x_m\| \right| \leq \|x_n - x_m\|$$

dunque, x_n è di Cauchy, ma siamo in \mathbb{R} , dunque essa è **convergente**. X non è uno spazio completo, ma x_n è di Cauchy su esso, e la successione delle sue norme è di Cauchy (essendo essa di Cauchy, come abbiamo appena dimostrato), dunque in \mathbb{R} la successione delle norme è di Cauchy, dunque convergente.

Questa funzione è una seminorma. Al fine di cercare uno spazio in cui essa è una norma a tutti gli effetti, dato N lo spazio nullo della pseudonorma:

$$N = \{x \in \mathcal{E} : p(x) = 0\}$$

questo è lo spazio delle successioni convergenti a 0! Infatti, la pseudonorma si valuta come il limite di una successione, dunque essa va a 0 se il limite va a 0, ma se il limite va a 0 significa che la successione converge a 0. Si definisce ora \hat{X} come il quoziente dello spazio \mathcal{E} rispetto a N :

$$\hat{X} = \mathcal{E}/N$$

A questo punto, dal momento che \hat{X} è definito come spazio quoziente, esso è uno spazio di classi di equivalenza tra successioni di Cauchy! Dunque:

$$\hat{X} = \mathcal{E}/N = \{[x] : x \in \mathcal{E}\}$$

dunque, si ha che:

$$[x] = [y], \iff x - y \in N, \iff x_n - y_n \rightarrow 0 \text{ in } X$$

Quindi, \hat{X} è definito, in altre parole, come il quoziente rispetto al caso in cui la differenza delle successioni di Cauchy, al limite, va a 0. Ora, ossia su \hat{X} ,

$$|[x]| = p(x) = \lim_{n \rightarrow \infty} x_n$$

è una **norma**.

Abbiamo costruito una norma per lo spazio; un altro ingrediente del quale si parlava nel teorema, era l'immersione, ossia l'isometria che permetteva di passare da uno spazio all'altro, T . Un modo per costruire T è quello di mappare l'elemento x in una classe di equivalenza. Un esempio di queste isometrie è:

$$Tx = [(x, x, x, \dots)]$$

ossia, si mappa x in una successione costante. A questo punto, verifichiamo che essa sia un'isometria:

$$|Tx|_{\hat{X}} = \lim_{n \rightarrow \infty} x_n$$

l'immagine Tx ha come rappresentante la (x, x, x, \dots) ; dunque:

$$|Tx| = \lim_{n \rightarrow \infty} |x| = |x|$$

dunque, effettivamente, applicando T a x e calcolando la norma, questa si mantiene uguale: è effettivamente un'isometria. A questo punto, per completare la dimostrazione del teorema, sarebbe necessario dimostrare che \hat{X} è uno spazio completo, e che Tx è denso in \hat{X} .

A questo punto, si dispone di un teorema che permette di immergere in uno spazio più vasto un certo spazio (a patto che quest'ultimo sia denso in quello *più grande*); il primo esempio che ci piacerebbe risolvere dunque potrebbe essere il completamento di \mathbb{Q} con \mathbb{R} ; questo, tuttavia, **non si può fare**: si sfrutta infatti nella dimostrazione l'ipotesi di completezza di \mathbb{R} (che si dimostra in altra maniera), al fine di usare il risultato con le norme, come visto precedentemente; di conseguenza, si avrebbe come un *cane che si morde la coda*.

Si è detto che la norma appena definita per lo spazio \mathcal{L}^p in realtà è solo una seminorma. Al fine di completare il discorso e ottenere dunque una norma, è necessario quozientare come segue. Si definisca prima di tutto $A \subset \mathbb{R}$ un insieme a misura nulla nel senso di Lebesgue se, $\forall \varepsilon > 0$, data I_n una successione di intervalli,

$$I_n = [a_n, b_n]$$

tali per cui:

$$A \subset \bigcup_n I_n$$

si ha che:

$$\sum_n (b_n - a_n) < \varepsilon$$

questa è la definizione di insieme a misura nulla. Dunque, per completare la definizione, si definiscono equivalenti, ossia $x \sim y$, quando esse sono uguali quasi ovunque, eccetto che su un insieme di misura nulla; si definisce dunque una classe di equivalenza $[x]$ data dall'insieme delle funzioni y equivalenti in questo senso alle x , e quindi troviamo N come lo spazio delle funzioni non uguali (quello di misura nulla); dunque, si trova \mathcal{L}^p/N , e in questa maniera, su questo spazio, si avranno delle norme.

2.5 Teorema di approssimazione di Weierstrass

Si vuole a questo punto introdurre un teorema di approssimazione: esso afferma che ogni funzione continua su un intervallo limitato può essere approssimata arbitrariamente bene mediante dei polinomi. Verrà dunque enunciato questo problema in alcuni modi, tutti equivalenti tra loro.

- Nello spazio delle funzioni continue $C([a, b])$ con

$$\|x\| = \sup_{t \in [a, b]} |x(t)|$$

il sottospazio delle funzioni polinomiali è denso.

- Dato $P(t)$ polinomio,

$$\forall x \in C([a, b]) \forall \varepsilon > 0 \exists P(t) : \|x - P(t)\| = \sup_{t \in [a, b]} |x(t) - P(t)| < \varepsilon$$

- Data $\{P_n(t)\}$ successione di polinomi,

$$\forall x \in C([a, b]) \exists \{P_n(t)\} : P_n(t) \rightarrow x(t), \text{ uniformemente su } [a, b]$$

La dimostrazione di questo teorema non è immediata dal momento che, per proporla, è necessario prima utilizzare un lemma, che verrà ora proposto.

2.5.1 Lemma dell'identità approssimata

Prima di tutto, date due funzioni $f(x)$ e $g(x)$, si ricorda che la loro convoluzione è data da:

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(x-y)g(y) dy = \int_{-\infty}^{+\infty} f(y)g(x-y) dy$$

In secondo luogo, si definisce una **identità approssimata** come segue. Date $K_n(x)$ delle funzioni integrabili su \mathbb{R} , tali da soddisfare le seguenti ipotesi, esse sono dette **identità approssimate**:

- essere sempre maggiore o uguale a 0

$$K_n(x) \geq 0 \forall x$$

- avere integrale sull'intero dominio pari a 1

$$\int_{-\infty}^{+\infty} K_n(y) dy = 1 \forall n$$

- avere integrale che tende ad annullarsi, per valori di y elevati:

$$\forall \delta > 0, \int_{|y| > \delta} K_n(y) dy \rightarrow 0, \quad n \rightarrow \infty$$

Dunque, se f è una funzione $C_c(\mathbb{R})$ (dove con questo simbolo si intende **continua e a supporto compatto**), ossia per cui

$$\exists R > 0 : f(x) = 0, |x| > R$$

allora, si ha che:

$$(K_n * f)(x) \rightarrow f$$

uniformemente su \mathbb{R} , per $n \rightarrow \infty$.

L'identità è approssimata tramite la convoluzione; questo significa, in altre parole, che K_n è l'elemento neutro per la convoluzione. Se si sceglie dunque un K_n opportuno, si può scegliere quale approssimante utilizzare per la funzione; dunque, è possibile fare in modo da ottenere, come si vedrà a breve, i polinomi. Prima di ciò, tuttavia, si dimostrerà questo lemma.

2.5.2 Dimostrazione del lemma

Il lemma permette semplicemente di avere garanzie per quanto concerne la convergenza di $K_n * f$ a f ; questo significa che, dunque, è possibile studiare la differenza di questi due elementi, non appena n supera una certa soglia. Si ha dunque:

$$|(K_n * f)(x) - f(x)| = \left| \int_{-\infty}^{+\infty} K_n(y) f(x-y) dy - f(x) \right|$$

A questo punto, una considerazione, dal momento che f è una funzione a supporto compatto, all'interno del segno di integrale f è non nulla solamente quando y è prossimo a x ; sfruttando la seconda proprietà delle identità approssimate, si ha che:

$$f(x) = \int_{-\infty}^{+\infty} K_n(y)f(x)dy$$

infatti, $f(x)$ è in una diversa variabile di integrazione. Ora si può riscrivere ciò che è contenuto nel valore assoluto come segue:

$$\left| \int_{-\infty}^{+\infty} K_n(y)f(x-y)dy - f(x) \right| = \left| \int_{-\infty}^{+\infty} K_n(y)[f(x-y) - f(x)]dy \right|$$

A questo punto, se y è piccolo, si ha che $f(x-y) \simeq f(x)$; l'integrale di K_n inoltre va a 1 (dunque non cresce eccessivamente); f è una funzione continua a supporto compatto, dunque per il teorema di Weierstrass (non quello che stiamo per dimostrare ovviamente!) essa è limitata; ciò che si può dunque fare è spezzare l'integrale in due, e fare alcune considerazioni su ciascun pezzo.

$$= \int_{|y| \leq \delta} K_n(y) |f(x-y) - f(x)| dy + \int_{|y| > \delta} K_n(y) |f(x-y) - f(x)| dy$$

A questo punto: per quanto riguarda il primo contributo, ossia quello per cui y è piccolo, si può dire che:

$$|f(x-y) - f(x)| < \varepsilon$$

grazie al fatto che la funzione è continua, e se l'incremento y rispetto a x è sufficientemente piccolo.

Potrei applicar Potrei applicar Heine-Cantor, ché la funzione è continua in un compatto, e ciò la rende immantinente continua uniformemente (non potrei farlo se $[a, b]$ fosse aperto).

Data dunque $f \in C_c$, si ha che:

$$|f(x-y) - f(x)| < \varepsilon \quad \text{se } |y| < \delta, \quad x \in \mathbb{R}$$

Per quanto riguarda la parte per $\delta > 0$, si sa che esiste un certo $n > N$ tale per cui:

$$\int_{|y|>\delta} K_n(y) dy < \varepsilon$$

Dato ε dunque si sceglie un certo δ , dato δ si sceglie una soglia N ; con δ si è spezzato l'integrale, e si fa in modo da vedere che in un caso la differenza è maggiorabile con ε ; nel secondo caso, si ha comunque la limitatezza, grazie a Heine-Cantor, quindi:

$$|f(x - y) - f(x)| \leq M$$

dove M è una certa costante. Quindi, si possono maggiorare i due integrali come segue:

$$\leq \varepsilon \int_{|y|\leq\delta} K_n(y) dy + M \int_{|y|>\delta} K_n(y) dy$$

il contenuto dell'integrale, per la terza ipotesi sulle identità approssimate, è minorabile con ε , per $n > N$; per la prima di queste proprietà, invece, il primo integrale va a 1; di conseguenza:

$$\leq \varepsilon + M\varepsilon = (M + 1)\varepsilon$$

dunque, essendo $M + 1$ una costante, è possibile approssimare bene quanto si vuole f con la K_n .

Il fatto che la funzione sia a supporto compatto è in verità una ipotesi sovrabbondante: in verità, è sufficiente che la funzione f sia uniformemente continua e limitata.

2.5.3 Dimostrazione del teorema di Weierstrass

Si considerino funzioni continue, $f \in C([0, 1])$, invece che sul generico intervallo $[a, b]$; la dimostrazione verrà svolta su funzioni in questo intervallo, invece che in quello generico. Ciò è possibile poiché, data $g \in C([a, b])$, è possibile definire:

$$x = a + t(b - a), \quad t \in [0, 1]$$

ossia, un mapping lineare tra un intervallo e l'altro; di conseguenza:

$$g[a + t(b - a)] = G(t)$$

è una funzione su $t \in [0, 1]$. Allo stesso modo, se $P_n(t) \rightarrow G(t)$, uniformemente, su $t \in [0, 1]$, si ha che, usando il mapping inverso:

$$P_n \left(\frac{x-a}{b-a} \right) \rightarrow G \left(\frac{x-a}{b-a} \right) = g(x)$$

è ancora un polinomio: infatti la x sta solo al numeratore!

Supponiamo inoltre che $f \in C([0, 1])$ soddisfi la condizione $f(0) = f(1) = 0$; questo potrebbe sembrare limitante, ma non è così, dal momento che è possibile fare ciò:

$$g(x) = g(x) + \alpha x + \beta - (\alpha x + \beta)$$

aggiungendo un certo $\alpha x + \beta$, si riesce a ottenere una funzione leggermente diversa da $g(x)$, ma ancora una volta riconducibile a essa, e soprattutto approssimabile dal teorema: si è aggiunto un solo termine polinomiale, che dunque può essere senza fatica approssimato mediante polinomi! Se α, β sono tali per cui la funzione $g(x) + \alpha x + \beta$ si annulli per $x = 0$ e $x = 1$, allora è risultati saranno ancora validi, aggiungendo a posteriori il termine.

Questo risultato serve per poter ottenere una funzione a dominio compatto, che si annulli agli estremi, e che dunque sia estensibile per continuità sull'intero \mathbb{R} : si ha infatti un raccordo continuo con lo zero.

A questo punto, si applichi il lemma dell'identità approssimata, usando la seguente funzione $K_n(x)$:

$$K_n(x) = \begin{cases} (1-x^2)^n a_n & \text{se } |x| \leq 1 \\ 0 & \text{altrove} \end{cases} \quad (2.1)$$

la successione di coefficienti a_n ha l'utilità di soddisfare la proprietà di normalizzazione delle identità approssimate; di conseguenza, esso si definisce come l'integrale:

$$a_n = \left(\int_{-1}^{+1} (1-x^2)^n dx \right)^{-1}$$

Dunque, le tre condizioni che fanno di $K_n(x)$ un'identità approssimata sono soddisfatte. Se $|\delta| > 1$, l'integrale è esattamente pari a 0; dobbiamo ora dimostrare che, per $|\delta| < 1$, esso è comunque tendente a 0. Per fare ciò, si consideri la Figura ???. L'integrale

$$\int_{|x| \leq \delta} K_n(x) dx$$

è dato dall'area delle due code. Questo si può migliorare considerando che, per $x = \pm\delta$, l'ordinata della funzione è:

$$(1 - \delta^2)^n a_n$$

dunque, scegliendo come base l'intero supporto della funzione (che sarà $[-1, +1]$), si ha che si può migliorare con:

$$\int_{|x| \leq \delta} K_n(x) dx \leq 2(1 - \delta^2)^n a_n$$

dove il 2 è $1 - (-1)$: la base del rettangolo.

A questo punto, rimane da stimare a_n : il calcolo analitico potrebbe essere effettuato, ma è complicato. Si usa un trucco simile a quello appena applicato, però per minorare un integrale (essendo a_n il reciproco di un integrale): considerando Figura ??, si consideri x la quota tale per cui la funzione vale $1/2$:

$$(1 - x_0^2)^n = \frac{1}{2} \implies x_0 = \left(1 - \frac{1}{2^{1/n}}\right)^{\frac{1}{2}}$$

è possibile minorare il reciproco di a_n con il rettangolo alto un mezzo e di quota $\pm x_0$: l'area chiusa da questo sarà senza dubbio più piccola dell'area sottesa dalla $K_n(x)$. Dunque, in questo caso, si può dire che:

$$\int_{-1}^1 (1 - x^2)^n dx \geq \frac{1}{2} 2 \left(1 - \frac{1}{2^{1/n}}\right)^{\frac{1}{2}} = \frac{2^{\frac{1}{n}} - 1}{2^{\frac{1}{2n}}}$$

Ora che abbiamo ottenuto questo risultato, è possibile considerare lo sviluppo asintotico di questa espressione per $n \rightarrow \infty$:

$$a^x - 1 \sim \ln(a)$$

dunque:

$$\rightarrow \sim \left(\frac{\log(2)}{n}\right)^{\frac{1}{2}}, n \rightarrow \infty$$

Dunque, essendo a_n l'inverso dell'integrale, si ha che:

$$a_n \leq b_n \sim \left(\frac{n}{\log 2} \right)^{\frac{1}{2}}$$

ossia, in sostanza, $b_n \sim n^{\frac{1}{2}}$. Questo serve per verificare che b_n , e dunque a_n , non disturbi la convergenza: non la influenza. Abbiamo dunque dimostrato che:

$$(K_n * f) \rightarrow f$$

uniformemente, su \mathbb{R} .

A questo punto, è necessario verificare che l'approssimante, ossia $(K_n * f)$, sia un polinomio, quando è ristretto a $[0, 1]$. Per fare ciò, si consideri quanto segue:

$$(K_n * f)(x) = \int_{-\infty}^{+\infty} K_n(x-y)f(y)dy$$

ora: $f(y)$ ha supporto in $[0, 1]$; se si fissa $x \in [0, 1]$, $y \in [0, 1]$; si ha che $x - y \in [-1, 1]$; è dunque possibile scrivere che l'espressione coincide con:

$$= a_n \int_0^1 [1 - (x-y)^2]^n f(y)dy$$

a questo punto, si dovrebbe scrivere, mediante il binomio di Newton, $[1 - (x-y)^2]^n$; si otterrebbe un'espressione del tipo:

$$[1 - (x-y)^2]^n = \sum_{h,k} c_{hk} x^h y^k$$

dove i c_{hk} sono i coefficienti del binomio di Newton. Dunque, è possibile sostituire questa nell'integrale:

$$= a_n \sum_{h,k} c_{hk} x^h \int_0^1 y^k f(y) dy$$

l'integrale ritorna un numero; il termine che lo moltiplica è un polinomio; di conseguenza, il risultato finale è un polinomio.

Esistono delle vie alternative di applicare il teorema, o anche solo di dimostrarlo; un esempio è applicare il **nucleo di Weierstrass**, ossia la $K_n(y)$ che veniva originariamente usata da Weierstrass:

$$K_n(y) = \frac{n}{\sqrt{\pi}} e^{-(nx)^2}$$

con questo, il risultato che si ottiene non è un polinomio; ciò che si ottiene, però, sarà analitico, di conseguenza espansibile mediante serie di Taylor; troncando, si ottiene un polinomio.

2.5.4 Cenni a variante del teorema di Weierstrass

Data una funzione continua su \mathbb{R} , e periodica di periodo 2π , essa è approssimabile uniformemente mediante **polinomi trigonometrici**, ossia mediante:

$$\sum_{k=-N}^N c_k e^{jkt}$$

Questo teorema afferma che, dato $C_{\text{per}}([0, 2\pi])$ lo spazio delle funzioni f continue su $[0, 2\pi]$ con $f(0) = f(2\pi)$, con la norma del sup, il sottospazio dei polinomi trigonometrici è denso.

La dimostrazione di questo teorema è analoga, dal momento che e^{jkt} assume gli stessi valori agli estremi, dunque, applicando la convoluzione periodica al posto di quella “classica”, è possibile trarre all’incirca le stesse considerazioni appena fatte.

Capitolo 3

Spazi con prodotto scalare

In questo capitolo verrà descritta quella che da un lato è la generalizzazione del concetto di spazio euclideo, e dall'altro è un'ulteriore specializzazione rispetto agli spazi metrici e normati: uno spazio euclideo è dotato di tre proprietà:

- esistenza di una metrica;
- struttura vettoriale;
- esistenza di un prodotto scalare.

Negli spazi normati appena descritti le prime due caratteristiche sono presenti, mentre la terza no. Al fine di procedere, si parte dunque con le definizioni del caso.

3.1 Definizione di prodotto scalare

Dato H uno spazio vettoriale su \mathbb{C} (in verità sarebbe possibile anche su \mathbb{R} , però non è nei casi di interesse). Un prodotto scalare è una applicazione da $H \times H \rightarrow \mathbb{C}$ ($|\cdot$) che rispetta le seguenti proprietà:

- linearità rispetto alla prima componente:

$$(\alpha x_1 + \beta x_2 | y) = \alpha(x_1 | y) + \beta(x_2 | y)$$

- hermitianità:

$$(y|x) = \overline{(x|y)}$$

- positiva definizione:

$$(x|x) \geq 0 \in \mathbb{R}, \quad (x|x) = 0 \iff x = 0$$

Da queste proprietà segue la proprietà di antilinearità:

$$(x|\alpha y_1 + \beta y_2) = \bar{\alpha}(x|y_1) + \bar{\beta}(x|y_2)$$

Una applicazione che segue quest'ultima proprietà e la prima è detta **se-squilineare**: la presenza dei coniugati infatti rende le cose diverse rispetto al caso in cui si ha la linearità in entrambi i casi: **bilinearità**.

Verranno a questo punto proposti alcuni esempi di prodotto scalare in vari spazi.

- dati $x, y \in \mathbb{R}^n$,

$$(x|y) = \sum_{k=1}^n x_k y_k$$

- dati $x, y \in \mathbb{C}^n$,

$$(x|y) = \sum_{k=1}^n x_k \bar{y}_k$$

- dati $x, y \in l^2$,

$$(x|y) = \sum_{k=1}^{\infty} x_k \bar{y}_k < \infty$$

dove si ha la **convergenza assoluta**; la convergenza assoluta deriva dalla disuguaglianza di Hölder:

$$\sum_{k=1}^{\infty} |x_k \bar{y}_k| \leq \|x\|_{l^2} \|y\|_{l^2}$$

questa, converge; dunque:

$$|(x|y)| \leq \|x\|_{L^2} \|y\|_{L^2}$$

- in $C([a, b])$, si prende:

$$(x|y) = \int_a^b x(t)\bar{y}(t) dt$$

si può applicare la Hölder anche in questo caso, e ottenere un risultato analogo a quello appena presentato per gli spazi di successioni.

- in $L^2[a, b]$, che è il completamento di $C([a, b])$ con la norma $\|\cdot\|_{L^2}$, si ha:

$$(x|y) = \int_a^b x(t)\bar{y}(t) dt$$

per ottenere ciò, ci sono sostanzialmente due vie, a partire dallo spazio delle funzioni continue:

- che il prodotto $x(t)\bar{y}(t)$ sia appartenente a L^1 ;
- estendere per continuità il prodotto $x(t)\bar{y}(t)$, dal momento che lo spazio delle funzioni continue è denso in quello delle funzioni L^2 .

Solitamente, nell'ambito dell'Analisi Funzionale, si considerano come spazi di lavoro l^2 per le successioni, L^2 per le funzioni, dal momento che sono spazi completi, e dunque più interessanti.

3.1.1 Diseguaglianza di Cauchy-Schwarz

In generale, è possibile mostrare che, se

$$\|x\| = (x, x)^{\frac{1}{2}}$$

questa è una norma, sempre. Al fine di verificare ciò, è necessario verificare le proprietà della norma; a questo scopo, è necessario introdurre la **diseguaglianza di Cauchy-Schwarz**. Si definisca (direttamente col simbolo di norma, anche se stiamo dimostrando che questa è effettivamente una norma):

$$\|x\| = (x|x)^{\frac{1}{2}}$$

in uno spazio H dotato di prodotto scalare, vale la seguente disuguaglianza:

$$|(x|y)| \leq \|x\| \|y\|$$

Dimostriamo questo fatto. Dato $\lambda \in \mathbb{R}$, si ha che:

$$0 \leq \|x + \lambda y\|^2$$

si sviluppi ciò, utilizzando la definizione appena introdotta:

$$\|x + \lambda y\|^2 = (x + \lambda y|x + \lambda y) = (x|x) + \lambda(x|y) + \lambda(y|x) + \lambda^2(y|y)$$

questo è:

$$= \|x\|^2 + 2\lambda \operatorname{Re}\{(x|y)\} + \lambda^2 \|y\|^2$$

Questa ultima espressione può essere vista come il quadrato di un trinomio nella variabile λ ; al fine di avere la funzione sempre maggiore di zero, non si devono avere soluzioni reali, dunque è necessario richiedere che il discriminante di questo trinomio sia minore di zero:

$$\operatorname{Re}\{(x|y)\}^2 - \|x\|^2 \|y\|^2 < 0$$

da cui:

$$|\operatorname{Re}\{(x|y)\}| < \|x\| \|y\|$$

Purtroppo questo risultato non è esattamente quello che ci aspettavamo di trovare: dal momento che non è detto che si lavori sempre in spazi reali, non è detto che si abbia a che fare con la sola parte reale; questa disuguaglianza dunque non verifica immediatamente quella che vogliamo ottenere. Al fine di ottenere esattamente la relazione che desideriamo, si consideri il seguente trucco: dati $x, y \in H$, il prodotto scalare $(x|y)$ è uno scalare che può non essere né reale né positivo; tuttavia, dato $\vartheta \in \mathbb{R}$, è possibile fare in modo che:

$$e^{j\vartheta}(x|y) > 0$$

ossia, trovare l'angolo nel piano complesso, ϑ , tale da “ruotare” il prodotto scalare fino a renderlo reale e positivo. Dunque, è possibile fare le seguenti considerazioni:

$$|(x|y)| = |e^{j\vartheta}(x|y)| = |(e^{j\vartheta}x|y)| = (e^{j\vartheta}x|y)$$

il modulo è stato rimosso nell'ultimo passaggio poiché, grazie alla rotazione, esso è divenuto superfluo. Dunque, si può dire che, ora:

$$|\Re \{(e^{j\vartheta}x|y)\}| = (e^{j\vartheta}x|y) \leq \|e^{j\vartheta}x\| \|y\|$$

Tuttavia:

$$\|e^{j\vartheta}x\| = (e^{j\vartheta}x|e^{j\vartheta}x) = e^{j\vartheta} e^{-j\vartheta}(x|x) = (x|x)$$

da cui:

$$\|e^{j\vartheta}x\| \|y\| = \|x\| \|y\|$$

Questo dimostra la disuguaglianza di Cauchy-Schwarz.

È ora possibile recuperare questo risultato al fine di dimostrare che

$$(x, x)^{\frac{1}{2}} = \|x\|$$

ossia, che è effettivamente una norma. Le prime due proprietà sono semplici, dunque non verranno riprese; quella più complicata è la disuguaglianza triangolare, che si può dimostrare come segue:

$$\|x + y\|^2 = (x + y|x + y) = \|x\|^2 + 2\Re \{(x|y)\} + \|y\|^2$$

ma è noto che, per un numero complesso z , si ha che:

$$\Re \{z\} \leq |\Re \{z\}| \leq |z|$$

e, applicando la disuguaglianza di Cauchy-Schwarz, si ha:

$$\|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

da cui dunque:

$$\|x + y\| \leq \|x\| + \|y\|$$

3.2 Ortogonalità

Una proprietà degli spazi di Hilbert che stiamo studiando in questo capitolo ma non degli spazi normati precedentemente introdotti è la **ortogonalità**. Due vettori $x, y \in H$ sono detti **ortogonali** se:

$$(x|y) = 0$$

Questo risultato ha diverse conseguenze, che verranno ora rapidamente proposte e dimostrate.

3.2.1 Continuità del prodotto scalare

Il prodotto scalare è continuo nei due argomenti. In altre parole, dato $x_n \rightarrow x, y_n \rightarrow y$, si ha che:

$$(x_n|y_n) \rightarrow (x|y)$$

Dimostriamo questo risultato: sommando e sottraendo, è possibile ottenere la seguente disuguaglianza:

$$\begin{aligned} |(x_n|y_n) - (x|y)| &= |(x_n - x|y_n) + (x|y_n - y)| \leq |(x_n - x|y_n)| + |(x|y_n - y)| \leq \\ &\leq \|x_n - x\| \|y_n\| + \|x\| \|y_n - y\| \end{aligned}$$

dal momento che, inoltre, $\|y_n\| \rightarrow \|y\|$, che è costante, e allo stesso modo $\|x_n\| \rightarrow \|x\|$, si ha che, dal momento che le norme delle differenze tendono a 0, la continuità è verificata.

3.2.2 Identità del parallelogramma

Un'altra proprietà che deriva dall'ortogonalità è la cosiddetta **identità del parallelogramma**, la cui espressione ricorda quella che si ha nel caso di spazi euclidei meno generici:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

Questa dimostrazione si può fare analogamente a quella di prima: si tira fuori anche in questo caso una parte reale, questa volta moltiplicata per 2, e in questo modo si effettuano i trucchi appena mostrati.

Questa identità è molto utile, dal momento che può essere utilizzata per verificare che la norma di certi spazi normati **non** proviene da un prodotto scalare. Per esempio, ci si può porre una domanda: la norma l^p può provenire da un prodotto scalare? La risposta è **no**: non esiste un prodotto scalare in l^p , $p \neq 2$, che induca la norma l^p :

$$\|\cdot\|_{l^p} \quad \|x\|_{l^p} = (x|x)^{\frac{1}{2}}$$

Se la norma esistesse, dovrebbe infatti soddisfare questa condizione. Al fine di vedere se ciò è possibile, si deve verificare l'identità del parallelogramma. Si consideri dunque:

$$x = (1, 1, 0, 0, \dots)$$

$$y = (1, -1, 0, 0, \dots)$$

dunque:

$$\|x + y\|_p = 2$$

$$\|x - y\|_p = 2$$

essendo in entrambi i casi $(2^p)^{\frac{1}{p}}$. Invece:

$$\|x\|_p = ((1^p) + (1^p))^{\frac{1}{p}} = 2^{\frac{1}{p}}$$

e idem:

$$\|y\|_p = ((1^p) + (1^p))^{\frac{1}{p}} = 2^{\frac{1}{p}}$$

Dunque,

$$2^2 + 2^2 = 4(2^{\frac{2}{p}} + 2^{\frac{2}{p}})$$

solo se $p = 2$.

In verità (ma questo è molto più complicato da dimostrare), si ha il seguente risultato: se in un generico spazio vale l'identità del parallelogramma, allora la norma viene indotta da un prodotto scalare: vale dunque anche il viceversa. Questo significa che questo è un criterio che permette di caratterizzare le proprietà delle norme.

3.2.3 Teorema di Pitagora e sistemi ortonormali

Esiste una generalizzazione del teorema di Pitagora per spazi di Hilbert. Dati x, y vettori ortogonali, si ha

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2$$

La dimostrazione è semplice:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\{(x|y)\}$$

ma per ipotesi x e y sono ortogonali, dunque il prodotto scalare è nullo.

Questo teorema può essere esteso a un numero indefinito di vettori appartenenti allo spazio di Hilbert: dato $\{e_k\}_{k=1}^n$ un sistema di vettori a due a due ortogonali e di norma pari a 1, allora vale la seguente relazione:

$$\left\| \sum_{k=1}^n \alpha_k e_k \right\|^2 = \sum_{k=1}^n |\alpha_k|^2$$

infatti:

$$\left\| \sum_{k=1}^n \alpha_k e_k \right\|^2 = \sum_{k=1}^n \|\alpha_k e_k\|^2 = \sum_{k=1}^n |\alpha_k|^2 \|e_k\|^2 = \sum_{k=1}^n |\alpha_k|^2$$

essendo la norma degli e_k pari a 1.

3.3 Teorema di proiezione ortogonale su spazi di dimensione finita

Molto spesso, in questo capitolo, si cercherà semplicemente di estendere i concetti presenti per spazi euclidei in situazioni più generiche. Uno di questi risultati è il teorema di proiezione. Dato H uno spazio con prodotto scalare, $x \in H$ un punto appartenente a esso, $\{e_k\}$ un sistema ortonormale, vogliamo trovare un certo y_n appartenente allo spazio $\operatorname{span}\{e_1, e_2, \dots, e_n\}$ (dunque di dimensione pari a n) tale per cui $x - y_n$ sia ortogonale allo spazio.

Prima di tutto, si vuole chiarire meglio quanto detto: H è un certo spazio, nel quale consideriamo un punto x ; per esempio, in Figura ??, si ha a che fare con $H = \mathbb{R}^3$, e si può avere come sottospazio quello composto da $\operatorname{span}\{e_1, e_2\}$. In questo caso, y_2 è la rappresentazione di x sullo span dei vettori e_1 e e_2 :

un punto dello spazio tridimensionale, rappresentato con due vettori, dunque su uno spazio bidimensionale (un piano). Si può utilizzare la Figura ?? per vedere qual è il significato di y_n : esso è la **proiezione** di x sullo span di e_1 e e_2 . Per **proiezione** di x si intende il punto y_n tale per cui il vettore $x - y_n$ è ortogonale a $\text{span}\{e_1, e_2\}$. Come vedremo, questo problema è interessante per vari motivi. Completando l'esempio nello spazio \mathbb{R}^3 , come è possibile ottenere y_n ? Si ha che:

$$x = (x_1, x_2, x_3)$$

come si possono mettere in relazione i due punti x e y_n ? Essendo la “proiezione”, diciamo a priori (e dimostreremo in seguito) che:

$$y_2 = (x|e_1)e_1 + (x|e_2)e_2 = x_1e_1 + x_2e_2 = (x_1, x_2)$$

ossia, i coefficienti della rappresentazione su $\text{span}\{e_1, e_2\}$ sono i prodotti scalari con ciascun elemento di cui si fa poi lo span: ciascun elemento del sistema ortonormale.

Una volta fatto questo esempio per chiarire, si consideri l'esempio di uno spazio di proiezione n -dimensionale: si ha che

$$y_n = \sum_{k=1}^n (x|e_k)e_k$$

In questa situazione, il teorema di proiezione ortogonale afferma che

1. $x - y_n$ è ortogonale al sottospazio (come già anticipato):

$$(x - y_n) \perp \text{span}\{e_1, \dots, e_n\}$$

in altre parole, si ha che:

$$((x - y_n|e_h) = 0 \forall h \in \{1, 2, \dots, n\})$$

2. Applicando il teorema di Pitagora generalizzato, è possibile vedere che:

$$\|x - y_n\|^2 = \|x\|^2 - \|y_n\|^2$$

essendo però

$$y_n = \sum_{k=1}^n (x|e_k)e_k$$

e avendo dimostrato la volta scorsa il risultato che segue:

$$\|y_n\|^2 = \left\| \sum_{k=1}^n (x|e_k)e_k \right\|^2 = \sum_{k=1}^n |(x|e_k)|^2$$

si ha che:

$$\|x - y_n\|^2 = \|x\|^2 - \|y_n\|^2 = \|x\|^2 - \sum_{k=1}^n |(x|e_k)|^2$$

questa è la stima dell'errore tra il punto esatto x e l'approssimante y_n appartenente al sottospazio in questione.

3. L'approssimazione y_n ottenuta mediante proiezione ortogonale è la migliore possibile. Anche questa implicazione può essere dimostrata utilizzando il teorema di Pitagora: si consideri al posto di y_n un certo z , che si suppone essere un altro punto appartenente a $\text{span}\{e_1, e_2, \dots, e_n\}$. Si ha che:

$$\|x - z\|^2 = \|x - y_n\|^2 + \|z - y_n\|^2 \geq \|x - y_n\|^2$$

essendo il termine $\|z - y_n\| \geq 0$, si ha la disuguaglianza, a meno che $z = y_n$. Dunque:

$$\|x - y_n\| \leq \|x - z\|$$

4. Vale la seguente disuguaglianza:

$$\sum_{k=1}^n |(x|e_k)|^2 \leq \|x\|^2$$

E questo vale per ogni n . Questo risultato si può dimostrare rapidamente:

$$\|x - y_n\|^2 = \|x\|^2 - \|y_n\|^2 \geq 0 = \|x\|^2 - \sum_{k=1}^n |(x|e_k)|^2$$

da cui:

$$\sum_{k=1}^n |(x|e_k)|^2 \leq \|x\|^2$$

valendo per ogni n , è possibile far tendere $n \rightarrow \infty$; la serie converge, poiché X è un punto limitato, e le ridotte n -esime della serie sono tutte limitate, dunque:

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n |(x|e_k)|^2 \leq \|x\|^2$$

questa è la cosiddetta **diseguaglianza di Bessel**.

Si osservi che per tutti i risultati finora proposti non abbiamo sfruttato la completezza dello spazio H : questo significa che essi prescindono da questa.

Tutto ciò che è stato finora detto dipende dal fatto che il primo risultato, ossia l'ortogonalità di $x - y_n$ dallo span dei e_k , sia verificata. Rimane dunque da dimostrare questo risultato, per garantire la validità di tutti gli altri.

Per dimostrare il primo punto, si calcoli il prodotto scalare tra il vettore differenza, $x - y_n$, e un generico e_h :

$$(x - y_n | e_h) = (x - e_h) - (y_n | e_h) = (x - e_h) - \left(\sum_{k=1}^n (x | e_k) e_k | e_h \right) = (x - e_h) - \sum_{k=1}^n (x | e_k) (e_k | e_h) = 0$$

quest'ultimo passaggio si può motivare così: prima di tutto si porta fuori la sommatoria per linearità del membro sinistro del prodotto scalare; poi, si deve dunque calcolare il prodotto scalare tra e_k della sommatoria, e e_h ; per definizione, questo è non nullo solo se $k = h$: altrimenti, essendo ortogonali, $(e_k | e_h) = 0$. Riassumendo:

$$(e_h, e_k) = \begin{cases} 1, & h = k \\ 0, & h \neq k \end{cases}$$

finalmente poi, i due termini uguali sottratti si annullano, dimostrando la validità del risultato.

Esempio 1

Si vuole a questo punto proporre un primo esempio, semplice, di sistemi ortonormali. A tal fine, si considerino, in l^2 , i vettori:

$$\{e_k\} : e_k = (0, 0, 0, \dots, 1, 0, \dots)$$

ossia, una successione costituita da tutti zeri, tranne nella k -esima posizione. Ricordando che il prodotto scalare è definito come:

$$\sum_{k=1}^{\infty} x_k \bar{y}_k$$

facendo il prodotto scalare, tutti i termini con $k \neq h$ si annullano, e solo uno sarà non nullo.

Esempio 2 : serie di Fourier

Si consideri lo spazio $L^2(0, 2\pi)$ come lo spazio

$$L^2(0, 2\pi) = \{f : [0, 2\pi] \rightarrow \mathbb{C} \text{ misurabili secondo Lebesgue } \}$$

con norma:

$$\|x\|_{L^2} = \left(\int_0^{2\pi} |x(t)|^2 dt \right)^{\frac{1}{2}}$$

finita, con identificazione di funzioni uguali quasi ovunque (o come completamento dello spazio delle funzioni continue). Si è già visto che in questo spazio è possibile identificare bene un prodotto scalare (ed esso induce la norma):

$$(x|y) = \int_0^{2\pi} x(t)\bar{y}(t) dt$$

A questo punto, come sistema ortonormale, si utilizzi:

$$e_k(t) = \frac{1}{\sqrt{2\pi}} e^{jkt}, \quad k \in \mathbb{Z}$$

Questo effettivamente costituisce, come stiamo per vedere, un sistema ortonormale, dal momento che:

$$(e_k | e_h) = \int_0^{2\pi} \frac{1}{\sqrt{2\pi}} e^{jkt} \frac{1}{\sqrt{2\pi}} e^{-jht} dt = \frac{1}{2\pi} \int_0^{2\pi} e^{j(k-h)t} dt = \frac{1}{2\pi} \left[\frac{e^{j(k-h)t}}{j(k-h)} \right]_0^{2\pi}$$

ora (tenendo conto che si ha a che fare con una discontinuità eliminabile per $k = h$), si ha che:

$$\frac{1}{2\pi} \left[\frac{e^{j(k-h)t}}{j(k-h)} \right]_0^{2\pi} = \begin{cases} 1, & k = h \\ 0, & k \neq h \end{cases} \quad (3.1)$$

Di conseguenza, ha senso pensare alla serie di Fourier in termini di applicazione del teorema di proiezione ortogonale: si proietta su un certo sottospazio di $L^2(0, 2\pi)$. Si ottiene dunque:

$$y_n = \sum_{k=-n}^n (x | e_k) e_k$$

dove:

$$(x | e_k) = \int_0^{2\pi} x(t) \frac{e^{-jkt}}{\sqrt{2\pi}} dt$$

Da qui, si ottiene:

$$y_n = \sum_{k=-n}^n c_k e^{jkt}$$

dove

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} x(t) e^{-jkt} dt$$

(il secondo $\sqrt{2\pi}$ deriva dal e_k nella sommatoria).

y_n è la ridotta n -esima della serie di Fourier: è una $S_n(t)$!

Grazie al teorema di proiezione ortogonale, inoltre, abbiamo dimostrato che è possibile stimare come segue l'errore:

$$\|x - S_n\|^2 = \|x\|^2 - \frac{1}{2\pi} \sum_{k=-n}^n |c_k|^2$$

essendo

$$(x|e_k) = \sqrt{2\pi}c_k$$

3.4 Procedimento di Gram-Schmidt: ortogonalizzazione di sistemi di vettori

L'ipotesi finora utilizzata per approssimare un certo punto è stata quella di avere, come approssimante, un sistema ortogonale. Non è tuttavia detto che lo spazio sul quale si intende effettuare la proiezione sia noto come span di vettori ortonormali. Tuttavia, esiste il seguente risultato: sia $\{x_k\}_{k=1}^{\infty}$ un sistema di vettori linearmente indipendenti (ossia, ogni suo sottoinsieme finito è costituito da vettori linearmente indipendenti: $\forall n, \{x_1, x_2, \dots, x_n\}$ è linearmente indipendente); dunque, esiste un sistema ortonormale $\{e_k\}_{k=1}^{\infty}$ tale per cui:

$$\text{span} \{e_1, e_2, \dots, e_n\} \equiv \text{span} \{x_1, x_2, \dots, x_n\} \quad \forall n$$

Questa dimostrazione può essere fatta per induzione, applicando il procedimento di Gram-Schmidt, ora richiamato. Si consideri, per incominciare:

$$e_1 = \frac{x_1}{\|x_1\|}$$

Per e_2 , si consideri la Figura ??: ora che si conosce e_1 dalla prima iterazione, si calcola la proiezione di x_2 su e_1 , dunque si sottrae il contributo di questa proiezione da x_2 ; in questo modo, quello che si otterrà sarà un vettore che conterrà solo componenti ortogonali a e_1 , essendo la componente lungo e_1 sottratta!

$$y_2 = x_2 - (x_2|e_1)e_1$$

poi questo si deve normalizzare:

$$e_2 = \frac{y_2}{\|y_2\|}$$

Il procedimento si ripete iterativamente, utilizzando dunque la seguente formula:

$$y_n = x_n - \sum_{k=1}^{n-1} (x_n | e_k) e_k, \quad e_n = \frac{y_n}{\|y_n\|}$$

Un esempio di applicazione su spazi di funzioni di questo procedimento è quella dei polinomi ortogonali, come i polinomi di Legendre, i polinomi di Chebyshev, i polinomi di Laguerre, di Hermite, e così via: applicando a $\text{span}\{1, t, t^2, t^3, \dots\}$ il procedimento di Gram-Schmidt (utilizzando eventualmente diversi “pesi” al momento di calcolare i prodotti scalari), è possibile ottenere queste funzioni speciali.

3.5 Rappresentazione di un vettore in termini di un sistema di vettori

Si vuole a questo punto provare a introdurre, in alcune situazioni, il concetto di base, già noto nell’ambito degli spazi euclidei. Un primo risultato è il seguente: ogni spazio vettoriale ha **base di Hamel**: ogni elemento di un insieme può essere espresso come combinazione lineare finita di elementi presi da un sottoinsieme. Il problema di questo risultato è che, di fatto, esso è inutilizzabile nell’ambito degli spazi di Banach: è possibile infatti che questi sottoinsiemi non siano numerabili, abbiano ossia la potenza del continuo, rendendo di fatto inutilizzabile questa nozione (volendo noi costruire basi mediante processi iterativi, o comunque utilizzarle in calcolatori).

Sfruttiamo la seguente definizione: in uno spazio H dotato di prodotto scalare, un sistema di vettori $\{x_k\}_{k=1}^{\infty}$ si dice **completo** se è denso in H :

$$\overline{\text{span}\{x_k\}} = H$$

Lo span è dato da somme **finite** di vettori; se dunque l’insieme delle combinazioni lineari finite di vettori è denso in H , allora significa che ogni vettore è rappresentabile con precisione arbitraria come una somma di vettori appartenenti allo span.

Si osservi che in questo ambito la completezza è ben diversa rispetto a quanto abbiamo finora detto: qua stiamo parlando di **completezza di uno spazio di vettori**, ciascuno dei quali appartiene a uno spazio con prodotto scalare.

Si consideri dunque un esempio:

$$M = \{t^k, k \geq 0\}$$

questo è un sistema completo in $L^2(a, b)$. Questo risultato è già noto: infatti, per il teorema di Weierstrass precedentemente discusso e dimostrato, lo spazio dei polinomi è denso nello spazio delle funzioni continue $C[(a, b)]$, il quale a sua volta è denso in $L^2(a, b)$.

Esempio simile è quello dei polinomi trigonometrici:

$$M = \{e^{ikt}, k \in \mathbb{Z}\}$$

dunque:

$$\text{span } \{M\}$$

è lo spazio dei polinomi trigonometrici. Questo, per la “variante” del teorema di Weierstrass, è denso in $L^2(0, 2\pi)$. M è dunque completo in $L^2(0, 2\pi)$.

3.6 Caratterizzazione dei sistemi completi

In questa sezione si proporrà un elenco di proprietà equivalenti alla completezza di un sistema ortogonale di vettori, e si dimostreranno le relazioni tra esse. Vale dunque il seguente teorema: dato H uno spazio dotato di prodotto scalare, per un sistema ortonormale numerabile $\{e_k\}_{k=1}^{\infty}$, sono equivalenti le seguenti affermazioni.

1. $\{e_k\}$ è completo, ossia:

$$\overline{\text{span } \{e_k\}} = H$$

2. Per ogni $x \in H$ è possibile scrivere:

$$x = \sum_{k=1}^{\infty} (x|e_k)e_k$$

ossia vale questa decomposizione, dove i coefficienti della rappresentazione non sono altro che le proiezioni. Questo è un tentativo di estendere la teoria classica degli spazi euclidei, dove si hanno dei vettori scritti come combinazioni lineari di altri; in questo caso, però, la combinazione lineare è sostituita da una serie.

3. Vale l'identità di Parseval $\forall x \in H$:

$$\|x\|^2 = \sum_{k=1}^{\infty} |(x|e_k)|^2$$

Si procede a questo punto con una dimostrazione delle varie implicazioni.

2 \implies 1

Se ogni $x \in H$ può essere scritto mediante la serie, allora è chiaro che il sistema è completo; infatti:

$$\sum_{k=1}^n (x|e_k)e_k$$

appartiene allo span degli $\{e_k\}$, ed è un vettore; inoltre, esso, per $n \rightarrow \infty$, tende a x ; dunque, comunque preso il vettore, esso tende a un $x \in H$; di conseguenza, la chiusura dello span è densa.

2 \iff 3

Dato lo span dei primi n vettori, se si considera la proiezione ortogonale di x su questi, si ottiene:

$$\left\| x - \sum_{k=1}^n (x|e_k)e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |(x|e_k)|^2$$

grazie a questa relazione, si ha che le implicazioni 2 e 3 sono equivalenti; infatti, per $n \rightarrow \infty$, se la serie converge, e ha somma pari a x , il membro sinistro deve andare a zero; di conseguenza, il membro destro è tale per cui:

$$\sum_{k=1}^n |(x|e_k)|^2 = \|x\|^2$$

quindi questo implica convergenza a x : l'implicazione vale per ogni x fissato, dal momento che, se per un certo x si ha che:

$$x = \sum_{k=1}^{\infty} (x|e_k)e_k$$

se si ha questa serie convergente, allora la serie dei moduli quadri sicuramente convergerà. Questo, incondizionatamente (dal momento che la serie di Parseval è una somma di termini positivi, l'ordine con il quale si effettua la somma è indifferente: non cambia la somma). Dunque, l'ultima espressione scritta, sicuramente convergerà, nella norma dello spazio considerato.

Riassumendo: sia partendo dall'identità di Parseval, sia partendo dalla possibilità di rappresentare x come visto precedentemente, è possibile vedere che le due espressioni sono equivalenti, grazie all'osservazione appena scritta, e ora riportata per chiarezza:

$$\left\| x - \sum_{k=1}^n (x|e_k)e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |(x|e_k)|^2$$

1 \implies 2

Come ipotesi di partenza abbiamo il fatto che $\{e_k\}$ è un insieme completo. Se esso è completo, per definizione, ogni vettore di H può essere approssimato arbitrariamente bene con una combinazione dei vettori appartenenti a $\{e_k\}$; questo significa, in altre parole, che esistono coefficienti $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{C}$ tali per cui:

$$\left\| x - \sum_{k=1}^n \alpha_k e_k \right\| \leq \varepsilon$$

Tuttavia, di una cosa siamo sicuri, ossia del fatto che:

$$\left\| x - \sum_{k=1}^n (x|e_k)e_k \right\| \leq \varepsilon \leq \left\| x - \sum_{k=1}^n \alpha_k e_k \right\| \leq \varepsilon$$

Infatti, come visto precedentemente la norma in questo ambito quantifica l'errore rispetto x dell'approssimazione ottenuta a partire dall'uso del sistema ortonormale; tuttavia, è noto che quella approssimazione che usa come coefficienti le proiezioni $(x|e_k)$, è la migliore: questo, per il teorema della proiezione ortogonale. Dunque, essendo l'approssimazione migliore, l'errore sarà più piccolo. Di conseguenza, essendo entrambe queste serie minori di un ε arbitrario, si ha che:

$$\alpha_k = (x|e_k)$$

Un ulteriore risultato: si immagini di avere $m > n$; in questa situazione, vale sempre la disuguaglianza:

$$\left\| x - \sum_{k=1}^n (x|e_k)e_k \right\| \geq \left\| x - \sum_{k=1}^m (x|e_k)e_k \right\|$$

questa si può motivare usando la seguente relazione:

$$\left\| x - \sum_{k=1}^n (x|e_k)e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^n |(x|e_k)|^2$$

questa è nota, ma questa, al crescere di n , è anche monotona decrescente: la serie a membro destro infatti è una serie crescente, essendo a termini positivi, di conseguenza l'intero membro sinistro decresce al crescere di n .

3.6.1 Esempio

Al fine di fissare i concetti, al solito, si propone un esempio; si consideri:

$$\left\{ \frac{1}{\sqrt{2\pi}} e^{jkt} \right\}, \quad k \in \mathbb{Z}$$

questo è un insieme ortonormale è completo in $L^2(0, 2\pi)$. Questo significa che, per ogni $x(t) \in L^2(0, 2\pi)$, si ha che:

$$x(t) = \sum_{k \in \mathbb{Z}} c_k e^{jkt}, \quad c_k = \frac{1}{2\pi} \int_0^{2\pi} x(t) e^{-jkt} dt$$

Il fatto di poter rappresentare la generica $x(t)$ in questo modo garantisce il fatto che la serie abbia **convergenza** a un elemento di $L^2(0, 2\pi)$, ossia a una funzione appartenente a questo spazio. In altre parole:

$$\lim_{N \rightarrow \infty} \int_0^{2\pi} \left| x(t) - \sum_{k=-N}^N c_k e^{jkt} \right|^2 dt \rightarrow 0$$

Si osservi che si è scelto di usare la sommatoria con k che va da $-N$ a N ; ai fini della convergenza incondizionata, in verità, questo non è eccessivamente importante, dal momento che la serie come già detto precedentemente è a termini positivi; in L^2 , dunque si può scegliere k nell'intervallo che si vuole. Volendo considerare tuttavia la convergenza puntuale, questo non è più vero:

in questo caso, per i teoremi di convergenza della serie di Fourier, è necessario avere k che va da $-N$ a N .

Ovviamente, vale anche l'identità di Parseval:

$$\int_0^{2\pi} |x(t)|^2 dt = 2\pi \sum_{k \in \mathbb{Z}} |c_k|^2$$

Due osservazioni riguardo a questa:

- il fatto che valga l'eguaglianza di Parseval garantisce che la serie dei moduli quadri converga; questo significa che $\{c_k\} \in l^2$; d'altra parte, essa **non** appartiene a l^1 : infatti, si può pensare che questa successione non vada a zero sufficientemente rapidamente da appartenere a l^1 .
- l'eguaglianza di Parseval non è una proprietà intrinseca delle funzioni appartenenti a L^2 : tutto il ragionamento appena fatto poteva essere applicato alle funzioni integrabili nel senso di Riemann, senza il minimo problema.

3.7 Spazi separabili

Finora si è visto che, dato un sistema ortonormale completo, è possibile utilizzarlo per sviluppare un certo vettore appartenente allo spazio in questione. Tuttavia, non è stata discussa una possibilità: potrebbe essere che, in uno spazio con prodotto scalare, non esista un sistema ortonormale completo (che abbia la cardinalità del numerabile!). Di conseguenza, è necessario identificare una classe di spazi in cui si abbia la certezza dell'esistenza di un sistema ortonormale completo. Questi spazi, sono gli **spazi separabili**.

Uno spazio normato si dice **separabile** se contiene un **sottoinsieme numerabile denso** in esso. Questa definizione, in un qualche senso, classifica la *dimensione* dello spazio: se uno spazio è sufficientemente *piccolo* da avere un sottoinsieme numerabile denso in esso, allora esso è definito **separabile**. Consideriamo alcuni esempi di spazi separabili.

- Lo spazio l^p , $1 \leq p < \infty$ è separabile.
- Lo spazio delle funzioni continue con norma del sup è separabile;
- $L^p(a, b)$, $1 \leq p < \infty$, è uno spazio separabile (per esempio $L^2(a, b)$ è un esempio di ciò!)

- lo spazio l^∞ delle successioni limitate **non** è uno spazio separabile!

3.7.1 Dimostrazione: separabilità di l^p

In questa sottosezione si dimostrerà il fatto che lo spazio l^p è separabile: come vedremo, è possibile identificare un sottoinsieme numerabile denso in esso.

Un'osservazione preliminare: si supponga di avere a che fare per esempio con la successione così definita:

$$e_1 = (1, 0, 0, \dots) \quad e_2 = (0, 1, 0, \dots) \quad e_3 = (0, 0, 1, 0, \dots)$$

Ossia, e_k , per ogni k -esima posizione vale 1, in tutti gli altri casi zero. Dato $x \in L^p$, dove $x = (x_1, x_2, x_3, \dots)$. Dunque, si può verificare che, $\forall \varepsilon > 0$, esiste un certo n tale per cui:

$$\left\| x - \sum_{k=1}^n x_k e_k \right\| < \frac{\varepsilon}{2}$$

Infatti, la differenza tra x e la serie è:

$$= (0, 0, 0, 0, \dots, x_{n+1}, x_{n+2}, \dots)$$

ossia, rimane solo la *coda* della serie, che è una serie convergente.

Questo ci ha permesso di trovare un sottoinsieme **denso**, ma **non numerabile**: infatti, ciascun elemento $x_k \in \mathbb{C}$: il problema sono dunque i **coefficienti**, che appartengono al campo complesso, che non è numerabile. Al fine di ottenere un insieme denso, è necessario approssimare questi coefficienti, mediante dei razionali; dunque, per $k = 1, 2, \dots, n$, si considerino degli α_k tali per cui:

$$\alpha_k \in \mathbb{Q} + j\mathbb{Q}$$

ossia, degli α_k complessi, ottenuti mediante una parte reale e una parte immaginaria razionali. Per questi, richiediamo che:

$$|\alpha_k - x_k| < \frac{\varepsilon}{2} \frac{1}{n^{\frac{1}{p}}}$$

A questo punto, consideriamo la norma di ciò:

$$\left(\sum_{k=1}^n |x_k - \alpha_k|^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n \left(\frac{\varepsilon}{2n^{\frac{1}{p}}} \right)^p \right)^{\frac{1}{p}} = \frac{\varepsilon}{2n^{\frac{1}{p}}} \left(\sum_{k=1}^n 1 \right)^{\frac{1}{p}} = \frac{\varepsilon}{2}$$

Dunque, a meno di ε , è possibile rappresentare il vettore x_k mediante questa somma finita. Si può per questo dire che l^p , per $1 \leq p < \infty$, sia:

$$l^p = \overline{\bigcup_{n \in \mathbb{N}} \text{span}_{\mathbb{Q}+j\mathbb{Q}} \{e_1, \dots, e_n\}}$$

E questo è quello che abbiamo verificato: esiste un certo n tale per cui questo span è in grado di generare l^p . Questa unione inoltre è **numerabile**; infatti, si ha che:

$$\text{span}_{\mathbb{Q}+j\mathbb{Q}} \{e_1, \dots, e_n\}$$

è una combinazione lineare di certi coefficienti, che sono razionali, dal momento che e_1 sta nello spazio $\mathbb{Q}+j\mathbb{Q}$, idem e_2 , idem e_3 , e così via; di conseguenza, essendo il piano complesso isomorfo a \mathbb{R}^2 , si può dire che per ciascuna componente si abbia una variazione dentro $\mathbb{Q} \times \mathbb{Q}$. Quindi, per n componenti, si ha che x_k sta in $(\mathbb{Q} \times \mathbb{Q})^n$.

3.7.2 Osservazioni aggiuntive sugli spazi separabili

I ragionamenti appena fatti possono essere esportati anche per quanto riguarda lo spazio delle funzioni continue con la norma del sup; in questo caso, è possibile scrivere che:

$$C([a, b]) = \overline{\bigcup_{n \in \mathbb{N}} \text{span}_{\mathbb{Q}+j\mathbb{Q}} \{1, t, t^2, \dots, t^n\}}$$

di conseguenza, è possibile applicare un trucco simile a quanto visto per quanto riguarda gli spazi di successioni, solamente considerando i polinomi, che sono funzioni dense, per il teorema di Weierstrass di approssimazione, nello spazio delle funzioni continue. A partire da qui è possibile anche dire che lo spazio L^p è separabile: partendo dal fatto che lo spazio delle funzioni continue è infatti separabile, è possibile ricordare che $C([a, b])$ è denso in $L^p(a, b)$, e dunque essendo uno denso nell'altro, le proprietà si esportano.

Caso completamente diverso è quello di l^∞ : in questo caso, non è possibile trovare un insieme denso e numerabile in esso. Verrà ora esposta una motivazione intuitiva, rafforzata da un cenno di dimostrazione. L'idea, come già accennato, è che lo spazio in questione sia *troppo grande* per avere questo insieme denso e numerabile.

Per capire di cosa si parla, si consideri X così definito:

$$X = \{0, 1\}^{\mathbb{N}} \subset l^\infty$$

la notazione A^B è una notazione compatta per indicare:

$$A^B = \{f : B \rightarrow A\}$$

ora stiamo nel dettaglio considerando l'insieme delle successioni i cui valori sono sempre zero e uno. Questo insieme X è **non numerabile**, dal momento che la cardinalità delle funzioni da B in A è quella di A , elevata a quella di B ; essendo nel caso in questione l'elevazione a potenza la cardinalità di \mathbb{N} , questa è la \aleph_0 : considerando 2 (cardinalità dell'insieme $\{0, 1\}$, elevato a \aleph_0 , si ha: 2^{\aleph_0} , che è la potenza, la cardinalità del continuo). Sullo spazio X , dati $x, y \in X$, si ha che, per $x \neq y$,

$$\|x - y\|_{l^\infty} = \sup_k |x_k - y_k|$$

dal momento che x può valere o zero o uno, si ha che:

$$\|x - y\|_{l^\infty} = 1$$

infatti, al massimo, la differenza può valere 1. Questo significa che ogni punto di X è **distanziato** 1 con il successivo.

Se a questo punto l^∞ avesse un sottoinsieme numerabile denso nell'intorno di un certo punto x , di raggio $1/2$, dovrei trovare un punto, appartenente all'insieme D ipoteticamente esistente, denso e numerabile. Per la definizione di densità, in ogni intorno di raggio $1/2$ di un generico punto di X dovrei trovare un diverso punto appartenente a D ; ma questo è in contraddizione col fatto che D sia numerabile: se per ogni punto $x_i \in X$ ho un punto di D nell'intorno, allora, essendo X non numerabile, D non può essere numerabile.

3.7.3 Teorema di esistenza di sistemi ortonormali completi in spazi separabili

Si propone a questo punto un importante teorema: dato uno spazio H dotato di prodotto scalare, di dimensione infinita, esso è separabile **se e solo se** esso ha un sistema di vettori numerabile completo ortonormale (il quale, ovviamente, può essere poi usato per sviluppare un generico elemento dello spazio in serie).

Dimostrazione: dato H separabile, significa che esiste un sistema numerabile denso: $\{x_k, k\mathbb{N}\}$. Dunque, si ha che:

$$\text{span} \{x_k\}$$

è un insieme denso; infatti, lo span contiene x_k , dunque è un insieme più grande, e dunque è denso anche esso.

La dimostrazione nel primo senso potrebbe già essere terminata, se non fosse per un dettaglio: si è trovato un sistema, ma non ortonormale; al fine di ottenere il risultato finale, dunque, è necessario applicare il processo di ortonormalizzazione di Gram-Schmidt. Per fare ciò, è necessario supporre che $\{x_k\}$ sia un insieme di elementi linearmente indipendenti (questo non è un problema: al fine di ottenere un sistema di questo genere è necessario eliminare tutti i vettori che sono dati da una combinazione lineare dei precedenti; lo span rimarrà uguale!). Ciò che rimane dalla *pulizia* dei vettori linearmente dipendenti, può essere ortonormalizzato alla solita maniera.

Per quanto riguarda la dimostrazione nel senso opposto, del *viceversa*, se H ha un sistema ortonormale completo numerabile, $\{e_k\}$, allora si ha un punto di partenza, ma non subito quello di arrivo: il fatto di disporre di un sistema numerabile non garantisce che i coefficienti siano appartenenti a un insieme numerabile. Di conseguenza, essendo essi complessi, è possibile sfruttare lo stesso trucco visto precedentemente, e definire H come:

$$H = \overline{\bigcup_n \text{span}_{\mathbb{Q}+j\mathbb{Q}} \{e_1, \dots, e_n\}}$$

Considerando dunque per esempio $L^2(a, b)$, esso è separabile, e per esempio come insieme ortonormalizzato è possibile utilizzare i polinomi di Legendre.

3.8 Spazi di Hilbert

3.8.1 Definizione e osservazioni

Si definisce **spazio di Hilbert** uno spazio H dotato di prodotto scalare completo rispetto alla norma indotta da esso. In uno spazio di Hilbert, un sistema ortonormale completo è detto **base ortonormale** (o **base di Hilbert**). Si intenderanno, a meno che non sia esplicitamente indicato, spazi di Hilbert **separabili**: la base in questione sarà dunque della forma $\{e_k, k \in \mathbb{N}\}$ (basi **numerabili**).

Il motivo per cui si chiama **base** ricorda il concetto noto negli spazi euclidei:

$$x = \sum_{k=1}^{\infty} (x|e_k)e_k$$

che converge in H . Alcuni esempi sono ora riportati.

- l^2, L^2 sono spazi di Hilbert separabili;
- $C([a, b])$ con la norma $\|\cdot\|_{L^2}$ è uno spazio con prodotto scalare, ma **non** è uno spazio completo: **non** è uno spazio di Hilbert.

Per chiudere il discorso degli spazi di Hilbert non separabili, si vuole proporre il fatto che **essi esistono**; un esempio sono gli spazi di funzioni **quasi periodiche**, per esempio,

$$\sin(x) + \sin(\sqrt{2}x)$$

è una funzione **non periodica**, essendo 1 e $\sqrt{2}$ non commensurabili, ma essa è detta **quasi periodica**.

Si consideri, come esempio:

$$E = \left\{ \text{funzioni continue } f : \mathbb{R} \rightarrow \mathbb{C} : p(f) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^{+T} |f(x)|^2 dx \exists, < \infty \right\}$$

p è una seminorma su questo spazio. Ciò che si può fare dunque è ricavare N , nucleo della seminorma:

$$N = \{f \in E : p(f) = 0\}$$

e quozientare E rispetto a N , quindi trovarne il completamento: X . Su X è possibile definire il prodotto scalare:

$$(f|g) = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^{+T} f(x) \bar{g}(x) dx$$

Un esempio di base per questi spazi sono funzioni del tipo:

$$\{e^{i\lambda x}\}, \lambda \in \mathbb{R}$$

con questo prodotto scalare è possibile vedere che il sistema appena scelto è ortonormale; essendo tuttavia λ variabile in \mathbb{R} e non in \mathbb{Z} , la base non è sicuramente numerabile, essendo λ variabile con la potenza del continuo.

3.8.2 Teorema di Riesz-Fischer

Rispetto a un più generale spazio dotato di prodotto scalare, uno spazio di Hilbert è anche **completo**. Si vuole dunque capire a cosa serve introdurre la completezza, e perché essa sia così importante. Si è detto precedentemente che

$$x(t) = \sum_{k=-N}^N c_k e^{ikt}$$

e grazie all'identità di Parseval, questa converge nella norma di L^2 . Questo teorema ha permesso di esprimere un certo vettore come funzione di altri vettori dello spazio, mediante una combinazione lineare dettata da coefficienti c_k .

Si consideri ora però un problema opposto: si supponga di disporre di una successione di coefficienti $\{c_k\} \in l^2$; ci si pone dunque la seguente domanda: esiste una certa x che si possa costruire a partire da quello sviluppo? Ha dunque senso costruire una x utilizzando i c_k come coefficienti di sviluppo? La risposta è: non con funzioni integrabili nel senso di Riemann, ma sì, con funzioni appartenenti a L^2 .

Il teorema di Riesz-Fischer è un risultato molto forte riguardante gli spazi di Hilbert: tutti gli spazi di Hilbert separabili e di dimensione infinita sono isomorfi a l^2 . L'isomorfismo è quello che associa a ogni elemento il suo coefficiente di Fourier.

Più formalmente: sia H uno spazio di Hilbert separabile di dimensione infinita, e $\{e_k\}_{k=1}^{\infty}$ una sua base di Hilbert. Allora, l'applicazione da H a l^2 tale per cui:

$$x \longrightarrow \{(x|e_k)\}$$

è un isomorfismo di spazi di Hilbert: tutte le proprietà di uno spazio di Hilbert, ossia

- linearità;
- biiettività;
- conservazione di norma e prodotto scalare (si ha a che fare con **isometrie**)

sono verificate.

Questo, in altre parole, ci dice che **esiste un unico spazio di Hilbert di dimensione infinita**; tutti gli altri, sono isomorfi a esso.

Dimostriamo passo passo questo teorema: prima di tutto, l'applicazione è ben definita, dal momento che, se $x \in H$, la successione $\{(x|e_k)\} \in l^2$. Per dimostrare ciò sarebbe possibile utilizzare l'identità di Parseval, ma, senza voler tirare fuori risultati così importanti, è sufficiente utilizzare la disuguaglianza di Bessel, e il fatto che $\{e_k\}$ dunque è un sistema ortonormale, senza aggiungere ulteriori informazioni.

Secondo passo è la linearità: ovviamente, l'applicazione riguarda una successione di prodotti scalari, i quali soddisfano le proprietà di linearità; è banale sostituire a x un generico $\alpha x + \beta y$ e vedere che tutto è verificato.

Il terzo passo sfrutta l'ortonormalità e la completezza di $\{e_k\}$: se si considera

$$\|x\|_H^2 = \sum_{k=1}^{\infty} |(x|e_k)|^2$$

si ha l'identità di Parseval, dunque tutto è regolare.

Il passo successivo riguarda la biiettività dell'applicazione; per quanto riguarda l'iniettività, è banale dimostrarla: essendo l'applicazione lineare, essa è iniettiva se e solo se il suo nucleo coincide con l'elemento nullo, ossia se l'elemento nullo viene mappato nell'elemento nullo; questo è verificato. Per

quanto riguarda la suriettività, per ogni $\{\alpha_k\} \in l^2$, deve esistere un certo $x \in H$ tale per cui:

$$(x|e_k) = \alpha_k$$

ossia, data una generica successione di α_k appartenente a l^2 , si deve poter trovare un x che sia mappato in questa, per ogni k . Al fine di fare ciò, si verifichi che:

$$\sum_{k=1}^{\infty} \alpha_k e_k$$

converga. Una volta fatto ciò, potremo usare questa serie come x . Al fine di dimostrare la convergenza, è necessario utilizzare l'ipotesi di completezza: dati $\{\alpha_k\}$, si troverà un certo x di cui i coefficienti sono proprio quelli di partenza. Essendo lo spazio completo, possiamo studiare la successione delle somme parziali della serie appena scritta; per $n > m$, si ha (si può fare anche l'opposto):

$$\|S_n - S_m\|^2 = \left\| \sum_{k=m+1}^n \alpha_k e_k \right\|^2 = \sum_{k=m+1}^n |\alpha_k|^2$$

data dunque:

$$s_n = \sum_{k=1}^n |\alpha_k|^2$$

si ha che:

$$\|S_n - S_m\|^2 = s_n - s_m$$

essendo tuttavia la successione degli α_k per ipotesi appartenente a l^2 , allora s_n converge in \mathbb{R} , e dunque è di Cauchy; questo permette di scrivere che, dato un certo ε , esiste un certo N_ε tale per cui, se n, m sono maggiori di esso,

$$s_n - s_m < \varepsilon$$

Ma dunque, se riusciamo a dire ciò per s_n , possiamo dirlo anche per S_n : anche essa sarà di Cauchy! Ma, essendo H uno spazio completo, e S_n una successione di Cauchy, allora essa è anche una successione convergente; per questo motivo, il limite di questa sarà:

$$x = \sum_{k=1}^{\infty} \alpha_k e_k$$

A questo punto, si legghino gli α_k con i prodotti scalari:

$$(x|e_k) = \left(\sum_{h=1}^{\infty} \alpha_h e_h \middle| e_k \right)$$

essendo il prodotto scalare continuo, è possibile scambiare i segni di sommatoria e prodotto scalare, ottenendo:

$$(x|e_k) = \sum_{h=1}^{\infty} \alpha_h (e_h|e_k) = \alpha_k$$

e grazie a ciò abbiamo ottenuto anche la suriettività.

Considerazione finale su prodotti scalari e norme

Il fatto che si abbia una conservazione dei prodotti scalari e delle norme è in realtà legato, ma ciò non dipende dal teorema di Riesz-Fischer: se si ha conservazione delle norme, si ha anche conservazione dei prodotti scalari. Si consideri una applicazione $A : H_1 \rightarrow H_2$, lineare, il fatto che conserva la norma dice che:

$$\|Ax\| = \|x\|$$

mentre, per i prodotti scalari:

$$(Ax|Ay) = (x|y)$$

la seconda implica la prima banalmente: se al posto di y si mette x , si ha:

$$(Ax|Ax) = (x|x)$$

ossia:

$$\|Ax\|^2 = \|x\|^2$$

La dimostrazione del viceversa è meno banale: si deve conoscere la seguente formula, che permette di scrivere un prodotto scalare in termini di norme:

$$(x|y) = \frac{1}{4} \sum_{k=0}^3 j^k \|x + j^k y\|^2 = \frac{1}{4} [\|x + y\|^2 - \|x - y\|^2 + j \|x + jy\|^2 - j \|x - jy\|^2]$$

questa è la cosiddetta **identità di polarizzazione**. Si applichi a questa Ax , invece di x :

$$\begin{aligned} (Ax|Ay) &= \frac{1}{4} [\|Ax + Ay\|^2 - \|Ax - Ay\|^2 + j \|Ax + jAy\|^2 - j \|Ax - jAy\|^2] = \\ &= \frac{1}{4} [\|A(x + y)\| - \|A(x - y)\| + j \|A(x + jy)\| - \|A(x - jy)\|] = \\ &= \frac{1}{4} [\|(x + y)\| - \|(x - y)\| + j \|(x + jy)\| - \|(x - jy)\|] = \\ &= (x|y) \end{aligned}$$

il primo passaggio sfrutta la linearità di A , raccogliendo l'operatore; il secondo, l'isometria; il terzo, semplicemente, è algebra.

Esempio veloce: applicazione di Riesz-Fischer

Un risultato del teorema di Riesz-Fischer è il seguente: si consideri una successione $\{\alpha_k\} \in l^2$; questa permette di garantire che esiste una funzione $x \in L^2$ tale per cui:

$$x = \sum_{k \in \mathbb{Z}} \alpha_k e^{jkt}$$

con convergenza in $L^2(0, 2\pi)$. A una x , dunque, è possibile associare i suoi coefficienti c_k .

3.8.3 Teorema di proiezione su sottospazi

Precedentemente è stato proposto un teorema di proiezione ortogonale su sottospazi; ora, questo teorema verrà esteso. Dato H uno spazio completo, è possibile effettuare l'operazione di proiezione su un qualsiasi sottoinsieme **chiuso e convesso**. Dato dunque $x \in H$, si dimostra che esiste un unico y appartenente a $M \subset H$ convesso e chiuso tale per cui la distanza è minima: il minimo **esiste** ed è **unico**.

Al fine di procedere, si parta dalla definizione di distanza, in uno spazio normato: la distanza tra un punto x e un insieme M è:

$$d(x, M) = \inf_{z \in M} \|x - z\|$$

Dunque, per ogni $x \in H$ esiste un $y \in M$ tale per cui:

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

Prima di mostrare la dimostrazione di questo teorema, si vuole proporre qualche osservazione preliminare riguardante la necessità delle ipotesi. Prima di tutto, dato un aperto (anziché un chiuso), si avrebbe sostanzialmente un problema legato all'**unicità**: non esisterebbe il minimo, poiché non apparterebbe all'insieme; è infatti evidente che il inf in questione debba essere sul *bordo*, e dunque se il bordo non appartenesse all'insieme (ossia se l'insieme fosse aperto) ciò non sarebbe possibile.

La convessità, in spazi di dimensione finita, riguarda soprattutto l'**unicità** del limite. Si consideri Figura ??: se l'insieme non fosse convesso, si avrebbero due punti di minimo, invece che uno: due punti per cui la distanza è minima ed è uguale. In realtà, in spazi di dimensione infinita, la convessità interviene anche nell'unicità.

Si procede ora con la dimostrazione dell'**esistenza** del punto. Per fare ciò, prima di tutto, si consideri:

$$\delta = d(x, M)$$

al fine di alleggerire la notazione. Dunque, vogliamo dimostrare che il punto a minima distanza esiste. Al fine di fare ciò, cerchiamo una successione di punti $\{y_n\} \in M$ tale per cui:

$$\|x - y_n\| \rightarrow \delta$$

questo, per la definizione di inf. Essendo lo spazio completo, y_n è di Cauchy (oltre che convergente). Dunque, vale l'identità del parallelogramma, applicata come segue:

$$2\|x - y_n\|^2 + 2\|x - y_m\|^2 = \|2x - (y_n + y_m)\|^2 + \|y_n + y_m\|^2$$

questo può essere riscritto isolando un membro e portando fuori un 2 (che va al quadrato essendoci la norma al quadrato):

$$\|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2$$

Si effettui a questo punto la seguente osservazione: y_n e y_m sono successioni in uno spazio convesso; sicuramente, dunque, si ha che la successione media tra le due appartiene a M :

$$\frac{y_n + y_m}{2} \in M$$

questo significa che, dunque, essa è sicuramente minorabile da δ :

$$x - \frac{y_n + y_m}{2} \geq \delta$$

infatti, essendo il punto all'interno di M , la sua distanza sarà di sicuro maggiore o uguale a δ ; δ è positivo, ed è al quadrato, dunque l'intera espressione potrà essere **maggiorata** come segue:

$$2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\delta^2$$

infatti, essendo δ più piccolo dell'espressione, ma essendo il termine minorato con segno **negativo**, l'espressione globale sarà **più grande**, mettendo δ !

A questo punto, per $n \rightarrow \infty$, si ha che:

$$x - y_n \rightarrow \delta$$

dunque, si ottiene:

$$\leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0, \quad n, m \rightarrow \infty$$

ossia, per $n, m \rightarrow \infty$, tutta questa espressione tende a zero. Dunque:

Siccome H è uno spazio completo, si può usare questo trucco:

$$\exists y = \lim_{n \rightarrow \infty} y_n \in H$$

ma $y \in M$, essendo M chiuso:

$$\|x - y\| = \lim_{n \rightarrow \infty} \|x - y_n\| = \delta$$

e la dimostrazione è completata.

Per quanto riguarda la dimostrazione di unicità, invece, si può supporre per assurdo che, dati due punti $y, y' \in M$, essi abbiano la stessa distanza da x :

$$\|x - y\| = \|x - y'\| = \delta$$

al fine di vedere ciò, è necessario trovare due successioni, $\{y_n\}, \{y'_m\} \in M$ che **minimizzino la distanza**: ossia, due successioni che tendano agli ipotetici due minimi, y e y' . Si definisce dunque una successione come:

$$(y_1, y'_1, y_2, y'_2, \dots)$$

ossia, una successione in cui si prendono alternativamente un termine della prima successione, un termine della seconda. Anche questa è una successione minimizzante: cercando di far tendere questa a δ , quello che si vede è che essa converge e che, siccome tutte le sottosuccessioni di una successione convergente convergono allo stesso limite della successione, è necessario che $y = y'$.

Quando si parla della proiezione y di un punto $x \in H$ su un sottospazio M , usualmente si trova una nomenclatura del tipo:

$$y = \mathcal{P}_M(x)$$

A questo teorema è possibile associare un ulteriore teorema, che fornisce una **caratterizzazione** di spazi. Date infatti le ipotesi appena proposte, vale:

$$y = \mathcal{P}_M(x) \iff x - y \perp M$$

ossia, dire che un certo punto è la proiezione su un sottospazio M , coincide con dire che la differenza tra il punto stesso e la proiezione è ortogonale al sottospazio di proiezione stesso. In altre parole:

$$(x - y|z) = 0 \forall z \in M$$

A questo punto si propone la dimostrazione in entrambi i sensi: prima di tutto, verifichiamo che il fatto che $x - y$ sia ortogonale a M implichi che y è la proiezione $\mathcal{P}_M(x)$. Per fare ciò, si applichi il teorema di Pitagora: $\forall z \in M$, si ha che:

$$\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2$$

essendoci il \geq , si ha che $z = y$ è il punto che minimizza la distanza.

Per quanto riguarda la dimostrazione del viceversa, ossia del fatto che, data y proiezione di un punto x sul sottospazio M , $x - y$ è ortogonale a M . Dato un generico $z \in M$, e un generico numero $\lambda \in \mathbb{R}$, si osserva che:

$$\|x - y\|^2 < \|x - (y + \lambda z)\|^2 = \|x - y\|^2 - 2\lambda \operatorname{Re}\{(x - y|z)\} + \lambda^2 \|z\|^2$$

il primo termine del membro destro si semplifica col membro sinistro, ottenendo dunque:

$$-2\lambda \operatorname{Re}\{(x - y|z)\} + \lambda^2 \|z\|^2 \geq 0$$

a questo punto, fissato un certo z , si vari λ ; si può notare che questa situazione coincide con quella di un trinomio di secondo grado in λ , ossia qualcosa nella forma

$$a\lambda + b\lambda^2 \geq 0$$

in questa situazione, al fine di garantire la validità della disequazione, è necessario richiedere che $a = 0$: infatti, osservando questa funzione in λ , si ha una parabola passante per l'origine, ma dunque una funzione che sicuramente cambierà segno, a meno che a non sia uguale a zero. Dunque, si deve richiedere che:

$$\operatorname{Re}\{(x - y|z)\} = 0$$

in generale non è detto che se questa espressione è vera, per un particolare z , allora si abbia anche che $(x - y|z) = 0$; dal momento che tuttavia è necessario che questa condizione sia verificata per ogni $z \in M$, allora utilizziamo il seguente trucco, già visto in precedenza: si consideri $\vartheta \in \mathbb{R}$ tale per cui:

$$e^{j\vartheta}(x - y|z) \in \mathbb{R}$$

ossia, il termine che elimina la parte immaginaria del segnale, ottenendo:

$$e^{j\vartheta}(x - y|z) = |(x - y|z)|$$

dunque, si ha:

$$(x - y|z) = e^{j\vartheta} e^{-j\vartheta} (x - y|z) = e^{j\vartheta} (x - y|z) e^{j\vartheta}$$

e questo è nullo: se si richiede che la parte reale di questo sia nulla, vuol dire che si richiede che $(x - y|ze^{j\theta})$ sia nullo, e così deve essere, essendo questo reale, e non potendo essere $e^{j\theta} = 0$.

3.8.4 Conseguenze dei risultati proposti

Un primo corollario dei teoremi appena proposti riguarda le caratteristiche dell'operatore di proiezione. Infatti, se L è un sottospazio vettoriale chiuso dello spazio di Hilbert H , allora:

$$\mathcal{P}_L : H \longrightarrow L$$

è un'applicazione lineare, e soddisfa la relazione:

$$\|\mathcal{P}_L(x)\| \leq \|x\|, \forall x \in H$$

ciò è ovvio: la norma della proiezione $\mathcal{P}_L(x)$ è infatti minore di quella di x , essendo presenti solo le componenti di x dello spazio di proiezione. Il fatto che sia lineare, invece, è dovuto al fatto che L è un sottospazio: al posto di x basta proiettare su un $\alpha x_1 + \beta x_2$, e vedere cosa succede:

$$\mathcal{P}_L(\alpha x_1 + \beta x_2) = \alpha \mathcal{P}_L(x_1) + \beta \mathcal{P}_L(x_2)$$

Grazie al teorema di caratterizzazione dimostrato nella sezione precedente, è possibile verificare che la differenza tra questo $\alpha x_1 + \beta x_2$ e la sua proiezione sia ortogonale al sottospazio; verificando l'ortogonalità, si verifica automaticamente che essa è una proiezione (essendoci, in quel teorema, la doppia implicazione). Vogliamo dunque verificare che:

$$\alpha(x_1 - \mathcal{P}_L(x_1)) + \beta(x_2 - \mathcal{P}_L(x_2)) \perp L$$

per fare ciò, si deve verificare che:

$$(\alpha(x_1 - \mathcal{P}_L(x_1)) + \beta(x_2 - \mathcal{P}_L(x_2)))|z) = 0 \forall z \in L$$

tuttavia, usando la linearità del prodotto scalare, si ha che:

$$= \alpha(x_1 - \mathcal{P}_L(x_1))|z) + \beta(x_2 - \mathcal{P}_L(x_2))|z) = 0$$

grazie alle solite osservazioni.

Questa proprietà permette di verificare una proprietà che, come si vedrà tra breve, sarà molto importante quando si studieranno operatori continui:

$$\|\mathcal{P}_L(x) - \mathcal{P}_L(y)\| \leq \|x - y\|$$

ossia, \mathcal{P}_L è un operatore continuo, e anzi Lipschitz, con costante di Lipschitz uguale a 1.

3.8.5 Teorema di decomposizione ortogonale

Dato $S \subset H$, un sottoinsieme (anche non sottospazio vettoriale), si definisce:

$$S^\perp = \{x \in H : (x|y) = 0 \forall y \in S\}$$

questo è detto **sottoinsieme ortogonale**. In uno spazio di H dotato di prodotto scalare, S^\perp è un **sottospazio vettoriale**, ed è sempre **chiuso**.

A partire da questa definizione è possibile proporre il seguente teorema: dato H uno spazio di Hilbert, L un sottospazio vettoriale chiuso, allora:

$$H = L + L^\perp, \quad L \cap L^\perp = \{0\}$$

ossia, ogni vettore dello spazio H può essere scritto come somma di un vettore appartenente al sottospazio L e di uno appartenente al complemento di L . Inoltre, questa decomposizione è **unica**:

$$\forall z \in H \exists! x \in L, y \in L^\perp, : z = x + y$$

in altre parole:

$$H = L \oplus L^\perp$$

ossia, si ha una **somma diretta ortonormale**.

Dimostriamo questo risultato: dato $x \in L \cap L^\perp$, allora:

$$\|x\|^2 = (x|x) = 0$$

infatti, essendo $x \in L \cap L^\perp$, significa che $x \in L^\perp$; di conseguenza, si ha che $(x|x) = 0$. Tuttavia, se $\|x\| = 0$, allora sicuramente $x = 0$, per la proprietà della norma. A questo punto, dato $z \in H$, si consideri:

$$x = \mathcal{P}_L(z)$$

e si consideri $y = z - x$, ossia $z = x + y$; si ha che $x \in L$ per ipotesi, $y \in L^\perp$, dal momento che vale il teorema di caratterizzazione precedentemente visto: si è detto che se x è una proiezione su un certo spazio, di un certo z , e y è definito come la differenza tra questo punto proiettato z e la proiezione x , allora esso appartiene senza dubbio a uno spazio ortogonale a L : il complemento ortogonale L^\perp .

Rimane da verificare l'unicità: dato $z = x_1 + y_1$, $z = x_2 + y_2$, dove $x_1, x_2 \in L$, e $y_1, y_2 \in L^\perp$. Si ha dunque che:

$$z = x_1 + y_1 = x_2 + y_2$$

ma dunque:

$$x_1 - x_2 = y_2 - y_1$$

a membro sinistro si ha la differenza di due elementi in L , dunque essa appartiene a L ; a destra si ha la differenza di due elementi in L^\perp , dunque essa appartiene a L^\perp ; questo significa che, a causa dell'eguaglianza tra i due membri, si ha qualcosa che appartiene sia a L , sia a L^\perp ; di conseguenza, questo è l'elemento nullo! Si ha unicità, dunque, grazie a questo fatto.

Si propongono ora alcuni risultati aggiuntivi a partire da questo.

Corollario

Dato H uno spazio di Hilbert, L un sottospazio, anche non chiuso, allora, dato L^\perp , e $(L^\perp)^\perp$, è possibile dire qualcosa su di esso?

$$(L^\perp)^\perp = \bar{L}$$

infatti, dato un sottoinsieme ortogonale di un altro insieme, esso sarà chiuso; di conseguenza, considerando due volte l'ortogonale di uno stesso insieme, si avrà la **chiusura** di quello di partenza. Al fine di dimostrare ciò, è necessario verificare che ci siano le due inclusioni.

Prima di tutto, dato un elemento appartenente a \bar{L} , si vuole verificare che esso è anche in $(L^\perp)^\perp$. Per fare ciò, dato $x \in \bar{L}$, allora significa che esiste una successione $x_n \in L$, tale per cui $x_n \rightarrow x$. Intendiamo dimostrare che ciò appartiene anche a $(L^\perp)^\perp$. Per fare ciò, si ha che:

$$(x|z) = 0 \forall z \in L^\perp$$

ma, d'altra parte, si ha che:

$$(x|z) = \lim_{n \rightarrow \infty} (x_n|z)$$

x_n è una successione appartenente a L , ma dunque, essendo $z \in L^\perp$, si ha che i prodotti scalari sono tutti nulli. Si può dunque passare al limite.

Per quanto riguarda l'inclusione nell'altro senso, dato $z \in (L^\perp)^\perp$, si vuole verificare che $z \in \overline{L}$. Per fare ciò, utilizziamo la decomposizione ortogonale:

$$\exists x, y : z = x + y$$

dove

$$x \in \overline{L}, y \in (\overline{L})^\perp$$

questo, per come è definito il teorema. Ma allora, bisogna dimostrare che $y = 0$; questo è verificato, dal momento che:

$$(z|y) = (x + y|y) = \|y\|^2$$

ma $z \in \overline{L}, y \in \overline{L}^\perp$, dunque il membro sinistro è nullo; discorso analogo vale per $(x|y)$. Dunque:

$$\|y\|^2 = 0$$

e per la proprietà della norma, $y = 0$.

Consideriamo a questo punto un esempio di sottospazio ortogonale. Si consideri per esempio uno spazio aperto su l^2 , come segue:

$$\text{span} \{e_k, k \in \mathbb{N}, k \neq 0\}$$

dove e_k è la successione che ha tutti zeri tranne nella k -esima posizione, dove ha un 1; si abbia x_k definita come:

$$x_k = \sum_{k=1}^N \alpha_k e_k$$

questa x_k è una successione che avrà, dato un certo N , tutti zeri da $k = N$ in poi; il sottospazio delle x_k è un aperto, dal momento che la sua chiusura è l^2 , e dunque essendo diverso dalla sua chiusura non può essere per definizione chiuso. Il complemento ortogonale di questo spazio è:

$$x \in L^\perp \iff (x|e_k) = 0 \forall k$$

$$x = (x_1, x_2, \dots) \implies (x|e_k) = x_k = 0 \forall k$$

l'insieme dei valori x tali per cui $x_k = 0 \forall k$ è $\{0\}$: l'insieme vuoto. In altre parole, la direzione ortogonale alle altre è la direzione nulla.

Criterio di approssimazione

Prima di procedere con la descrizione di questo criterio di approssimazione, si vuole proporre una premessa. Si supponga di disporre di un sottospazio L in generale non chiuso; ci si pone dunque la seguente domanda: quando un vettore può essere approssimato arbitrariamente bene con un vettore di L ? Ossia, quando esiste una successione a valori in L che approssimi x ? Perché ciò avvenga, è necessario che $x \in \bar{L}$.

Se $y \in L^\perp$, allora y ha prodotto scalare nullo con ogni vettore di L ; se $x \in \bar{L}$, allora y è ortogonale a x ? La risposta è sì: data $x_n \rightarrow x$, si ha che:

$$(x_n|y) = 0$$

allora, passando al limite,

$$\lim_{n \rightarrow \infty} (x_n|y) = (x|y) = 0$$

tuttavia, **vale anche il viceversa**, come stiamo per vedere.

Vale il seguente teorema: dato H uno spazio di Hilbert, L un sottospazio, $x_0 \in H$, quando $x_0 \in \bar{L}$? Ossia, quando esso è approssimabile bene quanto si vuole con un vettore di L ?

La risposta è: sia x_0 tale per cui $(x_0|y) = 0 \forall y \in L^\perp$; allora, si ha che $x_0 \in \bar{L}$.

Questo teorema, al posto di farci cercare una successione in \bar{L} tale per cui essa tenda a x , ci permette di verificare che, per ogni y , $(x_0|y) = 0$. Segue la dimostrazione dire che $(x_0|y) = 0 \forall y$, coincide con dire che x_0 è ortogonale a ogni y , dove ogni y appartiene al sottospazio ortogonale di L ; dunque:

$$x_0 \in (L^\perp)^\perp$$

ma abbiamo dimostrato poco fa che:

$$(L^\perp)^\perp = \bar{L}$$

da qui il teorema.

Criterio di densità

Si consideri uno spazio di Hilbert H , e L un suo sottospazio. Allora, si ha che L è denso se e solo se l'ortogonale è uguale a $\{0\}$:

$$L^\perp = \{0\}$$

in altre parole, se si prende una $y \in L$, si cerca $(x|y) = 0, \forall y \in L$, e si vede che $x = 0$. La dimostrazione di questo teorema è data dalla seguente osservazione:

$$\bar{L} = (L^\perp)^\perp$$

ma, se considero la definizione di densità, ossia se dico che L è denso in H , si ha che $\bar{L} = H$; dunque:

$$\bar{L} = H \implies H = (L^\perp)^\perp$$

a questo punto, si calcoli l'ortogonale di ambo i membri:

$$H^\perp = L^\perp$$

questo significa che un $x \in H^\perp$ deve essere ortogonale a sé stesso; questo significa, in altre parole, che deve essere a misura (norma) nulla, e dunque nullo.

Questo criterio è utile dal momento che permette di verificare che un sistema di vettori sia completo, ossia che il chiuso dello span sia lo span. Dato:

$$\{x_k\} \subset H$$

e

$$(x|x_k) = 0 \forall k \implies x = 0$$

ossia, si deve ha che, se $(x|x_k) = 0$ implica $x = 0$, allora:

$$\overline{\text{span}\{x_k\}} = H$$

in altre parole, il sistema $\{x_k\}$ è completo; infatti, l'unico vettore ortogonale allo span in questa situazione è il vettore nullo, dunque lo span è denso.

Capitolo 4

Funzionali lineari

4.1 Definizione di funzionale lineare su spazi vettoriali

Si consideri uno spazio vettoriale E (nel quale dunque non si abbia per ora una norma, bensì esclusivamente la struttura algebrica) in \mathbb{R} (o in \mathbb{C}); per **funzionale lineare** si intende una applicazione $f : E \rightarrow \mathbb{R}$ (o in \mathbb{C}).

Si propongono ora alcuni esempi di funzionale lineare.

1. Il primo esempio riguarda funzionali su $E = \mathbb{R}^n$. Un esempio di funzionale lineare f può essere:

$$f(x) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

Questo funzionale, nel linguaggio delle matrici, può essere rappresentato mediante una **matrice riga**, che, moltiplicata per un vettore $x \in \mathbb{R}^n$, restituisce uno scalare appartenente a \mathbb{R} . Gli insiemi di livello di questo funzionale $f(x) = \alpha$ sono come in Figura ??: l'insieme degli x tali per cui $f(x) = \alpha$. Dal momento che si ha un'applicazione che parte da uno spazio di dimensione n e si arriva in uno spazio di dimensione 1 (a meno che f non sia il funzionale identicamente nullo), si sa per certo che $\dim \ker(f) = n - 1$; quando $\alpha = 0$, lo spazio che si considera è proprio il nucleo del funzionale f ; al crescere di α , si avranno degli iperpiani paralleli a quello del nucleo; in altre parole, si può anche dire che lo spazio in questione ha codimensione pari a $n - 1$. In altre parole

ancora, unendo la direzione rispetto a cui gli iperpiani sono paralleli con quelle degli iperpiani, si genera l'intero \mathbb{R}^n .

2. Si consideri, sullo spazio $E = l^p$, il funzionale:

$$f(x) = \sum_{k=1}^{\infty} a_k x_k, \quad a_k \in \mathbb{R}(\mathbb{C})$$

dunque, dato un input $x \in l^p$, questo funzionale restituisce in uscita un singolo numero reale. Il funzionale in questione coinvolge il calcolo di una serie infinita; volendo avere la certezza che questa converga, la condizione sufficiente è che $a = (a_1, a_2, \dots, a_n)$ appartenga allo spazio l^q , dove p e q sono esponenti coniugati:

$$\frac{1}{p} + \frac{1}{q} = 1$$

questo, applicando la disuguaglianza di Hölder: se si sceglie q tale da soddisfarla, la serie finale converge assolutamente; inoltre, si ha che:

$$|f(x)| \leq \sum_{k=1}^{\infty} |a_k x_k| \leq \|a\|_{l^q} \|x\|_{l^p}$$

3. Dato lo spazio delle funzioni continue $C([0, 1])$, si consideri una funzione continua $x(t) \in C([0, 1])$ e un funzionale $f(x)$ definito come:

$$f(x) = \int_0^1 x(t)y(t) dt$$

dove $y(t)$ è una certa funzione fissata, scelta in maniera tale che l'integrale abbia senso. Perché l'integrale abbia senso, si deve richiedere che:

$$y(t) \in C([0, 1])$$

ma anche solo:

$$y(t) \in L^1(0, 1)$$

infatti, se la funzione appartiene a L^1 , si ha che:

$$\int_0^1 |y(t)| dt < \infty$$

dunque, si avrà che anche l'altro integrale sarà finito; nel dettaglio, si può dire che:

$$|f(x)| < \int_0^1 |x(t)| |y(t)| dt \leq C \int_0^1 |y(t)| dt$$

infatti, per il teorema di Weierstrass, una funzione continua in un intervallo limitato è limitata e ammette massimo e minimo; di conseguenza, essa è maggiorabile con C , dove C è il massimo della funzione stessa. Volendo essere più fini, si può sicuramente scrivere che:

$$|x(t)| \leq \sup_{t \in [0,1]} |x(t)|$$

e come C si può scegliere proprio questo sup:

$$|f(x)| \leq \sup_{t \in [0,1]} |x(t)| \|y\|_{L^1}$$

Ossia, è possibile in questo modo controllare il funzionale in termini della norma stessa del funzionale.

Un altro esempio, sempre lavorando sullo spazio delle funzioni continue, è il seguente: dato $\alpha \in [0, 1]$, $x \in C([0, 1])$, si può definire un funzionale $f(x)$ in un altro modo, come:

$$f(x) = x(\alpha)$$

dato α fissato, dunque, questo funzionale prende una funzione $x(t)$ e la valuta per $t = \alpha$, ritornando in uscita questo valore. Volendo fare una considerazione, il nucleo di questo funzionale è dato dall'insieme delle funzioni $x(t)$ che presentano uno zero per $t = \alpha$.

4. Si consideri uno spazio di Hilbert, H ; si definisca $x \in H$; e si consideri un funzionale $f(x)$ definito come:

$$f(x) = (x|y)$$

dove $y \in H$ è una certa funzione fissata. Si osservi che questo è un funzionale lineare grazie alla linearità dell'operazione di prodotto scalare, ma se avessimo scritto $(y|x)$ avremmo ancora avuto un funzionale, ma **non** lineare, dal momento che esso sarebbe stato antilineare (avremmo avuto il coniugato). Volendo per esempio possiamo considerare, su $L^2(a, b)$, il seguente funzionale:

$$f(x) = \int_a^b x(t)\overline{y}(t) dt$$

dove al solito y è fissata in L^2 . Questo esempio è simile a prima, ma in realtà si ha una fondamentale differenza: in questo caso la funzione y deve essere a quadrato integrabile, e non L^1 ; questo significa che, anche se il funzionale è quasi identico rispetto al caso precedente, lo spazio dal quale si deve prelevare la funzione y è ben diverso.

5. Un ultimo esempio, dato lo spazio delle funzioni differenziabili una volta, $C^{(1)}([0, 1])$, data $x \in C^{(1)}([0, 1])$, si definisca:

$$f(x) = \left. \frac{dx(t)}{dt} \right|_{t=1}$$

ossia, si consideri la derivata, valutata per $t = 1$.

4.2 Funzionali continui

Ora, si considerino funzionali lineari su di uno spazio normato; il motivo per cui introduciamo una norma, dipende dal fatto che intendiamo ricercare delle condizioni che caratterizzino la continuità di un certo funzionale. Nel dettaglio, ciò che ora mostreremo è che la continuità di un funzionale lineare è garantita dalla possibilità di **stimare**, di **controllare** il funzionale stesso, in termini di una norma.

Si consideri dunque il seguente teorema: dato X uno spazio normato, f un funzionale lineare su X , le seguenti proprietà sono equivalenti:

1. Esiste un $c > 0$, costante, tale per cui:

$$|f(x)| \leq c \|x\| \quad \forall x \in X$$

ossia, a parole, è possibile controllare f mediante una norma. In queste situazioni, si dice che f è un funzionale limitato. Si osservi che la costante c deve essere unica per ogni valore di $x \in X$. Si consideri che in genere un funzionale non è limitato nel senso di insieme limitato: la definizione di limitatezza di un funzionale non è quella di limitatezza di un insieme. Si ha che:

$$f(\alpha x) = \alpha f(x)$$

dunque, se esiste un $x_0 \in X$ in cui $f(x_0) \neq 0$, αx_0 è l'insieme dei punti in cui α cresce; α però non deve per forza assumere valori appartenenti a un insieme limitato: può assumere qualsiasi valore in \mathbb{R} , dunque l'insieme di uscita di un funzionale non è limitato. La limitatezza si intende considerando x in una palla: se l'argomento del funzionale è tale per cui $\|x\| < 1$ (considerando per esempio la palla unitaria), si ha che:

$$|f(x)| < \|x\|$$

è definizione di limitatezza del funzionale.

2. f è continuo $\forall x_0 \in X$
3. f è continuo in $x_0 = 0$

Si noti che l'implicazione da (2) a (3) è assolutamente banale; ciò che è meno banale osservare è il fatto che (3) implica (2): il fatto che un funzionale sia continuo in $x_0 = 0$ garantisce la continuità in ogni punto dell'insieme.

Dimostrazione del teorema

Prima di tutto si dimostra che (1) \implies (2): dato un punto x_0 , per vedere se la funzione è continua, si deve trovare un certo ε tale da controllare la funzione; dunque, il primo passo, è scrivere la differenza tra la funzione in un generico punto x e in x_0 :

$$|f(x) - f(x_0)| = |f(x - x_0)| \leq c \|x - x_0\|$$

infatti, per (1), si ha sostanzialmente che f in modulo può essere maggiorato dalla norma; applicando a ciò, si è dimostrato che f è non solo continuo, ma pure Lipschitz, dove c è la costante di Lipschitz. Scegliendo

$$\varepsilon = \frac{\delta}{c}$$

è possibile dimostrare l'implicazione.

Il secondo punto da dimostrare (saltando (2) \implies (3) che è ovvia) è che (3) \implies (1); questo, si fa per assurdo. Supponiamo che (1) non valga, ossia che non esista una costante c tale per cui:

$$\frac{|f(x)|}{\|x\|} \leq c$$

in altre parole, $\forall c$ si ha, $\forall n$, che esiste x_n tale per cui:

$$\frac{|f(x_n)|}{\|x_n\|} > n$$

$\|x_n\|$ è un numero, dunque può essere portato dentro il valore assoluto (essendo un numero positivo) e dentro f (per linearità del funzionale), ottenendo:

$$\frac{|f(x_n)|}{\|x_n\|} = \left| f\left(\frac{x_n}{n\|x_n\|}\right) \right| > 1$$

si definisca y_n come segue:

$$y_n = \frac{x_n}{n\|x_n\|}$$

si ha che:

$$\|y_n\| = \frac{\|x_n\|}{\|x_n\|n} = \frac{1}{n}$$

dunque, si ha una successione che tende a 0, ma i valori di y_n devono essere per la disequazione tutti maggiori di 1; questo è un assurdo. Se infatti $y_n \rightarrow 0$, $f(y_n) \not\rightarrow 0$, dovremmo avere che $f(y_n) \rightarrow f(0)$, ma ciò non capita.

4.2.1 Norma di un funzionale limitato

Per i funzionali lineari è possibile introdurre una definizione di **norma**. Dato X uno spazio normato, f un funzionale lineare limitato, allora si definisce:

$$\|f\| = \inf \{c > 0 : |f(x)| \leq c\|x\| \forall x \in X\}$$

Ossia, essa è **la più piccola costante che maggiori** $|f(x)|$. In altre parole, questo può essere definito come:

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\| \leq 1} |f(x)|$$

Le prime due valgono se $X \neq \{\emptyset\}$, ossia solo se X non coincide con l'insieme vuoto. Dimostriamo la validità di queste disequaglianze:

$$\frac{|f(x)|}{\|x\|} \leq c \forall x \in X$$

questo coincide con il dire che c è il sup di

$$c = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}$$

infatti, c non è altri che il minimo dei maggioranti dell'insieme $\frac{|f(x)|}{\|x\|}$; il minimo dei maggioranti è proprio il sup, per definizione. Si osservi che questo inf è anche un min, ossia è anche un minimo: l'inf dei maggioranti infatti è ancora un maggiorante, dal momento che appartiene all'insieme, dunque è un minimo; trovarlo, come vedremo tra breve, non è tuttavia banale. Data dunque $\|f\| = c$, si ha che:

$$|f(x)| \leq \|f\| \|x\| \quad \forall x$$

Discutiamo a questo punto la seconda eguaglianza, sfruttando ancora una volta il trucco precedentemente utilizzato, ossia portare la norma di x dentro all'argomento del funzionale:

$$\frac{|f(x)|}{\|x\|} = \left| f \left(\frac{x}{\|x\|} \right) \right| = |f(y)|, \|y\| = 1$$

quindi, dato:

$$\left\{ \frac{x}{\|x\|} : x \neq 0 \right\}$$

di fatto, è la stessa cosa di scrivere:

$$\{y : \|y\| = 1\}$$

dunque, sostituendo a x la y , si ottiene la seconda eguaglianza:

$$= \sup_{\|x\|=1} |f(x)|$$

Infine, si ha l'ultima eguaglianza: questa richiede solo una piccola osservazione. Se $\|x\| \leq 1$, invece che valere 1, si ha che:

$$|f(x)| < \left| f\left(\frac{x}{\|x\|}\right) \right|$$

essendo $\|x\| \leq 1$; di conseguenza, è come fare il sup dell'espressione di prima: tutti i punti interni non contribuiscono al calcolo del sup.

Si osservi che tutti questi ragionamenti sono stati fatti su $|f(x)|$ e non su $f(x)$: questo è necessario dal momento che, se i numeri in questione possono essere complessi, si ha la possibilità di dover maggiorare un numero complesso, e questo non ha assolutamente senso: non è possibile comparare due numeri complessi se non in modulo. Se invece fossimo in \mathbb{R} , sarebbe possibile trarre la seguente conclusione:

$$f(x) < c \|x\|$$

ma, essendo f lineare, se mettiamo $-x$ al posto di x , abbiamo:

$$f(-x) < c \|-x\| = \|x\|$$

ma

$$f(-x) = -f(x)$$

dunque:

$$f(x) > -c \|x\|$$

che, messa insieme all'altra, produce di nuovo la diseguaglianza col modulo:

$$|f(x)| < c \|x\|$$

4.2.2 Interpretazione geometrica della norma di un funzionale lineare

Si consideri un funzionale $f \neq 0$; ricordiamo che la definizione di norma è:

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}, \quad X \neq \{\emptyset\}$$

A questo punto, si applichi il seguente trucco: si dividano sia il numeratore sia il denominatore per $f(x)$:

$$= \sup_{x \neq 0} \frac{1}{\frac{\|x\|}{|f(x)|}}, \quad x \neq 0, f(x) \neq 0$$

sembrerebbe a prima vista che abbiamo aggiunto un vincolo: $f(x) \neq 0$; in verità, questo non è un problema, dal momento che $f(x) = 0$, se il funzionale f non è identicamente nullo, non può contribuire in alcun modo al calcolo del sup, dunque in verità non si sta escludendo alcun punto. Si definisca ora:

$$y = \frac{x}{f(x)}$$

si può osservare che:

$$f(y) = \frac{f(x)}{f(x)} = 1$$

infatti, al numeratore si ha x che è un vettore (su cui il funzionale può agire), a denominatore uno scalare (l'uscita di un funzionale), dunque il risultato è proprio 1. Questo significa che y di per sé non ha norma pari a 1, ma il valore del funzionale f applicato a y vale 1. Quindi, è possibile scrivere che questa norma è uguale a:

$$\|f\| = \sup \left\{ \frac{1}{\|y\|} : f(y) = 1 \right\}$$

d'altra parte, si può osservare che questo coincide con il ragionare esattamente allo stesso modo di prima, quando si era considerato $\frac{x}{\|x\|}$. Dunque, è possibile dire che questo coincide con lo scrivere:

$$\|f\| = \frac{1}{\inf \{ \|y\| : f(y) = 1 \}}$$

ossia, si sfrutta la proprietà per cui il sup di un numero coincide con il reciproco dell'inf del numero reciproco.

Si può tuttavia osservare che l'insieme dei punti y tali per cui $f(y) = 1$ non è altri che un insieme di livello del funzionale: un iperpiano. Questo

non è uno spazio vettoriale, dal momento che non passa per l'origine (si può pensare come uno spazio affine, ossia traslato); si considerino dunque tutti gli y appartenenti a questo iperpiano, come in Figura ???. Dati tutti gli y in questo iperpiano, calcolandone le norme, ciò che si fa sostanzialmente è calcolare le distanze dei vari punti dall'origine O dello spazio; data δ la distanza minima, essa coincide con l'inf di cui si stava parlando; di conseguenza, si può dire che:

$$\|f\| = \frac{1}{\delta}$$

Ossia, la norma del funzionale f non è altri che il reciproco della minima distanza tra l'insieme di livello e l'origine.

SI consideri ora per esempio la situazione di Figura ???, e si cerchi di trarre alcune conclusioni. In figura sono presentate le curve di livello dei funzionali e la sfera unitaria. Dati i funzionali $f(x)$ e $g(x)$, ciò che si può immediatamente dire a occhio è il fatto che:

$$\|f\| < 1$$

e

$$\|g\| > 1$$

infatti, nel primo caso, si ha che la distanza δ è sicuramente maggiore di 1, dal momento che il punto della prima curva di livello a minima distanza dall'origine dello spazio certamente non è nella sfera unitaria; discorso opposto vale per il secondo caso: essendo δ_g all'interno della sfera unitaria, certamente si ha che $\|g\| > 1$.

L'idea dietro alle norme di funzionali dunque è la seguente: se la norma del funzionale è piccola, significa che vi sono punti molto distanti dall'origine, dove il risultato ritornato dal funzionale vale 1; allo stesso modo, se le norme sono grandi, i punti sono vicini all'origine.

Esistenza del punto di massimo

Si è detto che:

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}$$

il sup può non essere un punto di massimo; questo significa che non è detto che esista un certo x tale per cui il valore dell'espressione sia proprio uguale alla norma. A questo punto dunque ci si pone la seguente domanda: come si fa per calcolare la norma di un certo funzionale lineare?

Prima di tutto, si ha che, se il funzionale è continuo:

$$|f(x)| \leq c \|x\| \quad \forall x \in X$$

questo significa che, se si richiede di calcolare la norma di un funzionale¹:

1. Si mostra prima di tutto che il funzionale è continuo; questo significa dunque che esiste una certa c tale da soddisfare l'espressione appena scritta:

$$|f(x)| < c \|x\|$$

2. Se la c è stimata *al minimo*, ossia se essa è stata scelta in maniera tale da non abbondare troppo, ossia se la maggiorazione è la "minima" possibile (la stima al ribasso), allora è possibile che c sia effettivamente la costante minima.

Per verificare che vale l'eguaglianza, dovrebbe essere necessario verificare che esista un certo x_0 tale per cui

$$|f(x_0)| = c \|x_0\|$$

se si riesce dunque a trovare questo particolare x_0 , esso è un punto di massimo; dunque, in questa situazione,

$$\|f\| = c$$

A questo punto, è necessario "andare a caccia" di questo x_0 ; certe volte può essere facile, certe volte meno. Può essere poi che si abbia un certo "candidato" c , ma che non si riesca a trovare x_0 . In queste situazioni, ciò che si può fare (verrà proposto a breve un esempio) è prendere un certo ε , e cercare un x_0 tale da trovarsi tra c e $c - \varepsilon$; se ciò si riesce a fare $\forall \varepsilon$, allora il gioco è comunque fatto.

¹cosa che non sempre si può fare, ma anzi di solito è fattibile solo in casi piuttosto particolari

4.2.3 Esempi di calcolo di norme

Si considera a questo punto un certo numero di esempi; questi avranno la stessa numerazione di quelli proposti a inizio capitolo, di conseguenza considerazioni aggiuntive possono essere fatte rivedendo quelli.

Esempio 1

Si richiede di calcolare la norma del funzionale definito come:

$$f(x) = a_1x_1 + a_2x_2 + \dots$$

Dal momento che questo funzionale era definito su \mathbb{R}^n , come norma associata allo spazio si può considerare la norma euclidea:

$$\left(\sum_{k=1}^n x_k^2 \right)^2$$

La prima domanda da porci a questo punto è: f è un funzionale continuo? La risposta è sì: grazie alla disuguaglianza di Cauchy-Schwarz, o alla disuguaglianza di Hölder per $p = 2$, è possibile dire che:

$$|f(x)| \leq \left(\sum_{k=1}^n a_k^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^n x_k^2 \right)^{\frac{1}{2}}$$

Da qui segue che f è un funzionale continuo, e che la norma del funzionale è:

$$c = \|a\|_{l^2}$$

questo almeno è un sospetto: la c appena scelta è una maggiorante a caso, o è anche la *migliore*? Dal momento che è stata una disuguaglianza molto “minimale”, ossia una maggiorazione molto blanda, può essere che c sia la “migliore costante”. In questo caso è piuttosto semplice verificare che l’eguaglianza è verificata, e determinare il x_0 in questione. Vogliamo dunque trovare:

$$|f(x_0)| = \|a\|_{l^2} \|x_0\|$$

da un lato si ha che:

$$|f(x_0)| = \left| \sum_{k=1}^n a_k x_k \right|$$

all'altro membro si ha:

$$\|a\|_{l^2} \|x_0\| = \left(\sum_{k=1}^n a_k^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^n x_k^2 \right)^{\frac{1}{2}}$$

ma se noi scegliamo:

$$x = a$$

si ha che:

$$|f(x)| = \sum_{k=1}^n a_k^2$$

ma dunque l'eguaglianza è soddisfatta, dove la costante c è quella decisa.

Esempio 2

Il secondo esempio riguarda un funzionale definito su l^p :

$$f(x) = \sum_{k=1}^{\infty} a_k x_k$$

dove $a = (a_1, a_2, \dots) \in l^q$, e

$$\frac{1}{p} + \frac{1}{q} = 1$$

era stato precedentemente verificato il fatto che:

$$|f(x)| \leq \|a\|_{l^q} \|x\|_{l^p} \quad \forall x \in l^p$$

questo ci garantisce la continuità del funzionale; ossia:

$$\|f\| \leq \|a\|_{l^q}$$

questa è la nostra candidata. Dal momento che, di nuovo, abbiamo usato la disuguaglianza di Hölder, che è "fine", allora possiamo ipotizzare che valga

anche l'eguaglianza. Per dimostrarlo, cerchiamo una successione $x \in l^p \neq 0$, sapendo che:

$$|f(x)| = \|a\|_{l^q} \|x\|_{l^p}$$

da una parte abbiamo dunque che:

$$|f(x)| = \left| \sum_{k=1}^{\infty} a_k x_k \right|$$

mentre, all'altro membro:

$$\|a\|_{l^q} = \left(\sum_{k=1}^{\infty} a_k^q \right)^{\frac{1}{q}}$$

questa successione, per $1 < p \leq \infty$, deve permetterci di ottenere l'eguaglianza.

Alcune considerazioni: quando si hanno delle situazioni in cui si hanno dei \leq , spesso capita che si abbiano delle compensazioni, dovute al fatto che le sommatorie sono a segni alterni; di conseguenza, supponiamo di avere sempre $a_k x_k \geq 0$. In secondo luogo, proponiamo una scelta per gli x_k :

$$x_k = \begin{cases} \frac{|a_k|^q}{a_k}, & a_k \neq 0 \\ 0, & a_k = 0 \end{cases}$$

Data questa successione x_k ; si ha infatti che:

$$|f(x)| = \left| \sum_{k=1}^{\infty} a_k \frac{|a_k|^q}{a_k} \right| = \|a\|_{l^q}^q$$

Dunque, vediamo cosa si ha a secondo membro:

$$\|x\|_{l^p} = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}}$$

dove gli x_k sono stati appena scritti; dunque:

$$\left(\sum_{k=1}^{\infty} |a_k|^{p(q-1)} \right)^{\frac{1}{p}}$$

ma, essendo p e q esponenti coniugati, si ha:

$$(q - 1)p = q$$

da cui:

$$= \|a\|_{l^q}^{\frac{q}{p}}$$

questo inoltre è finito, dal momento che $a \in l^q$. Dunque, è sicuramente vero che:

$$|f(x)| = \|a\|_{l^q} \|x\|_{l^p}$$

oppure no? Beh:

$$\|x\|_{l^p} = \|a\|_{l^q}^{\frac{q}{p}}$$

dunque, gli esponenti:

$$1 + \frac{q}{p} = q$$

da cui, si ottiene proprio l'eguaglianza desiderata.

Rimane non discusso il caso $p = 1$; in altre parole, è vero che:

$$|f| = \|a\|_{l^\infty}$$

oppure no? La risposta è: questa volta, non si riesce a trovare agevolmente una x tale per cui si veda proprio verificata l'eguaglianza. Al fine di verificare ciò, si provi con un particolare esempio:

$$a_k = 1 - \frac{1}{k}$$

in questo caso, si avrebbe:

$$f(x) = \sum_{k=1}^{\infty} a_k x_k = 0x_1 + \frac{1}{2}x_2 + \frac{2}{3}x_3 + \frac{3}{4}x_4 + \dots$$

questa, sta in l^∞ . Si ha inoltre che:

$$|f(x)| \leq \|a\|_{l^\infty} \|x\|_{l^1}$$

infatti:

$$|f(x)| = \left| \sum_{k=1}^{\infty} \left(1 - \frac{1}{k}\right) x_k \right| \leq \sum_{k=1}^{\infty} |x_k|$$

infatti, volendo maggiorare a_k con la sua norma, si ha che:

$$\|a\|_{l^\infty} = \sup \left| 1 - \frac{1}{k} \right| = 1$$

dunque, si ha, scrivendo per esteso membro sinistro e membro destro:

$$0x_1 + \frac{1}{2}x_2 + \frac{2}{3}x_3 + \dots \leq |x_1| + |x_2| + |x_3| + \dots$$

sicuramente, questa disequaglianza è vera; tuttavia, il $=$ non è assolutamente verificato: non esiste una x non nulla tale per cui questa possa diventare un uguale! Tutti i termini a sinistra sono senza dubbio più piccoli di quelli a destra, ma strettamente più piccoli!

Tuttavia, possiamo giocare con ε : definire una certa soglia, a cercare di trovare un punto che stia tra la c trovata e $c - \varepsilon$. Quello che si può dire è che:

$$\forall \varepsilon \exists k : a_k > 1 - \varepsilon$$

dunque, si cerca un k_0 tale per cui si cada tra 1 (che è la c : $c = 1$) e $1 - \varepsilon$. Dunque, se prendiamo come x la successione di tutti zeri tranne nella k -esima posizione, si ha che:

$$|f(x)| = \left| \sum_{k=1}^{\infty} |a_k x_{k_0}| \right| = |a_{k_0}|$$

dove k_0 è proprio quel k_0 tale per cui si ha tutto soddisfatto. Essendo x composta da un unico elemento non nullo,

$$\|x\|_{l^1} = 1$$

dunque:

$$\frac{|f(x)|}{\|x\|} > 1 - \varepsilon$$

e dunque, si ha che $\|f\|$ è effettivamente uguale a $c = 1$.

Riassumendo, in questo ultimo step è stato fatto ciò: si ha una successione come $1 - \frac{1}{k}$; questa, per un k_0 sufficientemente elevato, sarà tale da valere meno di $1 - \varepsilon$, per ogni ε . Dunque, scegliendo come x un vettore di zeri tranne che nel k_0 -esimo posto, dove k_0 è scelto in modo tale da essere oltre $1 - \varepsilon$, il problema è risolto.

Esempio 3

Si consideri a questo punto il funzionale:

$$f(x) = \int_0^a x(t)y(t) dt, y, y \in L^1(0, 1), x \in C([0, 1])$$

Abbiamo che:

$$|f(x)| \leq \|x\|_C \|y\|_{L^1}$$

ora: dal momento che il funzionale va da $C([0, 1])$ a \mathbb{R} è continuo, con costante

$$c = \|y\|_{L^1}$$

dunque, abbiamo che:

$$\|f\| \leq \|y\|_{L^1}$$

Anche in questo caso si può dimostrare che vale l'eguaglianza, tuttavia ciò non verrà fatto in quanto abbastanza complicato.

Si era discusso anche un secondo esempio, parlando di funzioni continue: $f(x) = x(\alpha)$. Prima di tutto, verifichiamo che questo funzionale sia continuo: questo si può fare come segue

$$|f(x)| = |x(\alpha)| \leq \sup_{t \in [0,1]} |x(t)| = \|x\|_C$$

dunque, f è continuo, con costante $c = 1$ (dal momento che a destra c'è solo la norma su C).

A questo punto, si consideri la seguente funzione $x(\alpha)$, al fine di verificare l'eguaglianza:

$$|x(\alpha)| = \sup_{t \in [0,1]} |x(t)|$$

ossia una funzione tale per cui in α assume il valore massimo. In questa situazione, si ha che:

$$|f(x)| = |x(\alpha)| = \|x\|_C$$

e il gioco è fatto.

Esempio 4

Dato H uno spazio di Hilbert, si consideri il funzionale definito dal prodotto scalare:

$$f(x) = (x|y)$$

in questo caso la continuità è garantita:

$$|f(x)| \leq \|y\| \|x\|$$

segue che la stima che dà la continuità è $\|y\|$. Anche in questo caso, è possibile ricavare l'eguaglianza, per

$$x = y$$

in questo caso, infatti,

$$f(x) = \|x\|^2$$

e dunque a secondo membro abbiamo, come a primo membro,

$$\|y\| \|x\| = \|y\|^2$$

Esempio 5

Finora abbiamo considerato tutti esempi di funzionali continui; non è assolutamente detto, tuttavia, che un funzionale debba per forza essere continuo. Ora si considererà un esempio di funzionale non continuo, e verranno proposte alcune affermazioni aggiuntive. Dato lo spazio delle funzioni derivabili una volta, $x \in C^{(1)}([0, 1])$, si ha:

$$f(x) = \left. \frac{dx(t)}{dt} \right|_{t=1}$$

come norma è naturale usare:

$$\|x\| = \sup_{t \in [0,1]} |x(t)|$$

in questo caso, f non è continuo; infatti, non è possibile trovare una certa $c > 0$ tale per cui:

$$|f(x)| \leq c \|x\|$$

infatti:

$$|x'(1)| \not\leq c \sup_{t \in [0,1]} |x(t)| \quad \forall x$$

Al fine di vedere ciò, supponiamo per assurdo che questa c esista. Dunque, cerchiamo una successione $x_n(t)$ per cui il membro sinistro sia molto grande, quello destro molto piccolo; se riuscissimo a ricadere in una situazione del genere, allora la disuguaglianza non avrebbe senso, e cadremmo in una situazione assurda.

Dunque, in questo caso, $x_n(t)$ deve essere una successione di funzioni limitate, la cui pendenza però aumenti al crescere di n . Un esempio è:

$$x_n = t^n$$

questa è una funzione limitata, ma la cui derivata cresce al crescere di n per $t = 1$:

$$\|x_n\| = 1 \quad x'_n(1) = n$$

perché tutto abbia senso, dovrei avere:

$$n \leq c \times 2 \forall n$$

questo è un assurdo: non esiste una c che possa essere maggiore di n , per ogni n .

Dunque, in questo caso, il funzionale non è certamente continuo.

A questo punto possiamo provare a capire a cosa serve qualificare la continuità di un funzionale: dato un funzionale continuo, perturbando di poco la funzione che gli viene data in ingresso, x , allora il valore del funzionale risulterà essere poco perturbato; la continuità garantisce dunque una certa

“stabilità” del funzionale al variare dell’input; se questa non si ha, questa stabilità allo stesso modo sparisce.

È possibile *rendere continuo* un funzionale: per fare ciò, è necessario **cambiare norma**; nell’ambito degli spazi topologici questo può essere visto come una sorta di cambio di topologia; è infatti noto che se si cambia la topologia, passando da una a una più fine, è “più probabile” che l’applicazione con la topologia più fine sia continua.

L’idea, in questo caso, è quella di utilizzare la seguente norma:

$$\|x\| = \sup_{t \in [0,1]} |x(t)| + \sup_{t \in [0,1]} |x'(t)|$$

in questo caso, è “più probabile” che il funzionale sia continuo; infatti, stiamo prendendo una “norma più grande”. Infatti, ora stiamo maggiorando anche con $x'(1)$, col suo sup, dunque stiamo “allargando” il bound. In questo caso, si ha che:

$$|x'(1)| \leq c + \|x'\|$$

ossia, con questa norma, è possibile maggiorare il funzionale. Non è sufficiente definire la norma solo come sup della derivata, dal momento che la derivata può annullarsi, di conseguenza se non vi si somma la norma della funzione, si finisce per avere solo una pseudonorma.

4.3 Spazi duali

Lo spazio dei funzionali continui è detto **spazio duale**. Si tratta di uno spazio normato, dove la norma è quella appena definita. Dato dunque X uno spazio normato, per **spazio duale** di X , denotato con X^* , si intende l’insieme dei funzionali definiti su X , dotato delle operazioni:

$$(f + g)(x) = f(x) + g(x)$$

e

$$(\alpha f)(x) = \alpha f(x)$$

se dunque f, g sono funzionali lineari continui, allora la loro somma è ancora un funzionale continuo. In questo spazio, si ha anche una norma $\|f\|$ definita come:

$$\|f\| = \inf \{c > 0 : |f(x)| \leq c\|x\| \forall x \in X\} = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|$$

si può dimostrare che questa è effettivamente una norma, confermando dunque il fatto che X^* è uno spazio normato.

Esiste un teorema che afferma che lo spazio duale di X , X^* , è **sempre** uno spazio di Banach, anche se lo spazio X è non completo.

4.3.1 Problema dell'estensione di funzionali lineari continui

Dato un funzionale lineare f_0 continuo definito su X_0 , il quale è un sottospazio di X ($X_0 \subset X$), un problema che ci si può porre può essere quello di estendere f_0 allo spazio ambiente; questo problema è molto simile a quello affrontato precedentemente per quanto riguarda le funzioni: data una funzione definita su un certo insieme, estenderla per continuità a un insieme più grande.

In questo caso si ha a che fare anche con l'esistenza di una norma, per questo funzionale; ciò che si intende fare è determinare l'esistenza di un funzionale che non solo sia un'estensione del precedente (ossia che, se valutato sull'insieme di partenza, restituisca valori che coincidano con quelli del funzionale da cui si è partiti), ma che abbia anche **norma minima**. Dal momento che si considera un insieme più grande rispetto a quello di partenza, ci si può aspettare che la norma possa essere maggiore o uguale della norma di partenza; il caso ottimo dunque è quello in cui la norma non varia rispetto a quella dello spazio di partenza. Si ha un primo risultato, che rappresenta un'estensione al classico teorema di estensione per continuità, e poi un secondo risultato, più importante e generale.

Si parta dal primo teorema: dato X uno spazio normato, dato $X_0 \subset X$ un suo sottospazio, dato $f_0 \in X_0^*$ (ossia, f_0 è un elemento dello spazio duale, il che significa che esso è un funzionale continuo), si vuole trovare un funzionale $f \in X^*$ tale che estenda f_0 , ossia tale per cui:

$$f(x) = f_0(x), \forall x \in X_0$$

in altre parole, si vuole che la restrizione del funzionale f su X_0 coincida con f_0 :

$$f|_{X_0} = f_0$$

inoltre, si ha qualcosa in più:

$$\|f\|_{X^*} = \|f_0\|_{X_0^*}$$

ossia, si ha che la norma del funzionale esteso coincide con la norma del funzionale di partenza.

Esiste un teorema che afferma che questo problema ha un'unica soluzione, quando X_0 è denso in X : in questo caso, infatti, il teorema rappresenta una banale estensione del teorema di estensione visto precedentemente. Vediamo la dimostrazione di questo fatto: dato $f_0 \in X_0^*$, si ha che:

$$\exists c > 0 : |f_0(x)| \leq c \|x\|, \forall x \in X_0$$

ossia, dati $x, y \in X$, $f_0(x)$ è Lipschitz; questo si può verificare come segue:

$$|f_0(x) - f_0(y)| = |f_0(x - y)| \leq c \|x - y\|$$

infatti, f_0 è un operatore lineare, dunque è possibile portare entrambi i termini dentro l'argomento; infine, si è usata la continuità, per il secondo passaggio. Dunque, f_0 è Lipschitz, ma non solo: si ha che $f_0 : X_0 \rightarrow \mathbb{R}$ (oppure, \mathbb{C}); applicando il teorema di estensione continua, dal momento che lo spazio di arrivo è sempre uno spazio completo, f_0 presenta un'unica estensione continua, $f : X \rightarrow \mathbb{R}$. Utilizzando la teoria precedentemente vista, inoltre, è possibile anche costruire il funzionale esteso. Infatti, analogamente a prima, si può dire che, data $x_n \in X_0$ una successione di Cauchy convergente a limite $x \in X$, si può definire f come:

$$f(x) = \lim_{n \rightarrow \infty} f_0(x_n), x_n \in X_0, x_n \rightarrow x$$

Grazie al fatto che X_0 è denso in X , rispetto a prima, questo funzionale è contenuto in X^* , e ovviamente è pure lineare (per verificare ciò è sufficiente sostituire a x un βx , e quindi a x_n una successione βx_n). Per quanto riguarda la norma, si vorrebbe che $\|f\| = \|f_0\|$; tuttavia, grazie al teorema di estensione continua, è noto che f_0 è Lipschitz, ma anche f è Lipschitz, per di più con la stessa costante c di prima; in altre parole, si ha che:

$$\|f\| = \|f_0\| = c$$

infatti:

$$|f(x) - f(y)| \leq \|f_0\| \|x - y\|$$

dal momento che non è possibile che la norma sia strettamente minore di $\|f_0\|$, essendo lo spazio X più grande di X_0 , si avrà forzatamente l'uguaglianza!

4.4 Teorema di Hahn-Banach

Il risultato presentato nella sezione precedente consente di garantire l'esistenza e l'unicità di un'estensione f su uno spazio X a partire da un funzionale f_0 definito su $X_0 \subset X$; questo, era basato sull'ipotesi di densità di X_0 in X . Tuttavia, questo teorema è sovrabbondante: infatti, non è necessario, per garantire l'esistenza di questo funzionale esteso, il fatto che lo spazio di partenza sia denso in quello di arrivo. Questo fatto è affermato dal teorema di Hahn-Banach.

Il teorema di Hahn-Banach garantisce che, dato X spazio normato, $X_0 \subset X$ un suo sottospazio, dato un funzionale $f_0 \in X_0^*$, esiste una sua estensione $f \in X^*$ dotata della stessa norma. Alcune osservazioni: il fatto che la norma sia la stessa garantisce che il funzionale *non cresce* eccessivamente. In secondo luogo, un'osservazione sull'unicità: questo teorema non garantisce l'unicità del funzionale esteso. Infine, una nota su questo teorema: esso cade, se non si accetta la validità dell'assioma della scelta.

Questo teorema non verrà dimostrato, ma verranno a questo punto discusse le sue principali conseguenze.

4.4.1 Conseguenze del teorema di Hahn-Banach

Esistenza di iperpiani di supporto

La prima conseguenza del teorema di Hahn-Banach è l'esistenza di iperpiani di supporto. Questo significa che, dato $x_0 \in X$, dove x_0 non è per ipotesi l'elemento nullo ($x_0 \neq 0$), allora esiste un funzionale f lineare continuo su X di norma 1, tale per cui:

$$\exists f : \|f\| = 1, f(x_0) = \|x_0\|$$

Alcune osservazioni riguardo a ciò: se $\|f\| = 1$, allora, dalla definizione di norma, si ha che:

$$|f(x)| \leq \|f\| \|x\| = 1 \times \|x\|$$

quindi, in modulo, si può dire che, se $f(x_0) = x_0$, allora $|f(x)|$ assume un massimo in x_0 , essendo il punto a norma unitaria: se $f(x_0) = \|x_0\|$, ma se

$$|f(x_0)| \leq \|f\| \|x_0\|$$

e in verità il \leq è un $=$, si ha che su questo punto si ha un massimo.

È possibile considerare anche una interpretazione geometrica: dato X spazio vettoriale reale, si chiami:

$$\delta = \|x_0\|$$

considerando l'esempio nel piano di Figura ??, l'insieme di livello del funzionale in x_0 sarà:

$$\{x \in X : f(x) = \delta\}$$

si può dire che l'iperpiano passi per x_0 : $|f(x_0)| = \|f\|$. Si può dire (e si può capire anche dal disegno) che l'iperpiano **supporta la palla**: infatti, tutta la palla è contenuta nel semispazio a sinistra dell'iperpiano.

Più formalmente: la palla è contenuta nel semispazio $\{x : f(x) \leq \delta\}$; infatti, nell'origine il funzionale si annulla, dunque si ha senza dubbio il \leq ; essendo contenuta, si può dire che la palla è contenuta nell'insieme:

$$\|x\| \leq \delta$$

ma dunque, essendo $|f(x)| \leq \|x\|$, si ha che:

$$f(x) \leq \delta$$

Riscriviamo tutti questi passi una volta per tutte:

$$f(x) \leq |f(x)| \leq \|f\| \|x\| = 1 \times \|x\| \leq \delta$$

infatti, il modulo maggiora il suo argomento, ma tutto ciò con la definizione di continuità di funzionale può essere maggiorato dalla norma del funzionale per la norma di x , e la norma di x al massimo è δ .

Ora, questa conseguenza verrà dimostrata. La dimostrazione è articolata in due passi:

- il primo passo consiste nella costruzione di un funzionale definito solo sullo span di X_0 ;
- il secondo passo è una banale applicazione del teorema di Hahn-Banach, per estendere X_0 allo spazio ambiente X .

Per quanto riguarda il primo passo, si consideri dunque:

$$X_0 = \text{span} \{x_0\} : \|f_0\| = 1, f_0(x_0) = \|x_0\|$$

si ha dunque che:

$$f_0 = \lambda x_0, \lambda \in \mathbb{R}$$

infatti, dire che si viaggia sullo span di x_0 , significa considerare un generico elemento λx_0 , dove λ può assumere qualsiasi valore. Questo significa viaggiare su un segmento come in Figura ??: al variare di λ ci si muove su una retta passante per l'origine (caso in cui $\lambda = 0$) e x_0 (caso in cui $\lambda = 1$, potendo assumere qualsiasi valore su di essa). Si richiede dunque che:

$$f_0(\lambda x_0) = \lambda \|x_0\|$$

se dunque $\lambda = 1$, si ha che:

$$f(\lambda x_0) = \lambda f(x_0)$$

ma:

$$f(x_0) = \|x_0\|$$

dunque, con $\lambda = 1$, si ha:

$$f(\lambda x_0) = \lambda \|x_0\| = \|x_0\|$$

e questo è quello che si voleva.

I funzionali lineari continui separano i punti di X

Si supponga di non poter conoscere precisamente per qualche motivo quali sono gli elementi appartenenti a un insieme, ma invece di poter usufruire con facilità dei funzionali di questi; per esempio, gli elementi degli insiemi possono essere successioni, dunque oggetti complicati da gestire, mentre i funzionali

sono semplicemente dei numeri reali. Questa conseguenza del teorema di Hahn-Banach permette di garantire il fatto che, dati due elementi distinti di un insieme X , esiste un funzionale lineare continuo che assume valori diversi nei due punti. Questo permette di **distinguere** i punti, ossia gli elementi dell'insieme: è possibile che due elementi siano infatti diversi, ma che il funzionale valutato in essi ritorni lo stesso valore, rendendo di fatto gli elementi indistinguibili, se si dispone solo dei funzionali, ossia se si *guarda lo spazio solo attraverso i funzionali*. Formalmente:

$$x, y \in X, x \neq y, \exists f \in X^* : f(x) \neq f(y)$$

si osservi che questo non è un risultato legato all'iniettività: ciò che ci è garantito da questo risultato è semplicemente il fatto che, dati due valori distinti, sicuramente esiste un funzionale che permetta di *distinguere* i due punti; ciò che non è detto, tuttavia, è che questo funzionale sia unico per ogni punto: può essere che dati due punti $x \neq y$, il funzionale $f_1 \in X^*$ sia in grado di distinguerli, ma che non sia in grado di distinguere altri due punti, $v \neq w \neq x \neq y$. D'altra parte, questo risultato garantisce che esisterà un qualche altro funzionale, f_2 , che possa distinguere i due punti.

Esistenza di funzionali lineari continui non banali

Dato $X \neq \{\emptyset\}$, allora sicuramente $X^* \not\equiv \{\emptyset\}$. Ciò può essere dimostrato facilmente a partire dalla prima delle conseguenze qui elencate, poiché se $x_0 \neq 0$, si ha che esiste sicuramente un certo $f \in X^*$ tale per cui $f(x_0) = \|x_0\|$, e questo deve essere diverso da zero; in altre parole, il funzionale f **non può essere identicamente nullo**; questo è garantito dal teorema di Hahn-Banach.

L'insieme duale di X è totale

Si dice che X^* è un insieme totale; questo significa che esso è *grande*: esistono abbastanza funzionali continui da poter distinguere ogni punto. Dunque, dato $x_0 \in X$, il fatto che $f(x_0) = 0 \forall f \in X^*$ garantisce che $x_0 = 0$. Questo si può anche scrivere in un altro modo:

$$\bigcup_{f \in X^*} \ker \{f\} = \{0\}$$

questo è molto simile a quanto visto precedentemente: si era infatti detto che esiste sicuramente un funzionale in grado di distinguere due punti; ciò ricorda questo, dal momento che si può pensare di distinguere un punto dall'insieme nullo; ciò che questo risultato dice è che solo se un valore è nullo, tutti i funzionali di X^* sono nulli. L'interpretazione mediante il nucleo dei funzionali è equivalente a ciò: il nucleo è l'insieme dei valori che inviano nello spazio nullo; se si considera l'intersezione di tutti questi nuclei dei funzionali, essa sarà equivalente all'elemento nullo, poiché solo nel caso dell'elemento nullo non ci sono funzionali in X^* che ritornino un valore non nullo.

Criterio di approssimabilità

Verrà ora introdotto un criterio di approssimabilità. Quando si parla di approssimabilità, si parla sostanzialmente di densità; dal momento che la densità viene caratterizzata confrontando uno spazio ambiente con la chiusura del sottospazio in analisi, si può dire che questo criterio sia un criterio di caratterizzazione della chiusura di un sottospazio.

Dato $L \subset X$, dato $f \in X^*$, dato $f(x) = 0 \forall x \in L$, e supposto che $x_0 \in \bar{L}$, allora se $x_0 \in \bar{L}$, ossia se x_0 è nella chiusura di L , a causa della continuità del funzionale è necessario che $f(x_0) = 0$. Il risultato interessante è il fatto che vale anche il viceversa: se $x_0 \in X$, $f(x_0) = 0$, allora si può dire che $x_0 \in \bar{L}$: vale la doppia implicazione.

La dimostrazione di questo fatto può essere effettuata per assurdo, ossia ipotizzando che $x_0 \notin \bar{L}$. In questa situazione, si può dire che, data la distanza x_0, L come:

$$d(x_0, L) = \delta, \delta > 0$$

vedremo che la tesi è violata. Questo significherebbe infatti che dovrebbe esistere un funzionale $f \in X^*$ tale per cui si possa annullare su L ma non su x_0 :

$$f(x) = 0 \forall x \in L, f(x_0) = \delta$$

dove $\|f\| = 1$.

Dobbiamo ora costruire questo funzionale f_0 : questo deve essere un funzionale nullo su L , e con un certo valore su x_0 ; possiamo costruirlo in maniera tale che $f_0(x_0) = \delta$. Per lavorare, si consideri X_0 come:

$$X_0 = \text{span} \{L \cup \{x_0\}\}$$

ossia, X_0 è dato dallo span di x_0 che sarebbe, come precedentemente visto, λx_0 , con $\lambda \in \mathbb{R}$, dunque un iperpiano passante per x_0 e per l'origine O , e L , anche in questo caso iperpiano passante per l'origine.

Il secondo passo sarà dunque costituito dall'estensione di questo funzionale grazie al teorema di Hahn-Banach allo spazio ambiente, ottenendo quindi un funzionale $f \in X^*$, tale per cui $\|f\| = 1$; la restrizione di f su X_0 sarà ovviamente uguale a quella del funzionale di partenza.

Criterio di densità

Un'ultima conseguenza del teorema di Hahn-Banach è un **criterio di densità**, ossia un criterio per la caratterizzazione della chiusura di un sottospazio. Come noto, un sottospazio è denso se la chiusura coincide con lo spazio ambiente, dunque se ogni punto nello spazio ambiente è rappresentabile mediante punti del sottospazio. Si supponga dunque che $f \in X^*$, e che $f = 0$ su L implichi per ipotesi che $f = 0$ identicamente; questo dimostra che L è denso in X , ossia che:

$$\overline{L} = X$$

La dimostrazione è una conseguenza del precedente criterio, ossia del criterio di approssimazione: dato $x_0 \in X$, si deve dunque dimostrare che $x_0 \in \overline{L}$. Abbiamo visto che l'ipotesi riguarda il fatto che dato un funzionale f appartenente al duale di X , il fatto che esso sia nullo su L implica che f sia identicamente nullo. Se è identicamente nullo, allora $f(x_0) = 0$, per forza: è semplicemente una conseguenza dell'ipotesi. Ma se $f(x_0) = 0$, è possibile applicare il criterio di approssimazione, e dire che se $f(x_0) = 0$, allora $x_0 \in \overline{L}$, e dunque la dimostrazione è ultimata.

Ci sono alcune somiglianze tra ciò che è stato appena descritto e ciò che era stato visto precedentemente parlando di spazi di Hilbert: il criterio di densità nell'ambito degli spazi di Hilbert. In quel contesto si era detto che, dato $x_0 \in H$, dato $(x_0|y) = 0, \forall y \in L$, allora $x_0 = 0$. Se supponiamo di definire il funzionale lineare come:

$$f(y) = (y|x_0)$$

(si osservi che y è messo al primo termine al fine di garantire la linearità, al posto dell'antilinearità, del funzionale). Riconoscendo nel prodotto scalare un funzionale lineare, si ha dunque questa somiglianza.

Tutto ciò serve per dire che per studiare la chiusura è possibile studiare i funzionali lineari: conoscendo le caratteristiche di X^* , è possibile caratterizzare la chiusura di uno spazio.

Conclusioni

Sono stati introdotti alcuni criteri per la caratterizzazione di un certo insieme di spazi e di loro sottospazi; ciò che si è dunque capito è che, una volta note le proprietà dei funzionali appartenenti al duale X^* di uno spazio X , è possibile caratterizzare i suoi sottospazi. Questi criteri dunque sono molto utili, ma solo a patto di essere in grado di descrivere in qualche maniera gli spazi dei funzionali; in seguito verranno introdotti dei *teoremi di rappresentazione*, il cui obiettivo è proprio quello di rappresentare le caratteristiche degli spazi, in modo da applicare facilmente i criteri.

4.5 Teoremi di rappresentazione

In questa sezione verranno introdotti i teoremi di rappresentazione per gli spazi. Prima di ciò, tuttavia, è necessario introdurre alcune definizioni aggiuntive.

4.5.1 Annullatore di un sottospazio

Si è precedentemente visto che gli spazi di Hilbert sono una generalizzazione degli spazi euclidei; essi infatti preservano proprietà molto simili a quelle degli spazi euclidei finito-dimensionali. Negli spazi normati, che rappresentano una categoria più generale rispetto a quella degli spazi di Hilbert, si hanno proprietà più *strane*, come per esempio il fatto che possono esserci vettori linearmente indipendenti tali per cui la norma della loro somma coincide con la somma delle loro norme, o cose di questo genere. Cercando di rimanere vincolati alle proprietà degli spazi di Hilbert, si provi ora a definire in qualche senso un **prodotto scalare**, tra un vettore nello spazio e un funzionale applicato allo spazio duale.

Sia X uno spazio normato, e sia $L \subset X$ un sottoinsieme (anche non per forza un sottospazio), quindi $F \subset X^*$ un sottoinsieme dello spazio duale. È

possibile definire per questi spazi un qualcosa di equivalente agli ortogonali che, in questo contesto, vengono chiamati **annullatori**: l'annullatore L^\perp è definito come:

$$L^\perp = \{f \in X^* : f(x) = 0 \forall x \in L\}$$

questo è l'annullatore di L ; si può dimostrare che esso è sempre un **sottospazio chiuso** di X^* . Si osservi che L è un sottoinsieme dello spazio di partenza, L^\perp un sottospazio del duale dello spazio di partenza. La linearità si può verificare prendendo due funzionali f e g , entrambi appartenenti allo spazio, e se entrambi sono nulli, la loro somma o moltiplicazione per scalare continuerà a esserlo; per quanto riguarda la chiusura, essa può essere verificata a partire dall'uso della continuità, propria di X^* . Il simbolo \perp è messo come **apice**, dal momento che esso denota il passaggio da un sottospazio dello spazio di partenza a un sottospazio dello spazio duale.

Come per L , anche per F è possibile definire l'annullatore, come:

$$F_\perp = \{x \in X : f(x) = 0 \forall f \in F\}$$

in questo caso, osserviamo subito che \perp è introdotto come pedice: infatti, in questo caso si passa da un sottoinsieme dello spazio duale a un sottospazio dello spazio naturale. In questo caso, F_\perp è l'insieme di tutti i punti in cui i funzionali si annullano, dunque esso è un **sottospazio vettoriale chiuso** di X . In verità, è evidente che:

$$F_\perp = \bigcap_{f \in F} \ker \{f\}$$

Questo dal momento che F_\perp è definito come l'insieme dei punti appartenenti a X tali per cui ogni funzionale in F si annulla; questo vuol dire, in altre parole, cercare i punti x per cui ogni funzionale si annulla: l'intersezione dei nuclei di ogni funzionale (se un funzionale si annulla per x ma non per un altro, allora x apparterrà al nucleo di un funzionale, ma non dell'altro, e quindi non all'intersezione). Essendo f continua, il nucleo di f è un insieme chiuso; inoltre, essendo l'intersezione arbitraria di chiusi sempre un chiuso, abbiamo così banalmente verificato che si tratta di un sottospazio chiuso.

L'annullatore ha una proprietà molto importante. Dato $L \subset X$ sottospazio vettoriale,

$$(L^\perp)_\perp = \overline{L}$$

Per dimostrare questo, è necessario dimostrare le inclusioni nei due sensi. Prima, si dimostri che $\overline{L} \subset (L^\perp)_\perp$; per fare ciò, si parta da un $x \in \overline{L}$: da ciò si vuole dimostrare che x è ortogonale a ogni funzionale di un funzionale ortogonale a L . Se dunque x appartiene alla chiusura di L , per definizione, si ha che esiste una successione $\{x_n\}$ di elementi appartenenti a L tali per cui $x_n \rightarrow x$: il fatto di pescare un x dalla chiusura infatti garantisce che il limite appartenga sempre a essa. Si vuole verificare che $x \in (L^\perp)_\perp$, ossia, in altre parole, che x sia ortogonale a (L^\perp) . Matematicamente, questo si può scrivere:

$$f(x) = 0 \forall f \in L^\perp$$

è verificato ciò? Sì: infatti, $f(x)$ può essere scritta come:

$$f(x) = \lim_{n \rightarrow \infty} f(x_n)$$

tuttavia, $x_n \in L$. Dal momento che $x_n \in L$, l'ortogonalità di x_n a L^\perp si verifica prendendo ogni funzionale $f \in L^\perp$ e applicandolo a x_n ; per definizione tuttavia si è appena visto che:

$$L^\perp = \{f \in X^* : f(x) = 0 \forall x \in L\}$$

ma, dunque, $f(x_n) = 0$ per definizione. Se $\forall n$ si ha che $f(x_n) = 0$, anche il limite andrà a 0:

$$f(x) = 0$$

ma dunque, se $f(x) = 0, \forall x \in \overline{L}$, significa che x è ortogonale a L^\perp .

Per quanto riguarda il viceversa, si vuole dimostrare che:

$$(L^\perp)_\perp \subset \overline{L}$$

in questo caso dunque si deve partire da un punto generico $x \in (L^\perp)_\perp$, e verificare che esso appartiene alla chiusura di L , \overline{L} . Questa dimostrazione può essere condotta per assurdo: se per assurdo $x \notin \overline{L}$, dovremmo trovare una contraddizione.

È possibile sfruttare la tesi, ossia il fatto che x come conseguenza deve appartenere a \overline{L} , ricordando il criterio di approssimabilità che era stato introdotto precedentemente; esso affermava che, se $f = 0$ su L , allora $f(x_0) = 0$, e viceversa; se dunque $f(x_0)$ è nullo, $x_0 \in \overline{L}$, e viceversa. Se non valesse

il criterio di approssimabilità, si avrebbe che $\exists f \in X^*$ tale per cui $f = 0$ su L , ma $f(x_0) \neq 0$, dove x_0 è un certo punto fissato: questo non potrebbe appartenere alla chiusura, poiché non soddisfa il criterio di approssimabilità.

La nostra ipotesi per questo punto è il fatto che il punto x_0 che consideriamo è ortogonale a ogni elemento in L^\perp ; di conseguenza:

$$f(x_0) = 0 \forall f \in L^*$$

ma dunque $f(x_0) = 0$, e dunque x_0 soddisfa il criterio di approssimabilità: se $x_0 \notin \overline{L}$, abbiamo una contraddizione.

Osservazione sui sottoinsiemi dello spazio duale

Precedentemente era stato definito F come sottoinsieme dello spazio duale, per definire poi F_\perp su di esso. Verrebbe dunque spontaneo dire che:

$$(F_\perp)^\perp = \overline{F}$$

Questo **non è vero**: solo in alcune condizioni particolari è possibile affermare che questa relazione sia verificata. Nel dettaglio, è necessario prendere una topologia **con meno aperti** di quella che si considera di solito: la **topologia debole stella**, ossia

$$\overline{F}^{\text{weak*}}$$

4.5.2 Spazi biduali

Precedentemente è stato definito il duale di uno spazio X , X^* , come:

$$X^* = \{f : X \rightarrow \mathbb{R} \text{ lineari continui} \}$$

(in alternativa ovviamente lo spazio di arrivo può essere \mathbb{C} , ogni volta che si scrive \mathbb{R}). Tra le varie considerazioni fatte su X^* , si è tuttavia osservato anche che esso è uno **spazio normato**; di conseguenza, può avere senso considerare il **duale dello spazio duale**:

$$X^{**} = \{\varphi : X^* \rightarrow \mathbb{R} \text{ lineari continui} \}$$

(solite considerazioni sullo spazio di arrivo che può essere sia dei reali sia dei complessi). X^{**} è detto **biduale** di X : il duale del duale.

Verrebbe da dire che X e X^{**} siano legati; in dimensione finita, infatti, si può dimostrare che il duale del duale è uno spazio isomorfo dello spazio di partenza. In generale, però, nel contesto degli spazi normati, questa proprietà non è verificata.

Al fine di vedere cosa si può osservare, tuttavia, è possibile effettuare la seguente considerazione: ogni punto dello spazio X identifica un *punto* dello spazio biduale, X^{**} ; un punto dello spazio biduale sarà dunque un **funzionale lineare continuo**. In altre parole, dato $x_0 \in X$, è definito un funzionale lineare continuo che agisce su X^* , che è stato chiamato φ nella definizione. Un esempio di funzionale di questo tipo può essere quello che associa, dato $x_0 \in X$, al funzionale f , la valutazione di questo su x_0 :

$$\varphi(f) = f(x_0) \quad \forall f \in X^*$$

Fissato x_0 in X , volendo ottenere un numero all'uscita di tutto, questo è la valutazione di f in x_0 (è un **esempio** di funzionale di funzionale che, come vedremo, è molto utile). Si può vedere facilmente che φ è un funzionale lineare, applicando φ a una combinazione lineare di operatori:

$$\varphi(\alpha f + \beta g) = \alpha f(x_0) + \beta g(x_0) = \varphi(f) + \varphi(g)$$

Il funzionale che stiamo considerando come esempio usualmente viene indicato come:

$$i(x_0)(f) \quad \text{o} \quad i(x_0)f$$

la seconda notazione non chiude f tra parentesi, dal momento che il funzionale i è lineare in f ; questo funzionale è la **valutazione del funzionale f** in x_0 .

Ci chiediamo se il funzionale $i(x_0)$ sia continuo (dal suo spazio di partenza a quello di arrivo, ossia da X^* a \mathbb{R}). Per vedere ciò, bisogna utilizzare la solita definizione di continuità per funzionali, ossia vedere che:

$$|i(x_0)f| \leq c \|f\|$$

Tuttavia, è banale vedere che:

$$|i(x_0)f| = |f(x_0)|$$

ma, essendo f appartenente a X^* , esso è certamente continuo, dunque per esso la relazione di controllo esiste:

$$|f(x_0)| \leq \|f\| \|x_0\|$$

dunque:

$$|i(x_0)f| \leq \|f\| \|x_0\|$$

quindi, la costante c è proprio $\|x_0\|$. Questo ci garantisce dunque che $i(x_0)$ è un funzionale continuo, e che dunque sta in X^{**} . In verità, si ha qualcosina in più: esso non solo è \leq , ma è addirittura $=$; infatti, basta verificare che esista un certo $f \neq 0$ tale per cui l'eguaglianza sia verificata. Se $x_0 \neq 0$, si ha che $\exists f \in X^*$ con $\|f\| = 1$, in cui $f(x_0) = \|x_0\|$; questo risultato è sicuramente valido, grazie alla prima conseguenza del teorema di Hahn-Banach. Calcolando su tale f il nostro $i(x_0)f$, è fatta.

Riassumendo:

- abbiamo introdotto una applicazione $x \rightarrow i(x)f$, definita da X a X^{**} ;
- dal momento che:

$$\|i(x)\|_{X^{**}} = \|X\|_X$$

questa è una **isometria lineare**, come abbiamo verificato;

- abbiamo verificato che:

$$\|i(x)f\| = \sup_{f \neq 0} \frac{|i(x)f|}{\|f\|_{X^*}} = \max_{f \neq 0} \frac{|i(x)f|}{\|f\|_{X^*}}$$

il sup è un max, grazie alla prima conseguenza di Hahn-Banach.

4.5.3 Spazi riflessivi

Lo spazio X è detto **spazio riflessivo** se l'applicazione i appena definita, $i : X \rightarrow X^{**}$ è **suriettiva**.

Alcune note: si parla di suriettività ma non di iniettività, dal momento che questa funzione è **sempre iniettiva**, essendo un'**isometria lineare**. Se dunque a punti distinti di X sicuramente corrispondono punti distinti di X^{**} , non è assolutamente detto che per ogni punto di X^{**} si abbia un corrispettivo in X , in generale.

Si noti che si ragiona solo su i : al fine di definire la riflessività, non si vuole che esista una generica applicazione φ che sia suriettiva, ma **esattamente** i : se esiste una isometria tra lo spazio di partenza e il suo biduale, suriettiva, ma questa non è i , se i non è suriettiva, allora lo spazio **non è riflessivo**; esempi di questo tipo sono molto complicati da dimostrare, ma esistono.

Il motivo per cui gli spazi riflessivi sono di grande interesse è il fatto che essi sono i più *simili* agli spazi di Hilbert.

Prima di passare al prossimo argomento, dunque, si vuole proporre qualche esempio di spazi riflessivi o non riflessivi.

- si dimostrerà che ogni spazio di Hilbert è riflessivo;
- gli spazi l^p , con $1 < p < \infty$, sono riflessivi;
- gli spazi L^p , con $1 < p < \infty$, sono riflessivi;
- gli spazi l^1 , l^∞ , L^1 , L^∞ **non** sono spazi riflessivi

4.6 Teoremi di rappresentazione

In questa sezione verranno introdotti i principali teoremi di rappresentazione per vari spazi.

4.6.1 Teorema di rappresentazione di Riesz-Frechet

Questo teorema permette di rappresentare il duale di uno spazio di Hilbert. Dato dunque H uno spazio di Hilbert, si può indicare con f_y il funzionale dipendente da y definito a partire dal prodotto scalare:

$$f_y(x) = (x|y)$$

Essendo il prodotto scalare continuo, si ha che $f_y \in H^*$, dove H^* è il duale dello spazio H di partenza. Si è inoltre dimostrato che:

$$\|f_y\| = \|y\|$$

Per ogni elemento $y \in H$ è possibile dunque identificare un funzionale. Il teorema di Riesz-Frechet che stiamo per introdurre garantisce che **ogni funzionale appartenente a H^* si può rappresentare mediante il prodotto scalare**.

Nel dettaglio, dato H uno spazio di Hilbert, l'applicazione $y \rightarrow f_y$ definita su H con immagine in H^* tale per cui:

$$f_y(x) = (x|y)$$

è una **isometria antilineare suriettiva**.

Dobbiamo dunque dimostrare che l'applicazione f_y così definita ha *valori* (ossia, *elementi*) in H^* , e dobbiamo dimostrare che è un'isometria suriettiva.

L'antilinearità è ovvia: dal momento che questa applicazione è definita a partire dal prodotto scalare, si ha che:

$$A(\alpha x + \beta y) = \bar{\alpha}A(x) + \bar{\beta}A(y)$$

e questo si ha dal momento che x è nel secondo termine del prodotto scalare. Purtroppo non è possibile fare altrimenti (mettere a membro sinistro e dunque ottenere la linearità), dal momento che, facendo così, $f_y \notin H^*$: non sarebbe più un funzionale lineare, e dunque non apparterebbe più al duale di H .

Per quanto riguarda la suriettività, vogliamo dimostrare che, dato $f \in H^*$, esiste sempre un $y \in H$ tale per cui $f = f_y$. Per fare ciò, però, è necessario trovare questo y . Questo coincide con il rappresentare il generico funzionale $f(x)$, appartenente a H^* , mediante un prodotto scalare, ossia nella forma $(x|y)$; questo, vale per un particolare y .

Si supponga che $z \in \ker \{f\}$; in questa situazione, y sarà certamente ortogonale a esso, dal momento che il prodotto scalare sarà nullo (ossia, il funzionale f valutato in x nel nucleo restituirà zero). Si può dunque identificare un sottospazio chiuso L definito come il nucleo del funzionale f :

$$L = \ker \{f\}$$

A questo punto: se $L \equiv H$, f si annulla ovunque, dunque f è identicamente nullo: tutto H fa parte del nucleo, dunque valutando il funzionale in ogni $x \in H$, avremo sempre zero.

Se $L \neq H$, ossia se L è solo un sottoinsieme di H , allora:

$$L^\perp \neq \{\emptyset\}$$

infatti, essendo L il nucleo, non essendo ora coincidente con H , significa che il range di H sarà non nullo. Si può utilizzare il teorema di decomposizione ortogonale, dicendo che:

$$H = L \oplus L^\perp$$

ogni elemento di H si può scrivere come somma diretta di un elemento di L e di uno del suo ortogonale. Dato dunque $z \in L^\perp$, $z \neq 0$, si verifica che, se $y = \alpha z$, con α opportuno, allora:

$$f(x) = (x|y) \forall x \in H$$

Rivediamo rapidamente: ora stiamo scegliendo come L il nucleo di f ; preso un punto $z \in L^\perp$, chiediamo che y sia normale al nucleo di f , ossia che $y = \alpha z$; α è un *grado di libertà*, nel senso che non si fissa esattamente quale elemento normale al nucleo si prende, ma se ne sceglie uno con una certa proporzionalità.

A questo punto, possiamo decomporre $x \in H$ come una parte lungo il nucleo e una parte lungo la porzione di spazio rimanente; questo si può fare, utilizzando il seguente trucco:

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

ossia, si somma e si sottrae quel termine. Scegliamo $f(y) \neq 0$; questo significa che $y \in L^\perp$, dunque che esso **non** può stare in L , dal momento che l'unico elemento che può stare sia in L^\perp sia in L è $y = 0$. y sarà un vettore, mentre $f(x)$ e $f(y)$ dei numeri, essendo i risultati dell'applicazione del funzionale al vettore. Dunque:

$$\frac{f(x)}{f(y)}y \in L^\perp$$

perché, come si è detto, $y = \alpha z$, $z \in L^\perp$. Per quanto riguarda la seconda parte, si può vedere che essa appartiene a L , ossia al nucleo; per dimostrare ciò, si deve applicare il funzionale f alla parte rimanente della decomposizione, e verificare che vada a zero. Ciò è verificato, poiché:

$$f\left(x - \frac{f(x)}{f(y)}y\right) = f(x) - \frac{f(x)}{f(y)}f(y) = 0$$

Infine, è necessario verificare che, effettivamente, il funzionale $f(x)$ sia scrivibile, per ogni x , come $(x|y)$. Per fare ciò, si deve calcolare esplicitamente il prodotto scalare $(x|y)$:

$$(x|y) = 0 + \frac{f(x)}{f(y)} \|y\|^2$$

la prima parte è zero, dal momento che il primo termine è nel nucleo di f e dunque restituirà sicuramente zero; per la seconda parte, gli scalari si portano fuori, e si finisce per avere $(y|y)$ che, come noto, è uguale a $\|y\|^2$. La domanda è: ora, questo, è uguale a $f(x)$? Ossia:

$$\frac{f(x)}{f(y)} \|y\|^2 \stackrel{?}{=} f(x)$$

si ricordi che $y = \alpha z$: non abbiamo ancora sfruttato in alcun modo il grado di libertà α ; sostituiamo αz nell'espressione, e otteniamo, semplificando $f(x)$ in ambo i membri:

$$\frac{\|\alpha z\|^2}{f(\alpha z)} = 1$$

ma, sfruttando la linearità di f e la proprietà della norma:

$$\frac{\|\alpha z\|^2}{f(\alpha z)} = \frac{\alpha^2 \|z\|^2}{\alpha f(z)}$$

se dunque:

$$\alpha = \frac{f(z)}{\|z\|^2}$$

si ha la garanzia che $f(x) = (x|y)$; questo coincide con il scegliere come y un **multiplo opportuno** di un generico elemento z appartenente all'ortogonale del ker.

Osservazioni conclusive

Uno dei corollari del teorema di Riesz-Frechet afferma che ogni spazio di Hilbert è riflessivo; tutti i funzionali continui sul duale infatti provengono da elementi dello spazio di Hilbert stesso (mediante una valutazione del funzionale su di un punto fissato). Questo garantisce che lo spazio l^2 , lo spazio $L^2(a, b)$ e così via sono tutti spazi riflessivi, essendo essi spazi di Hilbert.

4.6.2 Teorema di Riesz di rappresentazione del duale di l^p , $1 \leq p < \infty$

Come il titolo della sezione spiega, si vuole descrivere ora in maniera più *concreta* i funzionali lineari appartenenti al duale dello spazio l^p . In verità, precedentemente, era già stato introdotto qualche funzionale: considerando il caso $1 \leq p \leq \infty$, e q l'esponente coniugato a p , nel senso che soddisfa l'espressione

$$\frac{1}{p} + \frac{1}{q} = 1$$

e dove $y = (y_k) \in l^q$, allora si può dire che il funzionale $f_y(x)$ definito come segue:

$$f_y(x) = \sum_{k=1}^{\infty} x_k y_k$$

è un funzionale lineare continuo su l^p ; in altre parole, sicuramente, $f_y \in (l^p)^*$: dal momento che, per la disuguaglianza di Hölder, si era visto che la serie è assolutamente convergente, allora si ha che f_y senza dubbio è un funzionale lineare continuo. Avevamo dimostrato inoltre che:

$$\|f_y\|_{(l^p)^*} = \|y\|_{l^q}$$

il teorema di Riesz di rappresentazione garantisce che **tutti i funzionali lineari continui hanno questa forma**, nel duale di l^p .

Segue l'enunciato del teorema di Riesz: dato $1 \leq p < \infty$, q il suo esponente coniugato, l'applicazione che associa $l^1 \rightarrow (l^p)^*$, $y \rightarrow f_y$, è un'isometria lineare suriettiva.

In altre parole, ogni funzionale continuo su $(l^p)^*$ si può scrivere come visto precedentemente. Dimostriamo ora questo risultato.

Prima di tutto bisognerebbe studiare il fatto che l'applicazione sia ben definita; questo non verrà tuttavia fatto, dal momento che è già stato precedentemente verificato. Inoltre, si era già visto quanto vale la norma per questo operatore. Allo stesso modo non si vuole aggiungere nulla per quanto riguarda la linearità dell'applicazione.

Il punto chiave della dimostrazione è la dimostrazione della suriettività: dato $f \in (l^p)^*$, si vuole verificare che esiste $y \in l^q$ tale per cui $f = f_y$; questo significa avere suriettività. In altre parole, si vuole che:

$$f(x) = \sum_{k=1}^{\infty} x_k y_k \quad \forall x \in l^p$$

Se questo è vero, questa serie deve convergere, e il valore deve essere coincidente col funzionale f_y applicato al generico x in questione. L'elemento x sul quale si applica il funzionale appartiene a l^p , spazio del quale si vuole caratterizzare il duale. Dunque, si può scrivere x come:

$$x = (x_1, x_2, \dots, x_n, \dots)$$

questo, allo stesso modo, può essere scritto come:

$$x = \sum_{k=1}^{\infty} x_k e_k$$

dove e_k è il k -esimo elemento della base canonica ortonormale per l^p : si tratta semplicemente della *solita* successione con 1 al k -esimo posto e 0 in tutti gli altri posti. Essendo $x \in l^p$, sicuramente questa serie converge in l^p ; questo si era visto al momento di dimostrare la separabilità di l^p , in cui si era dimostrato che x converge a quella serie, facendo la successione delle ridotte n -esime. Tutto ciò è vero dal momento che $p \neq \infty$, altrimenti non sarebbe più vero che ogni successione limitata può essere scritta come una serie convergente in l^∞ .

Si ritorni alla dimostrazione, e ora si calcoli $f(x)$; per continuità, si ha che:

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k)$$

la continuità ha permesso di portare il segno di sommatoria al di fuori della valutazione del funzionale f .

Al fine di determinare il funzionale di interesse, è sufficiente definire il y_k di cui si parlava precedentemente come:

$$y_k = f(e_k)$$

Abbiamo dunque determinato la successione y_k tale per cui si può definire il funzionale che desideravamo. Rimane tuttavia da studiarla: si vuole verificare che la successione (y_k) appartenga effettivamente a l^q , al fine di avere

tutte le conferme che si volevano; se questa è verificata, la rappresentazione è effettivamente quella che cercavamo.

Precedentemente, si era detto che:

$$\|f_y\|_{(l^p)^*} = \|y\|_{l^q}$$

questa purtroppo non si può utilizzare, dal momento che ora non conosciamo nulla del membro destro: non abbiamo la certezza che y appartenga in l^q ; tuttavia, sappiamo invece che partiamo da un elemento appartenente allo spazio duale, dunque potremmo provare a ragionare *da sinistra verso destra*: conosciamo qualcosa del membro sinistro, e vogliamo informazioni sul membro destro.

Se $y = (y_k)$ appartiene a l^q , la sua norma deve convergere; dunque:

$$\|y\| = \sum_{k=1}^{\infty} |y_k|^q < \infty$$

Si osservi tuttavia che questa è una serie a **termini positivi**; questo significa che si deve verificare che le somme parziali costituiscono una successione limitata su \mathbb{R} . Al fine di fare ciò, si consideri L_n , un sottospazio definito come:

$$L_n = \text{span} \{e_1, \dots, e_n\}$$

se $x \in L_n$, si ha che:

$$f(x) = \sum_{k=1}^n x_k y_k$$

questo, ricordando che $y_k = f(e_k)$, e ricordando che lo span è una somma finita, di elementi appartenenti a uno spazio di dimensione finita. Il fatto di lavorare in uno spazio di dimensione finita, ci garantisce che:

$$\|f\| \geq \|f|_{L_n}\|$$

questo fatto si può dedurre dalla seguente osservazione: come noto, la norma di un funzionale è definita come

$$\|f\| = \sup_{x \in X} \frac{|f(x)|}{\|x\|}$$

se si restringe lo spazio X in cui si effettua la ricerca del sup, allora sicuramente non si può avere qualcosa di più grande: è possibile avere qualcosa di più piccolo, ma non di più grande; questo significa che la norma della restrizione del funzionale in un sottospazio senza dubbio sarà minore della norma nello spazio intero.

A questo punto, si valuti $\|f|_{L_n}\|$; si può dire che:

$$\|f|_{L_n}\| = \left(\sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}} = \left(\sum_{k=1}^{\infty} |y_k|^q \right)^{\frac{1}{q}}$$

è possibile introdurre ∞ nella sommatoria dal momento che, per $k \geq n + 1$, gli elementi y_k sono nulli, e dunque la somma si può effettivamente fare fino all'infinito. Considerando la successione al variare di n di $\|f|_{L_n}\|$, si ha dunque la successione delle ridotte n -esime; tuttavia, si è detto che:

$$\|f\| \geq \|f|_{L_n}\|, \forall n$$

la costante a sinistra non dipende da n , di conseguenza si ha la certezza che si possa migliorare tutto con una costante, e dunque che:

$$(l^p)^* = l^q$$

dal momento che dunque $y \in l^q$, e la dimostrazione è terminata. Si osservi che, essendo l^q uno spazio isomorfo al duale di uno spazio, completo, allora anche l^q è completo; la completezza di l^q era nota, ma ora è stata dimostrata una seconda volta.

Casi $p = 1, p = \infty$

A questo punto, verrebbe da dire che:

$$(l^1)^* = l^\infty$$

infatti, il teorema vale anche per $p = 1$. In effetti, il duale di l^1 è isomorfo a l^∞ .

Cosa si può invece dire per quanto riguarda il duale di l^∞ ? È vero che $(l^\infty)^* = l^1$, oppure no? La risposta è no: lo spazio duale di l^∞ è più grande di l^1 , dunque non è isomorfo a esso. In altre parole:

$$l^1 \subset (l^\infty)^*$$

ma

$$(l^\infty)^* \not\subset l^1$$

Quello che si vuole dire in altre parole è che, in questo caso, esistono funzionali appartenenti allo spazio duale di l^∞ , $(l^\infty)^*$, che non possono essere rappresentati mediante elementi $y \in l^1$ con:

$$f(x) = \sum_{k=1}^{\infty} x_k y_k$$

per ogni $x \in l^\infty$. In altre parole, in questa situazione, **non si ha la suriettività**. Per verificare questo fatto, si consideri, come funzionale lineare,

$$f_0 : c \rightarrow \mathbb{R}$$

dove c è lo spazio delle successioni convergenti, e \mathbb{R} può essere ovviamente sostituito con \mathbb{C} a seconda degli elementi della successione. Come funzionale f_0 si può definire il limite della successione:

$$f_0(x) = \lim_{k \rightarrow \infty} x_k$$

Un'osservazione preliminare che si può fare è:

$$|f_0(x)| \leq \sup_k |x_k| = \|x\|_{l^\infty}$$

infatti, il limite di una successione è sicuramente maggiorabile con il sup della successione. Il funzionale f_0 è lineare e continuo; se si pensa dunque a c come a un sottoinsieme di l^∞ , è possibile estenderlo, grazie al teorema di Hahn-Banach, a un funzionale f lineare continuo su l^∞ :

$$f \in (l^\infty)^*$$

Si può verificare che questo funzionale non ammette una rappresentazione del tipo di quella precedentemente usata per gli spazi con $p \neq \infty$. Al fine di dimostrarlo, si proceda per assurdo: si supponga che esista un $y = (y_1, y_2, \dots) \in l^1$ tale per cui:

$$f(x) = \sum_{k=1}^{\infty} x_k y_k$$

ossia:

$$\lim_{k \rightarrow \infty} x_k = \sum_{k=1}^{\infty} x_k y_k$$

ciò non può accadere, dal momento che ciò dovrebbe valere $\forall x \in l^\infty$. Se per esempio $(x_n) = (e_n)$, allora x_n ha un elemento non nullo al n -esimo posto, e tutti gli altri nulli; in altre parole, per $k \rightarrow \infty$, $x_k = 0$, essendo sempre nulla tranne che nella n -esima posizione. A secondo membro, invece, si ha che:

$$\sum_{k=1}^{\infty} x_k y_k = y_n, \forall n$$

ma dunque:

$$0 = y_n \forall n$$

questo è l'unico caso in cui i due membri sono uguali, e dunque l'unico caso è quello per cui tutti gli elementi di x siano nulli; questo significa che si deve avere un funzionale identicamente nullo. Essendo tuttavia f l'estensione di f_0 , non è detto che il funzionale debba essere identicamente nullo! Se si cambia x , considerando per esempio la successione costante, $f(x) = \lim_{k \rightarrow \infty} x_k = 1$: da una parte abbiamo trovato che, perché si abbia la verifica della proprietà, si deve avere per ogni x il funzionale nullo; d'altra parte abbiamo trovato che non è detto che, nella nostra situazione, f debba per forza essere nullo, dunque siamo caduti in un assurdo.

Riflessività di l^p

Viene proposto come corollario un ulteriore risultato. Dato l^p , con $1 < p < \infty$, esso è **riflessivo**.

Come abbiamo già detto in precedenza, riflessivo significa che il biduale è isomorfo allo spazio di partenza, tramite l'iniezione i precedentemente introdotta; dimostriamo brevemente alcuni punti di questo fatto:

$$(l^p)^* = l^q, \quad \frac{1}{q} + \frac{1}{p} = 1$$

calcolando il duale di questo, si ha:

$$(l^q)^* = (l^p)^{**} = l^p$$

il biduale di l^p effettivamente coincide con l^p . Questo in verità mostra solamente che **esiste un isomorfismo**; nessuno ha detto che l'isomorfismo in questione debba essere l'iniezione i . Questo è il punto mancante (e che non verrà mostrato) della dimostrazione.

4.6.3 Enunciato del teorema di rappresentazione del duale di $L^p(a, b)$

Esiste, per quanto riguarda gli spazi di funzioni integrabili nel senso di Lebesgue, L^p , un risultato analogo a quello appena discusso: il teorema di rappresentazione del duale di $L^p(a, b)$, ossia $(L^p(a, b))^*$. Anche in questo caso il teorema è attribuito a Riesz. Dato $1 \leq p < \infty$, q esponente coniugato di p , allora l'applicazione da $L^q(a, b) \rightarrow (L^p(a, b))^*$, $f \rightarrow f_y$,

$$f_y(x) = \int_a^b x(t)y(t) dt \quad \forall x \in L^p(a, b)$$

è una isometria lineare suriettiva. Si può verificare che l'integrale converge grazie alla disuguaglianza di Hölder, nel senso di Lebesgue; dalla Teoria della Misura si può poi sapere che la convergenza dell'integrale in questo senso coincide con la convergenza assoluta. f_y dunque è un funzionale lineare continuo.

Si osservi che:

$$(L^\infty(a, b))^2 \neq L^1(a, b)$$

ma

$$(L^1(a, b))^* = L^\infty(a, b)$$

idem per la riflessività: tutti i risultati sono assolutamente analoghi al caso precedente.

Si osservi che il risultato, per $1 < p < \infty$, si estende nel caso di spazi con misura qualsiasi (anche di Lebesgue); il caso $p = 1$, invece, vale solo su spazi con misura σ -finita, ossia in situazioni in cui lo spazio ambiente è un'unione numerabile di insiemi di misura finita.

4.6.4 Rappresentazione del duale di $C([a, b])$

Finora sono stati analizzati spazi dotati di proprietà *belle*: tutti spazi di Hilbert, riflessivi, separabili. In questo caso, considereremo uno spazio un poco *peggiore*: uno spazio di Banach, ma in generale non di Hilbert, separabile (grazie al teorema di Weierstrass, che ha permesso di identificare i polinomi, densi in esso), ma **non riflessivo**.

Prima di introdurre i risultati riguardanti questo spazio, tuttavia, è necessario introdurre due definizioni preliminari: quella di **funzioni a variazione limitata** e quella di **integrale di Riemann-Stieltjes**.

Funzioni a variazione limitata

Una funzione $x(t)$, dove $t \in [a, b]$, si dice **a variazione limitata** (BV: Bounded Variation) se:

$$\exists M > 0 : \forall \text{ partizione } p \text{ di } [a, b], : a = t_0 < t_1 < \dots < t_n = b,$$

$$\implies \sum_{k=1}^n |x(t_k) - x(t_{k-1})| \leq M$$

Si consideri, per comprendere la definizione, la Figura ???. Si partizioni l'asse t , e si considerino i valori che la funzione assume in ciascun punto considerato. Ciò che si fa per studiare la *variazione* della funzione è sommare tutte le differenze tra i k -esimo e $k-1$ -esimo valori della funzione (in valore assoluto). Se è possibile maggiorare tutto ciò con una certa M , allora la funzione è a variazione limitata. In altre parole, la *strada* che si percorre seguendo tutte le variazioni non deve superare la costante M di cui si parlava.

Si consideri per esempio in Figura ??? una funzione monotona crescente; in questo caso, che è un caso limite, la somma delle varie lunghezze è pari a $x(b) - x(a)$. Si può come pensare che la presenza delle oscillazioni faccia *aumentare* la variazione della funzione, un po' come se si *percorresse più strada*.

Esistono funzioni continue ma a variazione **non limitata**. Un esempio di ciò è la $x(t)$ definita come segue:

$$x(t) = \begin{cases} t \sin\left(\frac{1}{t}\right), & t \neq 0 \\ 0, & t = 0 \end{cases}$$

dove $t \in [0, 1]$. Per dimostrare che questa funzione non è a variazione limitata, si considerino per esempio come punti della partizione quelli in cui il seno assume valori pari a ± 1 (che, si osservi, **non sono** i massimi/minimi della funzione, dal momento che essa potrebbe non avere i massimi locali in corrispondenza dei tali punti); essendo il seno a valore ± 1 , si ottiene, dal calcolo delle variazioni, una cosa del tipo:

$$\sum_{n=1}^N \frac{1}{n}$$

questa, per $n \rightarrow \infty$, **non** si maggiora con una costante. Dunque, questa funzione è continua, ma non ha variazione limitata.

La definizione di funzione a variazione limitata si utilizza per studiare la **rettificabilità** delle curve: data una $x : [a, b] \rightarrow \mathbb{R}^2$, si può partizionare $[a, b]$, e si trova così la curva nel piano; se si infittiscono gli intervallini in t abbastanza, si può dire che il risultato delle somme non è altri che la lunghezza della curva. La rettificabilità è legata alla finitezza della lunghezza della curva. Si può definire la variazione di una funzione $x(t)$ su $t \in [a, b]$ come segue:

$$V_a^b(x) = \sup_P \sum_{k=1}^n |x(t_k) - x(t_{k-1})|$$

si osservi che questa **non è una norma**: essa infatti può essere nulla; un caso evidente in cui ciò capita è quello della funzione costante: le variazioni sono nulle, dunque la variazione è zero. Per definire una norma a partire dalla variazione, è possibile scrivere ciò:

$$\|x\|_{\text{BV}} = \|x\|_{\infty} + V_a^b(x)$$

in questo modo, la funzione si annulla sia se non varia (ossia, $V_a^b(x) = 0$), sia se la norma è nulla.

Integrale di Riemann-Stieltjes

L'integrale di Riemann-Stieltjes rappresenta una generalizzazione del concetto di integrale di Riemann. Si consideri una partizione di $[a, b]$ tale per cui, per ciascuna, si abbia sup e inf; un modo alternativo di vedere ciò è prendere un ξ_k appartenente a ciascun intervallino, e calcolare il limite:

$$\lim_{\Delta \rightarrow 0} \sum_{k=1}^{n_\Delta} (\xi_k)(t_k - t_{k-1})$$

il Δ del limite è la distanza tra due t_k ; n_Δ sarà dunque calcolato di conseguenza a Δ . Questo integrale può essere *generalizzato*, nel senso che si può attribuire un peso all'asse t . Si consideri una funzione $\Phi(t)$, e, invece di usare la precedente definizione, si usi:

$$S_p = \lim_{\Delta \rightarrow 0} \sum_{k=1}^{n_\Delta} (\xi_k)(\Phi(t_k) - \Phi(t_{k-1}))$$

in questo modo, la funzione Φ attribuisce un peso al punto considerato: intervalli differenti hanno pesi differenti.

Se Φ è una funzione a variazione limitata, e x è una funzione continua, questa sommatoria converge a un limite finito, all'infittirsi della partizione. In questa situazione, il valore del limite viene indicato come:

$$S_p = \int_a^b x(t) d\Phi(t)$$

Riassumiamo tutto ciò che è stato detto, al fine di comprenderlo meglio. Si consideri $A \in \mathbb{R}$ tale per cui, $\forall \varepsilon > 0$, esiste una $\delta > 0$ tale per cui, per ogni partizione p i cui intervallini hanno lunghezza minore di δ , e per ogni scelta dei punti intermedi ξ_k scelti, risulta che:

$$|S_p - A| < \varepsilon$$

dove:

$$S_p = \sum_{k=1}^n x(\xi_k)(\Phi(t_k) - \Phi(t_{k-1}))$$

Si può dimostrare che vale la seguente stima:

$$\left| \int_a^b x(t) d\Phi(t) \right| \leq \|x\|_\infty V_a^b(\Phi)$$

essendo Φ a variazione limitata, si ha che $V_a^b(\Phi)$ è senza dubbio finito. Per dimostrare questa relazione, è possibile maggiorare (ξ_k) con $\sup |x(\xi)|$, che equivale a $\|x\|_\infty$, mentre la seconda parte si può maggiorare con la variazione totale, che è ancora una volta un sup, come discusso in precedenza.

Considerazioni preliminari al teorema di rappresentazione di Riesz per le funzioni continue

Si definisca f il funzionale lineare $C([a, b]) \rightarrow \mathbb{R}$ definito come:

$$f(x) = \int_a^b x(t) \, r d\Phi(t)$$

questo è continuo, dal momento che è possibile scrivere, per le relazioni appena introdotte, che:

$$\|f\| \leq V_a^b(\Phi)$$

dunque, ogni funzione Φ a variazione limitata definisce un funzionale lineare continuo sulle funzioni continue. Non è tuttavia detto che valga l'uguaglianza in generale (almeno, per ora).

Al fine di fissare i concetti, si considerino dunque alcuni esempi

Il primo esempio è per $\Phi(t) = t$; in questo caso, l'integrale diventa il solito integrale di Riemann, come si può vedere applicando la definizione:

$$\int_a^b x(t) \, d\Phi(t) = \int_a^b x(t) \, dt$$

Un secondo esempio, più generale, riguarda il caso di funzioni $\Phi(t) \in C^{(1)}([a, b])$; in queste situazioni, si ha:

$$\int_a^b x(t) \, d\Phi(t) = \int_a^b x(t) \Phi'(t) \, dt$$

si osservi che l'ipotesi di derivabilità su una funzione Φ a variazione limitata non è minimale: una funzione può essere discontinua ma essere a variazione limitata.

Un terzo esempio riguarda la seguente funzione, definita su $t \in [0, 2]$:

$$\Phi(t) = \begin{cases} 0, & 0 \leq t \leq 1 \\ 1, & 1 < t \leq 2 \end{cases}$$

si può dimostrare che:

$$\int_a^b x(t) \, d\Phi(t) = x(1)$$

per fare ciò, ci sono due approcci possibili:

- utilizzare la definizione di integrale di Riemann-Stieltjes: si ha che $\Phi(t_k) - \Phi(t_{k-1}) = 0$, per ogni k in cui non si comprende il punto di discontinuità; nel punto di discontinuità, si avrà che $\Phi(t_k) - \Phi(t_{k-1}) = 1$, per come è definita la Φ ; inoltre, lo ξ_k in cui avrà senso calcolare la funzione sarà la discontinuità, dal momento che ci si troverà in un intorno della discontinuità: $\xi_k = 1$; di conseguenza, si ha il risultato;
- utilizzare la *sifting property* della delta di Dirac (questo è un ragionamento meno formale, usualmente considerato da fisici): si può definire il funzionale delta di Dirac come:

$$\delta_1(x) = x(1)$$

in questo modo, si arriva a dire che la derivata della funzione gradino è proprio la delta di Dirac, traslata nella discontinuità:

$$\int_a^b x(t) d\Phi(t) = \int_a^b x(t)\Phi'(t) dt = \int_a^b x(t)\delta(t-1)dt = x(1)$$

Teorema di rappresentazione di Riesz

Finalmente siamo in grado di enunciare questo teorema; dato funzionale $f \in (C([a, b]))^*$, si può dire che esista una funzione Φ a variazione limitata su $[a, b]$, tale per cui:

$$f(x) = \int_a^b x(t) d\Phi(t)$$

e

$$\|f\| = V_a^b(\Phi)$$

Quando si parla di funzionali lineari continui ricavati a partire da una funzione continua, ossia di funzionali appartenenti allo spazio duale dello spazio delle funzioni continue, il Φ che si utilizzerà per rappresentare il funzionale sarà tale da soddisfare **esattamente** l'eguaglianza: sarà il *funzionale ottimo*, ossia quello che minimizza la norma. Questo è ciò che il teorema di Riesz afferma *in più* rispetto a quanto detto prima.

Prima di concludere l'argomento dei teoremi di rappresentazione, si vuole aggiungere un breve commento: in questo capitolo si stanno studiando i

funzionali lineari, ma nel dettaglio ci si sta concentrando su funzionali lineari e continui; la continuità di un funzionale è importante dal momento che in generale un funzionale non continuo non è semplice da rappresentare, non essendovi teoremi di rappresentazione come nei casi dei funzionali continui; di conseguenza ha senso focalizzarsi su questi, essendo più semplici da studiare, da rappresentare.

4.7 Convergenza debole

Si consideri uno spazio normato X . È stata già più volte utilizzata la definizione di convergenza di una successione x_n a un elemento $x \in X$, in norma, dicendo che, se $x_n \rightarrow x$, allora $\|x_n - x\| \rightarrow 0$, per $n \rightarrow \infty$.

Esistono tuttavia definizioni alternative di convergenza; in questa sezione verrà discussa una di queste, studiandone le principali proprietà.

Una successione x_n si dice **debolmente convergente** a x , ossia $x_n \rightharpoonup x$, se e solo se:

$$f(x_n) \rightarrow f(x) \forall f \in X^*$$

questa convergenza è studiata sui numeri, dunque su successioni numeriche, dal momento che si studia la convergenza in senso tradizionale di funzionali applicati alle successioni in questione. Ancora una volta, è come studiare lo spazio normato X solo dal punto di vista dei funzionali lineari, dimenticandosi dell'esistenza di una norma.

Si propongono a questo punto alcune osservazioni. Prima di tutto, come il nome suggerisce, la convergenza debole è relazionata alla convergenza: se si ha convergenza, si ha anche convergenza debole, ma non il viceversa. In altre parole:

$$x_n \rightarrow x \implies x_n \rightharpoonup x$$

dimostriamo rapidamente: se $\|x_n - x\| \rightarrow 0$, si ha che:

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \|x_n - x\|$$

e, per ipotesi, $\|f\|$ è finita (consideriamo un funzionale appartenente allo spazio duale, dunque continuo, dunque limitato), e $\|x_n - x\| \rightarrow 0$ per ipotesi.

Esistono esempi di successioni che non convergono in norma, ma che convergono debolmente; si consideri per esempio in uno spazio di Hilbert

$\{e_n\}_{n \in \mathbb{N}}$ un sistema ortonormale; si ha allora che $e_n \rightharpoonup 0$, ma $\|e_n\| = 1$, essendo il sistema ortonormale. Per verificare ciò, si vuole vedere che:

$$f(e_n) \rightarrow f(0) \forall f \in X^*$$

la cosa vale per ogni $f \in X^*$, ma per nostra fortuna, essendo X uno spazio di Hilbert per ipotesi, è possibile utilizzare il teorema di rappresentazione:

$$f \in X^* \Rightarrow \exists y \in X : f(x) = (x|y)$$

dunque:

$$f(e_k) = (e_n, y) \rightarrow 0 = f(0)$$

se $n \rightarrow \infty$, infatti, è possibile applicare la diseguaglianza di Bessel:

$$\sum_{n=1}^{\infty} |(e_n|y)| \leq \|y\|$$

ma dunque, la serie a primo membro converge.

Un altro esempio (che in verità è solo un esempio pratico di quanto appena detto) è, in $L^2(0, 2\pi)$, il $e_n(t) = e^{jnt}$ (o anche solo la sua parte reale, tanto per provare); si consideri dunque $e_n(t) = \cos(nt)$. Se si disegna al variare di n il $\cos(nt)$, ciò che si ottiene è una funzione che non tende assolutamente a zero; ciò che tende a zero è la *media integrale* della successione. Si può dunque pensare, in questo contesto, alla convergenza debole come a una convergenza della media integrale.

Un'osservazione, o meglio, un'anticipazione: in questo caso, si considera $f(x_n)$: ciò che si fa per definire la convergenza debole è fissare il funzionale f , e far variare l'elemento x mediante pedice n . Questo è diverso da quanto si fa nelle distribuzioni, dal momento che in quel contesto quello che si fa è avere un qualcosa tipo $\langle u_n | \phi \rangle$, che è una convergenza puntuale (nel senso che ϕ , funzione test, è unica, fissata), ma al variare del funzionale u_n ; questo sarà limite a quanto si vedrà più avanti in un altro tipo di convergenza debole.

Si osservi che si hanno proprietà particolari in queste situazioni: in uno spazio di Hilbert H , $\{e_n\}$ è una successione limitata, che non dispone di sottosuccessioni convergenti; questa è una cosa impossibile in un caso euclideo. Per vedere ciò, sapendo che uno spazio di Hilbert è completo, dobbiamo vedere che non ci sono sottosuccessioni di Cauchy (essendo in uno spazio completo esse anche convergenti); se $m \neq n$, si ha:

$$\|e_m - e_n\|^2 = \|e_n\|^2 + \|e_m\|^2 = 2$$

dunque, $\|e_m - e_n\|$ non può tendere a 0, dal momento che esso vale $\sqrt{2}$; in altre parole, non è possibile avvicinare i due termini, e dunque le successioni non sono di Cauchy, ma dunque neanche convergenti. Questo fatto permette di dimostrare che:

$$\{x : \|x\| \leq 1\}$$

non è un compatto: le palle non sono un compatto; altrimenti, varrebbe il teorema di Bolzano-Weierstrass, e dunque si avrebbero sottosuccessioni convergenti.

Questo discorso serve per permettere di capire il senso di introdurre definizioni di convergenza più deboli, più lasche: convergenze deboli sono legate a topologie più grezze, ossia meno fini, dotate di meno aperti; in queste, è più facile trovare insiemi compatti, dal momento che ci sono meno ricoprimenti di aperti, e dunque è più facile che un insieme chiuso sia anche un insieme compatto. Per questo motivo, si cerca di lavorare in questa direzione: **facilitare la ricerca di compatti**. La compattezza è molto utile, dal momento che, per Bolzano-Weierstrass, essa garantisce l'esistenza di un minimo; questo può essere sfruttato in molti ambiti, per esempio nella risoluzione di equazioni alle derivate parziali, per mezzo del lemma di Lax-Milgram, in cui si può trovare la soluzione di un'equazione differenziale per mezzo della minimizzazione di un funzionale energetico.

Proprietà della convergenza debole

Si introduce a questo punto una proprietà della convergenza debole, che verrà dimostrata; per fare ciò vedremo che serviranno dei mezzi potenti di Analisi Funzionale, sui quali ci si dovrà soffermare. Si propone la seguente proprietà sotto forma di teorema.

Una successione debolmente convergente è limitata: data $x_n \rightharpoonup x$, allora $\exists M > 0 : \|x_n\| \leq M \forall n$.

La dimostrazione di questa proprietà non è assolutamente banale, e richiede strumenti che verranno ora messi a disposizione.

4.7.1 Lemma di Baire

Introduzione e definizioni preliminari

Quello che si sta per introdurre è un risultato generale per spazi metrici. In uno spazio metrico E , un suo sottoinsieme $A \subset E$ si dice **mai denso** se \overline{A} ha **interno vuoto**. Questo significa, in altre parole, che la chiusura **non contiene interamente alcuna palla**. In altre parole, l'insieme in questione è non vuoto, ma molto rarefatto.

Si considerano ora alcuni esempi.

- Su \mathbb{R} , si consideri:

$$A = \left\{ \frac{1}{n}, n \geq 1 \right\}$$

questo insieme ha 0 come punto di accumulazione, ma non è denso in \mathbb{R} ; dal momento che questo insieme non ha un intervallo, esso **non ha punti interni**: la chiusura di A è solamente lo stesso insieme, aggiungendo 0: non ci sono punti in più! Non è possibile avere una palla interamente contenuta!

- Su \mathbb{R} , si consideri:

$$\mathbb{Q} \cap [0, 1]$$

in questo caso, non è denso, però $\overline{A} = [0, 1]$, e questo **ha punti interni**, dunque questo non è **mai denso**.

- Su \mathbb{R}^2 , si consideri una retta; se si prende una palla a partire da un punto della retta, si va comunque fuori da essa: una retta è *mai densa* su \mathbb{R}^2 .

Fatti questi esempi introduttivi, è ora possibile passare al lemma di Baire, e alla sua dimostrazione.

4.7.2 Lemma di Baire e dimostrazione

Il lemma di Baire afferma che uno spazio metrico completo E **non** è un'unione numerabile di sottoinsiemi mai densi. In altre parole:

$$E \neq \bigcup_{n \in \mathbb{N}} A_n$$

dove A_n sono insiemi mai densi. Questo significa che, essendo gli A_n molto rarefatti, mettendo insieme una quantità numerabile di questi non è possibile costruire uno spazio completo, ossia senza buchi.

Un esempio di applicazione del lemma di Baire è la dimostrazione che \mathbb{R}^2 non è rappresentabile con un'unione numerabile di rette; se ciò è facile da vedere con rette parallele, non è possibile farlo con rette di inclinazione generica; con Baire diventa banale.

Per dimostrare il lemma, è necessario effettuare la seguente osservazione: se A è mai denso, allora significa che per ogni palla aperta B esiste un'altra palla aperta $B' \subset B$ tale per cui $B' \cap A = \emptyset$. Infatti, data A , essa **non** può mai stare tutta dentro A , essendo esso mai denso per ipotesi; questo significa che essa deve per forza stare un po' dentro e un po' fuori. Se B' è una palla aperta, $B' \setminus \overline{A}$ è aperto e non vuoto, dal momento che A è mai denso; se fosse vuoto, sarebbe tutto in \overline{A} ; dunque, questo $B' \setminus \overline{A}$ contiene una palla B' che è disgiunta da \overline{A} . Lo stesso trucco si può fare con le palle chiuse: se invece che farlo con le palle aperte lo faccio con le palle chiuse di dimensione appena più piccola della frontiera di quella della palla aperta di partenza, il trucco rimane in piedi.

La dimostrazione di questo risultato viene condotta per assurdo. Assurdo, significa che, dati A_n insiemi densi, $n \in \mathbb{N}$ (o comunque variabile in un insieme numerabile),

$$E = \bigcup_n A_n$$

l'obiettivo è mostrare che esiste un punto di E che **non appartiene all'unione**. Per fare ciò, si consideri la seguente costruzione: da A_1 , si trovi una palla B_1 non intersecata in A_1 : $A_1 \cap B_1 = \emptyset$; A_2 potrà poi essere intersecato con A_1 e con B_1 . Poi, per $B_2 \subset B_1$ si consideri una palla tale per cui $B_2 \cap A_2 = \emptyset$. Si procede in questa maniera. Ipotizziamo che il raggio r_1 associato alla palla B_1 sia minore a $1/2$:

$$r_1 < \frac{1}{2}$$

poi, si avrà per esempio che:

$$r_2 < \frac{1}{3}$$

il senso è che si ha una successione di palle incapsulate: $B_n \subset \dots \subset B_3 \subset B_2 \subset B_1$. I raggi, invece, sono numeri che vanno come $\frac{1}{n+1}$; dunque:

$$B_n \cap \left(\bigcup_n A_n \right) = \emptyset$$

infatti, B_n è per costruzione disgiunto dagli A_n . È possibile notare che i centri delle palle, chiamati per esempio x_n , costituiscono una successione di Cauchy, dal momento che $d(x_n, x_m) < \varepsilon$, se n, m sono al di sopra di una certa costante; infatti:

$$d(x_m, x_n) < \frac{1}{n+1}, \text{ se } n > m$$

siccome lo spazio è completo, dunque, essendo la successione di Cauchy, essa è anche convergente, ma dunque:

$$x = \lim_{n \rightarrow \infty} x_n$$

esiste. x appartiene a tutte le palle; inoltre, $\{x_k\}$ ha valori definitivamente in B_k , per ogni k (da un certo indice in poi; per esempio, in B_5 si hanno tutti i centri delle palle da 5 in poi). Dunque, $x \in B_k, \forall k$ (grazie al fatto che le palle sono chiuse). Dunque, se $x \in B_1$, esso **non** appartiene a A_1 ; se $x \in B_2$, esso non appartiene a A_1 e A_2 , e così via. Dunque, abbiamo trovato un assurdo.

Il lemma di Baire si utilizza di solito per **generare funzioni esotiche**, nella matematica: dal momento che esso è un teorema che garantisce la **non appartenenza**, è possibile applicarlo per cercare cose *strane*.

4.7.3 Teorema di Banach-Steinhaus

Questo teorema viene anche detto **principio della limitatezza uniforme**. Al fine di capire di cosa si parla, si considerino delle successioni di funzioni $x_n(t)$ su $t \in [0, 1]$. $x_n(t)$ si dice:

- puntualmente limitata, se:

$$\sup_n |x_n(t)| < \infty, \forall t$$

- uniformemente limitata, se:

$$\sup_n \sup_t |x_n(t)| < \infty$$

Al fine di fissare il concetto, si consideri la successione di funzioni in Figura ???: una successione di funzioni con un picco che vale n a $t = \frac{1}{2n}$, 0 per $t > \frac{1}{n}$, ed è a triangolo. La convergenza puntuale significa: focalizzarsi su un certo t , dunque fissarlo, e far crescere n ; si arriverà a un certo punto in cui la funzione, a forza di stringersi, farà valere 0 il punto, dunque essa sarà sotto questo punto di vista limitata. Al contrario, se si variano *assieme* sia n sia t , questa cosa non è vera: non esiste una soglia indipendente da n che non sia mai superata. In generale, dunque, la limitatezza puntuale **non implica la limitatezza uniforme**.

Tuttavia, il teorema di Banach-Steinhaus garantisce che, dato X uno spazio di Banach (si richiede Banach dal momento che la completezza serve per applicare il lemma di Baire), e $\{f_\alpha\} \subset X^*$ una famiglia (anche non numerabile) di funzionali lineari continui puntualmente limitata, ossia per cui:

$$\sup_\alpha |f_\alpha(x)| < \infty, \forall x \in X$$

allora, $\sup_\alpha \|f_\alpha\| < \infty$.

Questo equivale a dire:

- esiste M tale per cui $\|f_\alpha\| \leq M \forall \alpha$;
- $|f_\alpha(x)| \leq M \forall \alpha, \forall x : \|x\| < 1$
-

$$\sup_\alpha \sup_{x: \|x\| \leq 1} |f_\alpha(x)| < \infty$$

(si ricordi che un funzionale non è mai limitato, dunque la limitatezza va verificata sulla palla unitaria di x). Tutto ciò, grazie alla linearità.

Consideriamo a questo punto la dimostrazione di questo teorema. Si consideri il seguente insieme:

$$\{x \in X : |f_\alpha(x)| \leq n\}$$

ora, fissati α e n , si vogliono cercare gli x per cui la relazione è soddisfatta; l'insieme di questi x è un sottoinsieme **chiuso** di X ; infatti, si può usare il fatto che gli f_α sono continui, e dunque utilizzare i risultati noti riguardanti la controimmagine di palle chiuse: $f_\alpha^{-1}(\overline{B(0, n)})$. Dunque, si consideri:

$$A_n = \bigcap_{\alpha} \{x \in X : |f_\alpha(x)| \leq n\}$$

questo è l'insieme degli x per cui, per ogni α , è soddisfatta la relazione; dal momento che stiamo considerando un'intersezione (per quanto arbitraria) di chiusi, ciò sarà ancora un chiuso.

A questo punto, al fine di proseguire con la dimostrazione, è necessario scrivere l'ipotesi: essa è:

$$X = \bigcup_n A_n$$

ossia, $\forall x \in X$, si ha che il sup è finito, e dunque esiste una certa costante n tale per cui $|f_\alpha(x)| \leq n$, per ogni α (ossia, indipendentemente da α).

A questo punto è possibile applicare il lemma di Baire: esso garantisce che esiste un certo n_0 tale per cui A_{n_0} non è **mai denso**, ossia tale per cui la sua chiusura deve per forza contenere una palla; dunque:

$$\overline{A_{n_0}} = A_{n_0}$$

e questo ha una palla aperta, di centro x_0 e raggio r : $B(x_0, r) \subset A_{n_0}$. Dunque, se

$$\|x - x_0\| < r$$

allora,

$$f_\alpha(x) \leq n_0 \forall \alpha$$

dal momento che questo appartiene all'intersezione.

A questo punto, per proseguire, ci si riconduca al caso della palla unitaria: se $\|y\| \leq 1$, si ha che:

$$|f_\alpha(x_0 + ry)| \leq n_0 \forall \alpha$$

questo si può vedere nel seguente modo: abbiamo che $f_\alpha(y)$ è limitato nella palla unitaria; tuttavia esso è anche lineare, dunque si può scrivere:

$$|f_\alpha(y)| = \left| \frac{1}{r} f_\alpha(ry) \right| = \left| \frac{1}{r} f_\alpha(x_0 + ry) - \frac{1}{r} f_\alpha(x_0) \right| \leq \frac{1}{r} (|f_\alpha(x_0 + ry)| + |f_\alpha(x_0)|)$$

ora, in questi due casi, si ha che essi sono maggiorabili con n_0 : in un caso si ha quanto visto poche righe sopra, nell'altro si ha il caso $y = 0$ che ovviamente è dentro la palla unitaria in y , dunque il gioco è fatto. Dunque:

$$|f_\alpha(y)| \leq \frac{2n_0}{r}$$

Questo mostra che una certa costante riesce a effettuare la maggiorazione.

Dimostrazione della proprietà delle debole convergenza

A questo punto è possibile dimostrare la proprietà precedentemente introdotta: il fatto che una successione debolmente convergente è anche limitata. Questo risultato richiede il teorema di Banach-Steinhaus, che per essere dimostrato richiedeva il lemma di Baire. Questa proprietà può essere pensata come corollario di Banach-Steinhaus.

Questo corollario afferma che, se X è uno spazio normato, **anche non completo**, e $x_n \rightharpoonup x$, allora $\{x_n\}$ è limitata.

Per dimostrare questo fatto, come ipotesi abbiamo che $f(x_n) \rightarrow f(x)$, $\forall f \in X^*$. Tuttavia, dalla teoria precedentemente affrontata parlando di spazi riflessivi, si è visto che $f(x_n)$ può essere visto come:

$$f(x_n) = i(x_n)f$$

ossia, utilizzando l'iniezione i precedentemente vista. $i(x) \in X^{**}$, e $i(x)f = f(x) \forall f \in X^*$. Essendoci convergenza, dunque, si può dire che, per ogni f fissato, si ha che:

$$i(x_n)f \rightarrow i(x)f$$

dunque, essendo $i(x_n)f$ convergente, essa è anche limitata. Questo significa che:

$$\sup_n |i(x_n)f| < \infty \forall f \in X^*$$

ossia, abbiamo appena dimostrato che le $i(x_n)$ sono una successione limitata, per ogni funzionale f appartenente allo spazio duale di X .

A questo punto dobbiamo applicare il teorema di Banach-Steinhaus, ma allo spazio X^* ; questo significa che invece che lavorare su X , si lavora su X^* , dunque come elementi x abbiamo i funzionali f , che sono gli elementi degli spazi duali. Non è necessaria la completezza di X , dal momento che lo spazio su cui si lavora in verità è X^* , ed esso è sempre completo. Dunque, Banach-Steinhaus ci dice che:

$$\sup_n \|i(x_n)\|_{X^{**}} < \infty$$

essendo tuttavia $i(x)$ una isometria:

$$\sup_n \|i(x_n)\|_{X^{**}} = \sup_n \|x_n\| < \infty$$

isometria: norme uguali.

4.7.4 Criterio di la convergenza debole

È stata fornita la definizione di convergenza debole di una successione; tuttavia, per capire se una funzione è debolmente convergente, è possibile applicare un criterio di studio, più semplice rispetto alla definizione. L'idea è quella di sfruttare i risultati concernenti i teoremi di rappresentazione dei duali degli spazi studiati; si può dimostrare che se si ha convergenza in un sottospazio denso rispetto al duale dello spazio interessato, e la limitatezza di f , allora si ha convergenza debole.

Formalmente, dato X uno spazio normato, si ha che:

$$x_n \rightharpoonup x \iff \begin{cases} \{x_n\} & \text{è limitato} \\ f(x_n) \rightarrow f(x) & \forall f \in \Delta \end{cases}$$

dove la seconda condizione è verificata in Δ , che è un sottospazio dello spazio duale:

$$\Delta \subset X^* : \overline{\text{span}\{\Delta\}} = X^*$$

ossia, un sottoinsieme tale per cui la chiusura del suo span (combinazioni lineari finite) generi l'intero spazio duale X^* .

L'implicazione da sinistra verso destra si può dimostrare applicando il teorema di Banach-Steinhaus. Ora, ci si preoccuperà di dimostrare l'implicazione inversa. Data come ipotesi $f(x_n) \rightarrow f(x), \forall f \in X^*$ (come successione numerica), si vuole dunque dimostrare che si ha convergenza debole. Per linearità, prima di tutto, si può dire che, se questo vale $\forall f \in \Delta$, allora si ha anche che vale $\forall f \in \text{span}\{\Delta\}$; infatti, lo span di uno spazio è semplicemente l'insieme delle combinazioni lineari finite dei vettori nello spazio. Tuttavia, è nel nostro interesse verificare non che sia valido per ogni f nello span, ma nella chiusura dello span. Per fare ciò, si consideri $f \in X^*$, $f_0 \in \text{span}\{\Delta\}$ e si consideri un certo ε tale per cui:

$$f_0 \in \text{span}\{\Delta\}$$

dove:

$$\|f_0 - f\| < \varepsilon$$

per ipotesi, si ha che X^* si può approssimare con elementi appartenenti allo span di Δ (per la definizione di densità); dunque, quanto appena scritto è accettabile. A questo punto, si applichi a ciò la disuguaglianza triangolare generalizzata come segue:

$$|f(x_n) - f(x)| \leq |f(x_n) - f_0(x_n)| + |f_0(x_n) - f_0(x)| + |f_0(x) - f(x)|$$

a questo punto, si analizzi ciascuno di questi elementi, per maggiorarli:

- per quanto riguarda il primo elemento si ha che:

$$|f(x_n) - f_0(x_n)| \leq \|f - f_0\| \|x_n\| \leq \varepsilon M$$

infatti, $\|f - f_0\| \leq \varepsilon$ per ipotesi (approssimazione a meno di ε), mentre la norma di x_n è limitata per ipotesi del criterio (la prima ipotesi), a una costante che chiamiamo M ;

- per quanto riguarda il secondo termine:

$$|f_0(x_n) - f_0(x)| \leq \varepsilon$$

questo, per l'ipotesi per cui $f_0(x_n) \rightarrow f_0(x)$, ipotesi del criterio;

- per quanto riguarda l'ultimo termine, è possibile fare qualcosa di simile a quanto fatto per il primo:

$$|f_0(x) - f(x)| \leq \|f - f_0\| \|x\| \leq \varepsilon \|x\|$$

in questo caso, $\|x\|$ è un certo numero.

Unendo i tre pezzi, si ottiene:

$$\leq \varepsilon + \varepsilon + \varepsilon \|x\| = \varepsilon(M + 1 + \|x\|)$$

quindi: $f(x_n) \rightarrow f(x)$.

Si anticipa il fatto che questo principio è assolutamente generale: esso vale anche, come si vedrà, per gli operatori. Questo criterio è molto utile dal momento che se si ha per ipotesi la limitatezza, è possibile studiare la **convergenza puntuale** su un insieme denso invece che su ogni funzione; se per esempio si usa come insieme denso quello delle funzioni $C^{(\infty)}$, è molto facile lavorare.

A questo punto, al fine di chiarire e di proporre alcuni risultati applicativi, verranno proposti alcuni esempi.

Convergenza debole su l^p

Dato lo spazio l^p , con $1 < p < \infty$ (escludendo dunque gli estremi), cosa si intende per convergenza debole? Si consideri una successione $x^{(n)} \in l^p$ (ossia, una successione di successioni, dal momento che ogni elemento di l^p è una successione); si ha che:

$$x^{(n)} \rightharpoonup x \iff \begin{cases} \{x^{(n)}\} & \text{è limitata in } l^p \\ x_k^{(n)} \rightarrow x_k & \forall k \end{cases}$$

spieghiamo meglio: per ogni n , si ha un elemento di l^p , e dunque una successione, come già detto; $x_k^{(n)}$ è il k -esimo elemento della n -esima successione considerata su l^p ; l'ipotesi richiede che ciascuna k -esima componente converga puntualmente a un certo x_k , per ogni k .

Si motivi ora la seconda ipotesi, che non è visibilmente collegata a quella del criterio. Δ è un sottospazio dello spazio duale, la chiusura del cui span è isomorfa allo spazio duale. Come Δ si considera:

$$\Delta = \{e_k\} \subset (l^p)^* = l^q, \quad \frac{1}{p} + \frac{1}{q} = 1$$

dove e_k è la successione che ha 1 nella k -esima posizione e 0 negli altri punti. Definendo un funzionale $e_k(x)$ definito come:

$$e_k(x) = \sum_{k=1}^{\infty} x e_k = 0x_1 + 0x_2 + \dots + 0x_{k-1} + 1x_k + 0x_{k+1} + \dots = x_k$$

se f è uno di questi funzionali, si può dire che esso *selezioni* la k -esima componente di x . È necessario che la chiusura dello span dello spazio di questi vettori sia densa in $l^q = (l^p)^*$, ma è così, dal momento che $p > 1$, $q < \infty$, e quindi è vero.

Convergenza debole su $L^p(a, b)$

A questo punto si farà qualcosa di analogo rispetto a prima, considerando però lo spazio $L^p(a, b)$ delle funzioni integrabili in senso di Lebesgue. Si ha che:

$$x_n(t) \rightharpoonup x(t) \iff \begin{cases} \{x_n\} \\ \int_I x_n(t) dt \rightarrow \int_I x(t) dt \forall I \subset [a, b] \end{cases} \quad \text{è limitata in } L^p(a, b)$$

il I in questione è un **intervallo** di $[a, b]$. Prima di tutto, si ha la garanzia che gli integrali esistono; infatti:

$$L^p(I) \subset L^q(I), \quad p \geq q, \quad \text{mis}(I) < \infty$$

se dunque sta in L^p , sta anche in L^1 .

Per quanto riguarda la seconda condizione, come insieme Δ si considera l'insieme delle funzioni caratteristiche sull'intervallo I : χ_I ; dunque:

$$\Delta = \{\chi_I\} \subset (L^p)^* = L^1, \quad \frac{1}{p} + \frac{1}{q} = 1$$

A questo punto, si sfrutta il fatto che L^q è il duale di L^p ; quindi, si utilizza il teorema di rappresentazione dello spazio duale, scegliendo una famiglia di funzionali $y \in (L^p)^*$, e usando:

$$f_y(x) = \int_a^b x(t) y(t) dt$$

ma, se scegliamo come $y(t)$ la funzione caratteristica χ_I , otteniamo proprio:

$$\int_a^b x(t) \chi_I(t) dt = \int_I x(t) dt$$

Rimane a questo punto solo più un punto: dimostrare che lo di questo sottoinsieme sia denso in $(L^p)^*$. Questo è un risultato noto dalla teoria della misura, dal momento che:

$$\sum_{k=1}^N a_k \chi_{I_k}$$

ossia, le combinazioni lineari finite di intervalli, sono dense in L^q , $q < \infty$. I_k inoltre è un intervallo, non un generico sottoinsieme di misura finita, dunque è abbastanza ragionevole lavorare con esso.

Si osservi che, come già anticipato in precedenza, la convergenza in L^p è una convergenza **in media**. Avevamo precedentemente studiato ciò sull'esempio:

$$\cos(nt) \rightharpoonup 0$$

su $X = L^2$; per fare ciò avevamo sfruttato il fatto che ogni sistema ortonormale converge sempre debolmente a zero. Ora possiamo dire che converge a zero dal momento che la sua media su un qualunque intervallo fissato I tende a 0; infatti, fissato un certo intervallo, al crescere di n aumentano le oscillazioni su questo intervallo, ma dunque la media tende a zero, indipendentemente dal I scelto.

Convergenza debole su $C([a, b])$

In questo caso si presenterà il criterio di convergenza debole, per quanto esso non sia collegato agli altri. Nel contesto delle successioni di funzioni continue, si ha:

$$x_n(t) \rightharpoonup x(t) \iff \begin{cases} \{x_n\} & \text{è limitata} \\ x_n(t) \rightarrow x(t) & \forall t \in [a, b] \end{cases}$$

La prima condizione significa semplicemente che:

$$|x_n(t)| \leq M \forall n, \forall t$$

ossia, significa che $x_n(t)$ è **uniformemente limitata**; per quanto riguarda la seconda condizione, essa è una **convergenza puntuale**: fissato t , si richiede che converga $\forall n$; questo è più debole rispetto alla convergenza uniforme!

Il funzionale che realizza questo tipo di legame è la delta di Dirac; questo è quel funzionale δ_t tale per cui:

$$\delta_t(x_n) \rightarrow \delta_t(x)$$

e

$$\delta_t \in (C([a, b]))^*$$

dove:

$$\delta_t(x) = x(t)$$

Si osservi che lo span dei δ_t **non** è denso; questo esempio dunque **non segue** dal criterio precedente. Per dimostrare il criterio appena introdotto, sarebbe necessario fare uso del criterio di convergenza dominata, unito al teorema di rappresentazione di Riesz. Non vedremo i dettagli.

4.8 Convergenza debole* (convergenza debole stella)

La definizione di convergenza precedentemente introdotta riguardava gli spazi normati; quella che verrà introdotta ora riguarda i **duali** degli spazi normati. Sugli spazi duali è possibile applicare la definizione classica di convergenza, la convergenza debole (essendo i duali spazi normati, oltretutto completi), ma pure un'ulteriore concetto di convergenza: la convergenza debole stella.

Partiamo da cosa sappiamo già: se X è uno spazio normato, X^* è normato e completo (come già detto più volte), e, sullo spazio duale, si dice che $f_n \rightarrow f$ se:

$$f_n \rightarrow f \iff \|f_n - f\|_{X^*} \rightarrow 0$$

questa è la definizione classica di convergenza, ossia quella in norma, applicata però a elementi dello spazio duale, ossia a funzionali. La stessa cosa può essere fatta per quanto concerne la convergenza debole:

$$f_n \rightharpoonup f \iff \varphi(f_n) \rightarrow \varphi(f) \forall \varphi \in X^{**}$$

questa è del tutto analoga a quella vista sugli spazi di partenza, dove la differenza sta nel fatto che va valutata su funzionali appartenenti allo spazio duale a quello degli elementi: in questo caso, il bidual. Questa definizione dunque non è molto utile, dal momento che non è detto che il bidual di un certo spazio sia noto, a meno che lo spazio X sia riflessivo (caso in cui il bidual è isomorfo allo spazio di partenza). Dunque, ha senso definire una convergenza ancora più debole. Si dice che una successione $\{f_n\} \in X^*$ converga debolmente* a $f \in X^*$ se e solo se:

$$f_n(x) \xrightarrow{*} f(x) \forall x \in X$$

Volendo dunque confrontare la convergenza debole con la convergenza debole*, ciò che si può dire è che:

- nel caso della convergenza debole, in X , si ha che:

$$x_n \rightharpoonup x \iff f(x_n) \rightarrow f(x) \forall f \in X^*$$

- in X^* , nel caso della convergenza debole*, si ha che:

$$f_n(x) \xrightarrow{*} f(x) \forall x \in X$$

La differenza sta nel fatto che nel caso della convergenza debole si fissa il funzionale, si ha una successione di elementi $x_n \in X$, che deve tendere a x , e questa cosa deve essere soddisfatta per ogni funzionale. Nel caso della convergenza debole*, capita l'opposto: si fissa un certo punto x , e si considera una successione di funzionali; quest'ultima dovrà convergere a un funzionale f , puntualmente, e per ogni $x \in X$.

È possibile interpretare in modo simile la convergenza debole*, ma in maniera da capire esattamente quale sia il legame con la convergenza debole. È noto che dire che, dato il funzionale $f(x) \triangleq i(x)f$, si ha:

$$f_n(x) \rightarrow f \iff i(x)f_n \rightarrow i(x)f \forall x \in X$$

dove i è la solita iniezione: $i \in X^{**} : i(x)f \triangleq f(x)$.

L'applicazione $i(x)$ ricopre sostanzialmente il ruolo della $\varphi(x)$ precedentemente vista; questo significa che la convergenza debole* non è altri che la convergenza debole dove, invece che richiedere la validità per ogni possibile funzionale φ appartenente al biduale dello spazio di partenza, si richiede solo la validità per i funzionali nella forma di i : si richiede solo un caso specifico.

Se $f_n \rightharpoonup f$ debolmente, allora si ha che $f_n \xrightarrow{*} f$: la debole infatti richiede la validità $\forall \varphi$, tra cui anche i . Il viceversa non vale, a meno che non si introduca una particolare topologia, tale per cui le due convergenze possano coincidere; questa è una topologia con pochi aperti, ma molto ampi, non metrizzabile, non primo-numerabile, dunque non verrà approfondito il discorso.

Esiste un caso in cui si ha la biimplicazione tra convergenza debole e convergenza debole*: questo è il caso di X spazio riflessivo. Se dunque X è riflessivo:

$$f_n \rightharpoonup f \iff f_n \xrightarrow{*} f$$

dunque, l^p , L^p , con $1 < p < \infty$, ammettono questo tipo di relazione.

Un'ultima osservazione, prima di considerare un esempio: esiste un criterio che garantisce il fatto che una funzione sia convergente debolmente*, abbastanza simile a quello introdotto per la convergenza debole. Infatti:

$$f_n \xrightarrow{*} f \implies \{f_n\} \text{ è limitata in } X^*, \text{ se } X \text{ è completo}$$

questo segue dal teorema di Banach-Steinhaus: se si ha una successione di funzionali che converge puntualmente, ossia $\forall x$, allora f_n è puntualmente limitata; la convergenza infatti implica la limitatezza. Applicando Banach-Steinhaus grazie a questo fatto, il risultato è ottenuto.

Esempio nello spazio delle funzioni continue

Si consideri $X = C([-1, 1])$, e il funzionale $f_n(x)$ definito come:

$$f_n(x) = n \int_{-\frac{1}{2n}}^{+\frac{1}{2n}} x(t) dt$$

Questo funzionale è lineare per ogni n ; inoltre, definendo la funzione $K_n(t)$ come la funzione porta, è possibile riscrivere l'integrale come segue:

$$f_n(x) = n \int_{-\frac{1}{2n}}^{+\frac{1}{2n}} x(t) dt = \int_{-1}^{+1} K_n(t)x(t) dt$$

i funzionali sono continui; è possibile dimostrare che:

$$\|f_n\| \leq \|K_n\|_{L^1} = 1$$

Se $n \rightarrow \infty$, la porta tende a stringersi e ad alzarsi; questo significa che:

$$f_n \xrightarrow{*} \delta$$

ossia:

$$f_n(x) \rightarrow \delta(x)$$

questo si può dimostrare a partire dal teorema fondamentale del calcolo integrale. Tuttavia, in questo caso, lo spazio considerato non è riflessivo, e infatti f_n **non** tende debolmente a δ .

Criterio di convergenza debole*

Un criterio è già stato precedentemente proposto; esiste tuttavia un criterio completamente analogo a quello precedentemente mostrato per quanto riguarda la convergenza debole: dato X uno spazio normato, e $f_n, f \in X^*$, allora:

$$\begin{cases} \{f_n\} & \text{è limitata in } X^* \\ f_n(x) \rightarrow f(x) & \text{(conv. succ. numerica)} \end{cases} \implies f_n \xrightarrow{*} f$$

questo, deve valere $\forall x \in \Delta$, dove Δ è un sottospazio di X , tale per cui:

$$\overline{\text{span}\{\Delta\}} = X$$

Si osservi che l'implicazione è stata scritta solo in un senso, dal momento che in generale la biimplicazione non vale. Questa vale solo se X è uno spazio completo (potendo così applicare Banach-Steinhaus).

4.8.1 Teorema di Alaoglu (versione sequenziale)

Come già accennato precedentemente, l'accanimento nella ricerca di convergenze deboli riguarda la garanzia di aumentare il numero di insiemi compatti. Quando si ha a che fare con insiemi compatti, è possibile applicare su di essi il teorema di Weierstrass che garantisce la presenza di un minimo; questo è molto utile nell'ambito del calcolo delle variazioni. Date convergenze deboli, ci si basa su pochi aperti, e dunque è più facile che un insieme chiuso sia anche un compatto. Quello che si sta per mostrare è un teorema che garantisce l'esistenza di sottosuccessioni convergenti deboli*, estendendo dunque il teorema di Weierstrass.

Dato X spazio normato separabile, ogni successione $\{f_n\}$ limitata in X^* ammette sottosuccessione convergente debolmente*.

Questo è sostanzialmente quello che capita nell'ambito degli spazi euclidei, studiando la convergenza in norma. Dato H spazio di Hilbert e $\{e_k\}$ sistema ortonormale, per esempio, si era visto che la successione è limitata, ma **non ammette** sottosuccessioni convergenti nella convergenza classica; in questo caso però si sta considerando qualcosa di molto differente, dal momento che si parla dello spazio duale, e della convergenza debole. L'unico caso particolare riguarda proprio gli spazi di Hilbert, in cui H , generico spazio di Hilbert, è isomorfo al suo duale: $(L^2)^* = L^2$. La tesi del teorema dunque è: esiste una sottosuccessione $f_{n_k}(x)$, e un funzionale $f \in X^*$ tale per cui:

$$f_{n_k}(x) \rightarrow f(x) \forall x \in X$$

Si propone a questo punto una dimostrazione di questo teorema. Questa dimostrazione è sostanzialmente composta da due passi:

1. verificare che esista una sottosuccessione che converga in un sottoinsieme denso di X ; essendo X separabile, esso contiene un sottoinsieme denso numerabile D , definito come:

$$D = \{x_1, x_2, x_3, \dots\} \subset X$$

su questo, f_n deve avere una sottosuccessione f_{n_k} convergente su D , ossia tale per cui:

$$f_{n_k}(x) \rightarrow f(x), \text{ in } \mathbb{R}(\mathbb{C}), \forall x \in D$$

2. il secondo passo è verificare che esiste un $f \in X^*$ tale per cui:

$$f_{n_k}(x) \rightarrow f(x), \text{ in } \mathbb{R}(\mathbb{C}), \forall x \in X$$

ossia, tale per cui:

$$f_{n_k} \xrightarrow{*} f$$

Il primo passo viene svolto mediante l'applicazione del **procedimento diagonale di Cantor**: come ipotesi, si ha che la successione è limitata:

$$\exists M : \|f_n\|_{X^*} \leq M \forall n$$

questo implica che, $\forall x \in X^*$, si ha che $f_n(x)$ è limitata; infatti:

$$|f_n(x)| \leq \|f_n\| \|x\|$$

dove $\|f_n\|$ è limitata per ipotesi dal momento che si considerano funzionali dello spazio duale, mentre x è fissato. $f_n(x)$ è una successione di numeri reali/complessi; per Bolzano-Weierstrass, dunque, questa ha una sottosuccessione convergente in \mathbb{R} (o \mathbb{C}). Si vuole tuttavia che questa sottosuccessione sia convergente per tutti gli $x \in D$. Se D fosse finito, questa cosa sarebbe banale, dal momento che basterebbe considerare una successione e valutarla in x_1, x_2 , e così via. Quello che si fa qua, applicando Cantor, è costruire una successione di sottosuccessioni, prendendo gli elementi diagonali: il primo elemento della prima successione, il secondo elemento della seconda successione e così via; questa, convergerà in ogni punto.

Formalizziamo un poco meglio questa idea, prima di proporre il passaggio finale; dato D descritto come prima, $\{f_{n_k}(x)\}$ è una successione numerica; su questa dunque è possibile applicare Bolzano-Weierstrass. Per ipotesi, $\{f_n\}$ è limitata in X^* , dunque:

$$\exists M : \|f_n\|_{X^*} \leq M \forall n$$

quindi, si ha che $\{f_n(x)\}$ è limitata, ma in \mathbb{R} (o \mathbb{C}), essendo una successione numerica (di funzionali). Una volta fissato un certo $x \in D$, è dunque facile trovare una sottosuccessione convergente; il nostro obiettivo tuttavia è quello di trovare una sola sottosuccessione, che vada bene $\forall x \in D$.

Consideriamo la seguente notazione: se $n_1 < n_2 < n_3 < \dots$ è una successione di numeri naturali, ossia $n_i \in \mathbb{N}$, si definisca, per x_1 , una sottosuccessione f_{n_k} nel seguente modo: gli elementi di una successione sono identificabili mediante un insieme di indici n_i ; definito dunque $S_1 \subset \mathbb{N}$ l'insieme degli n_i tali per cui:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S_1}} f_n(x_1) \text{ esiste finito in } \mathbb{R}(\mathbb{C})$$

Questo è possibile, per il teorema di Bolzano-Weierstrass. S_1 è dunque l'insieme degli indici della sottosuccessione.

A questo punto, si consideri la sottosuccessione $\{f_n(x_2), n \in S_1\}$: esiste, per Bolzano-Weierstrass, un $S_2 \subset S_1$ infinito tale per cui:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S_2}} f_n(x_2) \text{ esiste finito in } \mathbb{R}(\mathbb{C})$$

E così via. Inoltre, si ha che:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S_2}} f_n(x_1) \text{ esiste finito in } \mathbb{R}(\mathbb{C})$$

dal momento che $S_2 \subset S_1$. Mediante questo procedimento è possibile costruire una catena di sottoinsiemi infiniti di \mathbb{N} :

$$S_k \subset \dots \subset S_3 \subset S_2 \subset S_1 \subset \mathbb{N}$$

tali per cui:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S_k}} f_n(x_j), j \leq k, \text{ esiste finito in } \mathbb{R}(\mathbb{C})$$

per costruzione e grazie a Bolzano-Weierstrass.

Tuttavia, l'obiettivo è quello di trovare una sottosuccessione che converga contemporaneamente per ogni $x \in D$; questo si può fare, definendo $S \subset \mathbb{N}$ tale da essere costituito dal primo elemento (dunque, un indice) di S_1 , dal secondo elemento di S_2 , dal terzo elemento di S_3 , e così via. Ciascuno di questi sottoinsiemi è un sottoinsieme di \mathbb{N} , dunque la relazione d'ordine è sempre la stessa; inoltre, anche nel caso sfortunato in cui il primo elemento di S_1 coincida con il secondo elemento di S_2 , grazie all'ordine, poiché non si prende il primo, ma il secondo elemento, di S_2 , non c'è rischio di prendere due volte lo stesso elemento.

S è un sottoinsieme di S_k (infinito, numerabile); $S \subset S_k$; questo è vero, tranne per i primi $k - 1$ termini, in cui nulla è garantito; questo significa dunque che la successione:

$$\{f_n, n \in S\}$$

è una sottosuccessione di $\{f_n, n \in S_k\}$, $\forall k$ (eccetto eventualmente nei primi $k - 1$ termini). In altre parole, si ha che:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S}} f_n(x_j), \forall j, \text{ esiste finito in } \mathbb{R}(\mathbb{C})$$

Dunque, questo limite converge, $\forall x_j \in D$.

A questo punto, abbiamo costruito la sottosuccessione, che converge solo nel sottoinsieme $D \subset X$; si vuole dunque dimostrare che questa sottosuccessione converge $\forall x \in X$, ma non solo: si vuole anche dimostrare che esiste un certo $f \in X^*$ tale per cui questa sottosuccessione converge a $f(x)$.

Si parte dunque con il **secondo passo** della dimostrazione: l'estensione della convergenza all'intero X . Si vuole dunque verificare che:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S}} f_n(x) \text{ esiste finito in } \mathbb{R}(\mathbb{C}) \forall x \in X$$

Per fare ciò, si sfrutta la definizione di densità: ogni punto x è approssimabile arbitrariamente bene (a meno dunque di un ε) con un elemento $\bar{x} \in D$; ossia:

$$\forall \varepsilon > 0 \forall x \in X \exists \bar{x} \in D : \|x - \bar{x}\| < \varepsilon$$

Su \bar{x} , abbiamo la certezza che la sottosuccessione converge; lo abbiamo dimostrato nel punto precedente. Dal momento che consideriamo successioni di funzionali, del tipo $\{f_n(x), n \in S\}$, esse sono successioni numeriche, la cui immagine è in \mathbb{R} (o in \mathbb{C}); possiamo sfruttare la completezza dello spazio dei numeri reali (o complessi) per dire che la sottosuccessione è convergente se e solo se è di Cauchy; in altre parole, per la convergenza, basta dimostrare che le $f_n(x)$ sono di Cauchy. Dunque, si deve stimare:

$$|f_n(x) - f_m(x)|$$

questo si può fare come segue:

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_n(\bar{x})| + |f_n(\bar{x}) - f_m(\bar{x})| + |f_m(\bar{x}) - f_m(x)|$$

A questo punto, è possibile maggiorare i tre termini come segue:

- per il primo termine:

$$|f_n(x) - f_m(\bar{x})| \leq \|f_m\| \|x - \bar{x}\|$$

dove la norma del funzionale è limitata a una certa M , essendo il funzionale limitato (continuo), mentre, $\|x - \bar{x}\|$ è minore di ε per ipotesi:

$$\leq M\varepsilon$$

- per quanto riguarda il secondo termine:

$$|f_n(\bar{x}) - f_m(\bar{x})| < \varepsilon, n, m > N_\varepsilon$$

questo, dal momento che stiamo valutando i funzionali in un \bar{x} dove si ha la certezza di avere convergenza.

- il terzo termine è analogo al primo.

Di conseguenza:

$$|f_n(x) - f_m(x)| \leq (2M + 1)\varepsilon$$

dove M è indipendente da ε ; dunque, abbiamo verificato che la successione è di Cauchy e, per completezza dei reali, pure convergente. Questo dimostra che:

$$\lim_{\substack{n \rightarrow \infty \\ n \in S}} f_n(x), \forall x \in X, \text{ esiste finito}$$

A questo punto, rimane ancora dimostrare che nello spazio duale esiste il *funzionale limite*; per fare ciò, si definisce $f(x)$ come segue:

$$f(x) := \lim_{\substack{n \rightarrow \infty \\ n \in S}} f_n(x)$$

Questo f senza dubbio è un funzionale lineare; inoltre, esso è continuo, dal momento che è possibile scrivere:

$$|f(x)| = \lim_{\substack{n \rightarrow \infty \\ n \in S}} |f_n(x)|$$

ma, per ipotesi, $|f_n(x)|$ ha norma limitata in X^* : limitata a M . Dunque:

$$|f(x)| \leq M \|x\| \quad \forall x \in X$$

Quindi, $\exists, f \in X^*$, essendo lineare e continuo. Per costruzione di f , dunque, si ha che:

$$\{f_n, n \in S\} \xrightarrow{*} f \in X^*$$

La limitatezza è sufficiente, per garantire questa convergenza.

Osservazione

Si osservi che, data una successione $\{f_n\} \in X^*$ di funzionali lineari e continui, ed esiste finito il limite di f_n , in modo tale da definire $f(x)$ come visto, il funzionale $f(x)$ certamente sarà lineare, ma non è detto che esso sia continuo. Perché sia anche continuo, è necessario che lo spazio X di partenza sia completo. Questo accade dal momento che $\{f_n\}$ in questa condizione è puntualmente limitata:

$$\sup_n |f_n(x)| < \infty, \quad \forall x$$

questo, grazie al fatto che $f_n(x)$ converge su \mathbb{R} . Questa $\{f_n\}$, essendo univoca e convergente, è anche limitata (in quanto convergente); per Banach-Steinhaus, dunque $\exists, M > 0$ tale per cui:

$$\|f_n\|_{X^*} \leq M$$

quindi, è chiaro che:

$$\{f\}_{X^*} < \infty : f \in X^*$$

Corollario del teorema di Alaoglu (applicazione per soluzione PDE)

In uno spazio di Banach riflessivo X , con X^* separabile, ogni successione $\{x_n\}$ limitata in X ha una sottosuccessione convergente debolmente.

Questo corollario ricorda il teorema di Bolzano-Weierstrass, dove però si forniscono garanzie sulla convergenza debole, invece che su quella in norma.

Dimostriamo questo corollario: grazie all'ipotesi di riflessività, è possibile dire che $X = X^{**}$; dunque, data $\{x_n\}$ limitata su X , si ha che:

$$\{i(x_n)\} \subset X^{**}$$

è limitata; questo, grazie al fatto che l'applicazione i è una isometria, ossia $\{i(x_n)\}_{X^{**}} = \|x_n\|_X$. A questo punto, si applichi il teorema di Alaoglu, utilizzando come spazio di partenza X^* invece che X . Il teorema di Alaoglu garantisce che esiste una sottosuccessione di $\{i(x_n)\}$, $\{i(x_{n_k})\}$ che converge debolmente* a un certo funzionale $\varphi \in X^{**}$. Ossia:

$$i(x_{n_k})f \rightarrow \varphi(f) \forall f \in X^*$$

si ribadisce: in questo caso lo spazio su cui si applica Alaoglu è X^* , ossia il duale di X , e lo spazio duale di questo sarà chiaramente il bidual, isomorfo (per riflessività) allo spazio di partenza. Tuttavia, si ha che:

$$i(x_{n_k})f := f(x_{n_k})$$

ma dunque:

$$f(x_{n_k}) \rightarrow \varphi(f)$$

A questo punto, è possibile usare ancora la riflessività di X , ricordando che il φ di cui si sta parlando deve provenire da un elemento $x \in X$, mappato mediante l'isometria i ; questo è garantito dalla suriettività della mappa. Siccome dunque X è riflessivo, $\exists x \in X$ tale per cui:

$$\varphi = i(x)$$

dunque:

$$\varphi(f) = f(x)$$

ma dunque:

$$f(x_{n_k}) \rightarrow f(x) \forall f \in X^*$$

e questo conclude la dimostrazione; infatti, quest'ultima frase semplicemente significa:

$$x_{n_k} \rightharpoonup x$$

Alcune osservazioni conclusive:

- la separabilità di X^* in verità non è necessaria come ipotesi; è possibile ometterla;
- si è scritto che X deve essere uno spazio di Banach; in verità, questa proprietà è una diretta conseguenza della riflessività, dal momento che se uno spazio è riflessivo allora esso è isomorfo a uno spazio completo, che è il duale del duale (ma dunque, essendo uno spazio duale di un altro, esso è sicuramente completo);
- se lo spazio X^* è separabile, allora X è separabile; il viceversa non è valido (esempio, L^1 è separabile, L^∞ no).

Capitolo 5

Operatori lineari

5.1 Introduzione ed esempi

Un operatore lineare è un'applicazione lineare tra due spazi normati. Dati dunque X, Y due spazi normati, definiti **sullo stesso campo** (entrambi reali o entrambi complessi), allora un'applicazione lineare $X \rightarrow Y$ si dice **operatore lineare**. Un funzionale è un particolare caso di operatore lineare, in cui $Y = \mathbb{R}$ (o \mathbb{C}). Dunque, questo significa che molti risultati ottenuti nell'ambito di studio della teoria dei funzionali valgono anche ora. Nel dettaglio, le seguenti proprietà sono equivalenti; dato A operatore:

1. A è limitato:

$$\exists C > 0 : \|Ax\|_Y \leq C \|x\|_X, \forall x \in X$$

2. A è continuo in $x_0 = 0$;
3. A è continuo $\forall x_0 \in X$.

Parlare di continuità ha senso: essendo X e Y spazi normati, essi hanno una metrica indotta dalla norma, e dunque esistono gli intorni e tutti gli altri concetti propri della continuità. Se dunque A è continuo, è possibile definire per A una norma, $\|A\|$, come:

$$\|A\| = \inf \{C > 0 : \|Ax\| \leq C \|x\| \forall x \in X\} = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\| \leq 1} \|Ax\|$$

questi, se $X \neq \{\emptyset\}$. In particolare, si ha che:

$$\|Ax\| \leq \|A\| \|x\| \quad \forall x \in X$$

Dato:

$$\ker \{A\} = \{x \in X : Ax = 0\}$$

detto **nucleo** dell'operatore A , esso segue le solite proprietà del nucleo: se A è lineare è un sottospazio, se A è continuo è chiuso (dal momento che è la controimmagine di un chiuso); si osservi che in generale non è detto che se il nucleo dell'operatore è chiuso, allora l'operatore A è limitato (a meno che A non sia un funzionale e non un operatore in senso più generale).

Nel caso di uno spazio di Hilbert H , data $A : H \rightarrow H$ un'applicazione limitata, allora:

$$\|A\| = \sup \{|(Ax|y)| : \|x\| \leq 1, \|y\| \leq 1\}$$

Questo è vero dal momento che, in uno spazio di Hilbert, la norma di un generico elemento z si può scrivere come:

$$\|z\| = \sup \{|(z|y)| : \|y\| \leq 1\}$$

Questo perché, per z fissato, questo è un funzionale che a y associa $(y|z)$, che è un funzionale da H in \mathbb{C} ; dunque, la norma di questo funzionale lineare è definita proprio come la norma di z : è il sup.

Modo alternativo:

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|$$

ma, su uno spazio di Hilbert, abbiamo appena visto che, sostituendo a z il Ax , si ha:

$$\|A\| = \sup_{\|x\| \leq 1} \sup_{\|y\| \leq 1} |(Ax|y)|$$

inoltre, i sup possono commutare.

5.1.1 Esempi di operatori lineari

Esempio 1: operatore di Volterra/Fredholm

Si considerino $X = Y = C([a, b])$, e un'applicazione $A : X \rightarrow Y$ definita come:

$$(Ax)(f) = \int_a^b K(t, \tau)x(\tau) d\tau$$

τ è una variabile di integrazione, K è detto **nucleo integrale** (integral kernel). $K(t, \tau)$ è in questo esempio una funzione continua su un intervallo $[a, b] \times [a, b]$; il risultato dell'integrazione in τ , è una funzione in t . Questo operatore è detto **operatore di Fredholm**.

Studiamo le proprietà di questo operatore: A è un operatore lineare (grazie alla linearità dell'integrale); inoltre, A è limitato, dal momento che si può dimostrare che $\exists C > 0 : \|Ax\|_{C([a,b])} \leq C \|x\|_{C([a,b])} \forall x \in C([a, b])$. La norma che si utilizza è la norma del sup.

A questo punto, è necessario valutare $|Ax(t)|$, per provare a maggiorarlo; si ha:

$$|Ax(t)| = \left| \int_a^b K(t, \tau)x(\tau) d\tau \right|$$

dal momento che si vuole ottenere una certa costante moltiplicata per la norma di x , dovremo maggiorare questo integrale; la prima mossa da fare è portare il modulo all'interno, e quindi maggiorare $x(\tau)$ con il sup di x :

$$\leq \int_a^b |K(t, \tau)| |x(\tau)| d\tau \leq \sup_{\tau} |x(\tau)| \int_a^b |K(t, \tau)| d\tau$$

ora, il sup è la norma, mentre l'altero termine è una certa funzione di t :

$$= \|x\|_{C([a,b])} \int_a^b |K(t, \tau)| d\tau$$

a questo punto:

$$\|Ax\| = \sup_{t \in [a,b]} |Ax(t)| \leq \|x\|_{C([a,b])} \sup_{t \in [a,b]} \int_a^b |K(t, \tau)| d\tau$$

questo sup è una costante. Dunque, A è un operatore continuo, e la sua norma è minore o uguale di questo integrale:

$$C \leq \sup_{t \in [a, b]} \int_a^b |K(t, \tau)| d\tau$$

Operatori di Hilbert-Schmidt

Si consideri a questo punto $X = Y = L^2(a, b)$; $A : X \rightarrow Y$, dove

$$(Ax)(f) = \int_a^b K(t, \tau)x(\tau) d\tau$$

dove il kernel integrale $K(t, \tau)$ è una funzione in $L^2([a, b] \times [a, b])$. Questo significa che:

$$\int_a^b \int_a^b |K(t, \tau)|^2 dt d\tau = \|K\|_{L^2(a, b) \times [a, b]} < \infty$$

ossia, K è una funzione misurabile (si osservi che l'ordine di integrazione si può scambiare, dal momento che, per il teorema di Tonelli, con funzioni integrande positive, è possibile scambiare l'ordine di integrazione, e calcolare l'integrale nei vari modi produrrà lo stesso risultato).

A questo punto, la domanda è: A è in L^2 ? È lineare, continuo?

La continuità richiede che:

$$\|Ax\|_{L^2(a, b)} \leq C \|x\|_{L^2(a, b)}$$

Tuttavia, ora dobbiamo scrivere $\|Ax\|_{L^2(a, b)}$; ma esso, è l'integrale del modulo quadro! Quindi, è possibile studiare in un colpo solo sia la continuità, sia l'appartenenza a $L^2(a, b)$!

$$\|Ax\|_{L^2(a, b)}^2 = \int_a^b |Ax(t)|^2 dt$$

(questo, elevando direttamente al quadrato tutto). A questo punto, è necessario stimare l'integranda:

$$|Ax(t)|^2 \leq \int_a^b |K(t, \tau)|^2 |x(\tau)|^2 d\tau$$

a ciò è possibile applicare la disuguaglianza di Hölder:

$$|Ax(t)|^2 \leq \int_a^b |K(t, \tau)|^2 d\tau \int_a^b |x(\tau)|^2 d\tau = \int_a^b |K(t, \tau)|^2 d\tau \|x\|_{L^2(a,b)}^2$$

Questa, era l'integranda; ora, è necessario integrare tutto ciò:

$$\int_a^b |Ax(t)|^2 dt \leq \|x\|_{L^2(a,b)}^2 \int_a^b \int_a^b |K(t, \tau)|^2 d\tau dt = \|x\|_{L^2(a,b)}^2 \|K\|_{L^2([a,b] \times [a,b])}$$

queste non dipendono più né da t né da τ . Quindi, si ha che:

$$\|A\| = C \leq \|K\|_{L^2([a,b] \times [a,b])}$$

Terzo esempio: operatore di moltiplicazione

Si considerino $X = Y = L^2(a, b)$, e $A : X \rightarrow Y$ definita come:

$$Ax(t) = K(t)x(t)$$

questo è un **operatore di moltiplicazione**. K è una funzione continua. Si vuole verificare che questo operatore sia lineare, continuo, e che invii effettivamente la funzione in $L^2(a, b)$. La linearità è banale; per quanto riguarda le altre due proprietà, si procede come prima:

$$\|Ax\|_{L^2} = \left(\int_a^b |Ax(t)|^2 dt \right)^{\frac{1}{2}} = \left(\int_a^b |K(t)|^2 |x(t)|^2 dt \right)^{\frac{1}{2}}$$

L'obiettivo è controllare questa espressione mediante la norma 2 di x ; per fare ciò, è sufficiente maggiorare il kernel integrale mediante il suo sup:

$$\|Ax\|_{L^2} \leq \sup_t |K(t)| \left(\int_a^b |x(t)|^2 dt \right)^{\frac{1}{2}} = \|x\|_{L^2} \sup_t |K(t)|$$

ciò verifica la limitatezza. In questo particolare caso, tuttavia, è possibile dimostrare che:

$$\|A\| = \sup_t |K(t)|$$

ossia, che questa è esattamente la norma, e non solamente una generica costante che può essere usata per maggiorare. Per fare ciò, si definisca:

$$C = \|K\|_{C([a,b])}$$

è necessario trovare una $x(t)$ non nulla tale per cui valga l'eguaglianza. Fare ciò non è banale, di conseguenza quello che faremo è esibire una successione di funzioni $x_n \in L^2$ tale per cui:

$$\frac{\|Ax_n\|}{\|x_n\|} \rightarrow C$$

Al fine di fare ciò, consideriamo una successione x_n tale da avere norma pari a 1 (questa è solo una normalizzazione di comodo per procedere); quindi, si vuole che $\|Ax_n\| \rightarrow C$. Essendo K continuo, il sup è un max; questo è il valore di $K(t)$ in un certo punto $t = t_0$. Di conseguenza, come x_n è possibile scegliere una successione che converga alla delta di Dirac, valutata nel punto; infatti, si vuole che:

$$|K(t_0)| = C$$

dunque, si sceglie:

$$x_n(t) = \begin{cases} \sqrt{\frac{n}{2}} & \text{se } t_0 - \frac{1}{n} < t < t_0 + \frac{1}{n} \\ 0 & \text{altrove} \end{cases}$$

si spieghi meglio ciò: fissato un certo n , si vuole che la funzione in questione sia sostanzialmente una porta di costante M ed esistente tra $-\frac{1}{n}$ e $\frac{1}{n}$, tale per cui la sua norma 2 valga 1; dunque:

$$\left(\int_{t_0 - \frac{1}{n}}^{t_0 + \frac{1}{n}} M dt \right)^{\frac{1}{2}} = 1$$

da cui:

$$\sqrt{\frac{2}{n}} M = 1$$

quindi:

$$M = \sqrt{\frac{n}{2}}$$

A questo punto, si ha che:

$$\|Ax_n\|_{L^2} = \left(\int_a^b |K(t)|^2 |x_n(t)|^2 dt \right)^{\frac{1}{2}} = \left(\frac{n}{2} \int_{t_0 - \frac{1}{n}}^{t_0 + \frac{1}{n}} |K(t)|^2 dt \right)^{\frac{1}{2}}$$

questa, per $n \rightarrow \infty$, tende a $|K(t_0)|$, dal momento che essa è la **media** della funzione la quale, stringendo l'intervallo, tende al valore della funzione stessa. Essendo la norma del sup (con la quale stiamo studiando il kernel integrale $K(t)$) il sup della funzione, ed essendo t_0 per ipotesi il punto di massimo allora il gioco è fatto:

$$\frac{\|Ax_n\|}{\|x_n\|} \rightarrow |K(t_0)| = C$$

Quarto esempio: operatore di shift

Si considerino ora $X = Y = l^2$, uguali a uno spazio di successioni; si consideri dunque il generico $x \in l^2$ definito come:

$$x = (x_1, x_2, x_3, \dots)$$

dunque, si definisce $A : X \rightarrow Y$ come:

$$Ax = (0, x_1, x_2, \dots)$$

questo consiste nel mettere 0 nella prima posizione, e spostare “verso destra” tutti gli altri elementi. Dunque, è banale vedere che se $x \in l^2$, anche $Ax \in l^2$; inoltre:

$$\|Ax\|_{l^2} = \|x\|_{l^2}$$

e questo, per ogni x ; dunque, questo significa, a parole, che l'operatore A è un'isometria; inoltre, $\|A\| = 1$ soddisfa esattamente l'eguaglianza.

Quinto esempio: ulteriore esempio sulle successioni

Si considerino a questo punto $X, Y \subset l^2$, e come $A : X \rightarrow Y$ l'operatore che applica su un generico $x \in X$, nella forma:

$$x = (x_1, x_2, x_3, \dots)$$

nel seguente modo:

$$Ax = \left(0, \left(1 - \frac{1}{2}\right) x_2, \left(1 - \frac{1}{3}\right) x_3, \left(1 - \frac{1}{4}\right) x_4, \dots \right)$$

ossia, la k -esima componente della successione finale è:

$$(Ax)_k = \left(1 - \frac{1}{k}\right) x_k$$

Presentiamo le solite verifiche: continuità e appartenenza a Y .

$$\|Ax\|_{l^2} \leq C \|x\|_{l^2}$$

dal momento che:

$$\|Ax\|_{l^2} = \left(\sum_{k=1}^{\infty} \left| \left(1 - \frac{1}{k}\right) x_k \right|^2 \right)^{\frac{1}{2}}$$

essendo poi:

$$1 - \frac{1}{k} < 1$$

si ha che:

$$\|Ax\|_{l^2} \leq \left(\sum_{k=1}^{\infty} |x_k|^2 \right)^{\frac{1}{2}} = \|x\|_{l^2}$$

Quindi:

$$C \leq 1$$

Dunque:

$$\|Ax\|_{l^2} \leq \|x\|_{l^2} \quad \forall x \in l^2$$

Dunque, A ha effettivamente valori in l^2 , ed è lineare, continuo, con $C \leq 1$.
Si vuole però verificare che:

$$\|A\| = 1$$

Per fare ciò, è necessario esibire una certa $\{x^{(n)}\}$, con norma unitaria (per facilità, la solita normalizzazione di comodo), e tale per cui:

$$\|Ax_n\| = \|x_n\|$$

dunque, si vuole che:

$$\|A\| = 1$$

ossia:

$$\frac{\|Ax^{(n)}\|}{\|x^{(n)}\|} \rightarrow 1$$

una $x^{(n)}$ che si può usare è la solita e_n :

$$x^{(n)} = e_n = (0, 0, 0, \dots, 1, 0, \dots)$$

dove si ha un 1 solo nella n -esima posizione. Qui, chiaramente:

$$\|x^{(n)}\|_{l^2} = 1, \forall n$$

inoltre:

$$\|Ax^{(n)}\| = \left\| \left(0, 0, 0, \dots, 1 - \frac{1}{n}, 0, \dots \right) \right\|_{l^2} = 1 - \frac{1}{n}$$

dunque, per $n \rightarrow \infty$, questa tende a 1. Ciò dimostra quello che si voleva ottenere.

Sesto esempio: versione discreta dell'operatore di Hilbert-Schmidt

Si considerino $X = Y = l^2$, e l'operatore $A : X \rightarrow Y$ definito come:

$$(Ax)_i = \sum_{j=1}^{\infty} \alpha_{ij} x_j$$

dove $\{\alpha_{ij}\}$ è una matrice infinita, tale per cui:

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |\alpha_{ij}|^2 < \infty$$

Questo operatore si può pensare come l'estensione in dimensione infinita del prodotto riga per colonna: $x \in l^2$ è un vettore colonna, e Ax è un altro vettore colonna, dato dalla moltiplicazione della matrice A , definita per componenti:

$$A = (\alpha_{ij}), i, j = 1 \dots n$$

dove $n \rightarrow \infty$. Questa era l'idea da cui Hilbert è partito per definire la sua visione dell'analisi funzionale: dall'algebra lineare è noto che gli operatori si possono scrivere mediante matrici; l'idea era quella di rappresentare un operatore in spazi di dimensione infinita come matrici infinite.

Se al posto di t si mette l'indice i discreto, al posto di τ si mette j , al posto di K si mettono gli α_{ij} , al posto dell'integrale la sommatoria, si ottiene un qualcosa di analogo all'operatore di Hilbert-Schmidt, prima visto in versione continua.

Si può dire, per questo, che:

$$\|\alpha\|_{l^2(\mathbb{N} \times \mathbb{N})}^2 = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |\alpha_{ij}|^2 < \infty$$

Dunque: A è un operatore continuo, e:

$$\|A\| \leq \|\alpha\|$$

Dove $\|\alpha\|$ è detta **norma di Frobenius**.

Settimo esempio: operatore di derivazione

Si consideri come spazio di partenza $X = C^{(1)}([a, b])$, e come spazio di arrivo $Y = C([a, b])$; si consideri su $C^{(1)}$ la seguente norma:

$$\|x\|_{C^{(1)}([a,b])} = \sup_{t \in [a,b]} |x(t)| + \sup_{t \in [a,b]} |x'(t)|$$

con questa, lo spazio diventa uno spazio di Banach; altrimenti, è un generico spazio normato, ma non più completo. Si definisca quindi l'operatore $A : X \rightarrow Y$ di derivazione:

$$Ax = x'$$

questo è un operatore lineare e continuo; infatti:

$$\|x'\|_{C([a,b])} \leq \|x\|_{C^{(1)}([a,b])}$$

questo è vero dal momento che:

$$\|x\|_{C^{(1)}([a,b])} = \sup_{t \in [a,b]} |x(t)| + \sup_{t \in [a,b]} |x'(t)|$$

da cui:

$$\|x'\|_{C([a,b])} \leq \sup_{t \in [a,b]} |x(t)| + \sup_{t \in [a,b]} |x'(t)|$$

e questo è sempre verificato. Si osservi che l'operatore di derivazione dunque è continuo, ma solo con questa norma. Si può verificare inoltre che, usando i soliti passi (cercando una opportuna successione di funzioni), la norma è esattamente quella usata (vale l'eguaglianza e non solo il minore uguale).

Se su $C^{(1)}$ si prende la norma del sup, senza sommare l'altro termine, allora si ha uno spazio non completo, e A inoltre non è più un operatore continuo; questo può essere verificato cercando una successione che violi una delle stime di limitatezza.

5.1.2 Proprietà aggiuntive dello spazio degli operatori continui

Si denoti con $\mathcal{L}(X, Y)$ lo spazio degli operatori continui (ossia limitati: \mathcal{L} sta proprio per limitati) da X a Y , dove X e Y sono spazi normati. Vale il seguente teorema: se lo spazio Y è uno spazio di Banach, allora $\mathcal{L}(X, Y)$ è uno spazio di Banach. Questo teorema è interessante, dal momento che ha alcuni corollari, come quello che segue: se X è uno spazio normato, X^* è uno spazio completo; in altre parole:

$$X^* = \mathcal{L}(X, \mathbb{R})$$

e dunque, in virtù del teorema, X^* è uno spazio completo.

Dimostriamo a questo punto il teorema: si consideri una certa successione di Cauchy, e si verifichi che essa è convergente; questa è la verifica della completezza. Sia dunque $\{A_n\} \subset \mathcal{L}(X, Y)$ una successione di Cauchy di operatori lineari continui; questo significa dunque che, $\forall \varepsilon \exists N_\varepsilon$ tale per cui, $\forall n, m > N_\varepsilon$, si ha:

$$\|A_n - A_m\| < \varepsilon$$

questa è una norma di operatori; quindi, è possibile scrivere ciò per esteso come segue:

$$\forall \varepsilon > 0 \exists N_\varepsilon : \forall n, m > N_\varepsilon \|A_n x - A_m x\| \leq \varepsilon \|x\| \quad \forall x \in X$$

questo segue semplicemente dalla proposizione precedente, e dalla definizione di norma di un operatore; si osservi che si è usato il \leq , dal momento che, altrimenti, la disuguaglianza non potrebbe valere nel caso $x = 0$, in cui gli operatori, essendo lineari, hanno immagine nulla.

A questo punto, per ogni $x \in X$ fissato, la successione $\{A_n x\}$ ha valori in Y , ed è di Cauchy in Y ; essendo tuttavia Y uno spazio di Banach per le ipotesi, ergo completo, la successione è anche convergente. A questo punto, dunque, si definisca il seguente limite:

$$Ax := \lim_{n \rightarrow \infty} A_n x$$

lo scopo di questa ultima definizione era quella di costruire un candidato limite; A dunque è sicuramente lineare, ma si può dire che sia anche continuo, ossia che $A \in \mathcal{L}(X, Y)$? Inoltre, $A_n \rightarrow A$, in $\mathcal{L}(X, Y)$? Si osservi che la risposta a questo domanda non discende banalmente dalla definizione, dal momento che essa afferma che si ha una convergenza puntuale all'operatore; invece, dire che A_n converge a A significa:

$$A_n \rightarrow A \iff \|A_n - A\| \rightarrow 0$$

In realtà, i due punti critici, ossia la continuità e la convergenza, si possono verificare assieme; si consideri n fissato in un primo momento, e ε fissato; m può variare, e si consideri:

$$\lim_{m \rightarrow \infty} \|A_n x - A_m x\| \leq \varepsilon \|x\|$$

dal momento che è lecito per linearità passare al limite dentro al segno di norma, si può dire che:

$$\forall \varepsilon > 0 \exists N_\varepsilon : \forall n, m > N_\varepsilon, \|A_n x - Ax\| \leq \varepsilon \|x\| \quad \forall x \in X$$

questo garantisce dunque che $A_n - A$ è un operatore limitato; dal momento che A_n è limitato per ipotesi, e $A_n - A$ lo è per quello che è stato appena detto, si può scrivere A come:

$$A = A_n - (A_n - A)$$

ma, dunque, anche A è limitato, per le ipotesi appena proposte. La proposizione scritta poche righe sopra inoltre si può leggere come segue:

$$\forall \varepsilon > 0 \exists N_\varepsilon : \forall n, m > N_\varepsilon, \|A_n - A\| \leq \varepsilon$$

che significa esattamente che:

$$\|A_n - A\| \rightarrow 0, n \rightarrow \infty$$

questo conclude la dimostrazione.

5.2 Operatori compatti

Finora sono state studiate le proprietà di operatori che possiedono solo la continuità; a questo punto, si chiede qualcosa in più, ossia qualche ipotesi aggiuntiva, al fine di ottenere qualcosa in più.

Siano dunque X, Y due spazi normati; un operatore lineare $A : X \rightarrow Y$ viene detto **compatto** se per ogni successione $\{x_n\}$ limitata in X , la successione $\{Ax_n\}$ ha una sottosuccessione convergente in Y . Date dunque successioni limitate nello spazio di partenza, queste vengono mandate in successioni in Y che ammettono una sottosuccessione convergente.

Questa nozione è più forte rispetto a quella di operatore continuo; infatti, se A è un operatore compatto, esso è anche un operatore continuo, ma il viceversa non è vero. Dato dunque A operatore compatto, si ha che:

$$A \in \mathcal{L}(X, Y)$$

Per dimostrare ciò, si ricorda che dire che $A \in \mathcal{L}(X, Y)$ significa che:

$$\exists C : \|Ax\| \leq C \|x\| \quad \forall x \in X$$

questo coincide con dire:

$$\|Ax\| \leq C \|x\| \leq 1$$

questo significa che l'immagine di A assume valori contenuti nella palla di centro pari all'origine e centro pari a C ; in altre parole ancora, questo è:

$$\sup_{\|x\| \leq 1} \|Ax\| \leq C$$

ossia, l'immagine della palla unitaria in X viene mappata su una regione limitata. In effetti, dunque, data l'ipotesi di operatore compatto, A è limitato sulla palla unitaria, dal momento che, se non lo fosse, esisterebbe una successione di punti $\{y_i\} \subset Y$ (appartenenti allo spazio immagine) la cui norma diverge all'infinito; dunque, esisterebbe $\{y_n\} = \{Ax_n\}$, con $\|x_n\| \leq 1$, e $\|y_n\| \rightarrow \infty$; questo significherebbe che l'immagine è non limitata, ma ciò andrebbe in contraddizione col fatto che l'operatore è compatto dal momento che, in questa situazione, nessuna sottosuccessione di $\{y_n\}$ potrebbe convergere.

L'aggettivo **compatto** per la qualificazione di un operatore fa pensare alla compattezza di insiemi; dato uno spazio metrico E , e un suo sottoinsieme M , il sottoinsieme M si dice **compatto** se ogni successione $\{x_n\}$ a valori in M ha una sottosuccessione convergente a un certo punto di M ; questa è la definizione di **compattezza sequenziale**. Equivalentemente (almeno, per quanto riguarda gli spazi metrici), si dice compatto se ogni ricoprimento aperto di M ammette un sottoricoprimento finito; se dunque M è ricoperto da una qualsiasi famiglia di aperti, anche non numerabile, basta un certo numero finito di aperti per ricoprire M . Dal momento che lavoriamo su spazi normati per ipotesi, essi sono metrici, e dunque queste due definizioni sono equivalenti; altrimenti, in spazi topologici, solo la seconda definizione ha senso.

Esiste una definizione di compattezza più debole: M si dice **relativamente compatto** se ogni successione $\{x_n\}$ ha una sottosuccessione con limite convergente in E , ma anche non in M ; equivalentemente, è come chiedere che \overline{M} , invece di M , sia compatta.

Esempi

Per fissare le idee sugli insiemi compatti, verranno ora proposti alcuni esempi.

- Per il teorema di Heine-Borel, in \mathbb{R}^n o \mathbb{C}^n , compatto significa **chiuso e limitato**; allo stesso modo, relativamente compatto vuol dire **solo limitato**.

- Dato uno spazio di Hilbert **di dimensione infinita**, i risultati del punto precedente saltano; la palla unitaria chiusa, definita come:

$$\{x : \|x\| \leq 1\}$$

è chiusa e limitata, ma non è compatta. Questa cosa si può dimostrare pensando a un sistema ortonormale infinito per l^2 ; usando questa, per esempio $\{e_n\}$, come base, allora non esistono sottosuccessioni convergenti, dal momento che due punti di questi elementi e_n distano almeno $\sqrt{2}$, dunque non possono avvicinarsi, **accumularsi**, verso un punto limite.

Si osservi tuttavia che, anche in dimensione infinita, se uno spazio è compatto, **allora** è chiuso e limitato; allo stesso modo, **relativamente compatto** implica **limitato**.

- Dato uno spazio ambiente E completo, si può dimostrare che ogni $M \subset E$ è relativamente compatto se e solo se M è totalmente limitato. Questo significa che, $\forall \varepsilon > 0$, il sottospazio M si può ricoprire con un numero finito di palle di raggio ε .

Osservazioni aggiuntive su compattezza e relativa compattezza

Verranno a questo punto introdotte alcune osservazioni.

1. Dato un insieme relativamente compatto, un suo sottoinsieme è a sua volta relativamente compatto; in altre parole, se la successione converge nell'insieme *grande*, a meno che il limite stia nello spazio ambiente, la stessa cosa vale anche per il sottoinsieme.
2. Dato $A \in \mathcal{L}(X, Y)$, e l'immagine della palla unitaria è relativamente compatta, allora l'immagine di ogni insieme limitato è relativamente compatto. Questo si dimostra nella seguente maniera: dato $M_1 \in X$ limitato, significa che:

$$\|x\| \leq r \quad \forall x \in M_1$$

tuttavia, si ha che $A(M_1)$, ossia l'immagine dell'operatore A applicata su M_1 , è relativamente compatta, dal momento che, se si prende una successione di $A(M_1)$, ossia $\{y_n\} = \{Ax_n\}$, dove $\{x_n\} \subset M_1$, si ha che:

$$\frac{y_n}{r} = A\left(\frac{x_n}{r}\right)$$

questo ultimo passaggio, per linearità, dal momento che si può portare il r dentro all'operatore; dunque, si ha che $\frac{x_n}{r}$ sta certamente nella palla unitaria, e quindi questa è una successione che sta nell'immagine della palla unitaria dell'operatore; di conseguenza, esiste di sicuro una sottosuccessione $\left\{\frac{y_{n_k}}{r}\right\}$ convergente a un certo y :

$$\left\{\frac{y_{n_k}}{r}\right\} \rightarrow y$$

ma dunque, moltiplicando ambo i membri per r , si ottiene:

$$\{y_{n_k}\} \rightarrow ry$$

3. Un operatore lineare $A : X \rightarrow Y$ è compatto se e solo se l'immagine della palla unitaria è un insieme relativamente compatto o, equivalentemente, l'immagine di ogni insieme limitato è un insieme relativamente compatto.

Da tutte queste osservazioni, è possibile proporre il seguente concetto:

- gli operatori limitati mandano insiemi limitati in insiemi limitati;
- gli operatori compatti mandano insiemi limitati in insiemi relativamente compatti.

questo, tuttavia, si può osservare solamente nel caso di spazi di dimensione infinita. Ad ogni modo, un'osservazione che verrà motivata in seguito, riguarda l'immagine degli operatori compatti: in un qualche senso, essa è **piccola**; per esempio, vedremo che l'operatore identità non è più compatto. In seguito si studierà questa cosa.

5.2.1 Teorema di Ascoli-Arzelà

Si consideri K uno spazio metrico compatto nel campo \mathbb{R} ; questo è un concetto più generale rispetto a quello di intervallo in \mathbb{R} ; il teorema di Ascoli-Arzelà, introducendo una caratterizzazione tra insiemi relativamente compatti, verrà dunque dimostrato in questo contesto.

Si considerino le funzioni continue su K , $C(K)$; si consideri quindi $M \subset C(K)$; M è relativamente compatto **se e solo se** (essendoci la biimplicazione questa è una caratterizzazione!):

1. M è limitato, ossia:

$$\exists C > 0 : \|x\|_{C(K)} \leq C \forall x \in M$$

2. M è **equicontinuo**. Questa è una condizione più forte rispetto alla continuità: ricordando la definizione di continuità col $\delta - \varepsilon$, in generale, nel caso della continuità, δ dipende dalla funzione x che si prende; nel caso dell'equicontinuit , invece, si richiede che:

$$\forall \varepsilon > 0 \exists \delta > 0 : t_1, t_2 \in K, d(t_1, t_2) < \delta, \implies |x(t_1) - x(t_2)| < \varepsilon \forall x \in M$$

la particolarit  in questa definizione sta nel fatto che il $\forall x \in M$ si mette alla fine: il δ che esiste si sceglie **indipendentemente** dalla funzione x che si considera.

Si propone a questo punto la dimostrazione del teorema di Ascoli-Arzel . In verit , delle due implicazioni, considereremo solo quella *se*, e non la *e solo se*; questo significa che le due ipotesi appena proposte saranno assunte, e si cercher  come risultato la relativa compattezza di M . Questa, scritta in altre parole, dice che:

$$\forall \{x_n\} \subset M \exists \{x_{n_k}\}, \text{ sottosuccessione convergente in } C(K)$$

ossia, convergente uniformemente.

Questa dimostrazione   piuttosto simile a quella del teorema di Alaoglu: come in quel caso, il primo passo consiste nella ricerca di una sottosuccessione che converga in un sottoinsieme denso di K ; dal momento che uno spazio metrico compatto   **sempre** separabile¹, allora sia $D \subset K$ il sottoinsieme numerabile denso. Per ipotesi, si ha che $\{x_n\}$   limitata, dunque:

$$\exists C > 0 : \sup_{t \in K} |x_n(t)| \leq C \forall n$$

¹risultato noto dalla topologia

questa è la definizione di **limitatezza uniforme**. Essendo uniformemente limitata, questa è anche puntualmente limitata; dunque:

$$\forall t \in K \quad |x_n(t)| \leq C \forall n$$

si ha che $\{x_n(t)\}$ è una successione numerica, ma dunque, per Bolzano-Weierstrass, essa ha una sottosuccessione convergente.

Il nostro obiettivo sarebbe però avere un'unica successione, che converga uniformemente per ogni punto di D ; questo può essere ottenuto, analogamente a quanto fatto per Alaoglu, applicando nuovamente il procedimento diagonale di Cantor; questo garantisce e trova la sottosuccessione $\{x_{n_k}\}$ che converge puntualmente su D . A questo punto, si deve verificare che $\{x_{n_k}\}$ converge uniformemente su tutto K . Dal momento che lo spazio delle funzioni continue su K , $C(K)$, è continuo, basta verificare che la successione sia di Cauchy, per avere automaticamente la sua convergenza. Fissato dunque t , si parte da un punto di D che disti meno di ε da esso. Dato dunque ε , e dato un δ che segua l'ipotesi di equicontinuit , si osservi che K , essendo uno spazio metrico compatto, pu  essere scritto come un ricoprimento di palle aperte B_j , di raggio pari a $\frac{\delta}{2}$:

$$K = \bigcup_{j=1}^N B_j$$

ossia, come unione di palle; questo, per la definizione canonica di compattezza. Dunque, in ciascuna palla, e dunque $\forall j, \exists t_j \in D \cap B_j$ (essendo D denso, un $t_j \in D$ appartiene a una delle palle, ossia a uno degli intorni, B_j). Dato $t \in K$, si ha che:

$$\exists j : t \in B_j$$

ma dunque, essendo K ricoperto come detto, $d(t, t_j) < \delta$ per un certo j , ossia per una particolare palla; dal momento che le palle sono di raggio $\frac{\delta}{2}$, se t e t_j sono nella stessa palla, ma nel caso estremo nei due punti opposti, la loro distanza sar  comunque maggiorabile col diametro della palla, ossia δ . Quindi, si pu  scrivere, detto tutto ci :

$$|x_{n_k}(t) - x_{n_h}(t)| \leq |x_{n_k}(t) - x_{n_k}(t_j)| + |x_{n_k}(t_j) - x_{n_h}(t_j)| + |x_{n_h}(t_j) - x_{n_h}(t)|$$

il primo e il terzo possono essere maggiorati con ε , dal momento che, dato il δ che soddisfa le ipotesi, l'immagine, ossia il valore della funzione, è minore di ε , per ipotesi; il secondo termine si può maggiorare con ε , per $h, k > N_\varepsilon$, dal momento che la sottosuccessione è per ipotesi di Cauchy.

Rimane un ultimo possibile problema: N_ε , ossia la soglia, può dipendere da j ? La risposta è sì: a seconda della palla che si considera, la soglia per avere arbitrariamente vicine le sottosuccessioni, può essere che si abbia un N_ε diverso; tuttavia, i t_j sono in numero finito, essendo le palle del sottoriprimento finite per ipotesi, di conseguenza come N_ε si può prendere il più grande tra tutti quelli delle situazioni possibili; quindi:

$$|x_{n_j}(t) - x_{n_h}(t)| < 3\varepsilon \forall t \in K, \text{ se } k, h > N_\varepsilon$$

dove, data l'osservazione appena fatta, N_ε dipende solamente dalla tolleranza ε desiderata. Essendo dunque $\{x_{n_k}(t)\}$ maggiorabile con 3ε , essa è di Cauchy e, per completezza di $C(K)$, è anche convergente.

A questo punto, il teorema di Ascoli-Arzelà verrà applicato su un certo numero di esempi, al fine di fissare le idee su di esso.

Esempio 1

Si considerino c_1, c_2 due costanti positive; si consideri l'insieme M definito come:

$$M = \{x \in C^{(1)}([a, b]) : |x(t)| \leq c_1, |x'(t)| \leq c_2 \forall t \in [a, b]\}$$

questo è un sottoinsieme delle funzioni derivabili $C^{(1)}$, dal momento che non tutte le funzioni derivabili hanno derivata limitata; applicando il teorema di Ascoli-Arzelà, verrà dimostrato che M è un insieme relativamente compatto, dal momento che soddisfa le due condizioni.

1. M è limitato, dal momento che c_1 maggiora le funzioni dell'insieme, ottenendo dunque la limitatezza;
2. le funzioni appartenenti a M sono equicontinue, dal momento che $|x'(t)| \leq c_2$; questo garantisce il fatto che la funzione sia *equilipschitz*.

Si discute ora il secondo punto: la seconda implicazione garantisce che

$$|x(t_1) - x(t_2)| \leq c_2 |t_1 - t_2| \forall t_1, t_2 \in [a, b]$$

infatti, è possibile applicare sul valore assoluto a membro sinistro il teorema di Lagrange, il quale garantisce l'esistenza di un punto ξ tale per cui:

$$|x(t_1) - x(t_2)| = |x'(\xi)| |t_1 - t_2|$$

il punto ξ appartiene all'intervallo: $\xi \in [t_1, t_2]$. Tuttavia, si ha che:

$$|x'(t)| \leq c_2 \forall t \in [t_1, t_2]$$

di conseguenza, la dimostrazione è conclusa. Inoltre, dato ε valore rispetto a cui si migliora il valore assoluto, si ha che, scegliendo come δ il valore:

$$\delta = \frac{\varepsilon}{c_2}$$

si ha che:

$$|x(t_1) - x(t_2)| \leq \varepsilon$$

(usando come distanza quella ordinaria: $d(t_1, t_2) = |t_1 - t_2|$).

Esempio 2

Si consideri l'operatore $A : C([0, 1]) \rightarrow C([0, 1])$ definito come:

$$Ax(t) = \int_0^t x(\tau) d\tau$$

Dimostrare che l'operatore è compatto. Si osservi, prima di procedere, che in questo esempio, nel precedente, e pure nei seguenti, si lavora con sottoinsiemi di funzioni continue; questo è necessario dal momento che si sta lavorando con il teorema di Ascoli-Arzelà, il quale garantisce risultati solo nell'ambito delle funzioni continue.

Si ricorda, prima di procedere, che dire che l'operatore è compatto significa che, per ogni successione $\{x_n\}$ di elementi dello spazio di partenza limitata, la successione delle immagini ha una sottosuccessione convergente. Equivalentemente, l'immagine mediante A di un insieme limitato è un insieme relativamente compatto. Ancora, equivalentemente, l'immagine mediante A della palla unitaria dello spazio di partenza è un insieme relativamente compatto. Utilizziamo quest'ultima condizione, al fine di verificare la compattezza dell'operatore.

Prima di tutto: qual è l'immagine della palla compatta? Semplicemente, è l'insieme M , definito come:

$$M = \{Ax : \|x\| \leq 1\}$$

Si consideri dunque un elemento $y \in M$; si ha che dunque esso deriva da un certo x , mediante A :

$$y \in M \rightarrow y = Ax \quad \|x\| \leq 1$$

dunque, le y sono uniformemente limitate; infatti:

$$|y(t)| = \left| \int_0^t x(\tau) \, d\tau \right| \leq \int_0^t |x(\tau)| \, d\tau \leq t \leq 1$$

l'ultima maggiorazione dipende dal fatto che x è nella palla unitaria, il modulo dell'elemento si maggiora con la norma, ma dunque è tutto minore di 1. Si può inoltre applicare il teorema fondamentale del calcolo integrale, ottenendo:

$$|y'(t)| = |x(t)| \leq 1$$

e verificare che le due condizioni sono soddisfatte; questo 1 rappresenta sia la c_1 sia la c_2 dell'esercizio precedente, che dunque può essere applicato, dimostrando che l'insieme M è relativamente compatto.

Questo significa che questo operatore di integrazione non solo è continuo, ma è anche **compatto**; in altre parole, l'immagine è in un qualche senso *più piccola*, dal momento che la palla unitaria viene *schacciata*; questo deriva dal fatto che le funzioni, in seguito all'integrazione, sono continue, ma anche integrabili, e dunque in seguito all'applicazione dell'integrazione vengono mappate in un sottospazio dello spazio di partenza.

Esempio 3

Si consideri l'operatore A così definito:

$$A : C^{(1)}([a, b]) \rightarrow C([a, b])$$

dato da:

$$Ax = x$$

questa è la cosiddetta **iniezione canonica**. Si osservi che, a causa degli spazi in cui è definita, essa non è l'identità. Questo operatore è compatto, utilizzando come norme quelle canoniche (tali per cui gli spazi siano spazi di Banach).

La verifica segue ancora una volta dal primo esempio; si consideri dunque l'immagine della palla unitaria, la quale è relativamente compatta, dal momento che, se:

$$y = Ax \quad \|x\|_{C(1)} \leq 1$$

allora, a maggior ragione, essendo:

$$|y(t)| = |x(t)|$$

si ha che:

$$\|x\|_{C(1)} = \|x\|_C + \|x'\|_{C(1)} \leq 1$$

dal momento che:

$$|y'| = |x'|$$

e tutto ciò è ≤ 1 .

Questo è l'esempio più classico di **immersione compatta**: un esempio di uno spazio immerso in un altro; determinare il fatto che un'immersione è compatta può essere interessante, per sapere a priori che esistono sottosuccessioni convergenti nello spazio di arrivo.

Si osservi che se A così definito fosse stato:

$$A : C^{(i)} \rightarrow C^{(i)}$$

allora questo operatore non sarebbe stato l'immersione, ma l'identità; si dimostrerà in seguito che questo **non è compatto**: infatti, l'identità mappa successioni limitate in successioni limitate, senza però garantire il fatto che queste nello spazio di arrivo (coincidente con quello di partenza) abbia sottosuccessioni convergenti.

Esempio 4

Dato l'operatore A definito come:

$$A : C([a, b]) \rightarrow C([a, b])$$

dove:

$$Ax(t) = \int_a^b K(t, \tau) x(\tau) d\tau$$

dove $K \in C([a, b] \times [a, b])$ è il **nucleo integrale** dell'operatore, si vuole dimostrare che A è un operatore compatto. Un esempio di questo operatore A può essere:

$$Ax(t) = \int_0^1 \sin(t + \tau)x(\tau) d\tau$$

Per quanto riguarda questo operatore, la continuità era già stata verificata; a questo punto, si vuole verificare se l'immagine della palla unitaria è un insieme relativamente compatto. Ancora una volta, si applica Ascoli-Arzelà, in modo diverso da prima però; si definisce dunque M come:

$$M = \left\{ y = Ax \mid \|x\|_{C([a,b])} \leq 1 \right\}$$

Come sempre, si devono verificare limitatezza e equicontinuità dell'operatore. Per quanto concerne la limitatezza, essa consegue dalla continuità dell'operatore (continuità è equivalente a limitatezza):

$$\|Ax\| \leq C \|x\|$$

Per quanto riguarda l'equicontinuità, sia $y = Ax \in M$, e si valuti la differenza di due funzioni in due punti distinti:

$$\begin{aligned} |y(t_1) - y(t_2)| &= |Ax(t_1) - Ax(t_2)| = \\ &= \left| \int_a^b K(t_1, \tau)x(\tau) d\tau - \int_a^b K(t_2, \tau)x(\tau) d\tau \right| \leq \\ &\leq \int_a^b |K(t_1, \tau) - K(t_2, \tau)| |x(\tau)| d\tau \end{aligned}$$

dato un ε , questo termine è minore o uguale di ε a patto di prendere t_1, t_2 sufficientemente vicini; infatti, $|x(\tau)| \leq 1$ per l'ipotesi solita; inoltre, essendo K continua, si ha che:

$$|K(t_1, \tau) - K(t_2, \tau)| \leq \varepsilon \text{ se } |t_1, t_2| < \delta_\varepsilon$$

per dimostrare ciò, è necessario disporre della uniforme continuità; tuttavia, dal momento che K è continua in $[a, b] \times [a, b]$, che è un insieme chiuso e limitato in \mathbb{R}^2 , ma dunque esso è compatto per Heine-Borel; essendo la funzione continua su un compatto, per Heine-Cantor essa è anche uniformemente continua. In più, dal momento che la funzione varia solo in t , ma non in τ (ossia, si può fissare τ a una certa quota), δ_ε non dipende dalla funzione: tutto ciò dunque è minore o uguale di:

$$\leq \varepsilon(b - a)$$

quindi, si è verificato che:

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : |y(t_1) - y(t_2)| \leq \varepsilon(b - 1) \forall y \in M$$

L'importante è che δ non dipenda da y , ma solo da ε , dunque tutto va bene.

5.3 Spazio degli operatori compatti

Si vuole a questo punto introdurre una struttura vettoriale all'insieme degli operatori compatti, al fine di introdurre per essi alcune proprietà generali. Siano dunque X, Y due spazi normati, e sia $\mathcal{K}(X, Y)$ l'insieme degli operatori compatti da X in Y . Dal momento che tutti gli operatori compatti sono anche continui, ma che il viceversa non vale, allora si può dire che $\mathcal{K}(X, Y) \subset \mathcal{L}(X, Y)$; in verità tuttavia $\mathcal{K}(X, Y)$ non è solo un sottoinsieme, ma anche un sottospazio vettoriale. Vale dunque il seguente teorema.

1. $\mathcal{K}(X, Y)$ è un sottospazio vettoriale di $\mathcal{L}(X, Y)$;
2. la composizione di un operatore limitato e di uno compatto produce un operatore compatto; dunque, l'operazione di composizione eredita le proprietà dell'operatore migliore:
 - se $A \in \mathcal{K}(X, Y)$, $B \in \mathcal{L}(Z, X)$, $C \in \mathcal{L}(Y, Z)$, allora:

$$AB \in \mathcal{K}(Z, Y)$$

infatti, prima si applica B , poi A , dunque dallo spazio Z si va allo spazio X mediante B , e dallo spazio X si va allo spazio Y mediante A ;

- con le definizioni del punto precedente,

$$CA \in \mathcal{K}(X, Z)$$

infatti, analogamente a prima, si passa da X a Y , e poi da Y a Z .

quando $X = Y$, si ha $\mathcal{K}(X, X)$, che è un **ideale bilatero** di $\mathcal{L}(X, X)$.

3. $\mathcal{K}(X, Y)$ è chiuso, se Y è completo; questo vuol dire che se $A_n \rightarrow A$, dove $\{A_n\}$ è una successione di operatori compatti, $A \in \mathcal{L}(X, Y)$, allora A è anche compatto: il limite della successione di operatori compatti è compatto:

$$\|A_n - A\|_{\mathcal{L}(X, Y)} \rightarrow 0$$

Si discutono ora brevemente alcuni passi di base delle dimostrazioni dei tre punti.

1. Per quanto riguarda il primo punto, dati A, B operatori compatti, si ha che:

$$\alpha A + \beta B$$

è ancora un compatto, per α, β scalari. Al fine di verificare ciò, è necessario vedere che l'operatore dato da una di queste combinazioni lineari mandi una successione limitata in una sottosuccessione convergente. Dal momento che $\{Ax_n\}$ ha sottosuccessioni convergenti essendo A compatto, esiste una sottosuccessione $\{x_{n_k}\}$ convergente; di $\{Bx_{n_k}\}$, per compattezza di B , esiste una $Bx_{n_{k_j}}$, anche essa convergente, e tale per cui anche $Ax_{n_{k_j}}$, essendo una sottosuccessione di una sottosuccessione convergente per A . Dunque, la combinazione lineare sarà convergente.

2. Per il secondo punto, essendo A compatto e B limitato, si ha che AB è compatto dal momento che si parte da una successione Bx_n che sarà limitata, ma applicando l'operatore A a Bx_n sicuramente si avrà una

sottosuccessione limitata. Per quanto riguarda il caso opposto, ossia A compatto, C limitato, applicando A a x_n si ottengono delle successioni che ammettono delle sottosuccessioni convergenti; se si identifica una di queste, AB_k è convergente dal momento che la sottosuccessione convergente rimane convergente anche dopo l'applicazione di un operatore tale da essere solo continuo.

3. Per quanto riguarda il punto 3, è necessario fare una precisazione. La convergenza:

$$A_n x \rightarrow Ax \quad \forall x \in X$$

è detta, in ambito operatoriale, **convergenza forte**; in verità, questa non è altri che la convergenza debole*, nell'ambito degli spazi duali e dunque dei funzionali; questa sarebbe la convergenza puntuale. Questa convergenza è più debole della convergenza in norma.

Per fissare l'ultimo punto, si consideri un esempio pratico: si consideri uno spazio di Hilbert H , separabile, di dimensione infinita; su di esso, sia $\{e_k\}$ una base hilbertiana; dato $x \in H$, dunque, è possibile scriverlo come:

$$x = \sum_{k=1}^{\infty} (x|e_k)e_k = \lim_{n \rightarrow \infty} \sum_{k=1}^n (x|e_k)e_k$$

dove:

$$\sum_{k=1}^n (x|e_k)e_k = \mathcal{P}_n(x)$$

ossia, la sommatoria di x non è altro che una proiezione ortogonale sullo spazio $\text{span}\{e_1, \dots, e_n\}$. Definito I l'**operatore identità**, si ha dunque che:

$$x = Ix = \lim_{n \rightarrow \infty} \mathcal{P}_n(x)$$

si può dimostrare che questi $\mathcal{P}_n(x)$ tendano fortemente (ossia **puntualmente**) a x ; inoltre, l'operatore di proiezione \mathcal{P}_n è compatto $\forall n$, tuttavia **l'identità non è un operatore compatto**; di conseguenza:

$$\|\mathcal{P}_n - I\| \not\rightarrow 0$$

nella *norma operatoriale*; questo significa che questa cosa vale nella norma su \mathcal{L} , non su quella operatoriale. Questa cosa si può vedere nel seguente modo:

$$x = \mathcal{P}_n(x) + (x - \mathcal{P}_n(x))$$

dal momento che vale il teorema di decomposizione ortogonale, si ha che $\mathcal{P}_n(x)$ sta nello spazio di proiezione, ma ovviamente dunque

$$(x - \mathcal{P}_n(x)) = (I - \mathcal{P}_n)x$$

sta nella proiezione ortogonale; l'operatore $(I - \mathcal{P}_n)$ è detto **proiettore ortogonale**, e questo, si può dimostrare, ha norma pari a 1.

5.3.1 Operatori di rango finito - Compattezza degli operatori di rango finito

L'operatore di proiezione \mathcal{P}_n è un operatore compatto, dal momento che la sua immagine ha dimensione finita. Questo discorso è assolutamente generale, e per introdurlo si propone a questo punto una definizione.

Dati X, Y spazi normati, un operatore lineare $A : X \rightarrow Y$ viene detto **di rango finito** se la sua immagine ha dimensione finita. Il **rango** di un operatore è dunque la dimensione della sua immagine:

$$\text{rango } \{A\} = \dim \{A(X)\}$$

essendo A lineare, inoltre, l'immagine è uno spazio vettoriale.

A questo punto si propone un esempio pratico, al fine di fissare le idee. Dato $A : C([a, b])$, si consideri il seguente operatore:

$$Ax(t) = \int_a^b K(t, \tau)x(\tau) d\tau$$

dove il kernel integrale è:

$$K(t, \tau) = \sum_{j=1}^N \varphi_j(t)\psi_j(\tau)$$

ossia, il kernel integrale è noto come prodotto tensoriale²; φ_j, ψ_j sono funzioni continue su $[a, b]$; in questo caso, A è di rango finito. Si osservi per esempio:

²in letteratura il prodotto tensoriale si trova anche con la notazione $(\varphi_j \otimes \psi_j)(t, \tau)$

$$Ax(t) = \int_0^1 \sin(t + \tau)x(\tau) d\tau$$

è possibile scrivere il seno come:

$$\sin(t + \tau) = \sin t \cos \tau + \cos t \sin \tau$$

ma dunque, l'integrale si può scrivere come:

$$Ax(t) = \sin(t) \int_0^1 \cos(\tau)x(\tau)d\tau + \cos(t) \int_0^1 \sin(\tau)x(\tau)d\tau$$

fissata una funzione x continua, gli integrali sono due numeri, a_1 e a_2 ; dunque, l'immagine ha dimensione finita e pari a 2, dal momento che è generata dalle combinazioni lineari di $\sin(t)$ e $\cos(t)$ per mezzo delle costanti a_1 e a_2 .

Tornando dunque al caso generale:

$$Ax(t) = \sum_{j=1}^N \varphi_j(t) \int_a^b \psi_j(\tau)x(\tau) d\tau$$

in questo caso, l'immagine è **contenuta** in $\text{span}\{\varphi_1, \dots, \varphi_n\}$; è contenuta e non coincidente dal momento che è possibile che, se $\psi_j = 0$, alcuni coefficienti siano nulli, dunque l'immagine può essere contenuta ma non coincidente con lo span.

Gli operatori limitati e di rango finito sono automaticamente operatori compatti; questo è il motivo per cui queste definizioni sono state introdotte. Non solo: se A è un operatore del tipo:

$$A : L^2(a, b) \rightarrow L^2(a, b)$$

dove le funzioni $\varphi_j, \psi_j \in L^2(a, b)$, allora A è **compatto**.

Riassumendo: se $A : X \rightarrow Y$, con X, Y spazi metrici è un operatore limitato e di rango finito, allora esso è compatto, dal momento che l'immagine della palla unitaria è relativamente compatta.

Questo fatto sussiste dal momento che, siccome A è limitato, l'immagine della palla unitaria sarà limitata, ma d'altra parte per ipotesi essa sarà anche contenuta in uno spazio di dimensione finita, essendo l'operatore di rango finito. Tuttavia, in uno spazio di dimensione finita, tutte le norme sono tra loro equivalenti; scelta una norma, nel dettaglio quella euclidea, è possibile

dimostrare che ogni spazio di dimensione finita è isomorfo a \mathbb{R}^n (le distanze si preservano); in \mathbb{R}^n , tutti gli insiemi limitati sono anche relativamente compatti, da cui il risultato è dimostrato.

Esempio 1

Si consideri su l^2 l'operatore A definito come:

$$Ax = (x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots)$$

l'operatore A così definito **non** è di rango finito, tuttavia è compatto, dal momento che esso è il limite di operatori di rango finito; se infatti si considerano operatori $A_N \in \mathcal{L}(l^2, l^2)$ definiti come:

$$A_N x = (x_1, \frac{1}{2}x_2, \dots, \frac{1}{N}x_N, 0, 0, 0, \dots)$$

dove dunque si tronca l'operatore alla N -esima componente, dunque l'immagine avrà dimensione al più pari a N . L'operatore A_N , per $N \rightarrow \infty$, tende a A ; per verificare ciò, è necessario stimare:

$$\|A - A_N\|_{\mathcal{L}(l^2, l^2)} \leq \frac{1}{N+1} \rightarrow 0$$

questo si può verificare formalmente vedendo che scrivendo questo termine, i primi termini si annullano, rimane dunque solo la coda, e quindi si ottiene ciò.

Esempio 2

Un secondo esempio che si può studiare è quello degli operatori integrali con nucleo in L^2 : dunque, su $L^2(a, b)$, si consideri

$$A : L^2(a, b) \rightarrow L^2(a, b)$$

definito come:

$$Ax(t) = \int_a^b K(t, \tau)x(\tau) d\tau$$

dove

$$K \in L^2([a, b] \times [a, b])$$

si è già dimostrato che un operatore di questo tipo è limitato; ora, se ne discuterà la compattezza.

Per dimostrare ciò, si deve verificare che questo operatore è il limite (in senso operatoriale, come nell'esempio precedente) in $\mathcal{L}(L^2, L^2)$ di operatori di rango finito.

Al fine di fare ciò, è necessario introdurre alcune nozioni preliminari: sia $\{e_n\}$ una base hilbertiana di $L^2(a, b)$; allora, si verifica che:

$$\{e_n(t)e_n(\tau)\}$$

(ossia, il prodotto tensoriale delle basi hilbertiane in questione) è una base di Hilbert per $L^2([a, b] \times [a, b])$; questo si può ottenere a partire da nozioni di Teoria della Misura, con le quali è necessario verificare la completezza. Se la suddetta base è di Hilbert, è possibile decomporre il nucleo integrale K come segue:

$$K(t, \tau) = \sum_{n,m=1}^{\infty} c_{n,m} e_n(t) e_m(\tau)$$

dove la convergenza e l'eguaglianza sono valide nello spazio $L^2([a, b] \times [a, b])$; l'eguaglianza è valida *quasi ovunque*.

A partire dalle considerazioni appena fatte, dunque, si consideri il kernel integrale dell'operatore, $K(t, \tau)$, scritto come:

$$K(t, \tau) = \lim_{N \rightarrow \infty} \sum_{n,m=1}^N c_{n,m} e_n(t) e_m(\tau)$$

sia dunque A_N l'operatore integrale con nucleo integrale K_N , dove

$$K_N = \sum_{n,m=1}^N c_{n,m} e_n(t) e_m(\tau)$$

A_N è chiaramente un operatore di rango finito; esso è poi compatto, dal momento che:

$$\|A_N - A\| \leq \|K - K_N\|_{L^2}$$

dove il termine destro tende a zero, per $N \rightarrow \infty$; questo dimostra l'affermazione.

5.4 Operatori aggiunti, operatori trasposti

Nell'ambito dell'Algebra Lineare, studiando dunque le matrici, si parla di matrici aggiunte, di matrici trasposte; a partire da essa si introduce la teoria spettrale, la quale studia operatori, autovalori e diagonalizzazione delle matrici; al fine di introdurre questa teoria anche per quanto riguarda gli operatori integrali, dunque, è necessario estendere queste definizioni di base anche per quanto concerne gli operatori.

Si consideri una matrice $\underline{\underline{A}}$ di dimensioni $m \times n$, a valori complessi:

$$(\underline{\underline{A}})_{jk} = (a_{jk}), j = 1 \dots n, k = 1 \dots m$$

La matrice trasposta si definisce come:

$$\underline{\underline{A}}^t : (\underline{\underline{A}}^t)_{jk} = (\underline{\underline{A}})_{kj}$$

ossia, si scambiano righe e colonne; per la matrice aggiunta, oltre a ciò, si considera:

$$\underline{\underline{A}}^* : (\underline{\underline{A}}^*)_{jk} = (\bar{a}_{kj})$$

ossia, si prendono le componenti, le si traspongono (scambiano righe e colonne), e si calcola il complesso coniugato di ciascuna componente.

5.4.1 Operatori aggiunti

Per quanto riguarda le matrici aggiunte, valgono le seguenti proprietà:

$$(\underline{\underline{A}} \underline{x} | \underline{y})_{\mathbb{C}^n} = (\underline{x} | \underline{\underline{A}}^* \underline{y})_{\mathbb{C}^n}$$

inoltre:

$$(\underline{\underline{A}} \underline{x} | \bar{\underline{y}}) = (\underline{x} | \overline{\underline{\underline{A}}^t \underline{y}})$$

dove:

$$(\underline{x}|\underline{y})_{\mathbb{C}^n} = \sum_{j=1}^n x_j \bar{y}_j$$

Si vuole estendere questa teoria per quanto riguarda gli operatori.

Prima di tutto, è necessario introdurre i prodotti scalari; per questo motivo, fin quando si parla di operatori aggiunti, è necessario lavorare su spazi di Hilbert. Dunque, siano H_1, H_2 due spazi di Hilbert; sia dunque:

$$A \in \mathcal{L}(H_1, H_2)$$

si definisce A^* l'operatore aggiunto da H_2 in H_1 :

$$A^* \in \mathcal{L}(H_2, H_1)$$

tale che:

$$(Ax|y) = (x|A^*y)$$

questo, $\forall x \in H_1, \forall y \in H_2$.

L'operatore aggiunto si definisce utilizzando il teorema di Riesz-Frechet: $\forall y \in H_2$, fissato, si consideri il funzionale che a x associa $(Ax|y)$, definito da H_1 a \mathbb{C} . Questo funzionale è lineare e continuo; la continuità si ha dal momento che è possibile maggiorare come segue:

$$|(Ax|y)| \leq \|Ax\| \|y\|$$

questo, per Cauchy-Schwartz; essendo tuttavia A limitato:

$$\leq \|A\| \|y\| \|x\|$$

dove y è fissato; di conseguenza, $\|A\| \|y\|$ è una costante. Quindi, applicando il teorema di Riesz-Frechet, è possibile rappresentare l'operatore in questione mediante un prodotto scalare; questo significa che:

$$\exists! z \in H_1 : (Ax|y) = (x|z)$$

si definisce dunque questo z come l'immagine di y a cui è stato applicato l'operatore aggiunto A^* :

$$A^*y \triangleq z$$

Si deve a questo punto verificare che questo operatore sia effettivamente limitato; per fare ciò, si ha che:

$$\|A^*\| = \sup \{(A^*y|x)\} : \|x\| \leq 1, \|y\| \leq 1$$

è finito; questo fatto vale solamente in spazi di Hilbert. Dal momento che il prodotto scalare è sotto valore assoluto, si ha che:

$$|A^*y|x| = |(x|A^*y)|$$

ma dunque:

$$|(x|A^*y)| = |(Ax|y)|$$

e questa è limitata. Con questo, la dimostrazione è conclusa.

Esempio

Si consideri a questo punto un esempio pratico in cui si calcola l'operatore aggiunto. Dato $A : L^2(a, b) \rightarrow L^2(a, b)$, definito come:

$$Ax(t) = \int_a^b K(t, \tau)x(\tau) d\tau$$

dove $K \in L^2([a, b] \times [a, b])$, qual è il suo operatore aggiunto?

Si ha:

$$(Ax|y) = (x|A^*y)$$

per fare ciò, si cerchi di isolare la funzione x nel prodotto scalare:

$$Ax = \int_a^b K(t, \tau)x(\tau) d\tau$$

il prodotto scalare sarà:

$$\int_a^b Ax\bar{y}(t) dt = \int_a^b \int_a^b K(t, \tau)x(\tau) d\tau\bar{y}(t) dt =$$

Vale la seguente eguaglianza: è possibile portare x fuori, e scambiare l'ordine di integrazione

$$= \int_a^b x(\tau) \int_a^b K(t, \tau) dt \bar{y}(t) d\tau$$

a questo punto, si può dire che:

$$\int_a^b x(\tau) \int_a^b K(t, \tau) dt \bar{y}(t) d\tau = \int_a^b x(\tau) \int_a^b \overline{K(t, \tau)} dt y(t) d\tau$$

ma dunque:

$$A^*y(\tau) = \int_a^b \overline{K(t, \tau)} dt$$

si osservi che y dipende da τ ; dunque, scambiando le variabili t e τ , per uniformare, si ha:

$$A^*y(t) = \int_a^b \overline{K(\tau, t)} d\tau$$

quindi, in questo modo, l'operatore aggiunto si ottiene scambiando le variabili del nucleo e facendone il coniugato, esattamente come nell'ambito delle matrici si scambiano gli indici delle componenti e quindi se ne calcola il coniugato.

Per concludere, vale il seguente importante **teorema**: dato un operatore A e il suo aggiunto A^* , A è compatto se e solo se A^* è compatto.

5.4.2 Operatori trasposti

Se l'aggiunto di un operatore si può definire solo nell'ambito degli spazi di Hilbert, il trasposto di un operatore si può anche definire nell'ambito degli spazi normati, anche non completi; la definizione passa attraverso la definizione di dualità.

Si considerino X, Y due spazi normati; dato A operatore limitato, dunque $A \in \mathcal{L}(X, Y)$, si definisce l'operatore trasposto A^t , che si dimostrerà essere limitato ($A^t \in \mathcal{L}(X^*, Y^*)$), come segue. Se $\varphi \in Y^*$, $x \in X$,

$$(A^t\varphi)(x) = \varphi(Ax)$$

quindi:

$$A : X \rightarrow Y$$

e

$$A^t : Y^* \rightarrow X^*$$

dunque, gli operatori trasposti sono definiti sugli **spazi duali**, e non sugli spazi di partenza, come si faceva nell'ambito degli operatori aggiunti. Quella appena scritta è la definizione di operatore trasposto; si può vedere che questo è lineare e limitato, dal momento che:

$$\|A^t\| = \|A\|$$

L'operatore trasposto in letteratura viene anche chiamato **operatore duale** o **operatore aggiunto**; quest'ultimo nome crea molta confusione, a meno che lo spazio in questione sia non-Hilbert, caso in cui non ha senso parlare di aggiunti.

Si può vedere facilmente che la definizione di trasposto ricorda la definizione di aggiunto; infatti, se $x \in X$, $f \in X^*$, si pone:

$$\langle f, x \rangle = f(x)$$

si può vedere che questa notazione ricorda molto il prodotto scalare; utilizzando questo tipo di notazione, è possibile scrivere la definizione di aggiunto e motivare la frase appena scritta:

$$\langle A^t \varphi, x \rangle = \langle \varphi, Ax \rangle$$

e questa definizione, se invece di questo tipo di notazione si utilizza quella con le parentesi tonde, tipiche del prodotto scalare, è proprio la definizione di aggiunto.

Si consideri a questo tipo $\varphi \in L^q$, e $x \in L^p$; p, q esponenti coniugati, $p \neq \infty$; si ha in questo caso che L^q è il duale di L^p : $(L^p)^* = L^q$; inoltre, si ha che:

$$\langle \varphi, x \rangle = \int_a^b \varphi(t)x(t) dt$$

A questo punto si giustifica questa espressione: ogni elemento di L^q definisce un elemento del duale di L^p ; di conseguenza però, è possibile rappresentare il duale di L^p mediante il teorema di rappresentazione di Riesz, introducendo un'applicazione che a y associa f_y , come:

$$f_y(x) = \int_a^b y(t)x(t) dt$$

ogni funzionale in $(L^p)^* = L^q$ si può rappresentare in questo modo; la differenza fondamentale rispetto al prodotto scalare, come si può notare, è l'assenza del simbolo di coniugazione. A questo punto, si verifica che il trasposto di un operatore è limitato, e che ha la stessa norma dell'operatore di partenza; per fare ciò, si ricorda che:

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

si è però visto che la norma in Y di un vettore si può esprimere in termini dei funzionali lineari definiti su Y :

$$\|Ax\|_Y = \sup_{\|\varphi\|=1} |\varphi(Ax)|$$

infatti, si può sfruttare il teorema di Hahn-Banach e le sue conseguenze, ottenendo che, se $x \in X$:

$$\|x\| = \sup \{|f(x)| : f \in X^* \|f\|_{X^*} = 1\}$$

inoltre, questo sup è pure un max, dal momento che esiste una funzione di norma unitaria tale per cui si assume quel valore; si hanno dunque due sup, che possono essere legittimamente scambiati di ordine:

$$\|Ax\|_Y = \sup_{\|\varphi\|=1} \sup_{\|x\|=1} |\varphi(Ax)|$$

tuttavia, è possibile vedere che:

$$\varphi(Ax) = (A^t\varphi)(x)$$

ossia, non è altri che l'operatore $(A^t\varphi)$ valutato in un punto x . Tuttavia:

$$\sup_{\|x\|=1} |(A^t\varphi)(x)| = \|A^t\varphi\|$$

ma, a sua volta:

$$\sup_{\|\varphi\|=1} \|A^t\varphi\| = \|A^t\|$$

e la dimostrazione è così ultimata.

Esempio di calcolo dell'operatore trasposto

Si propone a questo punto un esempio di calcolo del trasposto di un certo operatore; si consideri per esempio l'operatore $A : L^p(a, b) \rightarrow L^p(a, b)$, dove $1 \leq p < \infty$; si ha che:

$$Ax(t) = \int_a^b K(t, \tau)x(\tau) d\tau$$

$K \in C([a, b] \times [a, b])$ (per esempio, anche se si può dimostrare che la continuità non è una condizione vincolante, in quanto si può alleggerire). L'operatore trasposto deve essere un $A^t : L^q(a, b) \rightarrow L^q(a, b)$, con q esponente coniugato di p (al solito, si lavora sugli spazi duali). Usando la definizione:

$$A^t(x) = \int_a^b K(\tau, t)x(\tau) d\tau$$

ossia, simile all'aggiunto, senza avere la coniugazione; questo rappresenta l'estensione della matrice trasposta.

5.4.3 Teorema di Schauder

Si introduce a questo punto un altro teorema, analogo a un'osservazione precedentemente fatta nell'ambito degli operatori aggiunti. Sia $A \in \mathcal{L}(X, Y)$, compatto; allora, $A^t \in \mathcal{L}(Y^*, X^*)$ è compatto. Vale anche il viceversa, a patto che Y sia uno spazio completo.

Un'osservazione: nell'ambito degli spazi di Hilbert, è possibile calcolare sia il trasposto, sia l'aggiunto; esiste un legame, in questo ambito, tra i due. Siano H_1, H_2 due spazi di Hilbert; sia $A \in \mathcal{L}(H_1, H_2)$ un operatore; sia A^* il suo aggiunto, ossia $A^* : H_2 \rightarrow H_1$, e sia $A^t : H_2^* \rightarrow H_1^*$ il trasposto dell'operatore; esiste una qualche relazione tra i due? La risposta è sì: essendo gli spazi in questione spazi di Hilbert, H_2 è isomorfo a H_2^* , e così anche H_1 è isomorfo a H_1^* ; esiste un'applicazione J_2 canonica che associa a elementi di H_2 elementi di H_2^* ; questa è l'applicazione definita come:

$$J_2 y(x) = (x|y)_{H_2}$$

questo è semplicemente il prodotto scalare su H_2 ; tutto ciò deriva dal teorema di rappresentazione di Riesz-Frechet per quanto riguarda spazi di Hilbert: ogni funzionale appartenente al duale di uno spazio di Hilbert si può scrivere mediante un opportuno prodotto scalare.

Il legame tra operatori trasposti e aggiunti su spazi di Hilbert risiede nel diagramma in Figura ?? : esso è un **diagramma commutativo**. Questo significa, in altre parole, che:

$$J_1 A^* = A^t J_2$$

ossia, l'applicazione di A^* a un certo elemento di H_2 , e poi di J_1 a ciò, coincide con l'applicazione prima di J_2 a un elemento di H_2 , poi di A^t ; il risultato è lo stesso. In altre parole:

$$A^* = J_1^{-1} A^t J_2$$

o, ancora:

$$A^t = J_1 A^* J_2^{-1}$$

Queste eguaglianze si possono verificare come segue (se ne fa un esempio):

$$J_1 A^* = A^t J_2$$

si calcolano esplicitamente i due membri:

$$J_1 A^* : H_2 \rightarrow H_1^*$$

dove:

$$J_1 A^* y = J_1(A^* y) = \langle J_1(A^* y), x \rangle = (x | A^* y)$$

e $x \in H_1$; per quanto riguarda il secondo membro:

$$A^t J_2 : H_2 \rightarrow H_1^*$$

si può scrivere come:

$$A^t(J_2 y) = \langle A^t(J_2 y), x \rangle = \langle J_2 y, Ax \rangle$$

essendo però $J_2 = (J_2 y)(x) = (x | y)$, si ha:

$$= (Ax | y)$$

che, per definizione, è uguale all'aggiunto:

$$= (x | A^* y)$$

Questo significa che sarebbe possibile anche partire dall'operatore trasposto, e mediante la sua definizione arrivare anche a definire l'operatore aggiunto; tuttavia, in verità, quando si parla di operatori trasposti, per ottenere questi risultati è necessario utilizzare il teorema di Hahn-Banach (dunque l'assioma della scelta); di conseguenza, parlare separatamente di questi risultati permette di ottimizzare le ipotesi utilizzate.

Esercizio e alcune relazioni

Si può dimostrare che, se $A \in \mathcal{L}(X, Y)$, allora:

$$\ker \{A^t\} = (\text{range } \{A\})^\perp$$

dove per “range” si intende l'immagine dell'operatore; il simbolo \perp si riferisce all'annullatore dello spazio. Vale anche la seguente:

$$\ker \{A\} = (\text{range } \{A^t\})^\perp$$

Analogamente, se A è lineare, continuo tra due spazi di Hilbert, vale l'analogo anche per quanto concerne l'aggiunto (dove questa volta \perp indica l'ortogonale e non l'annullatore):

$$\ker \{A^*\} = (\text{range } \{A\})^\perp$$

$$\ker \{A\} = (\text{range } \{A^*\})^\perp$$

Si mostra un cenno di dimostrazione per una proprietà: essendo eguaglianze tra spazi, bisogna verificare l'inclusione nei due sensi; la dimostrazione consiste solo nello scrivere definizioni. Se si considera $\ker \{A^t\}$, significa che si sta cercando l'insieme dei φ (si lavora nello spazio duale) tali per cui:

$$A^t\varphi = 0$$

o, meglio:

$$A^t\varphi(x) = 0 \forall x$$

in altre parole ancora, questo è:

$$\varphi(Ax) = 0 \forall x$$

questo significa che φ sta nell'annullatore dell'immagine: infatti, $\forall x, \varphi(Ax)$ descrive l'immagine di A ; se questa è nulla, allora la dimostrazioncina è terminata, in un verso. Per l'altro verso, si ha semplicemente da ripetere i passaggi al contrario, verificando che φ sta nel ker del trasposto.

Esiste una relazione aggiuntiva, che vale solamente per spazi a dimensione finita:

$$\text{range } \{A\} = (\ker \{A^t\})_{\perp}$$

questo vale solo in spazi a dimensione finita poiché, nei casi più generali che si stanno studiando, in verità si ha:

$$\overline{\text{range } \{A\}} = (\ker \{A^t\})_{\perp}$$

ossia, questa informazione vale solo per la chiusura, quindi ciò non è utile; la relazione in questione vale solo nell'ambito degli operatori compatti, e questo è utile per discutere l'esistenza di soluzioni x a problemi del tipo:

$$Ax = y$$

5.5 Operatori invertibili

Un operatore è invertibile se è biiettivo, ossia iniettivo e suriettivo. Dati X, Y due spazi di Banach (normati sarebbe sufficiente, però si vuole introdurre la completezza da subito, dal momento che è utile per un risultato che verrà introdotto tra poco). Dato $A \in \mathcal{L}(X, Y)$, allora A è invertibile se è una biiezione, ossia:

$$\ker \{A\} = \{\emptyset\}$$

$$\text{range } \{A\} = Y$$

Se ciò è verificato, dunque, esiste un certo operatore B tale per cui:

$$AB = BA = I$$

in X . $B = A^{-1}$.

Valgono le seguenti due osservazioni:

- l'operatore B è lineare; ciò si dimostra in maniera analoga a quanto si può fare nella semplice Algebra Lineare in dimensione finita;
- B è un operatore continuo (ossia limitato).

Questo risultato segue dal teorema della mappa aperta.

5.5.1 Teorema della mappa aperta (teorema di Banach)

Dati X e Y due spazi di Banach, dato $A \in \mathcal{L}(X, Y)$, biiettivo, allora l'operatore inverso $B = A^{-1}$ è limitato.

Il motivo per cui questo operatore viene detto *della mappa aperta* è perché, se la mappa è una biiezione, dire che essa manda aperti in aperti, garantisce che l'inversa è continua; questo è più che la continuità: continuità significa che la controimmagine è un aperto, mentre ora si sta garantendo qualcosa in più.

Un'osservazione: se $Y = X$, A è iniettivo, allora, in spazi di dimensione finita, A è automaticamente pure suriettivo; infatti, il fatto di avere nucleo di dimensione pari a zero garantisce che l'immagine sia coincidente con l'intero spazio di arrivo; in dimensione infinita, questa cosa non è invece vera, dal momento che la sola iniettività non biimplica anche la suriettività. Infatti, se X ha dimensione infinita, $A \in \mathcal{L}(X, X)$ può essere iniettivo senza essere suriettivo; in altre parole, è possibile che:

$$\ker \{A\} = \emptyset \quad \text{range} \{A\} \neq X$$

inoltre, a volte l'immagine è densa nell'insieme X , ma a volte non è neanche garantito ciò.

Verranno a questo punto studiati alcuni esempi.

Esempio 1

Si consideri l'operatore $A : l^2 \rightarrow l^2$, definito come:

$$Ax = \left(x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right)$$

questo operatore, come è stato dimostrato precedentemente, certamente è compatto; si può tuttavia dire che esso è anche iniettivo? La risposta è sì:

l'unica possibilità di ottenere il vettore nullo è applicare a una successione x identicamente nulla l'operatore A in questione. Si può anche dire che esso è suriettivo? La risposta è no; infatti:

$$\text{range } \{A\} = \{y = Ax \mid x \in l^2\} = \left\{ y \in l^2 : y_k = \frac{1}{k} x_k(x_n) \in l^2 \right\}$$

l'immagine non genera tutto l^2 ; infatti, si ha che, se $y_k = \frac{1}{k} x_k$, allora:

$$x_k = k y_k$$

ma dunque:

$$\left\{ y \in l^2 : \sum_{k=0}^{\infty} |k y_k| < \infty \right\}$$

questa espressione non può essere usata per rappresentare l'intero spazio l^2 ; considerando per esempio:

$$y = \left(1, \frac{1}{2}, \frac{1}{3}, \dots \right)$$

allora, questo **non sta nell'immagine**; infatti, in questo caso:

$$\text{range } \{A\} = \{y \in l^2 : (y_1, 2y_2, 3y_3, \dots) \in l^2\}$$

sostituendo y appena definita, si ha:

$$Ax = (1, 1, 1, 1, \dots)$$

e questo non appartiene all'immagine dell'operatore. D'altra parte, questa è almeno densa: infatti, l'immagine contiene c_{00} , dal momento che:

$$c_{00} = \text{span } \{e_n\}$$

dal momento che l'insieme delle combinazioni lineari finite degli e_n possono essere rappresentate mediante elementi dell'immagine. Di conseguenza, dal momento che c_{00} è un sottoinsieme di l^2 , essendo c_{00} contenuto nell'immagine, essendo c_{00} denso in l^2 , allora l'immagine è densa in l^2 .

Esempio 2

Si considera a questo punto un ulteriore esempio: l'operatore di shift a destra, definito come:

$$A : l^2 \rightarrow l^2$$

definito come:

$$Ax = (0, x_1, x_2, \dots)$$

questo è iniettivo: $Ax = 0 \implies x = 0$, dal momento che di nuovo tutte le componenti devono essere nulle, per ottenere la successione identicamente nulla. Per quanto riguarda la suriettività,

$$\text{range } \{A\} \neq l^2$$

dal momento che il primo elemento è sempre nullo: questo insieme non è altri che l'ortogonale di e_1 , dal momento che:

$$\text{range } \{A\} = \{e_1\}^\perp$$

In questo caso, non è manco denso: il criterio di densità infatti dice che un insieme, per essere denso, deve avere l'ortogonale nullo, e questo non accade. In altre parole, non è possibile approssimare arbitrariamente bene ogni insieme, in questo caso, dal momento che gli elementi con la "prima componente" non sono approssimabili.

Esempio 3

Si consideri a questo punto l'operatore di shift a sinistra:

$$A : l^2 \rightarrow l^2$$

dove:

$$Ax = (x_2, x_3, x_4, \dots)$$

Questo non è iniettivo:

$$\ker \{A\} = \text{span } \{e_1\}$$

infatti, se $x = (\alpha e_1, 0, 0, \dots)$, l'operatore Ax è 0.

Questa è una situazione opposta rispetto a prima: l'operatore è suriettivo, ma non iniettivo.

Un'osservazione conclusiva: se l'operatore è continuo, il suo nucleo è chiuso; tuttavia, per quanto riguarda l'immagine, non si hanno garanzie di questo genere.

5.6 Serie di Neumann

Si introdurrà a questo punto un criterio di invertibilità, basato sulla serie di Neumann.

Sia X uno spazio di Banach, sia A un operatore lineare continuo, ossia $A \in \mathcal{L}(X, X)$; sia $I - A$ l'operatore identità I meno A , e sia la norma di A :

$$\|A\| = q < 1$$

allora, in queste condizioni, l'operatore $I - A$ è invertibile, e la norma dell'operatore inverso è stimabile come:

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

questo significa che se A mappa dei vettori in vettori con una norma più piccola, A non è in grado di perturbare troppo l'identità, e dunque l'operatore risultante sarà ancora invertibile.

Si procede ora con la dimostrazione di questo teorema; per fare ciò, si ricorda che:

$$\frac{1}{1 - q} = \sum_{n=0}^{\infty} q^n, \quad |q| < 1$$

Se al posto di q si mette l'operatore A , si ottiene una **serie di operatori**; si ha dunque:

$$B = \sum_{n=0}^{\infty} A^n$$

questa, deve convergere in $\mathcal{L}(X, X)$; data dunque B così definita, si vuole vedere che:

$$B(I - A) = (I - A)B = I$$

Dal momento che lo spazio X per ipotesi è uno spazio di Banach, dunque completo, è sufficiente verificare che la successione delle somme parziali sia di Cauchy, per verificare che è anche convergente. Si ha:

$$S_n = \sum_{k=0}^n A^k$$

se $m > n$, si può vedere che:

$$\|S_m - S_n\| = \left\| \sum_{k=n+1}^m A^k \right\| \leq \sum_{k=n+1}^m \|A^k\|$$

questo ultimo passaggio è una banale applicazione della disuguaglianza triangolare, la quale è valida, dal momento che lo spazio X è normato. Esiste un'altra relazione utile tra le norme: si può infatti verificare che la norma del prodotto degli argomenti è minore o uguale del prodotto delle norme:

$$\|A^k\| \leq \|A\|^k = q^k$$

ma dunque, la sommatoria si può controllare come:

$$\|S_m - S_n\| \leq \sum_{k=n+1}^m q^k$$

questo oggetto si mantiene minore di un certo ε , a patto di prendere n e m sufficientemente grandi; infatti, questo oggetto non è altri che la differenza tra le ridotte n -esime di una somma geometrica, e questo è minore di ε se $n, m > N_\varepsilon$, dove N_ε è una soglia sufficientemente grande.

Una volta verificata la convergenza della serie, si deve ancora verificare la relazione con l'inversa; data dunque B la somma della serie, si ha che:

$$\sum_{k=0}^n A^k (I - A) = I - A^{n+1}$$

questa è un'identità notevole delle somme geometriche; per $n \rightarrow \infty$, tuttavia, si ha che:

$$\sum_{k=0}^n A^k \rightarrow B$$

ma:

$$A^{n+1} \rightarrow 0$$

infatti:

$$\|A^{n+1}\| \leq \|A\|^{n+1} = q^{n+1}$$

e $q^{n+1} \rightarrow 0$, poiché $q \leq 1$.

Di conseguenza, per $n \rightarrow \infty$, si ha:

$$B(I - A) \rightarrow I$$

dunque, si può concludere che B è un **inverso sinistro** (e pure **inverso destro**, come si può verificare in maniera analoga).

Esempio

Si consideri l'equazione integrale:

$$x(t) - \lambda \int_0^a e^{\sin(t-\tau)} x(\tau) d\tau = y(t)$$

noto $y(t)$, si vuole verificare che, per λ sufficientemente piccolo, $\forall y \in C([0, 1])$, esiste unica $x \in C([0, 1])$, soluzione per l'equazione integrale (non si richiede di calcolarla).

Per risolvere questo problema, è sufficiente osservare che, dato A l'operatore:

$$A = \lambda \int_0^1 e^{\sin(t-\tau)} x(\tau) d\tau$$

si può vedere che l'equazione differenziale è nella forma:

$$(I - A)x = y$$

ma dunque, per trovare x , è sufficiente avere la possibilità di invertire l'operatore a sinistra; questo è possibile, in virtù del criterio appena introdotto, dal momento che, se si sceglie un λ opportuno, l'operatore A avrà norma minore di 1, e per questo motivo l'operatore è invertibile, a patto ovviamente che A sia continuo (come è, in questo caso). Ovviamente, l'esercizio finisce qua, dal momento che non è necessario calcolare esplicitamente la soluzione dell'equazione integrale.

la sostanziale conseguenza del criterio basato sulla serie di Neumann è il fatto che, se si perturba di poco un generico operatore invertibile, si avrà ancora a che fare con un operatore invertibile. Dato per esempio X uno spazio di Banach, $A \in \mathcal{L}(X, X)$ invertibile, $B \in \mathcal{L}(X, X)$ tale per cui:

$$\|B\| \leq \frac{1}{\|A^{-1}\|}$$

allora l'operatore $A + B$ è ancora invertibile. La verifica è immediata:

$$A + B = A(I + A^{-1}B)$$

questo operatore è nella forma richiesta dal criterio di Neumann (il segno è indifferente), dunque:

$$\|A^{-1}B\| \leq \|A^{-1}\| \|B\|$$

essendo però per ipotesi $\|B\| \leq \|A^{-1}\|$, la norma dell'operatore è minore di 1 per ipotesi.

Questo risultato può essere sintetizzato nella seguente affermazione: in $\mathcal{L}(X, X)$, l'insieme degli operatori invertibili è un aperto.

5.7 Lemma di Lax-Milgram

Si conclude a questo punto l'argomento dello studio degli operatori, proponendo un teorema molto importante per lo studio delle equazioni alle derivate parziali. Sia H uno spazio di Hilbert reale (non è necessario, ma semplifica le cose), e si consideri un'applicazione da H a H^* $y \rightarrow f_y$ definita come:

$$f_y(x) = \alpha(x, y)$$

si osservi che questa premessa ricorda molto la definizione di un prodotto scalare: se α ha le stesse proprietà di un prodotto scalare, sembrerebbe che si stiano ri-enunciando le ipotesi del teorema di rappresentazione di Riesz per il duale di spazi di Hilbert; invece, in questo caso si introdurranno meno proprietà a questa α rispetto a quelle che normalmente si introducono per un prodotto scalare, al fine di generalizzare il teorema di rappresentazione di Riesz. Quello che si vuole arrivare a dire è che, se α ha alcune (non tutte) proprietà del prodotto scalare, questa applicazione è biiettiva, permettendo una rappresentazione dello spazio duale di H .

Sia dunque H uno spazio di Hilbert reale, e sia α :

$$\alpha : H \times H \rightarrow \mathbb{R}$$

soddisfacente le seguenti tre proprietà:

1. bilineare, ossia:

$$\alpha(\lambda_1 x_1 + \lambda_2 x_2 + \dots, y) = \lambda_1 \alpha(x_1, y) + \lambda_2 \alpha(x_2, y) + \dots$$

e la stessa cosa per y , ossia **lineare** anche nel secondo fattore;

2. continua, dove per continua si intende che vale la seguente stima:

$$\exists c : |\alpha(x, y)| \leq c \|x\| \|y\|$$

3. coerciva: questo significa che esiste una certa $c > 0$ (questa ovviamente non deve coincidere con quella della precedente ipotesi) tale per cui:

$$\alpha(x, x) \leq c \|x\|^2$$

Si osservi che, nell'ambito reale, se:

$$\alpha(x, y) = (x|y)$$

allora tutto ciò è verificato: infatti, la seconda richiesta non è altri che Cauchy-Schwarz, mentre l'altra è l'eguaglianza, con $c = 1$. Ciò che non si ha in generale per α è la simmetria: non è possibile scambiare x e y ; se si aggiunge la simmetria, α è un prodotto scalare a tutti gli effetti.

Il lemma di Lax-Milgram afferma che, se α soddisfa le precedenti ipotesi, allora l'operatore $A : H \rightarrow H^*$ definito come:

$$(Ax)(y) = \alpha(x, y)$$

è invertibile (ossia biiettivo).

Questa conclusione può essere scritta in altri modi, assolutamente equivalenti a quello appena mostrato, che verranno ora elencati.

- Se è invertibile, allora è iniettivo e suriettivo:

$$\forall \varphi \in H^* \exists ! x \in H : (Ax) = \varphi$$

- Equivalentemente:

$$\forall \varphi \in H^* \exists ! x \in H : Ax(y) = \varphi(y), \forall y \in H$$

- Ancora: si può dire che $(Ax)(y) = \alpha(x, y)$:

$$\forall \varphi \in H^* \exists ! x \in H : \alpha(x, y) = \varphi(y) \forall y \in H$$

- Si sta usando continuamente φ , che è un generico funzionale dello spazio duale; per il teorema di rappresentazione di Riesz, tuttavia, si ha che ogni funzionale φ appartenente allo spazio duale si può, per un certo $z \in H$, scrivere mediante un prodotto scalare:

$$\varphi(y) = (y|z)$$

Usando queste ultime osservazioni, è possibile *dimenticarsi* il discorso delle applicazioni verso il duale, e quindi applicare direttamente il risultato del teorema di rappresentazione di Riesz e scrivere che:

$$\forall z \in H \exists ! x \in H : \alpha(x, y) = (y|z) \forall y \in H$$

questo è un modo alternativo, più *esplicito* di scrivere la conclusione, di scrivere la conclusione del teorema di Lax-Milgram.

Dimostrazione

Si deve dimostrare che A è un operatore invertibile; questo significa iniettivo e suriettivo. In effetti, si ha che A è lineare; fissato un $x \in H$, si ha che Ax è un elemento di H^* , dunque Ax è un'applicazione lineare continua su H ; infatti:

$$(Ax)(y) \leq |(Ax)(y)| \leq c \|x\| \|y\|$$

questo ultimo passaggio è possibile grazie all'ipotesi di continuità del teorema: fissato x , si ha che $c \|x\|$ è costante. Questo ci dice che A manda da H a H^* ; inoltre, bisogna dimostrare che A stesso è lineare e limitato, come applicazione da H in H^* ; questo si dimostra semplicemente vedendo che:

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

ma:

$$\|Ax\| \leq c \|x\|$$

dunque:

$$\|Ax\| \leq c$$

(essendo $\|x\| = 1$ l'insieme dei vettori in cui si cerca il sup).

Infine, bisogna dimostrare che A è biiettivo; questo segue dal fatto che:

- A è iniettivo e ha immagine chiusa;
- A ha immagine densa

Dimostriamo questi punti: dato un operatore A tale per cui:

$$\exists c > 0 : \|x\| \leq c \|Ax\|$$

allora, questo implica che è iniettivo con immagine chiusa dal momento che, se il Ax è nullo, allora $x = 0$; la chiusura è una conseguenza del fatto che, se si considera una successione $y_n \in \text{range}\{A\}$, dove $y_n \rightarrow y$, allora si dimostra che:

$$y \in \text{range}\{A\}$$

dal momento che y_n è convergente, essa è pure di Cauchy; dunque, si può definire:

$$y_n = Ax_n$$

dunque, siccome per ipotesi vale:

$$\|x\| \leq c \|Ax_n\|$$

si ha che:

$$\|x_n - x_m\| \leq c \|Ax_n - Ax_m\| \leq c \|y_n - y_m\|$$

quindi, essendo $\{y_n\}$ di Cauchy per ipotesi (essendo convergente a un certo y , è anche Cauchy), si ha che questa maggiora l'altra successione, quindi $x_n \rightarrow x$, e quindi:

$$Ax_n \rightarrow Ax$$

e, per unicità del limite, $y = Ax$.

Per quanto riguarda il secondo punto, ossia il fatto che l'immagine è densa, è possibile applicare il criterio di densità per uno spazio di Hilbert, al fine di verificare che:

$$\text{range } \{A\} = \{Ax : x \in H\} \text{ è densa in } H$$

Se un elemento che sta nel duale dello spazio H , dunque H^* , si annulla su tutti questi punti, allora deve essere nullo; tuttavia, si ha che H è uno spazio di Hilbert, e dunque riflessivo; questo permette di dire che un elemento di H^* deriva da un elemento di H per mezzo della solita isometria canonica:

$$i(y)(Ax) = (Ax)(y)$$

questo, per definizione; tuttavia, se questo è uguale a 0, allora $y = 0$. Questo fatto segue dalla coercività: se

$$(Ax)(y) = 0 \forall x$$

allora, esso sarà 0 anche quando $x = y$; in questo caso:

$$\alpha(x, x) \geq c \|x\|^2$$

ma, dunque, se $\alpha(x, x) = 0$, allora:

$$c \|x\|^2 \leq 0$$

e dunque $x = 0$, che implica, grazie a $y = x$, che $y = 0$.

Tutto ciò è stato fatto nell'ipotesi che:

$$\|x\| \leq c \|Ax\|$$

bisogna ancora dimostrare questo punto, al fine di poter applicare tutte le considerazioni appena proposte e quindi terminare la dimostrazione del lemma di Lax-Milgram. Si applica la coercività, dividendo ambo i membri per $\|x\|$:

$$\|x\| \leq c \frac{\alpha(x, x)}{\|x\|}$$

d'altra parte, si può dire che ciò è uguale a:

$$= c \frac{(Ax)(x)}{\|x\|} \leq \sup_{y \neq 0} c \frac{A(x)(y)}{\|y\|} = c \|Ax\|$$

quindi, Lax-Milgram è dimostrato.

Nota sulla coercività

Il termine *coerciva* deriva dal fatto che, in un certo senso, funzioni come la $\alpha(x, y)$ *costringono* x a variare in un certo dominio. Si consideri:

$$\alpha(x, x) \geq c \|x\|^2$$

questo ricorda una forma quadratica che, graficamente, si può disegnare come un paraboloide; l'argomento di α è dato dalle due variabili il cui span costituisce il piano; dunque, il grafico di questa funzione è un paraboloide, i cui autovalori sono dunque entrambi strettamente positivi. Se $x \in \mathbb{R}^2$, intersecando il piano con il paraboloide si ottiene una regione **limitata** dello spazio: la forma in questione **limita** x a poter variare solo in una regione ristretta dello spazio. Questo fatto è causato dal fatto che gli autovalori sono maggiori di zero.

Applicazione alle equazioni alle derivate parziali

Si consideri un esempio per mostrare come questo teorema sia utile al fine di studiare equazioni alle derivate parziali. Dato $\Omega \subset \mathbb{R}^n$, si vuole risolvere l'equazione alle derivate parziali:

$$\nabla^2 u = f \text{ in } \Omega$$

dove:

$$\nabla^2 = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$$

Questa equazione si può risolvere per mezzo della teoria classica, oppure applicando i risultati di analisi funzionale. L'idea è vedere questa equazione, ambientandola in opportuni spazi funzionali. Se si è interessati alle soluzioni che si annullano al bordo, ossia:

$$u|_{\partial\Omega} = 0, u \in C^{(2)}(\Omega)$$

(problema di Dirichlet), allora si possono ottenere risultati.

In verità, la cosa migliore è ambientare questa equazione nei seguenti spazi di funzioni:

$$u \in H_0^1(\Omega), f \in H^{-1}(\Omega)$$

questi sono spazi di Sobolev; essi verranno definiti più in dettaglio quando verrà introdotta la teoria delle Distribuzioni, tuttavia si introduce ora qualche cenno in merito a essi. Per ora, si propongono alcune “anticipazioni”:

$$H^{-1}(\Omega) = (H_0^1(\Omega))^*$$

dove:

$$H_0^1(\Omega) = \left\{ \text{complemento delle funzioni } C_C^{(\infty)}(\Omega) \right\}$$

questo è lo spazio delle funzioni $C^{(\infty)}$ a supporto compatto in Ω ; questo spazio è definito con la norma:

$$\|u\|^2 = \|u\|_{L^2(\Omega)}^2 + \sum_{j=1}^n \left\| \frac{\partial u}{\partial x_j} \right\|_{L^2(\Omega)}^2 = \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2$$

La conclusione di tutto ciò è:

$$\forall f \in H^{-1}(\Omega) \exists! u \in H_0^1(\Omega)$$

dove u è la soluzione; questo è il risultato di Lax-Milgram, applicando $A : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$, dove:

$$A = -\nabla^2$$

infatti:

$$(Au)(\varphi) = -(\nabla^2 u)(\varphi) = \sum_{j=1}^n \int_{\Omega} \frac{\partial u}{\partial x_j} \frac{\partial \varphi}{\partial x_j} dx$$

questo passaggio riguarda il calcolo della derivata in senso distribuzionale; questo verrà motivato in seguito, quando si introdurrà la teoria delle Distribuzioni. Dal momento che:

$$\sum_{j=1}^n \int_{\Omega} \frac{\partial u}{\partial x_j} \frac{\partial \varphi}{\partial x_j} dx = \alpha(u, \varphi)$$

la funzione α così definita è bilineare, coerciva, dal momento che:

$$\sum_{j=1}^n \int_{\Omega} \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_j} dx = \|\nabla u\|_{L^2(\Omega)}^2 \geq c \|u\|_{H_0^1(\Omega)}$$

e questo è vero, per la **diseguaglianza di Poincarè**: $\forall u \in H_0^1(\Omega)$, si può verificare che, per le $u \in C_C^{(\infty)}(\Omega)$, che questo è valido nel compatto, e quindi passare al limite questo risultato; la limitatezza, richiesta nelle ipotesi di Lax-Milgram, serve proprio per poter usare la diseguaglianza di Poincarè come stima.

Si osservi che, nel caso in cui $\alpha(x, y) = \alpha(y, x)$, si ha qualcosa in più: questa teoria, per come è stata finora proposta, fornisce semplicemente un buon ambiente di studio per gli spazi a cui le soluzioni dell'equazione differenziale appartengono; tuttavia, è possibile anche ricercare la soluzione mediante un principio variazionale, minimizzando il funzionale E definito come:

$$E = \sum_{j=1}^n \left\| \frac{\partial u}{\partial x_j} \right\|^2$$

se si trova la u (mediante tecniche variazionali) tale da minimizzare ciò, questa è la soluzione dell'equazione differenziale.

Capitolo 6

Teoria spettrale

6.1 Definizioni e proprietà fondamentali

La Teoria Spettrale è nota nell'ambito di applicazioni finito-dimensionali, dove questa teoria è nota. Nell'ambito degli spazi infinito-dimensionali, si ha qualche novità: se nell'ambito di applicazioni finito-dimensionali lo spettro coincide con l'insieme degli autovalori, nell'ambito più generale si hanno elementi in più. Questo è causato sostanzialmente dal fatto che un operatore in spazi di dimensione infinita può essere iniettivo ma non suriettivo.

Sia X uno spazio di Banach, e sia $A \in \mathcal{L}(X, X)$; un numero complesso λ si dice **punto regolare** di A se l'operatore

$$(A - \lambda I) : x \rightarrow X$$

è invertibile. Automaticamente dunque, grazie al teorema della mappa aperta, questo è pure limitato.

L'insieme dei punti regolari di A è detto **insieme risolvente** di A , ed è indicato con il simbolo $\rho(A)$. In altre parole:

$$\rho(A) = \{ \lambda \in \mathbb{C} : (A - \lambda I) : X \rightarrow X \text{ è invertibile} \}$$

Dunque, $\rho(A)$ è un sottoinsieme di \mathbb{C} ; lo **spettro**, indicato con $\sigma(A)$, è l'insieme complementare a \mathbb{C} dell'insieme risolvente:

$$\sigma(A) = \{ \lambda \in \mathbb{C} : (A - \lambda I) : X \rightarrow X \text{ non è invertibile} \}$$

ossia:

$$\sigma(A) = \mathbb{C} \setminus \varrho(A)$$

Lo spettro dispone di alcune proprietà: in dimensione finita, un operatore è invertibile se e solo se è iniettivo; se invece $\lambda \in \sigma(A)$, allora **non** è iniettivo. Nel contesto più generale, valgono le proprietà ora riportate.

1. $\varrho(A)$ è un insieme aperto in \mathbb{C} ; essendo $\sigma(A)$ il suo complementare, allora si ha che lo spettro è **chiuso**.
2. Se $|\lambda| > \|A\|$, allora λ sta nell'insieme risolvente; questo, in altre parole, significa che lo spettro è contenuto nella palla di centro pari all'origine, e di raggio pari a $\|A\|$:

$$\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|A\|\}$$

Osservazione: in dimensione finita, la prima proprietà è evidente, dal momento che lo spettro è al più costituito da n numeri: le radici dell'equazione caratteristica. In un contesto più generale, per A di dimensione infinita, si può avere qualcosa di molto diverso: si può avere qualcosa di infinito, o pure variabile con la potenza del continuo.

A questo punto, si verificano le proprietà.

Verifica della proprietà 1

Si consideri $\lambda_0 \in \varrho(A)$; dunque, in questa situazione, i numeri *vicini* a λ_0 sono ancora appartenenti a $\varrho(A)$, dal momento che, essendo $\varrho(A)$ un aperto, allora ogni λ_0 ha un intorno ancora appartenente all'insieme stesso. Dunque, se $\lambda_0 \in \varrho(A)$, e se $\lambda \in \mathbb{C}$, allora $A - \lambda I$ è certamente invertibile (il che significa che λ appartiene all'insieme risolvente); infatti:

$$A - \lambda I = A - \lambda_0 I + (\lambda_0 - \lambda)I$$

il primo operatore è invertibile, mentre il secondo ha norma pari a $|\lambda_0 - \lambda|$. Se si somma a un operatore invertibile un operatore di norma piccola allora l'operatore risultante dalla somma sarà ancora invertibile; questo implica che $\lambda \in \varrho(A)$. In altre parole, questo significa che l'insieme risolvente è stabile rispetto a piccole variazioni.

Verifica della proprietà 2

Se $|\lambda| > \|A\|$, allora si ha che l'operatore $A - \lambda I$ è certamente invertibile. Per dimostrare ciò, è sufficiente raccogliere:

$$A - \lambda I = -\lambda \left(I - \frac{A}{\lambda} \right)$$

ma si può vedere, grazie ancora una volta al criterio di Neumann, che l'operatore tra parentesi tonde è invertibile; infatti:

$$\left\| \frac{A}{\lambda} \right\| \leq \frac{\|A\|}{|\lambda|} < 1$$

infatti, abbiamo l'ipotesi:

$$|\lambda| > \|A\|$$

e dunque, lo spettro è effettivamente chiuso e limitato; questo significa, dal momento che si lavora con numeri in \mathbb{C} , che è compatto.

6.2 Classificazione dello spettro

Si consideri $\lambda \in \sigma(A)$; l'operatore $A - \lambda I$, dunque, sicuramente non è invertibile. Tuttavia, ci sono vari casi non equivalenti tra loro per cui si può ricadere in questa situazione: infatti, il fatto che l'operatore sia non invertibile può significare che esso sia non iniettivo, o non suriettivo. Studiamo i vari casi.

1. Il primo caso è quello per cui l'operatore $A - \lambda I$ è non iniettivo; questo significa dunque che il nucleo dell'operatore è diverso dal solo insieme vuoto \emptyset ; in questa situazione, si ha che il λ in questione è un **autovalore**; l'insieme degli autovalori si denota con σ_p : **spettro puntuale**. Nel caso a dimensione finita, questa è l'unica situazione possibile. In questo caso, dire che:

$$\ker \{A - \lambda I\} = 0$$

coincide con dire che

$$\exists x \neq 0 : (A - \lambda I)x = 0$$

tuttavia, questo, essendo $x \neq 0$, coincide con dire che:

$$Ax = \lambda x$$

esattamente come nel caso delle metrici; se ciò accade, x è detto **autovettore relativo a λ** : l'autovettore è una soluzione **non nulla** di questa equazione.

2. Il secondo caso è quello per cui l'operatore $A - \lambda I$ è iniettivo, ma non suriettivo; la suriettività è legata al fatto che l'immagine dell'operatore coincida con lo spazio di arrivo intero; se dunque l'immagine è **densa** in X , si dice che $\lambda \in \sigma_c(A)$, dove $\sigma_c(A)$ è detto **spettro continuo** dell'operatore; in questa situazione, dunque, si ha che:

$$\ker \{A - \lambda I\} = \emptyset$$

$$\text{range} \{A - \lambda I\} \neq X$$

ma

$$\overline{\text{range} \{A - \lambda I\}} = X$$

3. Il terzo caso è quello per cui $A - \lambda I$ è iniettivo, ma

$$\overline{\text{range} \{A - \lambda I\}} \neq X$$

ossia, l'immagine di $A - \lambda I$ non è nemmeno densa; in questo caso, il λ in considerazione è tale per cui $\lambda \in \sigma_r(A)$, dove $\sigma_r(A)$ è detto **spettro residuo** di A . Il termine *residuo* in un certo senso indica il fatto che questo *non è molto importante*: nei casi pratici in cui si applica la teoria spettrale, spesso questo coincide con l'insieme nullo.

Riassumendo:

$$\mathbb{C} = \varrho(A) \cup \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

dove lo spettro $\sigma(A)$ è dato da:

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

In verità, tutte queste considerazioni valgono dal momento che stiamo considerando spazi di Banach; tutto ciò, se lo spazio non fosse completo, sarebbe soggetto a condizioni aggiuntive.

Si osservi inoltre, prima di passare agli esempi, una peculiarità riguardante la teoria spettrale: tutti gli operatori che stiamo studiando partono da uno spazio X e arrivano allo spazio X stesso; questa cosa è giustificabile *a occhio* già solo dal fatto che, essendoci di mezzo un'identità I , la cosa è necessaria.

Esempio 1

Si consideri l'operatore $A \in \mathcal{L}(l^2, l^2)$, definito come:

$$Ax = \left(x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right)$$

calcoliamo lo spettro di questo operatore. Per prima cosa, si parta dallo spettro puntuale, ossia dall'insieme degli autovalori: si cerchino i λ tali per cui:

$$\lambda \in \mathbb{C} : \exists x \neq 0 : (A - \lambda I)x = 0$$

per fare ciò: Ax è stato appena scritto; per quanto riguarda l'altro termine:

$$\lambda Ix = (\lambda x_1, \lambda x_2, \lambda x_3, \dots)$$

dunque, unendo i due, si può scrivere che:

$$(A - \lambda I)x = \left(x_1 - \lambda x_1, \frac{1}{2}x_2 - \lambda x_2, \frac{1}{3}x_3 - \lambda x_3, \dots \right) = 0$$

si vuole che questo vettore sia zero, dato un certo λ e un vettore x che non sia il vettore nullo. Il vettore è infinito dimensionale (è in l^2), dunque ha infinite componenti; questo significa che per risolvere l'equazione, è necessario scrivere il seguente sistema "infinito":

$$\begin{cases} (1 - \lambda)x_1 = 0 \\ \left(\frac{1}{2} - \lambda\right)x_2 = 0 \\ \vdots \end{cases}$$

Dovendo trovare i λ tali per cui il sistema è risolto, si può operare come segue: prima di tutto, si individuano tutti i possibili λ che possono risolvere un'equazione per volta; questi sono:

$$\lambda \in \left\{ 1, \frac{1}{2}, \frac{1}{3}, \dots \right\}$$

a questo punto, se λ **non appartenesse** a questo insieme, $x = 0$ sarebbe l'unica soluzione possibile al problema; questo significa che i λ diversi da questi **di sicuro** non sono autovalori. Per quanto riguarda gli autovettori, un esempio è: dato $\lambda = \frac{1}{n_0}$, dove $n_0 \in \mathbb{N}$, $n_0 \geq 1$; in questo caso, un esempio di x in questione è:

$$x = (0, 0, 0, \dots, 1, \dots)$$

dove il 1 è nella n_0 -esima posizione. Non è detto che ci debba per forza essere un 1: la soluzione di un problema agli autovalori dà la forma dell'autovettore, ma non il suo modulo, e quindi è possibile mettere un numero qualsiasi in quel posto.

I λ finora discussi costituiscono lo spettro puntuale:

$$\sigma_p(A) = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \dots \right\}$$

dunque, questo è infinito, ma numerabile.

Cosa si può dire riguardo agli altri spettri? Prima di tutto, è possibile osservare che anche il limite sinistro, ossia $\lambda = 0$, appartiene allo spettro di A : lo spettro è chiuso, come noto dalle proprietà precedentemente analizzate, dunque è assolutamente necessario che il limite appartenga. Verifichiamo nel dettaglio che $0 \in \sigma_c(A)$. Per completare la verifica, è necessario vedere che:

$$(A - \lambda I)x = (A - 0I)x = Ax$$

abbia la chiusura dell'immagine uguale allo spazio di arrivo X . Ma questo, era stato verificato precedentemente, dunque la dimostrazione di questo punto è completata; infatti, si era visto che:

$$c_{00} \subset \left\{ y \in l^2 : \sum_{k=1}^{\infty} |ky_k|^2 < \infty \right\}$$

Rimane un punto: se

$$\lambda \notin \left\{ 0, 1, \frac{1}{2}, \frac{1}{3}, \dots \right\}$$

allora, $\lambda \in \rho(A)$; infatti, si può dimostrare che $A - \lambda I$, in questa situazione, è sempre invertibile; in altre parole,

$$\forall y \exists ! x : (A - \lambda I)x = y$$

si può invertire. Ancora una volta, si avrà a che fare con un sistema, ma questa volta non omogeneo:

$$\begin{cases} (1 - \lambda)x_1 = y_1 \\ \left(\frac{1}{2} - \lambda\right)x_2 = y_2 \\ \vdots \end{cases}$$

dal momento che λ non appartiene al suddetto insieme, allora si può scrivere:

$$x_1 = \frac{y_1}{1 - \lambda}, \quad x_2 = \frac{y_2}{\frac{1}{2} - \lambda}, \dots$$

e così via; in sostanza, per ogni y esiste una x , scritta come sopra:

$$x = \left(\frac{y_1}{1 - \lambda}, \frac{y_2}{\frac{1}{2} - \lambda}, \dots, \frac{y_n}{\frac{1}{n} - \lambda} \right) \in l^2$$

si può dire che effettivamente $x \in l^2$? In effetti, non è banale, quindi ciò richiede alcune discussioni in più. Si ha che:

$$\|x\|_{l^2}^2 = \sum_{k=1}^{\infty} \left| \frac{y_k}{\frac{1}{k} - \lambda} \right|^2$$

dove si sa che:

$$\sum_{k=1}^{\infty} |y_k|^2 < \infty$$

quello che si deve dunque verificare è che:

$$\|x\|_{l^2}^2 = \sum_{k=1}^{\infty} \left| \frac{1}{\frac{1}{k} - \lambda} \right|^2 \leq c$$

dove c è una costante indipendente da k . Si ha che:

$$\left| \frac{1}{k} - \lambda \right| \geq \frac{1}{\sqrt{c}}$$

è vero che questa cosa è verificata, per una certa c ? La risposta è sì: se infatti $\lambda \notin \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$, λ non appartiene a un chiuso, dunque appartiene a un aperto: questo significa che λ ha una *distanza non nulla* da questo insieme (sarebbe nulla, se λ fosse in un chiuso!), dunque:

$$\sigma_c(A) = \{0\}, \sigma_r(A) = \emptyset$$

e l'esempio è terminato; prima di passare al prossimo esempio, tuttavia, si vuole proporre un'osservazione aggiuntiva: l'operatore appena studiato si può vedere come:

$$\begin{bmatrix} 1 & 0 & \dots \\ 0 & \frac{1}{2} & \dots \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} x_1 \\ \frac{1}{2}x_2 \\ \vdots \end{bmatrix}$$

ossia, è un **operatore diagonale**: questo evidenzia il fatto che gli elementi in questa *matrice di dimensione infinita* siano gli autovalori, come in effetti abbiamo dimostrato.

Esempio 2

Si consideri l'operatore $A : l^2 \rightarrow l^2$ definito come:

$$Ax = (0, x_1, x_2, \dots)$$

si verifichi che 0 appartiene allo spettro residuo; per fare ciò, si deve vedere che:

$$(A - 0I)$$

è un operatore iniettivo, ma la cui immagine non è densa in X ; questo, tuttavia, è evidente:

$$\ker \{A\} = \emptyset$$

tuttavia:

$$\text{range } \{A\} = (\{e_1\})^\perp$$

questo, era stato precedentemente visto; dunque, essendo l'ortogonale non nullo, l'immagine non può essere densa.

Si osservi che A non è un operatore autoaggiunto, a differenza di quello dell'esercizio precedente (la cosa si può dimostrare: l'aggiunto di questo operatore è lo shift a sinistra, che è diverso da questo).

Esempio 3

Si propone a questo punto un esempio sugli spazi di funzioni, analogo al primo esempio. Si consideri l'operatore $A : L^2(0, 1) \rightarrow L^2(0, 1)$, dato da:

$$(Ax)(t) = tx(t)$$

si vuole verificare che:

$$\sigma(A) = \sigma_c(A) = [0, 1] \subset \mathbb{C}$$

Questo esempio presenta alcune analogie con il primo: infatti, anche in questo caso si ha una sorta di *operatore diagonale*, dal momento che si sta moltiplicando l'equivalente dell'*indice*, nella precedente situazione, con quello che lo rappresenta nel caso continuo: la variabile indipendente t . Vogliamo dimostrare che lo spettro, in questa situazione, è l'insieme dei valori che assume t .

Se $\lambda \notin [0, 1]$, allora λ sta nel risolvente; infatti:

$$\forall y \in l^2 \exists! x \in l^2 : (A - \lambda I)x = y$$

dove x è una soluzione del sistema; applicando l'operatore, si ha che:

$$Ax = tx$$

quindi:

$$(t - \lambda)x(t) = y(t)$$

questa, deve essere valida quasi ovunque in $L^2(0, 1)$ (dal momento che stiamo lavorando con spazi di Lebesgue). Dunque, è possibile calcolare questa $x(t)$ soluzione, come segue:

$$x(t) = \frac{y(t)}{t - \lambda}$$

ora, vogliamo vedere che questo sta in $L^2(a, b)$; questo in effetti è verificato, grazie al fatto che $\lambda \notin [0, 1]$. Si ha che, se $y(t) \in L^2(0, 1)$:

$$\int_0^1 \frac{|y(t)|^2}{|t - \lambda|^2} dt < \infty$$

questo, dal momento che:

$$\frac{1}{|t - \lambda|^2} \leq c$$

dunque:

$$\int_0^1 \frac{|y(t)|^2}{|t - \lambda|^2} dt \leq c \int_0^1 |y(t)|^2 dt = c \|y\|_{L^2(a,b)}^2$$

e questo non si può fare, se $\lambda \in [0, 1]$.

In questo caso, si ha che:

$$\lambda \in \sigma_c(A)$$

dal momento che l'operatore $(A - \lambda I)$ ha immagine densa, ma diversa da $L^2(0, 1)$. L'iniettività si ha perché il nucleo coincide con l'insieme vuoto \emptyset ; tuttavia, dire che:

$$(t - \lambda)x(t) = 0 \text{ quasi ovunque}$$

coincide con il dire che $x(t) = 0$ quasi ovunque; quindi, l'immagine è densa, ma:

$$\text{range } \{(A - \lambda I)\} = \left\{ y \in L^2(0, 1) : \int_0^1 \frac{y(t)}{|t - \lambda|} \in L^2(0, 1) \right\}$$

questa è densa, ma non coincidente con tutto $L^2(0, 1)$; infatti, se $\lambda = 0$, per esempio, esistono funzioni $y(t)$ che siano in L^2 ma tali per cui $\frac{y(t)}{t}$ non sia in L^2 . Per esempio, $y(t) = \sqrt{t}$ è di queste:

$$y(t) = \sqrt{t} \implies \frac{y(t)}{t} = \frac{1}{\sqrt{t}}$$

e:

$$\int_0^1 \left| \frac{1}{\sqrt{t}} \right|^2 dt$$

non esiste.

A questo punto, per terminare questo esempio, si deve dimostrare il fatto che lo spazio delle funzioni così definito è denso. Dato $\lambda \in [0, 1]$,

$$\left\{ y \in L^2(0, 1) : \frac{y(t)}{t - \lambda} \in L^2(0, 1) \right\}$$

è denso in $L^2(0, 1)$. Per dimostrare ciò, la tattica più semplice è quella di ricondursi a un sottospazio che è già noto per ipotesi: sapendo che per esempio $C([0, 1])$ è denso in $L^2(0, 1)$ (con la norma di L^2), si può prendere $y(t)$ continua, come segue:

$$X = \{y \in C([0, 1]) : y = 0 \text{ in qualche intorno di } \lambda\}$$

Una $y \in X$ è una funzione continua che in un certo intorno di λ si annulla. Si vuole verificare che:

- che X è contenuto nel sottoinsieme che si sta studiando;
- che questo X è denso in $L^2(0, 1)$; per questo secondo punto, si può utilizzare la teoria della misura.

Per dimostrare il primo punto si deve vedere che:

$$\forall \varepsilon > 0 \forall x \in C([0, 1]) \exists x_0 \in X : \|x - x_0\|_{L^2(a,b)} < \varepsilon$$

Si consideri dunque una funzione appartenente a $x_0 \in X$, che si annulla nell'intervallo $[\lambda - 2\delta, \lambda + 2\delta]$; questa $x_0(t)$ si può considerare uguale alla $x(t)$, tranne nell'intorno $[\lambda - 2\delta, \lambda + 2\delta]$, dove la x_0 andrà a 0; l'idea è vedere che, in norma $L^2(a, b)$, questa funzione converge a x . Questo si può fare dal momento che:

$$\int_0^1 |x(t) - x_0(t)|^2 dt = \|x_0 - x\|_{L^2}^2$$

$x(t)$ è continua in un intervallo chiuso, dunque è uniformemente limitata; si ha:

$$|x(t)| \leq M$$

inoltre, per costruzione:

$$|x_0(t)| \leq M$$

infatti, queste funzioni differiscono solo nell'intervallo $[\lambda - 2\delta; \lambda + 2\delta]$. Dunque:

$$|x(t) - x_0(t)| \leq 2M$$

quindi: dato che l'intervallo è ampio 4δ , si ha:

$$\int_0^1 |x(t) - x_0(t)|^2 dt \leq 4M^2 \times 4\delta$$

quindi, se δ è sufficientemente piccolo, è possibile dire che l'integrale sia maggiorabile con un ε arbitrario.

Si osservi che questa cosa è vera dal momento che si sta utilizzando la norma dello spazio $L^2(a, b)$, e **non** la norma del sup, rispetto a cui lo spazio delle funzioni continue sarebbe completo: infatti, se si avesse la norma del sup, la norma non potrebbe essere più piccola della distanza da 0 della funzione $x(t)$ per $t \sim \lambda$: la norma $L^2(a, b)$ *media* la funzione (media integrale), mentre quella del sup è un *caso peggiore*.

6.3 Teoria di Riesz-Fredholm

A questo punto verrà proposto un certo insieme di risultati concernenti gli operatori compatti.

6.3.1 Lemma di Riesz

Si consideri uno spazio normato E , e si consideri un suo sottospazio **proprio** e **chiuso** E_1 ; questo garantisce che esiste un vettore $y_0 \in E$, $y_0 \notin E_1$, di norma unitaria, che dista da E_1 più di $\frac{1}{2}$:

$$d(y_0, E_1) \geq \frac{1}{2}$$

dove per d si intende la funzione *distanza da un vettore a un insieme*.

Si procede a questo punto con la dimostrazione di questo risultato. Si consideri E_1 un insieme proprio, dunque $E_1 \neq E$, come per ipotesi. Dato $y \in E \setminus E_1$, si ha che certamente $d(y, E_1) > 0$; di conseguenza, si definisce δ come questa distanza:

$$\delta = d(y, E_1) > 0$$

ora, non si hanno naturalmente garanzie sul fatto che $\|y\| = 1$, e tanto meno sul fatto che $d(y, E_1) = \frac{1}{2}$; però, data δ , esiste certamente un punto $x_0 \in E_1$ tale per cui:

$$d(x_0, y) > 2\delta$$

di conseguenza:

$$\exists x_0 \in E_1 : \|y - x_0\| < 2\delta$$

A questo punto, si definisca un punto y_0 come segue:

$$y_0 = \frac{y - x_0}{\|y - x_0\|}$$

questo non è altro che normalizzare il vettore $y - x_0$. In questo caso, certamente $\|y_0\| = 1$; si deve verificare che:

$$d(y_0, E_1) \leq \frac{1}{2}$$

Si ha che:

$$d(y_0, x) = \|y_0 - x\| = \left\| \frac{y - x_0}{\|y - x_0\|} - x \right\| = \frac{\|y - (x\|y - x_0\| + x_0)\|}{\|y - x_0\|}$$

$x \in E_1$, e così x_0 , dunque di sicuro $x\|y - x_0\| - x_0 \in E_1$; di sicuro dunque, il numeratore sarà $\geq \delta$, dal momento che la distanza tra il numeratore e E_1 sarà maggiore o uguale della distanza minima δ prima definita. È inoltre possibile notare che:

$$\|y - x_0\| \leq 2\delta$$

e questo, per ipotesi (per costruzione); dunque:

$$\left\| \frac{y - x_0}{\|y - x_0\|} \right\| = \frac{\|y - (x\|y - x_0\| + x_0)\|}{\|y - x_0\|} \geq \frac{\delta}{2\delta} = \frac{1}{2}$$

Questo lemma è uno strumento tecnico che ci permette di trovare elementi che sostituiscono, almeno in parte, il concetto di sistema ortogonale; infatti, grazie a questo strumento, sarà possibile sfruttare una proprietà dei sistemi ortogonali, al momento di studiare la non-compattezza delle palle chiuse nell'ambito degli spazi normati di dimensione infinita (questa cosa è stata già dimostrata, solo nell'ambito degli spazi di Hilbert, sfruttando l'ortogonalità).

6.3.2 Corollario 1 del lemma di Riesz

Il primo corollario del lemma di Riesz che si vuole discutere a questo punto discutere è il seguente. In uno spazio normato di dimensione **infinita**, la palla unitaria

$$\{x \in X : \|x\| \leq 1\}$$

non è **mai** compatta.

La dimostrazione di questa proposizione nell'ambito degli spazi di Hilbert è già stata mostrata, considerando una base hilbertiana e verificando che la distanza tra una coppia di elementi distinti è sempre maggiore o uguale di una certa costante. Nell'ambito più generale degli spazi normati, non esiste il concetto di sistema ortogonale, dal momento che non esiste il concetto di ortogonalità, quindi sarà necessario costruire una successione che in qualche senso *prenda il posto* di questa base. Si consideri dunque una successione di sottospazi $E_1 \subset E_2 \subset E_3 \subset \dots$, di dimensione infinita (di conseguenza chiusi), e propri, ossia diversi l'uno dall'altro (ciascuno contenuto strettamente nel successivo). Questo è possibile, dal momento che lo spazio è di dimensione infinita, dunque è sempre possibile trovare un sottospazio di questo tipo. Si fissi un certo sottospazio, e si consideri il precedente, ossia quello che lo contiene; in questi, si può applicare il lemma di Riesz, che garantisce l'esistenza di un vettore nello spazio ambiente che disti almeno $\frac{1}{2}$ dal sottospazio:

$$\forall n \exists x_n \in E_n \setminus E_{n-1} : \|x_n\| = 1, d(x_n, E_{n-1}) \geq \frac{1}{2}$$

Quindi, si ha che certamente la successione (x_n) è contenuta nella palla unitaria:

$$\{x_n\} \subset \{x \in X : \|x\| = 1\}$$

Questa certamente non ha sottosuccessioni convergenti; infatti, dati due elementi di $\{x_n\}$, essi distano almeno $\frac{1}{2}$: se $n > m$, e si calcola la loro distanza, ossia $\|x_n - x_m\|$, si ha che $x_m \in E_m$, ma anche che $E_m \subset E_{n-1}$; di conseguenza il lemma si può applicare, e :

$$\|x_n - x_m\| \geq \frac{1}{2}$$

dal momento che $x_n \notin E_{n-1}$.

6.3.3 Corollario 2 del lemma di Riesz

Si consideri X uno spazio di Banach (di dimensione infinita), e A un operatore lineare continuo $A \in \mathcal{L}(X, X)$, compatto; allora,

$$0 \in \sigma(A)$$

Ossia, lo zero appartiene allo spettro dell'operatore di A .

Per dimostrare questa cosa, si consideri per assurdo che 0 non sia nello spettro dell'operatore; in questo caso, l'operatore

$$A - 0 \times I = A$$

dovrebbe essere invertibile, ottenendo un compatto A^{-1} . Dal momento che la composizione di operatori compatti deve dare un compatto, si ha una contraddizione rispetto al fatto che:

$$I = A^{-1}A$$

infatti, l'identità dovrebbe essere compatta, ma ciò contraddice il corollario precedente: mappando la palla unitaria con l'identità si otterrebbe ancora la palla unitaria, e questo non è vero. Se dunque un operatore è invertibile, certamente esso non può essere compatto.

Questo corollario non fornisce informazioni riguardo a quale parte dello spettro contenga 0: è possibile che esso stia nello spettro puntuale σ_p , o nello spettro continuo σ_c , o nello spettro residuo σ_r . Si considerino ora tre esempi delle tre situazioni.

1. Sullo spazio l^2 , si consideri:

$$A \in \mathcal{L}(l^2, l^2)$$

dove A è l'operatore identicamente nullo. Questo è certamente compatto, dal momento che l'intera immagine è convergente a 0; questo significa che lo spettro sta nello zero; nel dettaglio, 0 è l'autovalore dell'operatore, il cui autospazio coincide con l'intero l^2 . Si ricorda che:

$$\text{autospazio} = \ker \{A - \lambda I\}$$

2. Si consideri, ancora sullo spazio l^2 , l'operatore

$$A = \left(x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right)$$

in questo caso, abbiamo già dimostrato (precedentemente, discutendo lo spettro di questo operatore) che:

$$0 \in \sigma_r(A)$$

3. Un terzo esempio si può ricavare modificando l'operatore al punto precedente in l^2 :

$$Ax = \left(0, x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right)$$

si vuole verificare che:

$$0 \in \sigma_r(A)$$

Per verificare questo fatto, si consideri il fatto che $(A - 0 \times I)$ è iniettivo, ma che la sua immagine non è densa in l^2 ; per fare ciò, si può dimostrare che l'immagine è un sottospazio dell'ortogonale (non l'intero ortogonale, dal momento che si ha il vincolo sulle componenti: $\frac{1}{k}x_k$); tuttavia, si può dimostrare che la chiusura dell'immagine è:

$$\overline{\text{range}\{A\}} = \{e_1\}^\perp$$

che non coincide ovviamente con l^2 , dal momento che *manca un elemento*: e_1 .

6.3.4 Proposizione: molteplicità autovalori di un operatore compatto

Vale la seguente proposizione: se A è un operatore, ogni suo autovalore $\lambda \neq 0$ ha molteplicità (geometrica) finita, ossia:

$$\dim \{\ker \{A - \lambda I\}\} < \infty$$

Ossia, la dimensione dell'autospazio relativa a λ è finita; si ribadisce che questo risultato non vale per $\lambda = 0$.

Si vuole dimostrare *indirettamente* questa proposizione, ossia dimostrando un altro fatto: si vuole verificare che, alle condizioni scritte, la palla unitaria è **compatta**, in questo spazio. Se la palla unitaria è compatta, allora essa certamente il sottospazio in cui la si valuta è di dimensione finita. Si vuole dunque verificare che:

$$\{x \in \ker \{A - \lambda I\}, : \|x\| \leq 1\}$$

è compatto.

Si consideri $\{x_n\}$ una successione in questo spazio; si vuole verificare che questa è convergente. Sicuramente, si ha che, $\forall n, \|x_n\| \leq 1$, per ipotesi; inoltre, si ha che:

$$(A - \lambda I)x_n = 0, \forall n$$

dal momento che le x_n appartengono al kernel dell'operatore in questione; un modo diverso di scrivere questa affermazione è:

$$Ax_n = \lambda x_n$$

ma dunque, essendo $\lambda \neq 0$ per ipotesi:

$$x_n = \frac{1}{\lambda} Ax_n$$

l'operatore A è compatto per ipotesi, di conseguenza, data x_n limitata (e lo è, per ipotesi, essendo nella palla unitaria), allora Ax_n ha sicuramente una sottosuccessione convergente; ma, essendoci la proporzionalità a meno di una costante λ^{-1} tra Ax_n e x_n , come appena scritto, allora anche x_n ha una sottosuccessione convergente; questa cosa è valida solo dal momento che non si sta considerando $\lambda = 0$, e che A è compatto.

Si potrebbe dimostrare che oltretutto esiste l'elemento a cui si converge, ossia che questo sottospazio è chiuso, ma, dal momento che $\|x\| \leq 1$ è chiuso, e che il \ker è dato dalla controimmagine di un chiuso, per compattezza si ha la chiusura.

6.4 Spettro di un operatore compatto

A questo punto verranno proposti alcuni risultati riguardanti lo spettro di operatori compatti, continuando sull'onda di quelli appena ottenuti grazie al lemma di Riesz.

6.4.1 Teorema: proprietà dello spettro di un operatore compatto

Sia X uno spazio di Banach di dimensione infinita, e sia $A \in \mathcal{L}(X, X)$ un operatore compatto; allora:

- $0 \in \sigma(A)$;
- $\sigma(A) \setminus \{0\}$ è un insieme costituito da autovalori di molteplicità finita.

In altre parole, ogni punto dello spettro che non sia $\lambda = 0$ è un autovalore di un operatore compatto; inoltre, l'insieme $\sigma(A) \setminus \{0\}$, ossia l'insieme degli autovalori, può essere l'insieme vuoto, oppure un insieme finito, oppure una successione che tende a zero (convergente a 0). Al fine di fissare questo concetto, si propongono alcuni esempi delle varie situazioni.

1. Un esempio dell'ultimo caso era già stato analizzato precedentemente:

$$Ax = \left(x_1, \frac{1}{2}x_2, \frac{1}{3}x_3, \dots \right)$$

in questo caso dunque, la successione degli autovalori è una successione che converge a 0.

2. Un secondo esempio già analizzato in precedenza è $A = 0$: in questo caso, si ha un insieme nullo di autovalori.

3. Un esempio in cui si ha un insieme finito di autovalori è l'operatore in l^2 definito come:

$$Ax = (\lambda_1 x_1, \lambda_2 x_2, \lambda_3 x_3, \dots, \lambda_n x_n, 0, 0, 0, \dots)$$

questo è un *operatore diagonale troncato*, nel senso che dalla n -esima componente in poi si hanno solo zeri; questo operatore ha n autovalori; inoltre $\lambda = 0 \in \sigma_p(A)$, e ha molteplicità infinita; quindi:

$$\sigma(A) = \sigma_p(A) = \{0, \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n\}$$

e lo spettro quindi coincide con lo spettro puntuale.

4. Su $C([0, 1])$, si consideri l'operatore $Ax(t)$ definito come:

$$Ax(t) = \int_0^1 x(s) ds$$

questo è un operatore tipo Volterra: si tratta di operatori integrali in cui come estremo di integrazione si ha la variabile stessa; in altre parole, questa è la primitiva. Questo operatore è compatto, e lo spettro di questo operatore è:

$$\sigma(A) = \sigma_c(A) = \{0\}$$

6.4.2 Teorema di Fredholm

Si propone a questo punto un altro teorema importante per la caratterizzazione dello spettro degli operatori compatti. Sia X uno spazio di Banach, e sia $A \in \mathcal{L}(X, X)$ un operatore compatto, considerando $\lambda \neq 0$. Allora:

1. la dimensione dell'autospazio associato a λ è finita:

$$\dim \{\ker \{A - \lambda I\}\} < \infty$$

si osservi che questa vale per $\lambda = 0$, ma in generale non solo per gli autovalori; d'altra parte, se λ non fosse un autovalore, l'operatore avrebbe nucleo nullo, e dunque questa relazione sarebbe banale;

2. l'immagine dell'operatore

$$\text{range } \{A - \lambda I\}$$

è chiusa, e di codimensione finita; questo si ha solamente perché stiamo considerando un operatore compatto a cui si somma un multiplo (non nullo) dell'identità: questo significa che, dal momento che quando si parla di operatori compatti si parla di operatori in un certo senso *piccoli*, allora l'identità non viene perturbata di molto, dunque la proprietà di chiusura rimane mantenuta. Per quanto riguarda la codimensione, si può ricordare da quanto fatto in precedenza che essa sia la dimensione dello spazio ambiente X rispetto allo spazio dell'immagine; questo significa che l'immagine è *molto grande*: la codimensione si può pensare come a una sorta di *numero di dimensioni che mancano allo spazio da cui si fa il quoziente*; di conseguenza, se la codimensione dell'immagine è finita, significa che la dimensione dello spazio dell'immagine è *grande*: si sta *togliendo poco* a X . Questa è una proprietà degli operatori Fredholm: hanno il **nucleo piccolo**, e la **immagine grande**.

3. Da qui in poi si proporranno alcune proprietà, che verranno quindi discusse.

$$\ker \{A - \lambda I\} = (\text{range } \{A^t - \lambda I\})_{\perp}$$

e

$$\ker \{A^t - \lambda I\} = (\text{range } \{A - \lambda I\})^{\perp}$$

discutiamo la prima: a membro sinistro si ha l'insieme di X in cui l'immagine si annulla a destra, si ha l'insieme *ortogonale* all'immagine, ossia l'annullatore dell'immagine.

4. Un secondo gruppo di proprietà è:

$$\text{range } \{A - \lambda I\} = (\ker \{A^t - \lambda I\})_{\perp}$$

$$\text{range } \{A^t - \lambda I\} = (\ker \{A - \lambda I\})_{\perp}$$

5. Vale inoltre:

$$\ker \{A - \lambda I\} = 0 \iff \text{range} \{A - \lambda I\} = X$$

questa è una proprietà molto interessante, dal momento che dice che, per questa situazione (operatori compatti, $\lambda \neq 0$), l'operatore è iniettivo se e solo se è suriettivo, esattamente come accade in dimensione finita.

6. Un'ulteriore proprietà che si discuterà è:

$$\dim \{\ker \{A - \lambda I\}\} = \dim \{\ker \{A^t - \lambda I\}\} = \text{codim} \{\text{range} \{A - \lambda I\}\} < \infty$$

ossia, il numero di autovettori linearmente indipendenti di A rispetto a λ coincide con il numero di autovettori del trasposto rispetto a λ , e ciò coincide con la codimensione dell'immagine; tutto ciò, è finito.

A questo punto, al fine di approfondire, alcune di queste proprietà verranno spiegate meglio, al fine di comprenderne le applicazioni e il significato. Dato dunque A un operatore compatto, e $\lambda = 0$, allora si hanno le seguenti.

- La prima proprietà che si vuole discutere è:

$$\dim \{\ker \{A - \lambda I\}\} < \infty$$

questa, in altre parole, significa che l'equazione:

$$Ax - \lambda x$$

ha un numero finito di soluzioni linearmente indipendenti.

- La seconda proprietà che si vuole discutere è:

$$\text{range} \{A - \lambda I\} = (\ker \{A^t - \lambda I\})_{\perp}$$

questo significa che un elemento sta nell'immagine di $A - \lambda I$ se e solo se esso sta nell'ortogonale del nucleo dell'operatore trasposto; questo significa, dato $y \in X$, che l'equazione:

$$(A - \lambda I)x = y$$

ha almeno una soluzione (ossia, che y sta nell'immagine), se e solo se:

$$\varphi(y) = 0 \forall y \in X^*$$

tale per cui φ sia una soluzione dell'equazione omogenea trasposta:

$$A^t\varphi - \lambda\varphi = 0$$

questo, significa dunque che y sta nell'annullatore del \ker , e questo garantisce che la soluzione esiste, allo stesso tempo.

- L'ultima proprietà che si vuole studiare è la seguente:

$$\ker \{A - \lambda I\} = \{\emptyset\} \iff \text{range} \{A - \lambda I\} = X$$

questo significa che $A - \lambda I$ è iniettivo se e solo se è suriettivo; in altre parole, questo significa che l'equazione

$$(A - \lambda I)x = y$$

è risolubile per ogni y se e solo se l'equazione omogenea associata,

$$(A - \lambda I)x = 0$$

ha solo la soluzione banale, ossia $x = 0$. Solo in questo caso, la soluzione dell'equazione non omogenea esiste.

Questa è la cosiddetta **alternativa di Fredholm**; il termine *alternativa* deriva dal fatto che essa propone il fatto che una delle seguenti due possibilità sussiste:

– o

$$(A - \lambda I)x = y \text{ ha una soluzione } \forall y \in X$$

– oppure

$$Ax = \lambda x \text{ ha una soluzione non banale, ossia } x \neq 0$$

L'alternativa di Fredholm è molto importante, dal momento che essa permette di studiare suriettività e iniettività con un'unica azione; inoltre, capire che un operatore è iniettivo è molto più facile rispetto a capire che esso è suriettivo.

Questo risultato è molto importante, dal momento che può essere applicato per la determinare l'esistenza di soluzioni di equazioni alle derivate parziali. Si consideri, dato $\Omega \in \mathbb{R}^n$, l'equazione definita su Ω :

$$-\nabla^2 u = f$$

noi avevamo discusso questa equazione applicando il lemma di Lax-Milgram e altri strumenti *avanzati* dell'analisi funzionale; tuttavia, questi non c'erano, ai tempi di Fredholm; egli, dunque, dovette affrontare questo problema riconducendolo (con passaggi che non verranno riportati qui) a un'equazione integrale. La filosofia di ciò era: dato Ω aperto, con un certo bordo $\gamma(t)$, Fredholm riconduceva il problema a un problema sul bordo, del tipo:

$$x(t) - \int_a^b K(t, s)x(s) ds = y(t)$$

definito A un operatore:

$$Ax(t) = \int_a^b K(t, s)x(s) ds$$

dove $K(t, s)$ è una funzione continua, allora questa equazione è proprio nella forma:

$$(A - I)x = y$$

dato dunque questo operatore, il teorema dell'alternativa garantisce le soluzioni di questo problema.

Cenni conclusivi

Da quest'idea è nata la cosiddetta *teoria del potenziale*: questa è la possibilità di rappresentare le soluzioni di un'equazione differenziale, $u(t)$, mediante integrali sul contorno, e così via.

La prosecuzione naturale di questo tipo di studi è lo studio degli operatori di Fredholm: finora sono sempre stati considerati operatori tipo $I + A$, dove

I è l'identità e A è un operatore che perturba l'identità; dati più in generali operatori F tali per cui:

$$F \in \mathcal{L}(X, X) : \dim \{\ker \{F\}\} < \infty, \text{range } \{F\} \text{ è chiusa, } \text{codim } \{\text{range } \{F\}\} < \infty$$

se ne sono studiate le proprietà. Per esempio, si è definito **l'indice** per questi operatori: un numero reale, che rappresenta la dimensione del nucleo, meno la codimensione dell'immagine:

$$\text{ind } \{F\} = \dim \{\ker \{F\}\} - \text{codim } \{\text{range } \{F\}\}$$

questo è un numero. Si può dimostrare, per esempio nel caso degli operatori nella forma $I + A$, che:

$$\text{ind } \{I + A\} = \text{ind } \{I\} = 0$$

infatti, l'indice non è perturbato dall'operatore A , e si ha che, per I , sia il \ker sia la codimensione dell'immagine è zero, dunque la differenza è nulla.

Esempi di operatori di Fredholm sono gli operatori ellittici su varietà compatte, del tipo:

$$P = \sum_{|\alpha| \leq m} c_\alpha(x) \partial^\alpha$$

localmente; questi sono operatori P del tipo:

$$P : C^{(\infty)}(M) \rightarrow C^{(\infty)}(M)$$

Prima di concludere, si vuole proporre un cenno riguardante uno dei teoremi più importanti di questo filone dell'analisi funzionale: il teorema di Atiyah-Singer (1963). Questo afferma che, dato P operatore, si ha che:

$$\text{ind } \{P\} = \text{indice topologico } (P)$$

dove l'indice topologico è legato per esempio alla caratteristica di Eulero-Poincaré e ad altri concetti più geometrico/topologici.

Una nota finale: tutti i teoremi e le proposizioni mostrate su $(A^t - \lambda I)$, nel caso di **spazi di Hilbert**, valgono allo stesso modo, sostituendo a A^t l'aggiunto, A^* , e a λ il suo coniugato $\bar{\lambda}$:

$$(A^t - \lambda I) \implies (A^* - \bar{\lambda} I)$$

6.5 Operatori compatti e autoaggiunti

6.5.1 Operatori autoaggiunti

Prima di tutto è necessario introdurre una definizione per gli operatori autoaggiunti, per poi studiare le proprietà del loro spettro. Sia H uno spazio di Hilbert; un operatore limitato $A \in \mathcal{L}(H, H)$ è detto **autoaggiunto**, o **simmetrico**, o **hermitiano** se:

$$A = A^*$$

ossia, se:

$$(Ax|y) = (x|Ay)$$

si osservi: si lavora sempre nello stesso spazio: lo spazio di arrivo coincide con lo spazio di partenza.

Prima di enunciare una serie di proprietà, se ne vuole proporre una fondamentale: se l'operatore è autoaggiunto, allora $(Ax|x) \in \mathbb{R}$: è un numero reale. Per mostrare questa cosa, si ricordi che, in genere:

$$(a|b) = \overline{(b|a)}$$

tuttavia, qua, si prenda $a = Ax$, $b = x$:

$$(Ax|x) = \overline{(x|Ax)} = (x|Ax)$$

ma dunque, il numero è uguale al coniugato, e dunque è reale. Data per buona questa proprietà, dunque, se ne introducono altre. Dato un operatore A autoaggiunto, seguono le proprietà.

1. Lo spettro puntuale è contenuto nei reali:

$$\sigma_p(A) \subset \mathbb{R}$$

in verità, si ha che tutto lo spettro $\sigma(A)$ è contenuto in \mathbb{R} , e lo spettro residuo è l'insieme vuoto.

2. Dati λ_1, λ_2 due autovalori distinti, dunque $\lambda_1 \neq \lambda_2$, dati x_1 e x_2 autovettori relativi a essi:

$$Ax_1 = \lambda_1 x_1 \quad Ax_2 = \lambda_2 x_2$$

allora, $x_1 \perp x_2$, ossia $(x_1|x_2) = 0$. Per comprendere intuitivamente ciò (la dimostrazione verrà fatta dopo), si pensi che, quando si vuole costruire una base e poi ortogonalizzare, si utilizzano gli autovettori del problema aggiunto; in questo caso, dal momento che il problema aggiunto coincide col problema di partenza, intuitivamente gli autovettori devono essere ortogonali!

3. Se $L \subset H$ è un **sottospazio invariante**, ossia tale per cui:

$$A(L) \subset L$$

allora, anche L^\perp è un sottospazio invariante.

4. La norma dell'operatore A si può calcolare con la seguente espressione:

$$\|A\| = \sup_{\|x\| \neq 0} \frac{|(Ax|x)|}{\|x\|^2} = \sup_{\|x\|=1} |(Ax|x)|$$

questo, se $H \neq \emptyset$.

A questo punto, queste proprietà verranno dimostrate, una a una.

1. Per quanto riguarda la dimostrazione della prima proprietà, si ha che, se $Ax = \lambda x$, per $x \neq 0$, allora si ha che:

$$(Ax|x) = (\lambda x|x) = \lambda \|x\|^2$$

tuttavia, è stato precedentemente dimostrato il fatto che $(Ax|x) \in \mathbb{R}$; essendo inoltre $\|x\|^2 \in \mathbb{R}$, allora è per forza necessario che $\lambda \in \mathbb{R}$: gli autovalori sono reali, sempre! A meno che, $x = 0$.

2. Per quanto riguarda la seconda dimostrazione, concernente il fatto che autovettori relativi a diversi autovalori sono ortogonali, si consideri quanto segue:

$$\lambda_1(x_1|x_2) = (\lambda_1 x_1|x_2) = (Ax_1|x_2) = (x_1|Ax_2) = \lambda_2(x_1|x_2)$$

quindi:

$$(\lambda_1 - \lambda_2)(x_1|x_2) = 0$$

essendo però $\lambda_1 \neq \lambda_2$, è necessario che $(x_1|x_2) = 0$.

3. Per quanto riguarda la terza affermazione, è necessario verificare che, se $x \in L^\perp$, allora $Ax \in L^\perp$; per fare ciò, si consideri:

$$y \in L \implies (Ax|y) = (x|Ay)$$

ma, se $y \in L$, $Ay \in L$, essendo L uno spazio invariante; di conseguenza, se $x \in L^\perp$, si ha che:

$$(x|Ay) = 0$$

4. La quarta proprietà è quella più complicata da dimostrare, dal momento che è necessario introdurre alcuni artifici. Prima di tutto, per alleggerire la notazione, si chiami con C la seguente espressione:

$$C = \sup_{\|x\| \neq 0} \frac{|(Ax|x)|}{\|x\|^2}$$

a questo punto, si vuole verificare che $C = \|A\|$; questo significa, in altre parole, che si deve verificare che, contemporaneamente, $\|A\| \leq C$, $C \leq \|A\|$. Per quanto riguarda la prima, è semplice: applicando Cauchy-Schwarz, si ottiene:

$$C \leq \sup_{\|x\| \neq 0} \frac{\|Ax\| \|x\|}{\|x\|^2} = \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|$$

Per quanto riguarda la dimostrazione nell'altro verso, si deve vedere che $\|A\| \leq C$. Si può dunque scrivere ciò:

$$(A(x+y)|(x+y)) - (A(x-y)|(x-y)) = 2[(Ax|y) + (Ay|x)]$$

questa si dimostra svolgendo i conti usando le proprietà dei prodotti scalari. Si può a questo punto calcolare il valore assoluto per entrambi i membri, e scrivere che:

$$\begin{aligned} 2 |(Ax|y) + (Ay|x)| &= |(A(x+y)|(x+y)) - (A(x-y)|(x-y))| \leq \\ &\leq |(A(x+y)|(x+y))| + |(A(x-y)|(x-y))| \end{aligned}$$

da cui:

$$|(Ax|y) + (Ay|x)| \leq \frac{1}{2} |(A(x+y)|(x+y))| + |(A(x-y)|(x-y))|$$

Tuttavia, è possibile dire che:

$$|(A(x \pm y)|(x \pm y))| \leq C \|x \pm y\|^2$$

questo, usando la definizione di C , basata sulla norma. Dunque:

$$|(Ax|y) + (Ay|x)| \leq \frac{C}{2} [\|x+y\|^2 + \|x-y\|^2]$$

ora, si può usare l'identità del parallelogramma:

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

da cui:

$$= \frac{C}{2} (2\|x\|^2 + 2\|y\|^2)$$

ora che si ha ciò, si può scrivere che:

$$|(Ax|y) + (Ay|x)| \leq C (\|x\|^2 + \|y\|^2)$$

Ora, si consideri x di norma unitaria, e y definito come:

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

questo, per $Ax \neq 0$; si ha:

$$\left| \left(Ax \left| \frac{Ax}{\|Ax\|} \right. \right) + \left(\frac{A(Ax)}{\|Ax\|} |x \right) \right| = \frac{1}{\|Ax\|} |(Ax|Ax) + (Ax|Ax)| \leq C(1+1) = 2C$$

da qui:

$$\frac{2}{\|Ax\|} \|Ax\|^2 \leq 2C$$

quindi:

$$\|Ax\| \leq C$$

questo, $\forall x : \|x\| = 1$, che coincide con dire che:

$$\|A\| \leq C$$

nel caso $Ax = 0$, questo non varrebbe; tuttavia, dal momento che, nel caso $Ax = 0$, la norma di A sarebbe nulla, allora l'eguaglianza è banale. Si è dunque dimostrato che:

$$\|A\| = \sup \{|(Ax|x)| : \|x\| = 1\}$$

6.5.2 Lemma: esistenza di un autovalore

Prima di introdurre il teorema fondamentale per lo studio dello spettro di un operatore autoaggiunto, è buona cosa introdurre un lemma, che sarà utile per dimostrarlo. Sia $A \in \mathcal{L}(H, H)$ un operatore compatto e autoaggiunto; allora, A ha un autovalore λ tale per cui:

$$|\lambda| = \|A\|$$

inoltre,

$$\sup \{|(Ax|x)| : \|x\| = 1\}$$

è *raggiunto* da un certo x , che sarebbe l'autovettore relativo al λ appena citato.

Si procede a questo punto con la dimostrazione di questo lemma. Prima di tutto, si supponga che A non sia identicamente nullo; dal momento che A è un operatore compatto, sarà sufficiente dimostrare, per dimostrare questo lemma, che:

$$\exists \lambda \in \sigma(A) : |\lambda| = \|A\|$$

Giustificiamo questa frase: come mai è sufficiente dimostrare questo? La risposta è: noi vogliamo un autovalore che in modulo coincida con $\|A\|$; dunque, se $\|A\|$ non è nullo, allora anche questo λ sarà non nullo; di conseguenza, esso starà nello spettro puntuale:

$$\lambda \in \sigma_p(A)$$

infatti, se λ è non nullo, e l'operatore A è compatto, allora si ha che λ deve per forza essere un autovalore, in virtù del teorema riguardante lo spettro degli operatori compatti discusso precedentemente¹. Quindi, se $\|x_0\| = 1$, e $Ax_0 = \lambda x_0$, si ha che:

$$\begin{aligned} |(Ax_0|x_0)| &= |(\lambda x_0|x_0)| = |\lambda| |(x_0|x_0)| = \\ &= |\lambda| \|x_0\|^2 = |\lambda| = \|A\| = \sup \{|(Ax|x)| : \|x\| = 1\} \end{aligned}$$

Con questo, abbiamo dimostrato che effettivamente dimostrare quella condizione è sufficiente; a questo punto, però, bisogna verificare che quella proposizione sia soddisfatta: si vuole dimostrare che esiste un λ nello spettro tale per cui il modulo coincida con la norma dell'operatore. Per dimostrare ciò, sappiamo che esiste una successione $\{x_n\} : \|x_n\| = 1$, e tale per cui:

$$|(Ax_n|x_n)| \rightarrow \|A\|$$

questo, dal momento che la norma di A coincide con il sup, e quindi questo significa semplicemente aver preso e applicato la definizione di sup. Nell'espressione appena scritta, la convergenza è in modulo, dunque non è detto, se la successione è molto oscillante, che essa converga anche senza il valore assoluto; tuttavia, essendo questa successione di prodotti scalari sicuramente limitata (dal momento che assolutamente convergente, come appena detto), si ha che esiste, per Bolzano-Weierstrass, una sottosuccessione $\{x_{n_k}\}$ tale per cui:

$$(Ax_{n_k}|x_{n_j}) \rightarrow \lambda$$

dove questo λ è semplicemente il limite della sottosuccessione. Considerando i moduli di queste sottosuccessioni si avrà ovviamente che:

¹in verità non sarebbe necessario utilizzare quel teorema, ma sarebbe possibile dimostrare anche senza di esso questa proposizione, ma ciò non sarà fatto.

$$|(Ax_{n_k}|x_{n_j})| \rightarrow |\lambda|$$

ma, avendo tutte le sottosuccessioni lo stesso limite della successione, si avrà anche che:

$$|(Ax_{n_k}|x_{n_j})| \rightarrow \|A\|$$

ossia:

$$|\lambda| = \|A\|$$

Rimane l'ultimo punto: dato questo λ , ad ora semplicemente definito come il limite della sottosuccessione (x_{n_k}) , esso appartiene allo spettro dell'operatore A , $\sigma(A)$? Per verificare ciò, è necessario vedere che:

$$\|Ax_{n_k} - \lambda x_{n_k}\| \rightarrow 0$$

se questa successione fosse identicamente nulla, significherebbe che λ dovrebbe esattamente essere un autovalore (dal momento che soddisferebbe l'equazione agli autovalori, e dunque ciò che sta dentro la norma sarebbe identicamente nullo); tuttavia, non vogliamo verificare ciò, ma solo che appartiene allo spettro, quindi procediamo in modo diverso. Supponiamo per assurdo che $\lambda \notin \sigma(A)$; allora, in questo caso, $\lambda \in \rho(A)$, e dunque λ sarebbe tale per cui $A - \lambda I$ sarebbe un operatore invertibile; di conseguenza, sarebbe possibile scrivere:

$$\|x_n\| = \|(A - \lambda I)^{-1}(A - \lambda I)x_n\| \leq C \|(A - \lambda I)x\|$$

leggiamo quest'ultimo passaggio al contrario:

$$\|(A - \lambda I)x\| \geq \frac{1}{C} \|x_n\|$$

dove C è una certa costante, e $\|x_n\|$ è costante; questo significherebbe che $\|(A - \lambda I)x\|$ non potrebbe assolutamente tendere a zero, dal momento che sarebbe sempre maggiore di questa costante. Di conseguenza, è significativo verificare che:

$$\|Ax_{n_k} - \lambda x_{n_k}\| \rightarrow 0$$

dunque, lo si dimostri. Si ha, svolgendo il prodotto scalare:

$$\|Ax_{n_k} - \lambda x_{n_k}\|^2 = (Ax_{n_k}|Ax_{n_k}) - 2\lambda(Ax_{n_k}|x_{n_k}) + \lambda^2 \|x_{n_k}\|^2$$

però: $\|x_{n_k}\|^2 = 1$, per ipotesi; inoltre, sempre per costruzione, $(Ax_{n_k}|x_{n_k}) \rightarrow \lambda$ (è il limite della sottosuccessione, come definito precedentemente!), e infine:

$$(Ax_{n_k}|Ax_{n_k}) = \|Ax_{n_k}\|^2 \leq \|A\|^2 \|x_{n_k}\|^2 = 1 \times \lambda^2$$

da cui:

$$\leq \lambda^2 - 2\lambda^2 + \lambda^2 = 0$$

quindi, questo effettivamente tende a zero, e la dimostrazione del lemma è ultimata.

6.5.3 Teorema di Hilbert-Schmidt

Questo teorema risale circa al 1910, e rappresenta uno dei primi grandi successi dell'Analisi Funzionale. Questo teorema afferma che, dato H uno spazio di Hilbert, e dato A un operatore tale per cui $A \in \mathcal{L}(H, H)$, compatto e autoaggiunto, dove $A \neq 0$, allora esiste un insieme di autovalori $\{\lambda_i\}$ non nulli che può essere o finito, o numerabile (convergente a zero), e un sistema ortonormale di autovettori $\{e_i\}$,

$$Ae_i = \lambda_i e_i$$

tali per cui:

1. $\forall x \in H \exists y \in \ker \{A\}$, tale per cui:

$$x = y + \sum_{i=1}^{\infty} (x|e_i)e_i$$

2. $\forall x \in H$,

$$Ax = \sum_{i=1}^{\infty} \lambda_i (x|e_i)e_i$$

3. $\forall z \in \overline{\text{range}\{A\}}$, si ha che:

$$z = \sum_{i=1}^{\infty} (z|e_i)e_i$$

ossia, $\{e_i\}$ è una base di Hilbert per la chiusura dell'immagine dell'operatore A .

Prima di dimostrare questi punti, si vuole discutere un momento quale sia il loro significato.

1. Dato A operatore compatto e autoaggiunto, si ha un sistema di autovalori/autovettori; non solo: gli autovalori, al peggio, costituiscono una successione infinitesima; essi possono anche non essere distinti tra loro. Inoltre, ogni elemento di H si può scrivere come una combinazione degli elementi e_i , nel dettaglio la **proiezione ortogonale** sullo span degli e_i , sommando anche una y , appartenente al kernel dell'operatore; questo significa che per rappresentare l'intero H , non sono sufficienti gli e_i , se l'operatore non fosse iniettivo (e avesse dunque kernel non vuoto).
2. La seconda proprietà rappresenta la proprietà di **diagonalizzazione** dell'operatore: si parla di diagonalizzazione, dal momento che l'azione dell'operatore coincide con la sola moltiplicazione per i λ_i .
3. La terza proprietà afferma che gli e_i , sistema ortonormale, sono una base della chiusura dell'immagine.

Dimostrazione punti 1 e 2

Il punto 2 segue banalmente dal punto 1; dato $y \in \ker\{A\}$, applicando A a $x + y$, si ottiene:

$$A(y + x) = Ay + Ax = 0 + Ax = \sum_{i=1}^{\infty} (x|e_i)e_i$$

e così, il secondo punto è molto semplice.

Per quanto riguarda il primo punto, la dimostrazione può essere condotta applicando il lemma precedentemente dimostrato. Poiché abbiamo tutte le ipotesi per applicare questo lemma, si può dire che esiste un autovalore λ_1 tale per cui:

$$|\lambda_1| = \|A\|$$

Di conseguenza, a questo autovalore corrisponde un autovettore e_1 tale per cui:

$$Ae_1 = \lambda_1 e_1$$

dove è possibile normalizzare e_1 :

$$e_1 : \|e_1\| = 1$$

In questo modo, abbiamo garantito l'esistenza di un autovalore; tuttavia, si consideri ora il sottospazio E_1 come:

$$E_1 = \text{span} \{e_1\}$$

dunque, dato E_1^\perp il complemento ortogonale dello span di e_1 , si consideri la restrizione di A su questo:

$$A|_{E_1^\perp}$$

a questo punto, si supponga che questo spazio sia non nullo: $A|_{E_1^\perp} \neq 0$; di conseguenza, è di nuovo possibile applicare a questo spazio il lemma precedente, e dunque si può dire che esista λ_2 tale per cui si abbia un e_2 (autovalore e autovettore rispettivamente),

$$|\lambda_2| = \left\| A|_{E_1^\perp} \right\|$$

dove:

$$Ae_2 = \lambda_2 e_2$$

e

$$\|e_2\| = 1$$

Questo, esattamente come fatto poco fa per A invece che per $A|_{E_1^\perp}$. Si osservi che λ_1, λ_2 non possono essere nulli: infatti, su $A|_{E_1^\perp}$, per ipotesi, la norma non è nulla. Inoltre, dal momento che vale il lemma precedente, si ha che $A|_{E_1^\perp}$ deve essere compatto e autoaggiunto; inoltre, dal momento che E_1 è lo span di un autovettore, si ha che:

$$Ae_1 = \lambda_1 e_1$$

proprio come appena scritto; ma dunque E_1 è un sottospazio invariante e, come abbiamo dimostrato precedentemente, anche E_1^\perp è un sottospazio invariante. Di conseguenza a tutto ciò, si ha inoltre che:

$$0 < |\lambda_2| \leq |\lambda_1|$$

Questo si può motivare osservando che:

$$|\lambda_1| = \|A\|$$

e

$$|\lambda_2| = \left\| A|_{E_1^\perp} \right\|$$

nel secondo caso l'indagine è su di una restrizione, dunque la norma può essere minore o uguale, non di sicuro maggiore rispetto a quella dell'operatore non ristretto.

È possibile proseguire su questa falsa riga, definire E_2 come:

$$E_2 = \text{span} \{e_1, e_2\}$$

quindi, se $A|_{E_2^\perp}$ non è nullo, allora si può applicare il lemma, trovare λ_3, e_3 , e così via. Questo permette di costruire una successione di autovalori uno più piccolo dell'altro. Se a un certo passo n_0 si ha che:

$$A|_{E_{n_0}^\perp} = 0$$

allora, si sarebbe costruito un numero finito di autovalori $\{\lambda_1, \lambda_2, \dots, \lambda_{n_0}\}$, non nulli, con un sistema di autovettori ortogonali. In caso contrario, si otterrebbe una successione di autovalori tutti non nulli, sempre con autovettori ortonormali.

Nel primo caso, si chiami:

$$y_{n_0} = x - \sum_{i=1}^{n_0} (x|e_i)e_i$$

questo elemento, appartiene sicuramente al nucleo di A ; infatti, se si calcola x (dove $x \in H$), meno la proiezione ortogonale, allora

$$x - \mathcal{P}_n(x)$$

appartiene certamente allo spazio ortogonale allo span degli e_i , sui quali si sta proiettando x ; dunque:

$$y_{n_0} \in (\text{span} \{e_i\})^\perp$$

tuttavia, si è appena detto che:

$$A|_{E_{n_0}^\perp} = 0$$

quindi, se A nella restrizione è nullo, significa che tutti gli elementi applicabili a esso (ossia appartenenti alla restrizione, ossia appartenenti all'ortogonale dello span degli e_i) sarà nucleo: saranno mappati in zero. Quindi, y_{n_0} è certamente nel nucleo di A . In altre parole, a forza di restringere, si è ottenuto solo più il nucleo dell'operatore.

Questa era solo una delle possibilità: non è detto che ci sia questo n_0 tale per cui l'operatore ristretto al complemento ortogonale di n_0 sia sempre nullo; non è dunque detto che, a forza di restringere, si arrivi a ottenere il nucleo! In questo caso, si consideri:

$$y_n = x - \sum_{i=1}^n (x|e_i)e_i$$

il nostro desiderio sarebbe che, per $n \rightarrow \infty$, $y_n \rightarrow y$, e che questo y fosse nel nucleo di A .

Il motivo per cui y_n converge è perché x è fissato, e la serie converge, grazie a Riesz-Fischer; dalla dimostrazione del teorema di Riesz-Fischer, infatti, è possibile dedurre che, considerando l'applicazione che a x associa i prodotti scalari, si ottiene una successione $\{\alpha_i\}$; se questa è in l^2 , allora la serie converge. Mediante la disuguaglianza di Bessel è possibile dimostrare che la nostra successione di prodotti scalari, $(x|e_i)$, appartiene a l^2 , e dunque data una serie che moltiplica i vettori (tra loro ortogonali), si può mostrare che le somme parziali sono di Cauchy, e quindi per la completezza di l^2 , esse sono pure convergenti. Quindi:

$$y_n \rightarrow y \in H$$

e però rimane da chiedersi: è y in $\ker \{A\}$? La risposta è sì: dire che esso è nel nucleo significa dire che:

$$Ay = 0$$

questo si ha dal momento che:

$$\|Ay\| = \lim_{n \rightarrow \infty} \|Ay_n\|$$

e si ha che $\|Ay_n\|$ è nello spazio ortogonale dei $\{e_1, e_2, \dots, e_n\}$. Di conseguenza, si avrebbe, al n -esimo termine, che A dovrebbe essere ristretto a E_n^\perp , per come abbiamo costruito tutto; di conseguenza:

$$\|Ay_n\| \leq \left\| A|_{E_n^\perp} \right\| \|y_n\|$$

ma, si ha che $\|y_n\| \rightarrow \|y\|$, e che:

$$\left\| A|_{E_n^\perp} \right\| = |\lambda_{n+1}|$$

quest'ultima affermazione è semplice: per come abbiamo costruito gli autovallori, si ha che il modulo di λ_{n+1} è uguale alla norma dell'operatore, ristretto al complemento ortogonale di E_n (per esempio, $|\lambda_2| = \left\| A|_{E_2^\perp} \right\|$). Di conseguenza, dal momento che abbiamo già dimostrato che $\{\lambda_{n+1}\} \rightarrow 0$, per la compattezza dell'operatore (solito teorema), si ha che y sta nel nucleo di A (dal momento che l'immagine tende a 0). Quindi, abbiamo dimostrato che:

$$x = y + \sum_{i=1}^{\infty} (x|e_i)e_i$$

dove:

$$y \in \ker \{A\}$$

Dimostrazione del punto 3

Si osservi che $\{e_i\}$ è una base hilbertiana della chiusura dell'immagine; di conseguenza, ogni elemento nella chiusura dell'immagine si può scrivere mediante questa base. Dal momento che un sottospazio chiuso di uno spazio di Hilbert è completo, si vuole verificare se effettivamente quindi $\{e_i\}$ è un sistema completo, ossia se:

$$\overline{\text{span} \{e_i\}} = \overline{\text{range} \{A\}}$$

Per verificare che si ha eguaglianza tra i due insiemi, è necessario verificare le inclusioni nei due sensi.

Per quanto riguarda una inclusione si sa che $\text{range}\{A\}$ è contenuto nella chiusura dello span, grazie al punto 2 del teorema, precedentemente dimostrato; infatti:

$$Ax = \sum_{i=1}^n \lambda_i(x|e_i)e_i$$

di conseguenza,

$$\text{range}\{A\} \subset \text{span}\{e_i\}$$

questo significa che se la somma è finita (n finito), l'immagine di x attraverso A si può scrivere come combinazione lineare per coefficienti opportuni degli e_i , e dunque in effetti l'immagine di A sarà sicuramente un sottoinsieme di queste combinazioni lineari. Se poi $n \rightarrow \infty$, si deve considerare il limite della serie; dal momento che non è detto che il limite appartenga a $\text{span}\{e_i\}$, sicuramente però apparterrà alla sua chiusura:

$$\text{range}\{A\} = \overline{\text{span}\{e_i\}}$$

(il limite se non converge a un elemento dell'insieme, converge a un elemento della chiusura; se poi converge a un elemento della chiusura, vuol dire che l'insieme è chiuso, poiché coincide con la propria chiusura). Di conseguenza, la chiusura dell'immagine è contenuta nella chiusura dello span: infatti, la chiusura è il più piccolo sottoinsieme chiuso che contiene il sottoinsieme in questione, dunque deve essere contenuto in quello a destra: a destra si ha un chiuso (poiché la chiusura di un insieme è sempre chiusa), quindi a sinistra si dovrà avere un chiuso: la chiusura.

Per quanto riguarda l'inclusione nel verso opposto si deve verificare che gli $\{e_i\}$ stiano nell'immagine; questo tuttavia è banale da dimostrare, dal momento che gli e_i sono autovettori per l'operatore:

$$Ae_i = \lambda_i e_i$$

ma dunque, per linearità:

$$A \begin{pmatrix} e_i \\ \lambda_i \end{pmatrix} = e_i$$

dal momento che $\lambda_i \neq \{0\}$, sicuramente $\{e_i\} \in \text{range } \{A\}$
 di conseguenza, vale anche nel verso opposto l'inclusione:

$$\text{span } \{e_i\} \subset \text{range } \{A\}$$

dunque, questa cosa vale anche per i chiusi:

$$\overline{\text{span } \{e_i\}} = \overline{\text{range } \{A\}}$$

Questo conclude la verifica del punto 3.

6.5.4 Corollario del teorema di Hilbert-Schmidt

Si vuole a questo punto considerare un corollario di questo teorema. Il teorema di Hilbert-Schmidt esclude il caso $\lambda_i = 0$; questo non significa che esso non può essere un autovalore per l'operatore, bensì semplicemente il teorema non studia questo caso.

Si consideri H uno spazio di Hilbert separabile, e un operatore $A \in \mathcal{L}(H, H)$ compatto e autoaggiunto. Allora, esiste una base hilbertiana di H costituita da autovettori di A .

Questa è una generalizzazione del teorema di geometria: questo corollario garantisce che una matrice autoaggiunta è sempre diagonalizzabile, e che ha autovalori a due a due ortogonali; ovviamente in geometria non è necessario che una matrice sia autoaggiunta per essere diagonalizzabile: non è detto che una matrice non autoaggiunta non sia diagonalizzabile; tuttavia, questa è solo una condizione sufficiente, e che garantisce anche che gli autovettori siano ortogonali.

Nel teorema precedente, era stata trovata una base hilbertiana per la chiusura dell'immagine, ma non per l'intero spazio; tuttavia, si può decomporre lo spazio H come la chiusura dell'immagine, più l'ortogonale della chiusura dell'immagine, in virtù del teorema di decomposizione ortogonale, e quindi aggiungere solo più una base hilbertiana per lo spazio ortogonale. Si può tuttavia vedere che l'ortogonale della chiusura dell'immagine non è altri che il nucleo dell'operatore, e questo può essere tranquillamente rappresentato mediante degli autovettori. Il nucleo dell'operatore A è l'insieme degli x tali per cui:

$$Ax = 0$$

questo però può essere visto come:

$$Ax - \lambda x = 0$$

di conseguenza, se si considera l'autovalore $\lambda = 0$, si hanno proprio gli autovettori relativi a esso, come base del kernel. Essendo l'operatore A per ipotesi autoaggiunto, si ha:

$$\ker \{A\} = \ker \{A^*\} = (\text{range } \{A\})^\perp = \overline{(\text{range } \{A\})}^\perp$$

(infatti, l'ortogonale di un sottospazio è uguale all'ortogonale della chiusura di un sottospazio).

Di conseguenza, è possibile scrivere l'intero spazio H come:

$$\begin{aligned} H &= \overline{(\text{range } \{A\})}^\perp \oplus \overline{\text{range } \{A\}} = \\ &= \ker \{A\} \oplus \overline{\text{range } \{A\}} \end{aligned}$$

A questo punto si supponga che l'operatore A non sia identicamente nullo, e, in virtù del teorema precedente, già si ha una base della chiusura dell'immagine, costituita dagli autovettori di A . Quindi, è sufficiente *completare* il tutto con una base del nucleo di A , che abbiamo detto esistere; essa sarà costituita dagli autovalori di A corrispondenti a $\lambda = 0$. Se poi $\lambda = 0$ non fosse autovalore dell'operatore, allora $\ker \{A\} = \emptyset$.

Esempio applicativo

Si vuole proporre l'esempio sul quale Hilbert lavorava, in seguito alla formulazione di questo teorema: la diagonalizzazione di operatori integrali. Dato $A : L^2(a, b) \rightarrow L^2(a, b)$, definito come:

$$Ax(t) = \int_a^b K(t, s)x(s) ds$$

dove:

$$K(t, s) \in L^2([a, b] \times [a, b])$$

ossia:

$$\int_a^b \int_a^b |K(t, s)|^2 dt ds < \infty$$

Gli operatori di questo tipo sono detti di **Hilbert-Schmidt**. Si può dimostrare che questo operatore, dal momento che è limite di una successione di operatori di rango finito, è compatto; di conseguenza, lo spettro di questo operatore ha $\lambda = 0$ (che può essere nello spettro continuo o in quello puntuale), più uno spettro puntuale che può essere nullo, finito o costituito da una successione infinitesima di autovalori. In generale, non è però detto che operatori di questo tipo siano autoaggiunti: lo sono solamente se:

$$K(t, s) = \overline{K(s, t)}$$

per il teorema precedente (Hilbert-Schmidt insieme al corollario), si ha che esiste una base hilbertiana $\{e_i\}$ di $L^2(a, b)$ costituita da autovettori:

$$Ae_i = \lambda_i e_i$$

di conseguenza, per ogni $x \in L^2(a, b)$, è possibile scrivere:

$$x = \sum_{i=1}^{\infty} (x|e_i) e_i$$

a questo punto, per il secondo punto:

$$\begin{aligned} Ax(t) &= \sum_{i=1}^{\infty} \lambda_i (x|e_i) e_i = \sum_{i=1}^{\infty} \lambda_i e_i(t) \int_a^b x(s) \overline{e_i(s)} ds = \\ &= \int_a^b \sum_{i=1}^{\infty} \lambda_i e_i(t) \overline{e_i(s)} x(s) ds \end{aligned}$$

quindi, si può riconoscere che:

$$K(t, s) = \sum_{i=1}^{\infty} \lambda_i e_i(t) \overline{e_i(s)}$$

ossia che, dato un operatore compatto autoaggiunto, è possibile espandere il nucleo dell'operatore come somma di prodotti tensoriali $e_i(t) \overline{e_i(s)}$.

Capitolo 7

Teoria delle distribuzioni

Si vuole proporre qualche nozione storica riguardante la teoria delle distribuzioni. Le origini vengono sostanzialmente dalla fisica, negli inizi del 1900. Il primo lavoro che ha richiesto le distribuzioni potrebbe essere quello di Oliver Heaviside, nell'introduzione del calcolo operativo: la necessità di trattare equazioni differenziali come equazioni algebriche (cosa giustificabile mediante la trasformata di Laplace); tra i vari tipi di segnali di eccitazione per circuiti, uno era il **gradino** $U(t)$, o *funzione di Heaviside*; questo, doveva essere in qualche modo derivabile, ma chiaramente non è possibile, utilizzando le funzioni nel senso ordinario, effettuare questa derivazione.

Il secondo contributo fondamentale alla nascita della teoria delle distribuzioni fu nell'intorno del 1930, da parte di Paul Adrien Maurice Dirac, che aveva, per problemi legati alla meccanica quantistica, bisogno di definire una funzione **impulsiva**, $\delta(x)$, tale per cui:

$$\int_{-\infty}^{+\infty} \delta(x)\varphi(x) dx = \varphi(0)$$

ossia, serviva una funzione $\delta(x)$ tale per cui si restituisse $\varphi(0)$: questa deve essere una funzione nulla per $x < 0$, ma anche per $x > 0$, ma il comportamento in $x = 0$ non era ben determinato.

La prima definizione formale di distribuzione venne introdotta nel 1945 da Lauren Schwartz, mostrando come la teoria delle distribuzioni da lui inventata potesse essere applicata alla risoluzione di problemi come quello di Heaviside o quello di Dirac. L'idea è basata sulla ridefinizione del concetto di funzione: quando si parla di funzioni, si è interessati a sapere quale valore puntuale, numerico una certa *scatola* restituisca, dopo aver introdotto

un certo valore numerico. Qui, l'idea è completamente diversa: nell'ambito delle distribuzioni non si considera più il valore puntuale della funzione, ma se ne considera un qualche funzionale, per esempio integrale: non conta più quale valore una funzione assume in un punto, ma conta l'integrale! Quello che si fa, nel dettaglio, è interpretare, invece che come **funzione**, come **funzionale** un oggetto matematico: questo significa dunque *leggere una certa funzione come funzionale contro un insieme di funzioni test*: è un po' come **osservare uno spazio sotto il punto di vista dei suoi funzionali**. Data per esempio $\rho(\underline{r})$, con $\underline{r} \in \mathbb{R}^3$ una distribuzione, date φ delle funzioni *test*, appartenenti a un certo spazio di funzioni, si associa alle φ il funzionale integrale:

$$\varphi \rightarrow \int \rho(\underline{r})\varphi(\underline{r}) \, dx dy dz$$

questa era l'idea di Schwartz.

7.1 Definizioni fondamentali

Fondamentale è definire correttamente le funzioni test mediante le quali si fa ciò. Prima di tutto, dunque si definisce lo spazio delle funzioni test: dato Ω un sottoinsieme **aperto** di \mathbb{R}^n ,

$$\Omega \subset \mathbb{R}^n$$

allora, si considerino $C^{(k)}$ le funzioni derivabili k volte, definite come segue:

$$C^{(k)} = \{f : \Omega \rightarrow \mathbb{C} : \text{abbiano derivata continua fino all'ordine } k\}$$

al fine di trattare correttamente e in maniera sintetica queste derivate, si utilizza una **notazione multiindice**. Dato $x \in \mathbb{R}^n$, ossia un vettore n -dimensionale a valori reali, nella forma:

$$x = (x_1, x_2, x_3, \dots, x_n) \in \mathbb{R}^n$$

per indicare la derivata di un certo ordine rispetto a varie componenti, si utilizza la seguente notazione. Si definisce α come:

$$\alpha = (\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n) \in \mathbb{N}^n, \mathbb{N} = \{0, 1, 2, \dots\}$$

e, poi, si definisce la derivata α -esima di una funzione φ con la seguente definizione:

$$\partial^\alpha \varphi = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$$

dove $|\alpha|$ è detto **lunghezza del multiindice**, e non è il modulo, bensì la somma delle componenti:

$$|\alpha| = \sum_{i=1}^n \alpha_i$$

Al fine di fissare le idee sulla notazione, si considera un esempio pratico. Dato $x \in \mathbb{R}^3$, scritto come $x = (x_1, x_2, x_3)$, dato $\alpha = (5, 2, 4)$, si ha:

$$|\alpha| = 5 + 2 + 4 = 11$$

e quindi:

$$\partial^\alpha \varphi(x_1, x_2, x_3) = \frac{\partial \varphi^{11}}{\partial x_1^5 \partial x_2^2 \partial x_3^4}$$

ossia, si deriva cinque volte rispetto a x_1 , si deriva due volte rispetto a x_2 , si deriva quattro volte rispetto a x_3 . Se si suppone, in questo esempio, che $\varphi \in C^{(11)}(\Omega)$, allora l'ordine di derivazione è ininfluenza, dal momento che è possibile applicare ripetutamente il teorema di Schwarz.

In generale quindi, $\partial^\alpha \varphi$ non dipende dall'ordine con cui si deriva se $\varphi \in C^{(k)}(\Omega)$, $k \geq |\alpha|$.

A questo punto, per proseguire, si introducono definizioni aggiuntive. Si definisce lo spazio $C^{(\infty)}(\Omega)$ come segue:

$$C^{(\infty)}(\Omega) = \bigcap_{k \in \mathbb{N}} C^{(k)}(\Omega)$$

Quindi, questo significa che una funzione appartiene allo spazio $C^{(\infty)}(\Omega)$ se ha le derivate continue per ogni ordine.

Un'altra definizione che si vuole introdurre riguarda il **supporto** di una funzione. Data $f \in C^{(0)}(\Omega)$, si definisce il suo supporto come segue:

$$\text{supp } \{f\} = \overline{\{x \in \Omega : f(x) \neq 0\}}$$

si osservi che la chiusura in questione è rispetto a Ω , non rispetto a \mathbb{R}^n ; questa osservazione è molto importante, dal momento che a causa di questo fatto la chiusura è un sottoinsieme di Ω ; tuttavia, Ω è un aperto. Il fatto di fare la chiusura in questo senso, significa che è possibile ottenere, rispetto a \mathbb{R}^n , un insieme aperto. Si consideri, per comprendere meglio ciò, il seguente esempio:

$$f(x) = \begin{cases} 0, & 0 < x \leq \frac{1}{2} \\ x - \frac{1}{2}, & \frac{1}{2} \leq x < 1 \end{cases}$$

Questa funzione, per $x = 1$, non è definita. Questo significa che il supporto di questa funzione, ossia l'insieme dei valori per cui questa non è nulla, è:

$$\text{supp } \{f\} = \left[\frac{1}{2}, 1 \right)$$

ossia, sebbene questo sia la chiusura, dal momento che il Ω in questione non comprende $x = 1$, questo supporto non è un chiuso rispetto a \mathbb{R} . Questo supporto dunque non è compatto, dal momento che non è chiuso.

Si indichi a questo punto:

$$C_C^{(\infty)}(\Omega) = C_0^{(\infty)}(\Omega)$$

come lo spazio delle funzioni di classe $C^{(\infty)}(\Omega)$ tale per cui il loro supporto sia compatto; questo significa, dal momento che stiamo considerando spazi reali, che il supporto di f è chiuso e limitato. Lo spazio $C_C^{(\infty)}(\Omega)$ è lo **spazio delle funzioni test**, e spesso viene anche indicato come $\mathcal{D}(\Omega)$.

Si vuole a questo punto proporre un metodo per costruire una funzione test: si consideri, in \mathbb{R} , la funzione:

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Questa funzione non è a supporto compatto, tuttavia essa è utile, dal momento che viene fornito un **attaccamento di tipo** $C^{(\infty)}$, per $x = 0$; infatti, in questo punto, si ha un “incollamento” tale per cui la funzione è derivabile infinite volte; questo è dovuto al fatto che la funzione va a zero così rapidamente, per $x = 0$, che tutte le derivate sono continue; infatti, è possibile dimostrare che:

$$f^{(n)}(x) = \frac{P_n(x)}{x^{2n}} e^{-\frac{1}{x}}$$

dove $P_n(x)$ è un polinomio di grado n . Si può dimostrare questa formula per induzione, e questa è chiaramente continua.

Si è detto che questa non è una funzione test; tuttavia, prendendo due funzioni di questo tipo, traslandole, e ricavando la simmetrica rispetto a y di una di queste, si arriva a ottenere:

$$f(x+a)f(a-x) = \varphi(x) \in C_C^{(\infty)}(\Omega) = \mathcal{D}(\Omega)$$

infatti, in questo caso la funzione è a supporto compatto, e infinitamente derivabile in ogni punto. Per estendere questa idea a \mathbb{R}^2 , è possibile semplicemente ruotare rispetto all'asse y la curva, ottenendo così un solido di rotazione che rappresenta ancora una funzione test, ma questa volta per lo spazio euclideo di dimensione 2.

Lo spazio delle funzioni test ha proprietà molto interessanti; per esempio, è possibile dimostrare che questo spazio è denso in $L^p(\Omega)$, se $1 \leq p < \infty$. Questo è interessante dal momento che dice che, sebbene queste funzioni soddisfino proprietà molto particolari, esse, per la densità, sono anche in numero molto elevato, di conseguenza le funzioni integrabili possono essere approssimate mediante combinazioni di funzioni infinitamente derivabili.

Una seconda proprietà di $C_C^{(\infty)}(\Omega)$ riguarda la sua struttura algebrica: esso infatti è uno spazio vettoriale; infatti, sommando due funzioni appartenenti a questo spazio si ottiene una funzione appartenente a esso, e la stessa cosa moltiplicandola per uno scalare.

A questo punto, definiamo su questo spazio una certa nozione di convergenza: sia $\{\varphi_n\} \in C_C^{(\infty)}(\Omega)$; si dice che essa converge a $\varphi \in C_C^{(\infty)}(\Omega)$ se:

1. Esiste un compatto $K \in \Omega$ tale per cui:

$$\text{supp} \{\varphi_n\} \subset K, \forall n$$

ossia, se le varie funzioni hanno un supporto compatto, ma non solo: questo deve essere contenuto in questo K , in maniera tale che i supporti non si accumulino verso il bordo.

2. $\forall \alpha \in \mathbb{N}^n, \partial^\alpha \varphi_n \rightarrow \partial^\alpha \varphi$, uniformemente, su K . Questo significa che:

$$\|\partial^\alpha \varphi_n - \partial^\alpha \varphi\|_{L^\infty} \rightarrow 0$$

Questa nozione di convergenza è estremamente forte: essa infatti implica tutte le definizioni di convergenza precedentemente definite, come la convergenza debole, e così via.

7.1.1 Definizione di distribuzione

A questo punto, siamo finalmente pronti per fornire la definizione di distribuzione. Una distribuzione su Ω è un funzionale lineare di $C_C^{(\infty)}(\Omega)$ in \mathbb{C} , continuo:

$$u : C_C^{(\infty)}(\Omega) \rightarrow \mathbb{C}$$

dove la continuità è nel seguente senso: se $\{\varphi_n\}$ è una successione di funzioni test, se $\varphi_n \rightarrow 0$ in $C_C^{(\infty)}(\Omega)$, allora:

$$u(\varphi_n) \rightarrow 0 \text{ in } \mathbb{C}$$

ossia, il funzionale applicato alla successione diventa una successione di numeri complessi, tendente a zero. Nel dettaglio, si ha che $\varphi_j \in C_C^{(\infty)}(\Omega)$ converge a $\varphi \in C_0^{(\infty)}(\Omega)$ se e solo se:

- Esiste un compatto $K \subset \Omega$ tale per cui:

$$\text{supp } \{\varphi\} \subset K, \text{supp } \{\varphi_j\} \subset K \forall j$$

- $\forall \alpha, \partial^\alpha \varphi_j \rightarrow \partial^\alpha \varphi$, uniformemente, su K ; si deve specificare che la convergenza sia su K , dal momento che il limite potrebbe anche essere su Ω .

Esiste una definizione equivalente di continuità, basata sulla seguente stima. Questa afferma che, $\forall K \in \Omega$, dove K è un compatto, esistono costanti $C > 0$, $m \in \mathbb{N}$, tali per cui:

$$|u(\varphi)| \leq c \sum_{|\alpha| \leq m} \sup_{x \in K} |\partial^\alpha \varphi(x)| \quad \forall \varphi \in C_C^{(\infty)}(\Omega), \text{supp } \{\varphi\} \subset K$$

questo, con φ con supporto contenuto in K . Le costanti c e m , dipendono dal compatto K che si va a prendere. Questa stima controlla $u(\varphi)$ in termini di una costante, moltiplicata per un altro oggetto, che non è una norma: è una somma dei sup delle derivate di φ .

Si noti che:

$$\sup_{x \in K} |\partial^\alpha \varphi(x)| = \|\partial^\alpha \varphi\|_{L^\infty}$$

si può dunque intravedere un primo collegamento con i funzionali lineari e la loro teoria. Detto ciò, si vuole dimostrare che le due definizioni sono in verità equivalenti.

Dimostrazione prima parte

La prima implicazione che si vuole dimostrare è quella “inversa”: data come ipotesi la stima, si vuole dimostrare che la definizione di convergenza è valida. Questo è ovvio, dal momento che, se vale la stima, allora data una successione di φ_j che vanno a 0, per definizione di convergenza esiste un compatto fissato in cui tutte queste φ_j convergono; di conseguenza, è possibile effettuare, come vedremo tra breve, una maggiorazione; infatti, se, $\forall j$, esiste un K tale per cui $\text{supp} \{\varphi_j\} \subset K$, allora si ha che:

$$|u(\varphi_j)| \leq \sum_{|\alpha| \leq m} \sup_{x \in K} |\partial^\alpha \varphi_j(x)|$$

questa è una somma finita, dal momento che m è un numero finito; inoltre, per ipotesi, $\sup_{x \in K} |\partial^\alpha \varphi_j(x)|$ tende a 0, dal momento che le φ_j tendono a 0, così come le loro derivate; di conseguenza, il membro sinistro tende a 0.

Dimostrazione dell'implicazione diretta

Si vuole dimostrare a questo punto che, prendendo per ipotesi la definizione, si ha anche la stima. Questa dimostrazione verrà condotta per assurdo; dal momento che si ha una certa proposizione, al fine di negarla, è necessario scambiare i quantificatori \forall e \exists . La proposizione negata sarà:

$$\exists K \subset \Omega : \forall c > 0 \forall m \in \mathbb{N}, \exists \varphi_j \in C_C^{(\infty)}(\Omega), \text{supp} \{\varphi_j\} \subset K :$$

$$|u(\varphi_j)| > c \sum_{|\alpha| \leq j} \sup_{x \in K} |\partial^\alpha \varphi_j(x)|$$

questa deve valere per ogni c , dunque, si può scegliere un c , e per questo di sicuro dovrà essere corretto; si sceglie $c = j$:

$$|u(\varphi_j)| > j \sum_{|\alpha| \leq j} \sup_{x \in K} |\partial^\alpha \varphi_j(x)|$$

a questo punto, ha senso ipotizzare che $|u(\varphi_j)| = 1$; questo non limita in alcun modo la dimostrazione, dal momento che, se così non fosse, è possibile sostituire alle φ_j delle $\lambda\varphi_j$; questo λ può essere scelto in modo tale da normalizzare a 1 il tutto. Fatta questa ipotesi, si può portare il j all'altro membro:

$$\sum_{|\alpha| \leq j} \sup_{x \in K} |\partial^\alpha \varphi_j(x)| \leq \frac{1}{j}$$

ciascun termine della somma è positivo, dunque, se l'intera somma è minore di qualcosa, a maggior ragione ciascun termine dovrà esserlo:

$$\sup_{x \in K} |\partial^\alpha \varphi_j(x)| \leq \frac{1}{j}$$

ora, è ragionevole dire che, appena j supera $|\alpha|$, si avrà che j sarà sufficientemente grande da avere il sup molto piccolo; dunque, l'argomento sarà piccolo: $\varphi_j \rightarrow 0$. Tuttavia, $|u(\varphi_j)| = 1$, e quindi questo **non** tende a 0: questa è la contraddizione! Infatti, nonostante i φ_j vadano a 0, gli $|u(\varphi_j)| = 1$ non fanno altrettanto, e quindi si ha una contraddizione, dal momento che l'ipotesi stessa non è soddisfatta.

Esempio 1

A questo punto, si vuole proporre un primo esempio di distribuzione. Nel dettaglio, si consideri $L^1_{\text{loc}}(\Omega)$, definito come:

$$L^1_{\text{loc}}(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{C} \text{ misurabile e tale per cui } \int_K |f| < \infty \forall K \in \Omega \text{ compatto} \right\}$$

questa condizione non impone alcuna limitazione nella crescita al bordo: è sufficiente che la funzione f sia **localmente integrabile**. Data f in questo spazio, è possibile definire una distribuzione T_f come:

$$T_f(\varphi) = \int_{\Omega} f(x)\varphi(x) dx$$

ossia, ogni funzione localmente integrabile sui compatti definisce una distribuzione, in questa maniera. La φ si annulla fuori dal supporto compatto, e su quel supporto è infinitamente derivabile, dal momento che è una funzione test; $f(x)$ è integrabile, dunque il prodotto delle due è a supporto compatto, ed è integrabile. Di conseguenza, si può dire che $L^1_{\text{loc}}(\Omega)$ sia un sottoinsieme di tutte le distribuzioni: tutte le funzioni appartenenti a esso sono distribuzioni. Definito $\mathcal{D}'(\Omega)$ lo spazio delle distribuzioni, dunque:

$$L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$$

la novità è che funzioni appartenenti a questo spazio possono essere definite su tutto \mathbb{R} , ma, grazie alla funzione test, a patto che esse siano definite nel compatto K e integrabili in esso, tutto funziona.

Dato tutto ciò, ossia $f \in L^1_{\text{loc}}(\Omega)$, allora f è anche una distribuzione. Per dimostrare ciò, utilizziamo la caratterizzazione basata sulla stima:

$$|T_f(\varphi)| = \left| \int_{\Omega} f(x)\varphi(x) dx \right| \leq \int_K |f(x)| |\varphi(x)| dx$$

infatti, $\varphi(x)$ è non nulla solo su un compatto K , che rappresenta il suo supporto: $K = \text{supp}(\varphi)$. Si può ora maggiorare la funzione test $\varphi(x)$ con il suo sup (che, in verità, è pure un massimo, dal momento che K è un compatto):

$$\leq \sup_{x \in K} |\varphi(x)| \int_K |f(x)| dx$$

se si sceglie $c = \int_K |f(x)| dx$, la dimostrazione è completata; di conseguenza, f è una distribuzione.

Questo significa che si può pensare a $L^1_{\text{loc}}(\Omega)$ come a un sottoinsieme dello spazio $\mathcal{D}'(\Omega)$; in verità, prima di dire ciò, sarebbe ancora necessario vedere che due funzioni distinte possono essere viste come due distribuzioni distinte; in altre parole, se $T_f = 0$, allora $f = 0$, in L^1_{loc} (ossia, nulla quasi ovunque); questo però non è banale da dimostrare, dunque si salta.

Esempio 2

Nel precedente esempio si è discusso il fatto che un intero spazio di funzioni sia anche parte dello spazio delle distribuzioni; ora, si vuole studiare un esempio di distribuzione che non stia in questo spazio: la delta di Dirac.

$$\delta_0 \in \mathcal{D}'(\mathbb{R})$$

ossia, la delta di Dirac in $x = 0$ (precedentemente anche rappresentata come $\delta(x)$). Questa distribuzione agisce sulla funzione test come segue:

$$\delta_0(\varphi) = \varphi(0)$$

questa è una distribuzione; infatti, se $\text{supp}\{\varphi\} \subset K$, dove K è un certo compatto, allora:

$$|\delta_0(\varphi)| \leq \sup_{x \in K} |\varphi(x)|$$

Esempio 3

Come terzo esempio si vuole studiare, come distribuzione, il **valore principale**; nel dettaglio, data *V.P.* la funzione *valore principale*, si vuole studiare:

$$V.P. \frac{1}{x} \text{ in } \mathbb{R}$$

Si osservi che $\frac{1}{x}$ non è localmente integrabile; infatti, se K è un compatto che contiene l'origine, allora questa funzione non è integrabile. Tuttavia, si può vedere che, in senso distribuzionale, ha senso calcolare:

$$\left(V.P. \frac{1}{x} \right) (\varphi) = \lim_{\varepsilon \rightarrow 0} \int_{|x| > \varepsilon} \frac{\varphi(x)}{x} dx$$

Per ragionare su ciò, un metodo sarebbe quello di utilizzare la classica espansione di Taylor; utilizzando il teorema di Lagrange, si ha che:

$$\varphi(x) = \varphi(0) + x\varphi'(\xi)$$

dove ξ è un certo punto che sta nell'intervallo di valutazione; questo modo di lavorare non ha molto senso, dal momento che si potrebbe sostituire nell'integranda, e ottenere:

$$\frac{\varphi(0)}{x} + x \frac{\varphi'(\xi)}{x}$$

tuttavia, ξ può essere una funzione di x ; di conseguenza, non si hanno garanzie sull'integrabilità di questa cosa.

Al fine di lavorare correttamente, è necessario utilizzare la **formula di Taylor con resto in forma integrale**. Per dimostrarla rapidamente, si consideri la funzione di due variabili $\varphi(tx)$; si può dire che:

$$\int_0^1 \frac{d}{dt} [\varphi(tx)] dt = \varphi(x \times 1) - \varphi(x \times 0)$$

questo, per il teorema fondamentale del calcolo integrale; tuttavia, è possibile scrivere esplicitamente la funzione integranda, ottenendo:

$$\frac{d}{dt} \varphi(tx) = x \varphi'(tx)$$

dunque:

$$x \int_0^1 \varphi'(tx) dt = \varphi(x) - \varphi(0)$$

Da cui:

$$\varphi(x) = \varphi(0) + x \int_0^1 \varphi'(tx) dx$$

Questa formula non è troppo diversa da quella di prima, se non fosse che in questo caso è stata esplicitata la $\varphi'(\xi)$: a questo punto, infatti, si ha una funzione di x di classe $C^{(\infty)}$.

A questo punto, si utilizzi questa formula invece di quella precedentemente mostrata; data φ una funzione test su \mathbb{R} , si avrà che, dato un certo $A \in \mathbb{R}$, $A > 0$, allora:

$$\text{supp } \{\varphi\} \subset [-A; A]$$

dunque, riprendendo la formula del valore principale:

$$\left(V.P. \frac{1}{x} \right) (\varphi) = \lim_{\varepsilon \rightarrow 0} \int_{|x| > \varepsilon} \frac{\varphi(x)}{x} dx$$

dal momento che, come appena scritto, φ è a supporto limitato, si può dire da un lato che $|x| > \varepsilon$, ma anche che $|x| \leq A$; di conseguenza:

$$= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon < |x| \leq A} \frac{\varphi(0)}{x} dx + \int_{\varepsilon < |x| \leq A} \int_0^1 \frac{x\varphi'(x)}{x} dt dx = 0 + \int_{-A}^A \int_0^1 \varphi'(tx) dt dx$$

l'ultimo passaggio si motiva come segue: il primo integrale è nullo, dal momento che l'integrale in $[-A; -\varepsilon]$ è uguale e opposto a $[\varepsilon; A]$, dunque i due contributi si annullano. Rimane il secondo integrale, in cui il limite si può svolgere semplicemente *attaccando* i due contributi, ossia ponendo $\varepsilon = 0$, dal momento che non si ha più una discontinuità: questa si è semplificata grazie alla formula utilizzata. Dunque:

$$\left(V.P. \frac{1}{x} \right) (\varphi) = \int_{-A}^A \int_0^1 \varphi'(tx) dt dx$$

A questo punto, si vuole verificare che la stima sia soddisfatta:

$$\left| \left(V.P. \frac{1}{x} \right) (\varphi) \right| \leq \int_{-A}^A \int_0^1 |\varphi'(x)| dt dx \leq \sup_{x \in K} |\varphi'(x)| \int_{-A}^A \int_0^1 dt dx = 2A \sup_{x \in K} |\varphi'(x)|$$

7.1.2 Ordine di una distribuzione

Sono stati studiati alcuni esempi pratici, dove si è dimostrato il fatto che alcuni oggetti matematici sono effettivamente distribuzioni. Per effettuare queste dimostrazioni, è stata utilizzata la stima precedentemente dimostrata. Nei primi due esempi, questa stima è stata effettuata con $m = 0$, mentre nel secondo esempio con $m = 1$; m sta a indicare il massimo ordine di derivazione utilizzato per la stima. Si ha che:

$$\forall K \subset \Omega \exists c > 0, m \in \mathbb{N} : |u(\varphi)| \leq c \sum_{|\alpha| \leq m} \sup_{x \in K} |\partial^\alpha \varphi(x)| \quad \forall \varphi \in C_C^{(\infty)}(\Omega) : \text{supp} \{\varphi\} \subset K$$

Se si pensa ai funzionali, il fatto che un funzionale fosse *piccolo*, nel senso di mappare sottospazi piccoli in sottospazi piccoli, si definiva mediante la norma:

$$|f(x)| \leq c \|x\|$$

quando si pensa a spazi di funzioni derivabili, volendo lavorare sulle derivate, la norma era calcolata come somma delle varie derivate, da ordine (funzione stessa) a un certo ordine, esattamente come fatto qui; qui però si ha qualcosa di differente: in questo caso, $C_C^{(\infty)}(\Omega)$ non è uno spazio normato; di conseguenza, non si sta parlando esattamente di questo tipo di situazione, altrimenti sarebbe semplicemente stato possibile applicare la teoria dei funzionali lineari e degli spazi duali.

Ad ogni modo, se si può scegliere il massimo ordine m indipendentemente dal compatto K in questione, allora u si dice **distribuzione di ordine finito**, dove l'ordine è il più piccolo m per cui la stima è soddisfatta. Dunque, δ_0 è una distribuzione di ordine 0, mentre il valore principale è una distribuzione di ordine 1; man mano che l'ordine cresce, si può pensare che le distribuzioni sono sempre più irregolari.

Le distribuzioni di ordine 0 sono le cosiddette **misure**; questo significa che è possibile scriverle tutte nella forma:

$$u(\varphi) = \int_{\Omega} \varphi(x) d\mu$$

dove μ è una **misura** opportuna.

7.1.3 Derivata di una distribuzione

Una volta capito che le funzioni possono essere viste come distribuzioni, ha senso provare a riscrivere, in senso distribuzionale, alcune proprietà e operazioni che si possono applicare alle funzioni. Un esempio di operazione che si può fare è la **derivazione**: data u una funzione di classe $C^{(1)}$, è possibile derivarla, rispetto a una certa j -esima componente x_j :

$$\partial_j u = \frac{\partial u}{\partial x_j}$$

dunque:

$$\partial_j u \in C^{(0)}(\Omega)$$

a questo punto, ha senso studiare la derivata della distribuzione u :

$$T_{\partial_j u}(\varphi) = \int_{\Omega} \partial_j u \varphi$$

in questa rappresentazione, l'integrazione sarebbe nelle n variabili $(x_1, x_2, x_3, \dots, x_n)$ rispetto a cui u e φ possono variare; ciò è stato sottinteso. φ è a supporto compatto su $[-A, A] \times [-A, A] \times \dots \times [-A, A]$, ossia sul prodotto cartesiano di n intervalli $[-A, A]$. Si può capire come la derivata della distribuzione funziona, lavorando sulle funzioni test, applicando un procedimento di integrazione per parti; per esteso, si avrebbe:

$$\int_{-A}^A \int_{-A}^A \int_{-A}^A \dots \partial_j u \varphi dx_1 dx_2 \dots dx_n = \int_{-A}^A \int_{-A}^A \dots \left[- \int_{-A}^A u \partial_j \varphi dx_j + u \varphi \Big|_{-A}^A \right] dx_1 dx_2 \dots dx_n$$

questo si può dunque scrivere:

$$T_{\partial_j u}(\varphi) = \int_{\Omega} (\partial_j u) \varphi = - \int_{\Omega} u \partial_j \varphi = -T_u(\partial_j \varphi)$$

Dunque, se $u \in \mathcal{D}'(\Omega)$, si definisce:

$$\partial_j u(\varphi) = -u(\partial_j \varphi)$$

che, in generale, diventa:

$$\partial^\alpha u(\varphi) = (-1)^{|\alpha|} u(\partial^\alpha \varphi)$$

Un concetto fondamentale è il seguente: le distribuzioni si osservano attraverso dei funzionali lineari continui; questo è un integrale, una sorta di media, pesato mediante funzioni test. Una distribuzione può essere derivata tante volte quanto si vuole, dal momento che come spazio di funzioni test si considera C^∞ : questo è il motivo per cui si è scelto uno spazio così particolare.

In verità, si può fare un'osservazione: un sottoinsieme delle distribuzioni, come si è visto precedentemente, è dato dalle distribuzioni ottenute a partire dalle funzioni. Se u è una distribuzione, ma viene da una certa funzione di classe $C^{(1)}(\Omega)$, ossia:

$$u(\varphi) = \int v \varphi dx$$

ossia, $u = T_v$, allora:

$$\partial_j u = T_{\partial_j v}$$

Si dimostra con calma questa cosa:

$$\partial_j u(\varphi) = -u(\partial_j \varphi) = - \int v \partial_j \varphi \, dx = \int \partial_j v \varphi \, dx = T_{\partial_j v}(\varphi)$$

Esempio

Si consideri, in $\Omega = \mathbb{R}$, la distribuzione $\delta'(\varphi)$; questa è la derivata della δ di Dirac. Questa agisce nel seguente modo:

$$\delta'(\varphi) = -\delta(\varphi') = -\varphi'(0)$$

È possibile attribuire un significato fisico a questa distribuzione; dato un dipolo, dunque due cariche, questo è un sistema modellabile come limite della seguente situazione:

$$u_h = -\frac{1}{h}\delta_0 + \frac{1}{h}\delta_h$$

ossia, si può pensare a un dipolo come alla differenza di due delta di Dirac, una centrata in $x = h$, una centrata in $x = 0$; dunque, volendo vedere come questa distribuzione agisce su una funzione test, si può vedere:

$$u_h(\varphi) = \frac{\varphi(h) - \varphi(0)}{h}$$

a questo punto, il dipolo può essere pensato come questa situazione, in cui $h \rightarrow 0$; dunque:

$$\lim_{h \rightarrow 0} u_h(\varphi) = \lim_{h \rightarrow 0} \frac{\varphi(h) - \varphi(0)}{h} = -\delta'_0(\varphi)$$

7.1.4 Prodotto di una distribuzione per una funzione infinitamente derivabile

A questo punto si vuole discutere la distribuzione data dal prodotto di una distribuzione $u \in \mathcal{D}'(\Omega)$ per una funzione $f \in C^{(\infty)}(\Omega)$.

Al fine di introdurre l'argomento, si consideri una distribuzione T_u derivata a partire da una certa funzione $u \in C^{(0)}(\Omega)$; si ottiene:

$$(fu)(\varphi) = \int (fu)\varphi \, dx = \int u(f\varphi) \, dx$$

cosa succede: vale la proprietà associativa, di conseguenza è possibile sfruttarla, all'interno dell'integrale, per associare f a φ , invece che a u . Ora, si può definire dunque, per quanto riguarda una distribuzione $u \in \mathcal{D}'(\Omega)$, $f \in C_c^{(\infty)}(\Omega)$:

$$(fu)(\varphi) = u(f\varphi), \forall \varphi \in C_c^{(\infty)}$$

in altre parole: dal momento che φ è una funzione infinitamente derivabile a supporto compatto, se vi si moltiplica un'altra funzione infinitamente derivabile, allora si otterrà di nuovo una funzione infinitamente derivabile a supporto compatto, dove il supporto è quello della φ ; questo significa che la funzione sarà ancora una volta una funzione test. Per completare il discorso, e dimostrare che da ciò si ottiene effettivamente una distribuzione, sarebbe a questo punto verificare o mediante la definizione di distribuzione, o mediante la stima ad essa equivalente, il fatto che stiamo parlando di distribuzioni. Per esempio, si può procedere così: se $\varphi_j \rightarrow 0$ in $C_c^{(\infty)}(\Omega)$, allora certamente $fC_c^{(\infty)}(\Omega) \rightarrow 0$ in $C_c^{(\infty)}(\Omega)$; essendo per definizione:

$$(fu)(\varphi_j) = (fu)(\varphi) = u(f\varphi)$$

allora questo tende a 0 in \mathbb{C} .

In generale, non ha senso moltiplicare una distribuzione per una f che non sia infinitamente derivabile, dal momento che, alla fine, $f\varphi$ non sarebbe più una funzione test (non apparterebbe più allo spazio $\mathcal{D} = C_c^{(\infty)}(\Omega)$). Tuttavia, in alcune situazioni, questo può essere fatto: per esempio, distribuzioni come δ_0 , possono essere moltiplicate per una $f \in C^{(0)}(\Omega)$, la δ'_0 per una $f \in C^{(1)}(\Omega)$, e così via; si consideri il seguente esempio, sulla δ :

$$(f\delta_0)(\varphi) = \delta_0(f\varphi) = f(0)\varphi(0) = f(0)\delta_0(\varphi)(f(0)\delta_0)(\varphi)$$

nel penultimo passaggio si è sfruttata la definizione della δ_0 : $\varphi(0) = \delta_0(\varphi)$; nell'ultimo passaggio, si è usata la formula del prodotto di una funzione per una distribuzione.

Si osservi che **non** si devono confondere le due seguenti formule:

$$\delta_0(\varphi) = \varphi(0)$$

e

$$f\delta_0 = f(0)\delta_0$$

nel primo caso, si ha a membro destro un numero, ossia $\varphi(0)$, e nel membro sinistro una distribuzione applicata a una funzione test; nel secondo caso, si ha qualcosa di completamente diverso: qui si ha infatti una distribuzione a membro sinistro, e anche a membro destro; questo, dal momento che la δ si **estende** anche a $C^{(0)}(\Omega)$; **estende**, significa che, per **guardare** la δ , per studiarla, non è necessario per forza avere come spazio di funzioni test $C_c^{(\infty)}(\Omega)$, bensì è sufficiente $C_c^{(0)}(\Omega)$. Infatti, è possibile dimostrare che, anche con questo spazio di funzioni, la stima è soddisfatta:

$$|\delta(\varphi)| = |\varphi(0)| \leq \sup_{x \in \Omega} |\varphi(x)|$$

Dunque, alla fine della fiera, anche in questo modo si ottiene un funzionale lineare continuo: anche solo avendo $\varphi(x)$ continua! Si può parlare di estensione, dal momento che è possibile estendere lo spazio $C_c^{(\infty)}(\Omega)$ a $C_c^{(0)}(\Omega)$, che è più largo; non è possibile usare i teoremi di estensione precedentemente proposti, dal momento che in questo ambito non si ha a che fare con compatti, e neanche con spazi normati; tuttavia, esistono in letteratura teoremi che possono aiutarci.

L'ordine di derivabilità della funzione test coincide con m , ordine della distribuzione:

$$\varphi \in C_c^{(m)}(\Omega)$$

cosa significa ciò: è possibile estendere lo spazio delle funzioni test al minimo fino a $C_c^{(m)}(\Omega)$ (non si possono usare m minori dell'ordine della distribuzione). Questo ci aiuta a capire ancora meglio quale sia il significato dell'ordine di una distribuzione: maggiore è m , minore è la **regolarità** della distribuzione; questo dal momento che, per ottenere un funzionale lineare continuo, è necessario avere funzioni test con una regolarità maggiore, al fine di *compensare* l'irregolarità della distribuzione: distribuzioni molto regolari non hanno bisogno di funzioni test estremamente regolari, e viceversa in generale le distribuzioni hanno bisogno di funzioni test appartenenti a $C_c^{(\infty)}(\Omega)$, dal momento che, se non si conosce la regolarità della distribuzione, si deve supporre che essa sia pessima, e che dunque servano funzioni test infinitamente regolari per **pesare** le distribuzioni, al fine di ottenere, al termine di tutto, un funzionale lineare che sia **continuo**: questa è la base della teoria delle distribuzioni.

Impossibilità di definire il prodotto tra distribuzioni

Non è possibile definire il prodotto tra distribuzioni; questo, dal momento che un'ipotetica operazione di prodotto, non sarebbe commutativa e/o associativa.

Si supponga per assurdo di poter definire un prodotto tra due distribuzioni che sia associativo e commutativo; da questa ipotesi, è possibile ottenere un assurdo. Si consideri il seguente esempio, basato sull'uso di una $f(x) = x$:

$$(x\delta)(VP\frac{1}{x}) = \delta(xVP\frac{1}{x}) = \lim_{\varepsilon \rightarrow 0} \int x\frac{1}{x}\varphi dx = 1(\varphi)$$

$$(x\delta)VP\frac{1}{x} = x|_{x=0} \delta_0 VP\frac{1}{x} = 0$$

usando la proprietà associativa, si arriva a una contraddizione: due risultati diversi. Questo è un assurdo.

7.1.5 Convergenza di distribuzioni

Data u_j una successione di distribuzioni, si vuole definire la convergenza di questa a una distribuzione $u \in \mathcal{D}'(\Omega)$; in questo contesto, la convergenza si ha se si ha convergenza **su ogni funzione test fissata**:

$$u_j(\varphi) \rightarrow u(\varphi) \text{ in } \mathbb{C}, \forall \varphi \in C_c^{(\infty)}(\Omega)$$

Questa convergenza in ambito distribuzionale viene chiamata o semplicemente convergenza, o **convergenza debole**; tuttavia, nell'ambito degli spazi duali, questa definizione ricorda molto la convergenza debole*: quella che si studia attraverso i funzionali.

Esempio

Si consideri il seguente esempio, per chiarire il concetto di convergenza; si consideri $\psi \in C_c^{(\infty)}(\Omega)$, $\psi \geq 0$, tale per cui:

$$\int \psi dx = 1$$

si applichi, a ciò, una coppia di dilatazioni tali da preservare l'integrale:

$$\psi_\varepsilon(x) = \frac{1}{\varepsilon^n} \psi\left(\frac{x}{\varepsilon}\right) \text{ in } \mathbb{R}^n$$

il termine ε^{-n} viene introdotto al fine di compensare lo jacobiano di \mathbb{R}^n . Ricordando l'inizio della trattazione, $\psi_\varepsilon(x)$ soddisfa la proprietà dell'identità approssimata; di conseguenza, applicando il lemma dell'identità approssimata, è possibile dire che:

$$\int \varphi \psi_\varepsilon \, dx \xrightarrow{\varepsilon \rightarrow 0} \varphi(0)$$

Tuttavia, si può riconoscere che $\varphi(0) = \delta_0(\varphi)$; questo significa che, se la relazione vale $\forall \varphi$ continua e a supporto compatto in \mathbb{R}^n , $\psi_\varepsilon \rightarrow \delta_0 \in \mathcal{D}'(\mathbb{R}^n)$; in altre parole, dal momento che l'azione di ciascuna distribuzione ψ_ε su funzioni test φ fissate tende all'azione di δ_0 sulle stesse φ , allora si ha che $\psi_\varepsilon \rightarrow \delta_0$. Spesso in Fisica si usa questo tipo di definizione per introdurre la delta di Dirac, per l'efficacia didattica; tuttavia, questo è molto difficile da formalizzare.

Esempio ulteriore

Si vuole provare a studiare un secondo esempio molto interessante: data u_n :

$$u_n = \sin(nx)$$

ammette limite, in senso distribuzionale? Ossia, considerando:

$$u_n(\varphi) \rightarrow u(\varphi) \quad \forall \varphi$$

questa $u \in \mathcal{D}'$, esiste? La risposta è sì, e lo dimostreremo nell'immediato seguito: dato $[-A, A]$ il supporto della funzione test,

$$\begin{aligned} u_n(\varphi) &= \int_{-A}^{+A} \sin(nx) \varphi(x) \, dx = - \frac{\cos(nx) \varphi(x)}{n} \Big|_{-A}^{+A} - \int_{-A}^{+A} \varphi'(x) \int \sin(x) \, dx = \\ &= 0 - \frac{1}{n} \int_{-A}^{+A} \cos(nx) \varphi'(x) \, dx \leq \frac{1}{n} \sup |\cos(nx)| \int |\varphi'| \, dx = \frac{1}{n} \int |\varphi'| \, dx \end{aligned}$$

essendo φ derivabile, dunque, è possibile dire che quell'integrale è uguale a una costante finita, e dunque, per $n \rightarrow \infty$, $u_n(\varphi) \rightarrow 0$: questa distribuzione converge a 0.

7.1.6 Supporto di una distribuzione

Come fatto precedentemente, al fine di fissare il concetto e capire da dove esso derivi, una buona idea è studiare il caso in cui u è semplicemente una funzione continua, per poi estendere al caso di u distribuzione. Si consideri dunque, in un primo momento, $u \in C^{(0)}(\Omega)$; in questo contesto, è possibile definire il supporto di u come:

$$\text{supp } \{u\} = \overline{\{x \in \Omega : u(x) \neq 0\}} \text{ in } \Omega$$

si ricordi e si specifichi sempre che la chiusura non è rispetto a \mathbb{R}^n , bensì rispetto a Ω . Questa definizione va molto bene per le funzioni continue, ma non è altrettanto efficace per quanto riguarda funzioni non continue, come quelle integrabili. In questa definizione, si definisce ciò che sta nel supporto; un modo alternativo di procedere, è definire ciò che **non** sta nel supporto: dato $x_0 \in \Omega$, ma $x_0 \notin \text{supp } \{u\}$, si ha che la funzione u in x_0 sarà certamente nulla; si ha in verità però qualcosa in più: dal momento che il supporto è stato definito come chiusura rispetto a Ω , il complementare a Ω del supporto sarà certamente aperto in Ω (e non solo: anche aperto in \mathbb{R}^n , essendo Ω un sottoinsieme aperto di un aperto, e dal momento che stiamo considerando un sottoinsieme aperto di Ω); questo significa che ogni x_0 appartenente al complementare del supporto certamente è un punto interno di questo stesso insieme (definizione di aperto: ogni punto in un aperto è un punto interno, ossia ha una distanza di sicurezza dal bordo, ossia ogni punto ha un intorno interamente contenuto nell'insieme in questione). Dal momento che tutto ciò è vero, per ogni x_0 nel supplementare, vi sarà un intorno di x_0 in cui u si dovrà annullare.

Passiamo alle distribuzioni: dal momento che le distribuzioni sono definite a partire da funzioni test, a loro volte definite a partire da un certo aperto, si dice che una distribuzione $u \in \mathcal{D}'(\Omega)$ si annulla in un aperto $\omega \subset \Omega$, se $u(\varphi) = 0, \forall \varphi \in C_c^{(\infty)}(\omega)$. Dunque, in generale, dato $\omega \subset \Omega$, si può definire la **restrizione** di una distribuzione:

$$u|_{\omega}(\varphi) = (u)(\varphi) \forall \varphi \in C_c^{(\infty)}(\omega)$$

restringere una distribuzione in ω significa **osservarla**, mediante il solito funzionale, pesandola mediante funzioni test definite in sottoinsiemi compatti di ω , invece che di Ω . Si osservi che **non ha senso** parlare di distribuzioni nulle in un chiuso, o in particolare in un punto: i nostri funzionali sono

definiti solo grazie a funzioni test, che sono definite a partire da un qualche aperto; la distribuzione si deve poter annullare anche su un intorno, e non solo localmente su un punto, dunque, se si lavora con insiemi chiusi, non si può garantire l'esistenza di questo intorno: in questo modo, non staremmo estendendo correttamente quanto visto in forma preliminare sulle funzioni continue, all'inizio di questa sottosezione.

Una volta chiarito cosa significa che una distribuzione $u \in \mathcal{D}'(\Omega)$ si annulla, ossia una volta detto che $x_0 \notin \text{supp}\{u\}$ se esiste un intorno aperto ω_i di x_0 tale per cui $u = 0$ in ω , ossia:

$$u(\varphi) = 0 \forall \varphi \in C_c^{(\infty)}(\omega_i)$$

dire che la funzione non appartiene al supporto, coincide con dire che essa appartiene al complementare del supporto a Ω , che si può definire come segue:

$$\Omega \setminus \text{supp}\{u\} = \bigcup_i \omega_i$$

dove gli ω_i sono gli aperti in cui u si annulla.

Esempio

Si può vedere che, in \mathbb{R} ,

$$\text{supp}\{\delta_0\} = \{0\}$$

infatti, i punti diversi dell'origine **non stanno nel supporto**; dato $x_0 \neq 0$, esiste sicuramente un intorno aperto ω in cui non si incrocia lo 0. Questo accade dal momento che:

$$\delta(\varphi) = \varphi(0), \forall \varphi \in C_c^{(\infty)}(\omega)$$

ma ω non contiene $x = 0$, che è l'unico punto in cui la delta valuta la funzione test; di conseguenza, dal momento che le varie φ hanno supporto compatto in ω (ossia, sono non nulle solo per un insieme di punti interni a ω , che non interseca $x = 0$), allora la $\varphi(0)$ sarà sempre uguale a 0, per $x_0 \neq 0$, con un intorno ω di x_0 che non intersechi $x = 0$.

Osservazione conclusiva

Si consideri uno zero isolato di una funzione, x_0 ; in questo caso, $x_0 \in \text{supp}\{u\}$; in altre parole, gli zeri isolati appartengono al supporto di una distribuzione.

7.1.7 Distribuzioni a supporto compatto

La definizione di supporto è molto utile, dal momento che ci permette di introdurre un'ulteriore classificazione delle distribuzioni; nel dettaglio, una volta definito il supporto delle distribuzioni, è possibile identificare le distribuzioni a supporto compatto; il grosso vantaggio di queste distribuzioni risiede nel fatto che, con esse, è possibile utilizzare uno spazio di funzioni test differente da quello finora adottato.

Sia $\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega)$ il sottospazio delle distribuzioni a supporto compatto in Ω ; questo tipo di notazione deriva dal fatto che Laurent Schwartz chiamò $\mathcal{D}(\Omega) = C_c^{(\infty)}(\Omega)$ e $\mathcal{E}(\Omega) = C^{(\infty)}(\Omega)$; è possibile vedere che le distribuzioni a supporto compatto si identificano mediante funzionali lineari continui di funzioni $C^{(\infty)}$; per questo motivo, \mathcal{E}' si può pensare come il duale di \mathcal{E} .

Data dunque u una distribuzione a supporto compatto, ossia $u \in \mathcal{E}'(\Omega)$, allora è possibile definire $u(\varphi)$, dove $\varphi \in \mathcal{E}(\Omega)$: in questo contesto, dunque, φ non deve più essere necessariamente a supporto compatto.

Al fine di fissare le idee, si pensi ancora una volta al caso delle funzioni continue; si immagini che u sia una funzione continua, ma a supporto compatto (che dunque si annulla al di fuori dell'intervallo $[-A, A]$), e che si abbia una φ , funzione infinitamente derivabile ma a supporto non compatto; volendo calcolare:

$$\int_{-\infty}^{+\infty} u\varphi \, dx$$

quello che si fa “in pratica” è moltiplicar φ per una funzione caratteristica $\chi_{[-A, A]}$, la quale è una “finestra” che vale 1 nell'intervallo $[-A, A]$ e 0 altrove. Il risultato finale sarà:

$$\int_{-\infty}^{+\infty} u\varphi\chi_{[-A, A]} \, dx = \int_{-A}^{+A} u\varphi \, dx$$

Ossia, la finestra limita in maniera naturale l'intervallo di integrazione alla regione in cui la funzione u è compatta.

Se si ha a che fare con $u \in \mathcal{E}'(\Omega)$, la situazione è molto diversa: le distribuzioni sono molto irregolari, ma al fine di studiarle, è necessario avere a che fare con funzionali lineari continui; utilizzare esattamente lo stesso trucco in questo ambito non è possibile, dal momento che, moltiplicando la funzione test per una funzione come la finestra (che non è nemmeno continua), non si otterrebbe più una funzione test; questo significa che essa potrebbe non essere sufficientemente regolare da ripristinare la regolarità all'interno dell'integrale. L'idea di fondo in verità è valida, a patto di tenere in conto il seguente accorgimento: invece di utilizzare la classica funzione caratteristica, è necessario utilizzare una χ che si raccordi in maniera regolare con l'asse delle x :

$$\chi \in C^{(\infty)}(\Omega)$$

questa χ è tale per cui $\chi = 1$ in un intorno del supporto della distribuzione u : $\chi \in C_c^{(\infty)}(\Omega)$. Si definisce dunque $u(\varphi)$, nel caso di $u \in \mathcal{E}'(\Omega)$, $\varphi \in \mathcal{E}(\Omega)$, come:

$$u(\varphi) = u(\chi\varphi), \forall \varphi \in C^{(\infty)}(\Omega)$$

il risultato sarà:

$$\chi\varphi \in C_c^{(\infty)}(\Omega)$$

e dunque, intuitivamente, è stato fatto esattamente ciò che è stato fatto nel caso di u funzione continua, però tenendo conto dell'accorgimento in questione. Si osservi che questa definizione è indipendente da χ : se al posto di χ si prendesse una $\tilde{\chi}$ (sempre appartenente a $C_c^{(\infty)}(\Omega)$ e con le stesse proprietà), allora il risultato è lo stesso:

$$u(\chi\varphi) = u(\tilde{\chi}\varphi)$$

infatti:

$$u((\chi - \tilde{\chi})\varphi) = 0$$

dal momento che, se χ e $\tilde{\chi}$ sono uguali a 1 nell'intorno del supporto della distribuzione u , allora, in un intorno del supporto di u , si ha che $\chi - \tilde{\chi} = 0$ di conseguenza:

$$\text{supp } \{(\chi - \tilde{\chi})\varphi\} \subset (\Omega \setminus \text{supp } \{u\})$$

Tutto ciò può essere risolto con il metodo della **partizione dell'unità**: la φ , che è una funzione che può avere un supporto anche molto grande, può essere scritta come una somma finita di φ_j , ciascuna delle quali ha un supporto piccolo; utilizzando la compattezza, è possibile identificare dei ricoprimenti che ammetteranno sottoricoprimenti finiti.

In generale, è possibile dire qualcosa in più, sulla definizione di distribuzioni mediante spazi di funzioni test particolari: se

$$\text{supp } \{u\} \cap \text{supp } \{\varphi\}$$

è compatto, $\forall \varphi \in \mathcal{D}$, allora è possibile definire la distribuzione $u(\varphi)$, anche con $u \in \mathcal{D}'$.

7.2 Convoluzione

7.2.1 Convoluzione di funzioni

Si considerino $u, v \in C^{(0)}(\mathbb{R}^n)$; si vuole definire, su \mathbb{R}^n , la convoluzione. Questa operazione non si può definire su Ω , ossia su un generico aperto di \mathbb{R}^n , dal momento che non è detto che la convoluzione “rimanga” dentro allo spazio: essendo un'operazione che estende il supporto della funzione, non si ha la garanzia di stare dentro a Ω (questa cosa verrà discussa in seguito). Si supponga inoltre che una tra u e v sia a supporto compatto; fatte tutte queste premesse, si può definire la convoluzione come:

$$(u * v)(x) = \int_{\mathbb{R}^n} u(x - y)v(y) \, dy$$

Questo integrale sarà su un insieme certamente limitato: si supponga che v sia la funzione a supporto compatto; in questo caso, y , variabile di integrazione, non può variare in un insieme eccessivamente grande; di conseguenza, l'integrale sarà su un dominio certamente limitato.

Per discutere i risultati concernenti la convoluzione, merita discutere il seguente risultato: data $f(x, y)$ continua su $[a, b] \times [c, d]$, definita:

$$F(x) = \int_c^d f(x, y) \, dy$$

allora, $F(x)$ è certamente continua su $[a, b]$. Non solo: se esiste

$$\frac{\partial f}{\partial x}$$

continua, su $[a, b] \times [c, d]$, allora F è derivabile e la derivata si ottiene portandola sotto il segno di integrale:

$$F'(x) = \int_c^d \frac{\partial f}{\partial x}(x, y) dy$$

Questo risultato sarà utile nei prossimi conti.

Un primo risultato concernente invece la convoluzione è il seguente: la convoluzione è un'operazione commutativa, dal momento che, facendo un cambio di variabili, è possibile ottenere:

$$(v * u)(x) = \int u(y)v(x - y)dy = \int u(x - y)v(y)dy = (u * v)(x)$$

La convoluzione è anche associativa: date tre funzioni u, v, w , delle quali almeno due sono a supporto compatto¹.

$$(u * v) * w = w * (v * w)$$

7.2.2 Proprietà del supporto

È possibile dimostrare che il supporto della convoluzione è contenuto nella somma dei supporti:

$$\text{supp}\{u * v\} \subset \text{supp}\{u\} + \text{supp}\{v\}$$

è necessario specificare cosa si intende per ciò che sta scritto a membro destro:

$$\text{supp}\{u\} + \text{supp}\{v\} = \{x + y \in \mathbb{R}^n : x \in \text{supp}\{u\}, y \in \text{supp}\{v\}\}$$

Ossia, dato un punto x appartenente al supporto di u , un punto y appartenente al supporto di v , in $u * v$, questo punto sarà mandato nella somma

¹in caso contrario, si rischierebbe di dover calcolare un integrale di convoluzione tra due funzioni, delle quali nessuna sarebbe a supporto compatto, e dunque la convoluzione non potrebbe essere definita

vettoriale dei due supporti; questa è la somma nel senso della regola del parallelogramma.

In questo caso, si ha un'ulteriore particolarità: la somma di due chiusi, in generale, non è un chiuso; in questo caso, però, lo è, dal momento che un supporto sarà chiuso, ma l'altro sarà compatto, e “chiuso + compatto=chiuso”.

7.2.3 Proprietà di regolarità

Se invece che la situazione analizzata precedentemente, si ha che $u \in C^1(\mathbb{R}^n)$, $v \in C^0(\mathbb{R}^n)$, con una delle due a supporto compatto, si può dire qualcosa di più: la convoluzione, alla fine, sarà di classe $C^1(\mathbb{R}^n)$. Questo si vede utilizzando il secondo risultato precedentemente proposto per quanto concerne la funzione integrale $F(x)$: essendo x la variabile in cui varia il risultato della convoluzione, e y semplicemente un parametro di integrazione, applicando il risultato appena citato.

Questo ci suggerisce che la convoluzione eredita le proprietà “della funzione più bella”, nel senso di quella più regolare; infatti, grazie alla commutatività della convoluzione, questo vale se una a caso delle due funzioni è di classe $C^1(\mathbb{R}^n)$.

In verità, si ha addirittura qualcosa di più: non solo si “ereditano le proprietà della funzione più bella”, ma, se in generale $u \in C^k(\mathbb{R}^n)$, $v \in C^j(\mathbb{R}^n)$, allora:

$$(u * v)(x) \in C^{(j+k)}(\mathbb{R}^n)$$

e:

$$\partial^\alpha(u * v) = (\partial^\alpha u) * v$$

e

$$\partial^\alpha(u * v) = u * (\partial^\alpha v)$$

7.2.4 Convoluzione tra una distribuzione e una funzione

Al solito, ispirandoci a quanto fatto precedentemente, si estende ora al caso in cui $u \in \mathcal{D}'(\mathbb{R}^n)$, $\psi \in C_c^{(\infty)}(\mathbb{R}^n)$; in questo caso, l'integrale di convoluzione,

può essere interpretato come una distribuzione, dove ψ è trattato come una funzione test:

$$(u * \psi)(x) = u(\psi(x - \cdot))$$

il \cdot sarebbe la variabile di integrazione: il risultato della convoluzione varierà in x , esattamente come visto precedentemente, e però è necessario integrare in una qualche variabile, che sarà indicata con \cdot ; in altre parole, nel caso delle funzioni, $\cdot = y$. Il risultato di questa operazione è una funzione appartenente a $C_c^{(\infty)}(\mathbb{R}^n)$.

Questo è vero se ψ è a supporto compatto; in verità, il risultato è valido anche se è la distribuzione u a essere a supporto compatto, e $\psi \in C^{(\infty)}(\mathbb{R}^n)$; in questa situazione, infatti:

$$(u * v)(x) = u(\psi(x - \cdot))$$

infatti qui la cosa ha senso dal momento che, essendo x fissato, essendo la u a supporto compatto, allora $\psi(x - \cdot)$ è ancora una funzione appartenente a $C^{(\infty)}(\mathbb{R}^n)$ (infatti, è solo specchiata e traslata, ma dunque rimane regolare).

Nota sulla regolarità

Si consideri la seguente situazione: $\psi \in C_c^{(\infty)}(\mathbb{R}^n)$, $\psi \geq 0$, $\int \psi dx = 1$; si definisca dunque, analogamente a quanto fatto precedentemente:

$$\psi_\varepsilon(x) = \frac{1}{\varepsilon^n} \psi\left(\frac{x}{\varepsilon}\right)$$

allora, data $u \in \mathcal{D}'(\mathbb{R}^n)$, **ha senso calcolare:**

$$(u * \psi_\varepsilon)(x)$$

infatti, ψ_ε è una funzione test. Il risultato di ciò, sarà una funzione $C^{(\infty)}$, che converge a una $u \in \mathcal{D}'$. Ogni ψ_ε , per $\varepsilon \rightarrow 0$, tende a una delta di Dirac; dal momento che la delta di Dirac è l'elemento neutro della convoluzione, è possibile usare questa idea. Questa idea è molto importante dal momento che si può usare per regolarizzare le distribuzioni, facendo le convoluzioni con identità approssimate del tipo di ψ_ε , e quello che si ottiene come risultato è che la classe delle funzioni $C^{(\infty)}$ è densa nell'insieme delle distribuzioni; questo apre la strada alle dimostrazioni di vari risultati sulle distribuzioni.

7.2.5 Convoluzione tra due distribuzioni

È possibile definire la convoluzione anche tra due distribuzioni; al fine di fare ciò, come al solito, si considerino u, v due funzioni continue; inoltre, si consideri v a supporto compatto (una qualsiasi tra u e v andrebbe bene). Dunque, date $u, v \in C^{(0)}(\mathbb{R}^n)$,

$$(u * v)(x) = \int u(y)v(x - y) dy$$

l'azione di ciò su una funzione test è:

$$\int (u * v)(x)\varphi(x) dx = \int \int u(y)v(x - y)dy\varphi(x)dx$$

due di queste tre funzioni sono a supporto compatto (anche le φ lo sono); di conseguenza, per Fubini, è possibile scambiare gli integrali; inoltre, è possibile fare un cambio di variabili: $z = x - y$, ottenendo:

$$= \int \int u(y)v(z)\varphi(z + y)dzdy \xrightarrow{x \leftrightarrow z} \int \int u(y)v(x)\varphi(x + y)dx dy$$

di conseguenza, questo si può leggere come un integrale iterato: si può vedere $v(x)$ come una distribuzione, e applicarla alla funzione test $\varphi(x + y)$; dunque prendere il risultato di questa prima applicazione, trattando ciò da funzione test, applicandovi u .

Si proceda a questo punto definendo la convoluzione tra distribuzioni: se $u \in \mathcal{D}'(\mathbb{R}^n)$, $v \in \mathcal{E}'(\mathbb{R}^n)$, allora:

$$(u * v)(\varphi) = u_y(v_x(\varphi(x + y)))$$

ciò ha significato. Infatti, data $\varphi(x + y)$, si incomincia applicando v a $\varphi(x + y)$; nel dettaglio, v si comporterà come una v_x , dal momento che la distribuzione v agisce sulla funzione test coinvolgendo esclusivamente la variabile x ; y sarà un parametro esterno a ciò. Si ottiene dunque $v_x(\varphi(x + y))(y)$: questa è una funzione che appartiene a $C^{(\infty)}(\mathbb{R}^n)$, e che varia nella variabile y . Dal momento che φ è a supporto compatto, così come la distribuzione v , vale la formula del supporto, che garantisce che il supporto della convoluzione è un sottoinsieme della somma dei due supporti; di conseguenza, essendo la somma di due supporti compatti ancora compatta, questa funzione non solo appartiene a $C^{(\infty)}(\mathbb{R}^n)$, ma anche a :

$$v_x(\varphi(x+y))(y) \in C_c^{(\infty)}(\mathbb{R}^n)$$

Quindi, questa è una funzione test; di conseguenza, applicarvi u_y , ossia la distribuzione u che agisce solo sulla variabile y , ha senso.

La stessa definizione può essere applicata se è u ad appartenere a $\mathcal{E}'(\mathbb{R}^n)$, ossia a essere a supporto compatto, dal momento che, anche se $v \in \mathcal{D}'(\mathbb{R}^n)$, e dunque

$$v_x(\varphi(x+y))(y) \in C^{(\infty)}(\mathbb{R}^n)$$

la compattezza di u mette le cose a posto, permettendo di definire correttamente il tutto.

Esempio 1

Si consideri $u \in \mathcal{D}'(\mathbb{R}^n)$, e $\delta_0^{(x)}$ la delta di Dirac che agisce sulla variabile x , centrata in $x = 0$; si ha:

$$(u * \delta)(\varphi) = u_y(\delta(\varphi(x+y))) = u_y(\varphi(y)) = u(\varphi)$$

infatti, nel terzo passaggio, la delta agisce valutando x in $x = 0$. Ciò che è stato appena dimostrato in questo punto è il fatto che la δ potrebbe essere vista come l'identità in una qualche algebra. È proprio così? La risposta è no: in \mathcal{D}' , almeno, non è così, dal momento che \mathcal{D}' non è un'algebra; il discorso cambia se si considera \mathcal{E}' : in questo caso effettivamente è un'algebra, e δ rappresenta l'elemento neutro della convoluzione.

Un'osservazione: se $u, v \in \mathcal{D}'(\mathbb{R})$, con supporti tali per cui:

$$\text{supp } \{u\}, \text{supp } \{v\} \subset [0, +\infty)$$

allora, la definizione di prima va comunque bene; in questo caso, si identifica questo spazio di distribuzioni come $\mathcal{D}'_+(\mathbb{R}^n)$; anche questo spazio è un'algebra di convoluzione, dal momento che la convoluzione di due elementi appartenenti a questo spazio rimane ancora in esso; anche in questo contesto, la δ è l'elemento neutro.

Esempio 2

In \mathbb{R} , si consideri:

$$(u * \delta_0^{(x)})'(\varphi) = u_y(\delta'(\varphi(x+y))) = -u_y(\varphi'(y)) = u'(\varphi)$$

quindi, la convoluzione della derivata della δ coincide con la derivata della distribuzione u ; in generale:

$$u * (\partial^\alpha \delta) = \partial^\alpha u$$

quindi, questo ci dice che gli operatori di derivazione possono essere scritti come convoluzioni di derivate di δ .

7.2.6 Proprietà della convoluzione

Verranno a questo punto enunciate alcune proprietà della convoluzione.

1. Se $u, v \in \mathcal{D}'(\mathbb{R}^n)$, con una delle due a supporto compatto, allora:

$$u * v = v * u$$

2. Se $u, v, w \in \mathcal{D}'(\mathbb{R}^n)$, due delle quali sono a supporto compatto, allora:

$$(u * v) * w = u * (v * w)$$

È possibile dimostrare in diversi modi queste proprietà; un modo, è direttamente; un altro modo, è mediante approssimazione, sfruttando il concetto precedentemente enunciato, per cui è possibile vedere una distribuzione come limite del prodotto di una distribuzione per l'identità approssimata; di conseguenza, è possibile dire che:

- se $u_j \in \mathcal{D}'(\mathbb{R}^n)$, $u_j \rightarrow u$ in $\mathcal{D}'(\mathbb{R}^n)$;
- se $v_j \in \mathcal{E}'(\mathbb{R}^n)$, $v_j \rightarrow v$ in $\mathcal{E}'(\mathbb{R}^n)$

allora:

$$u_j * v_j \rightarrow u * v \text{ in } \mathcal{D}'(\mathbb{R}^n)$$

7.2.7 Applicazione: ruolo della convoluzione nelle PDE

Verrà a questo punto proposto un esempio applicativo di convoluzione, nell'ambito delle equazioni alle derivate parziali. Dato un generico operatore P differenziale, nella forma:

$$P = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha$$

come, per esempio, l'operatore di Laplace, che si può scrivere come:

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \dots + \frac{\partial^2}{\partial x_n^2}$$

i c_α sono delle costanti complesse: $c_\alpha \in \mathbb{C}$. Siamo interessati all'equazione $Pu = f$, dove f è una funzione nota; questa si può scrivere come:

$$\sum_{|\alpha| \leq m} c_\alpha \partial^\alpha u = f$$

Ha un ruolo fondamentale il caso in cui $f = \delta$: le E che soddisfano ciò sono dette **soluzioni fondamentali delle equazioni**. Una soluzione fondamentale di P è una distribuzione $E \in \mathcal{D}'(\mathbb{R}^n)$ tale per cui:

$$P(E) = \delta$$

Queste sono molto importanti dal momento che, una volta che le soluzioni fondamentali E sono note, allora, data $f \in \mathcal{E}'(\mathbb{R}^n)$, si può calcolare u come:

$$u = E * f$$

questa, infatti, è soluzione dell'equazione $Pu = f$.

Dimostriamo questo fatto:

$$Pu = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha (E * f) = \sum_{|\alpha| \leq m} c_\alpha (\partial^\alpha E) * f = f$$

infatti, è noto che $PE = \delta$, e che $\delta * f = f$.

Un'osservazione: se u_0 è la soluzione trovata con la convoluzione f , tutte le soluzioni sono date da:

$$u = u_0 + v$$

dove v è una soluzione omogenea del problema, ossia una soluzione tale per cui $Pv = 0$. Questo è ovvio: si può dire che v sia parte del nucleo di P , essendo una soluzione omogenea, di conseguenza se si calcola $P(u)$, dove $u = u_0 + v$, si ha:

$$P(u) = P(u_0 + v) = P(u_0) + P(v) = P(u_0)$$

Esempio teorico/applicativo

Si consideri, come esempio, $P = \Delta$, in \mathbb{R}^n . Si vuole risolvere $\Delta E = \delta$; si può dimostrare che:

$$E(x) = \begin{cases} -\frac{1}{2\pi} \log |x|, & n = 2 \\ \frac{1}{c_n(2-n)} \frac{1}{|x|^{n-2}}, & n \neq 2 \end{cases}$$

dove $x = (x_1, x_2, \dots, x_n)$, e c_n è la misura della sfera $(n - 1)$ -dimensionale in \mathbb{R}^n .

$E(x)$ è una funzione localmente integrabile, di conseguenza essa può essere vista anche come distribuzione: $E \in L^1_{\text{loc}}(\mathbb{R}^n)$; l'unica singolarità di questa infatti è in $x = 0$, e, dato l'integrale in $[-1, 1]$ (un esempio di intervallo che comprenda la singolarità), per n generico:

$$\int \frac{1}{|x|^\alpha} dx \begin{cases} \text{converge,} & \alpha < n \\ \text{diverge,} & \alpha \geq n \end{cases}$$

converge, significa che il volume proiettato sul compatto è finito (definizione di "locale integrabilità"). Data una soluzione non omogenea u dell'equazione:

$$\Delta u = f$$

dove $f \in \mathcal{E}'(\mathbb{R}^n)$, una soluzione sarà data da:

$$u = E * f$$

Si hanno delle proprietà buone: Δ è un operatore ellittico, dunque le soluzioni dell'equazione omogenea sono sempre in $C^{(\infty)}$; di conseguenza, queste sono soluzioni anche in senso classico, e non solo in senso distribuzionale. Le soluzioni dell'equazione $\Delta E = 0$ sono dette **funzioni armoniche**; queste sono interessanti dal momento che, data f a supporto compatto essa, fuori dal compatto, vale 0; le funzioni armoniche sono le uniche soluzioni di questo

problema che decadono a 0 all'infinito; di conseguenza, le soluzioni ottenute partendo dalla convoluzione di funzioni armoniche, sono le uniche che abbiano fisicamente senso, grazie a questa proprietà.

Il fatto che l'operatore sia ellittico è molto importante, dal momento che le soluzioni sono $C^{(\text{inf})}$ grazie al fatto che, nel caso della classe degli operatori ellittici, le *singolarità delle distribuzioni vengono eliminate*; questo non capita negli operatori **ipoellittici**.

Per applicare ciò, si consideri f come una funzione che rappresenta una densità di carica disposta su una certa superficie Σ ; si ha dunque:

$$f(\varphi) = \int_{\Sigma} \rho \varphi \, d\sigma$$

questa è detta **densità di strato semplice**; in questo caso, ρ è per esempio continua su Σ . In maniera abbastanza simile, è possibile definire la **densità di strato doppia**:

$$f(\varphi) = \int_{\Sigma} \rho \frac{\partial \varphi}{\partial \nu} \, d\sigma$$

quest'ultima rappresenta una sorta di densità di dipoli disposti su una superficie. Si consideri il caso $n = 3$, dunque \mathbb{R}^3 : se f è come nel caso della densità di strato singola, allora si dimostra che:

$$u = E * f(x) = -\frac{1}{4\pi} \int \frac{\rho(y)}{|x - y|} \, d\sigma$$

questa, si può dimostrare, è una funzione $C^{(\infty)}$, ma nulla al di fuori della superficie Σ ; 4π sarebbe il c_3 (la misura della sfera 2-dimensionale). Questa espressione è nota in Fisica, e viene di solito ricavata con altri ragionamenti.

Qualcosa di simile si può ottenere anche per quanto concerne le densità di strato doppie:

$$E * f(x) = -\frac{1}{4\pi} \int_{\Sigma} \rho(y) \frac{\partial}{\partial \nu} \frac{1}{|x - y|} \, d\sigma$$

7.3 Distribuzioni temperate

A questo punto si vuole introdurre una sottoclasse delle distribuzioni molto significativa, dal momento che permette di introdurre agevolmente il concetto

di trasformata di Fourier. Per introdurre queste, è necessario introdurre una nuova classe di funzioni test, detta **classe di Schwartz**. Per fare ciò, si definisce lo spazio di Schwartz, $S(\mathbb{R}^n)$, come le funzioni $f \in C^{(\infty)}(\mathbb{R}^n)$ tali per cui:

$$\forall \alpha, \beta \in \mathbb{N}^n, \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty$$

dove dunque $x = (x_1, x_2, \dots, x_n)$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, e:

$$x^\alpha = x_1^{\alpha_1} \times x_2^{\alpha_2} \times \dots \times x_n^{\alpha_n}$$

(ossia, è dato dal prodotto (di numeri) $x_i^{\alpha_i}$, dove i sta a indicare la i -esima componente del vettore x in \mathbb{R}^n).

Questo significa che la funzione e le sue derivate, ancora moltiplicate per dei polinomi, devono essere limitate.

Si considerino alcuni esempi.

1. Il primo esempio è:

$$f(x) = e^{-x^2} \in S(\mathbb{R}^n)$$

infatti, per induzione, si può vedere che ciò è sempre limitato.

2. Il secondo esempio è:

$$f(x) = e^{-x^2} \cos(e^{x^4})$$

in questo caso la funzione è limitata, ma le derivate no, dal momento che continuano a crescere; questo perché il termine x^4 compensa ciò che sta fuori.

3. Il terzo esempio garantisce che:

$$C_0^{(\infty)}(\mathbb{R}^n) \subset S(\mathbb{R}^n)$$

ovvio: se da un certo punto in poi sono nulle (essendo a supporto compatto), queste funzioni saranno anche in S .

Da ora in poi, $S(\mathbb{R}^n)$ sarà il nostro nuovo spazio di funzioni test; dunque, lo spazio delle distribuzioni temperate sarà il duale di questo spazio; dal momento che questo spazio di funzioni test è più grande di quello di prima, il suo duale sarà più piccolo (racchiuderà meno funzioni).

S ha una struttura lineare, ma non è uno spazio normato; su di esso si potrebbe definire una topologia, a partire dalla quale si può estrarre la seguente definizione di convergenza, che viene ora fornita in forma di definizione.

Si dice che una $f_j \in S(\mathbb{R}^n)$ converge a una $f \in S(\mathbb{R}^n)$ in $S(\mathbb{R}^n)$ se e solo se

$$\forall \alpha, \beta \in \mathbb{N}^n, x^\alpha \partial^\beta f_j \rightarrow x^\alpha \partial^\beta f$$

uniformemente su \mathbb{R}^n ; in altre parole:

$$\|x^\alpha \partial^\beta f_j - x^\alpha \partial^\beta f\|_{L^\infty(\mathbb{R}^n)} \rightarrow 0, \text{ per } j \rightarrow \infty$$

Questo significa che non solo le f_j , ma anche le derivate dovranno tendere a 0. In questo caso, non si ha uno spazio normato, ma qualcosa in meno: questo spazio è di Frechet, nel senso che, invece che una norma, si ha una famiglia di seminorme, ossia le:

$$\sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)|$$

Questa nozione di convergenza è più debole di quella in $C_c^{(\infty)}(\mathbb{R}^n)$; infatti, se $f_j \rightarrow f$ in $C_c^{(\infty)}$, allora questo vale anche su $S(\mathbb{R}^n)$, ma non il viceversa.

A questo punto, è possibile introdurre la definizione di distribuzione temperata.

Una distribuzione temperata è un funzionale lineare da $S(\mathbb{R}^n) \rightarrow \mathbb{C}$ continuo nel senso che, se φ_j è una successione di funzioni di Schwartz in $S(\mathbb{R}^n)$, $\varphi \in S(\mathbb{R}^n)$, allora, se $\varphi_j \rightarrow \varphi$, si ha:

$$u(\varphi_j) \rightarrow u(\varphi) \text{ in } \mathbb{C}$$

Equivalentemente, questo si può scrivere come una stima, similmente a quanto fatto precedentemente:

$$\exists m \in \mathbb{N}, c > 0 : |u(\varphi)| \leq c \sum_{|\alpha|+|\beta| \leq m} \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta \varphi(x)|$$

se φ è *piccolo*, $u(\varphi)$ deve essere piccolo. Dire che φ è piccolo significa che il suo range deve essere contenuto in una striscia molto piccola.

Si osservi che tutto ciò si fa su \mathbb{R}^n e non su un qualche aperto Ω di \mathbb{R}^n ; questo, dal momento che, con le distribuzioni temperate, quello che conta è il comportamento all'infinito, e non si studiano solo compatti di Ω come fatto precedentemente; di conseguenza, si studieranno sempre distribuzioni definite in \mathbb{R}^n .

Lo spazio delle distribuzioni temperate, analogamente a prima, viene indicato aggiungendo un apice a S : $S'(\mathbb{R}^n)$.

Esempio

Si consideri f funzione misurabile in \mathbb{R}^n , che cresce al più polinomialmente, ossia:

$$|f(x)| \leq (1 + |x|)^N$$

questa è una funzione a crescita lenta; in questa situazione f definisce una distribuzione temperata:

$$T_f(\varphi) = \int_{\mathbb{R}^n} f\varphi \, dx, \quad \varphi \in S(\mathbb{R}^n)$$

Note conclusive sulle distribuzioni temperate

In generale, ogni funzione in L^p può essere vista come una funzione temperata:

$$f \in L^p \implies T_f(\varphi) = \int f\varphi \, dx, \quad \varphi \in S(\mathbb{R}^n)$$

(si può dimostrare che questa è veramente una distribuzione applicando la disuguaglianza di Hölder). In generale, si può vedere che:

$$\mathcal{E}'(\mathbb{R}^n) \subset S'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$$

questa osservazione non è banale, e bisogna spiegare esattamente cosa significa. Dire che $S'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$ è un'affermazione che deve partire da come vengono definiti i funzionali che realizzano la distribuzione, ossia gli spazi da cui partono; $\mathcal{D}'(\mathbb{R}^n)$ ha funzionali lineari continui definiti a partire da $C_c^{(\infty)}(\mathbb{R}^n)$; $S'(\mathbb{R}^n)$ ha funzionali definiti da $S(\mathbb{R}^n)$; quest'ultimo, è uno spazio

di funzioni più vasto di $C_c^{(\infty)}(\mathbb{R}^n)$; quello che si può fare, dunque, è cercare una applicazione che associ a una certa u la restrizione in $C_c^{(\infty)}(\mathbb{R}^n)$; questa sarà nel senso più generale in \mathcal{D}' , dal momento che \mathcal{D}' è il duale di $C_c^{(\infty)}$; considerando dunque u invece che su S , su $C_c^{(\infty)}$, si ottiene una distribuzione.

Anche per quanto concerne le distribuzioni temperate, è possibile definire un concetto di convergenza. Si dice che $u_j \rightarrow u$ in $S'(\mathbb{R}^n)$ se:

$$u_j(\varphi) \rightarrow u(\varphi) \forall \varphi \in S(\mathbb{R}^n)$$

Esempio

Si considerino delle $f_j(x)$ definite come:

$$f_j(x) = e^{-\frac{x^2}{j}}$$

queste distribuzioni, per $j \rightarrow \infty$, tendono ad allargarsi, mantenendo il massimo costantemente a 1; di conseguenza, queste funzioni tendono, intuitivamente, alla costante:

$$f_j(x) \rightarrow 1 \text{ in } S'(\mathbb{R})$$

Si provi a verificare ciò, studiando l'azione di queste funzioni sulle funzioni di Schwartz, come distribuzioni. Si può vedere che:

$$\lim_{j \rightarrow \infty} \int e^{-\frac{x^2}{j}} \varphi(x) dx$$

si può trattare portando il limite sotto il segno di integrale, grazie al teorema della convergenza dominata di Lebesgue; questo teorema dice che, se la funzione integranda può essere maggiorata con una funzione in L^1 che non dipenda dal parametro j in questione, è possibile effettuare lo scambio di limite e derivata; in questo caso, è possibile dire, $\forall j$, che:

$$\left| e^{-\frac{x^2}{j}} \varphi(x) \right| \leq |\varphi(x)| \in L^1$$

Da qui, dunque:

$$\lim_{j \rightarrow \infty} \int e^{-\frac{x^2}{j}} \varphi(x) dx = \int \lim_{j \rightarrow \infty} e^{-\frac{x^2}{j}} \varphi(x) dx = \int 1 \varphi(x) dx$$

dunque, effettivamente, il limite di questa serie di gaussiane agisce come l'unità, sulle funzioni test.

7.4 Trasformata di Fourier

7.4.1 Trasformata di Fourier di funzioni: teoria L^1

Al fine di introdurre le idee in questione, si vuole definire la trasformata di Fourier nel contesto di funzioni; questo contesto verrà esteso a quello più generico di distribuzioni, in seguito.

La trasformata di Fourier $\hat{f}(\xi)$ di una funzione $f(x)$ viene definita come segue:

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-jx\xi} dx$$

qui, siamo in \mathbb{R}^n ; di conseguenza, $x = (x_1, x_2, x_3, \dots, x_n)$; ξ , detta **variabile duale**, si scrive come $\xi = (\xi_1, \xi_2, \dots, \xi_n)$. Il termine all'esponente è un prodotto scalare:

$$\xi x = (x_1 \xi_1, x_2 \xi_2, \dots, x_n \xi_n)$$

Lo spazio in cui ha senso definire questo integrale è L^1 : solo in questo caso, ha senso definirlo. Dunque, se $f \in L^1(\mathbb{R}^n)$, si ha che anche $e^{-jx\xi} f(x)$ sarà in $L^1(\mathbb{R}^n)$, $\forall \xi \in \mathbb{R}^n$; infatti:

$$\int |f(x) e^{-jx\xi}| dx = \int |f(x)| dx = \|f\|_{L^1}$$

se per ipotesi $f \in L^1$ come abbiamo detto, allora l'integrale è ben definito.

Vale il seguente risultato:

$$\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1}$$

infatti:

$$|\hat{f}(\xi)| \leq \int |e^{-jx\xi} f(x)| dx = \|f\|_{L^1}$$

Alcuni risultati tecnici

Al fine di procedere con le proprietà della trasformata di Fourier, è buona cosa riprendere alcuni risultati tecnici.

Il primo risultato che si vuole proporre è il seguente (valido per 1 variabile, ma la cosa poi è estensibile in molte variabili); date due funzioni $f, g \in C^{(1)}(\mathbb{R})$, si ha che:

$$\int_{\mathbb{R}} f'g dx = - \int_{\mathbb{R}} fg' dx \text{ se } fg \rightarrow 0 \text{ per } x \rightarrow \pm\infty, \text{ e } f'g, fg' \in L^1, x$$

Il secondo risultato che si vuole discutere è: data $f \in C^{(1)}(\mathbb{R} \times \mathbb{R})$, allora:

$$\frac{d}{d\xi} \int f(x, \xi) dx = \int \frac{d}{d\xi} f(x, \xi) dx$$

questo, è lecito per la condizione sufficiente per cui:

$$f(\cdot, \xi), \frac{\partial f}{\partial \xi}(\cdot, \xi) \in L^1, \forall \xi$$

dove l'appartenenza a L^1 è rispetto alla \cdot , che sarebbe, nell'integrale di esempio appena riportata, la x ; inoltre, si ha che:

$$\frac{\partial f}{\partial \xi} \in L^1(\mathbb{R} \times [a, b])$$

e questo è vero su ogni $[a, b]$.

Proprietà della trasformata di Fourier

La trasformata di Fourier è uno strumento matematico estremamente potente, dal momento che permette di trasformare operatori algebrici in operatori di derivazione, e viceversa. Si consideri $f \in S(\mathbb{R}^n)$; sia definita:

$$\partial_i = \frac{\partial}{\partial x_i}, \text{ oppure } \partial_i = \frac{\partial}{\partial \xi_i}, \text{ a seconda del contesto}$$

allora:

$$\partial_i \hat{f}(\xi) = \mathcal{F} \{-jx_i f(x)\}$$

per dimostrare ciò, è sufficiente recuperare il secondo dei due risultati tecnici appena proposti:

$$\begin{aligned} \frac{\partial}{\partial x_i} \int_{\mathbb{R}^n} e^{-jx\xi} f(x) dx &= \int \frac{\partial}{\partial \xi_i} e^{-jx\xi} f(x) dx = \\ &= \int (-jx_i) e^{-jx\xi} f(x) dx \implies \mathcal{F} \{(-jx_i) f(x)\} \end{aligned}$$

Una seconda proprietà molto rilevante è la seguente:

$$\partial_i \hat{f}(\xi) = \mathcal{F} \{-j\partial_i f(x)\}$$

per dimostrare ciò:

$$\begin{aligned} \xi_i \int e^{-jx\xi} f(x) dx &= \int \xi_i e^{-jx\xi} f(x) dx = \int \xi_i \frac{-j}{-j} e^{-jx\xi} f(x) dx = \\ &= \int j(-j\xi_i) e^{-jx\xi} f(x) dx = \int j\partial_i (e^{-jx\xi}) f(x) dx \end{aligned}$$

a questo punto, si applica l'integrazione per parti, ossia il primo dei risultati tecnici prima discussi, ottenendo:

$$= - \int j e^{-jx\xi} \partial_i f(x) dx = \mathcal{F} -j\partial_i f(x)$$

Queste due dimostrazioni possono essere usate in maniera iterativa, ottenendo:

$$\begin{aligned} \partial^\alpha \hat{f}(\xi) &= \mathcal{F} \{(-jx)^\alpha f(x)\} \\ \xi^\alpha \hat{f}(\xi) &= \mathcal{F} \{(-j\partial)^\alpha f(x)\} \end{aligned}$$

dove:

$$(-j\partial)^\alpha = (-j)^{|\alpha|} \partial^\alpha$$

Esempio applicativo: equazione del calore

Come esempio applicativo della trasformata di Fourier, si vuole impostare la soluzione del seguente problema differenziale:

$$\begin{cases} \partial_t u - \Delta u = 0, & (t, x) \in [0, +\infty) \times \mathbb{R}^n \\ u(0, x) = u_0(x) \end{cases}$$

Applicando la trasformata di Fourier a x , definita come:

$$\hat{u}(t, \xi) = \int e^{-jx\xi} u(t, x) dx$$

si può ottenere:

$$\mathcal{F} \partial_t u(t, \xi) = \partial_t \hat{u}(t, \xi)$$

infatti, ∂_t agisce solo sulla dipendenza in t , non su quella in x ; per quanto riguarda la seconda parte, quella contenente il laplaciano, è necessario effettuare la seguente considerazione:

$$\mathcal{F} \left\{ \frac{\partial^2 u}{\partial x_i^2} \right\} = (j\xi_i)^2 \hat{u}(t, \xi_i) = -\xi_i^2 \hat{u}(t, \xi)$$

dunque, facendo la stessa cosa per ogni componente, e sommandole, si ottiene:

$$\mathcal{F} \{ \nabla^2 u \} = -|\xi|^2 \hat{u}(t, \xi)$$

dunque, l'equazione differenziale diventa:

$$\begin{cases} \partial_t \hat{u}(t, \xi) + |\xi|^2 \hat{u}(t, \xi) = 0 \\ \hat{u}(0, \xi) = \hat{u}_0(\xi) \end{cases}$$

Questa equazione è nella forma:

$$\begin{cases} y' + ay = 0 \\ y(0) = y_0 \end{cases}$$

la soluzione di questa equazione differenziale è:

$$y(t) = y_0 e^{-at}$$

dove, nel nostro caso, $a = |\xi|^2$, e c'è la condizione iniziale $\hat{u}_0(\xi) = y_0$; dunque:

$$\hat{u}(t, \xi) = \hat{u}_0(\xi) e^{-t|\xi|^2}$$

dimostreremo, parlando di antitrasformata, che la soluzione dell'equazione differenziale, dunque, è:

$$u(t, x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} u_0(y) dy$$

Nota sullo spazio di Schwartz

La trasformata di Fourier manda lo spazio di Schwartz $S(\mathbb{R}^n)$ in $S(\mathbb{R}^n)$ stesso; ossia:

$$\mathcal{F} : S(\mathbb{R}^n) \rightarrow S(\mathbb{R}^n)$$

questa cosa non sarà verificata; la stessa cosa vale con l'antitrasformata. Questo ci aiuta a capire come mai Schwartz abbia scelto proprio queste funzioni come funzioni test per le distribuzioni temperate: il fatto di convertire la moltiplicazione in derivazione e viceversa. Schwartz, a partire da questa osservazione, ha avuto la seguente idea: considerare uno spazio di funzioni che viene mandato in sé stesso, dall'operatore *trasformata di Fourier*; siccome la trasformata di Fourier scambia le operazioni di derivazione e prodotto per la variabile, si ha che, se una funzione è molto derivabile, la sua trasformata deve decadere molto rapidamente all'infinito, e viceversa. Se lo spazio deve mandare uno spazio in sé stesso, ossia se spazio di partenza e spazio di arrivo devono essere uguali, le funzioni all'interno dovranno essere sia molto regolari, sia decadenti molto rapidamente all'infinito; questo è lo spazio di Schwartz.

7.4.2 Formula di inversione

A questo punto, data $f \in L^1(\mathbb{R}^n)$, si è visto come calcolare la sua trasformata di Fourier $\hat{f}(\xi)$; tuttavia, come si può ritrovare f a partire dalla trasformata?

In generale, $\hat{f}(\xi) \in L^\infty$, dunque è dura aggiungere qualcosa in questa situazione, a meno che non si aggiungano ipotesi²; ciò che si può fare, è richiedere che anche \hat{f} sia in L^1 , ossia integrabile; considerata questa ipotesi, si ha che la trasformata inversa è:

²in verità si può vedere che \hat{f} non solo è limitata, ma anche continua

$$f(x) = (2\pi)^{-n} \int e^{ix\xi} \hat{f}(\xi) d\xi$$

dove questa $f(x)$ è uguale **quasi ovunque** alla $f(x)$ di partenza. A partire da questa idea, applicata sulle funzioni, è dunque possibile definire l'operatore di *antitrasformata*, \mathcal{F}^{-1} , come:

$$\mathcal{F}^{-1} \{f(x)\} = (2\pi)^{-n} \int e^{ix\xi} f(\xi) d\xi$$

dunque:

$$f = \mathcal{F}^{-1} \{ \mathcal{F} \{f\} \}$$

7.4.3 Teoria L^2 della trasformata di Fourier

Finora, come ambiente naturale della trasformata di Fourier, è stato scelto L^1 , dal momento che la trasformata in questione è integrale, e per avere certezze sull'integrabilità è necessario richiedere che la funzione trasformanda sia integrabile; vedremo ora che \mathcal{F} può essere applicata anche in funzioni in L^2 . Dal momento che si sta lavorando su \mathbb{R}^n , dunque sull'intero n -spazio (non su un qualche sottointervallo/palla),

$$L^2(\mathbb{R}^n) \not\subset L^1(\mathbb{R}^n), \quad L^1(\mathbb{R}^n) \not\subset L^2(\mathbb{R}^n)$$

Al fine di poter motivare questo fatto, si consideri il seguente integrale: date $f, g \in L^1(\mathbb{R}^n)$, si può vedere che:

$$\begin{aligned} \iint e^{-ix\xi} f(x)g(\xi) dx d\xi &= \int \hat{f}(\xi)g(\xi) d\xi = \\ &= \int f(x)\hat{g}(x) dx \end{aligned}$$

Nel primo caso, viene prima fatto l'integrale in dx ; nel secondo caso, prima l'integrale in $d\xi$; l'ordine di integrazione è indifferente, dal momento che si ricade nelle ipotesi del teorema di Fubini, che può dunque essere applicato.

In entrambi i casi, si ha qualcosa di molto simile a un prodotto scalare: manca solo il coniugato. Dunque, è possibile partire da un altro integrale, per ottenere le relazioni che ci interessano:

$$(f|g)_{L^2(\mathbb{R}^n)} = \int f(x)\overline{g(x)}dx$$

a partire da ciò, dunque, si consideri quanto segue:

$$\begin{aligned} \iint e^{-jx\xi} f(x)\hat{g}(\xi)dxd\xi &= \int \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi = \\ &= \int f(x) \overline{\int e^{jx\xi} g(\xi)d\xi} dx = (f|\overline{\mathcal{F}g}) \end{aligned}$$

il passaggio in cui l'integrale è stato raggruppato semplicemente è basato sul coniugare tutti quei termini; come risultato, invece di $e^{-jx\xi}$ si ha il suo coniugato, $e^{jx\xi}$, e al posto di $\overline{\hat{g}(\xi)}$ si ha $g(\xi)$ (regola di coniugazione). Alcune note: è stato appena definito $\overline{\mathcal{F}}$ come:

$$\overline{\mathcal{F}}\{g(\xi)\} = \int e^{jx\xi} g(\xi)d\xi$$

questo è molto imparentato con la trasformata inversa di Fourier; come si può vedere:

$$\overline{\mathcal{F}}\{g(\xi)\} = (2\pi)^n \mathcal{F}\{g(\xi)\}$$

Dunque: dal momento che, in seguito al primo passaggio della dimostrazione, si è visto che:

$$\iint e^{-jx\xi} f(x)\hat{g}(\xi)dxd\xi = \int \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi = (\hat{f}|g)_{L^2}$$

e che, dal secondo passaggio,

$$\iint e^{-jx\xi} f(x)\hat{g}(\xi)dxd\xi = (f|\overline{\mathcal{F}g}) = (2\pi)^n (f|\mathcal{F}^{-1}\{g\})_{L^2}$$

si ha che:

$$(\hat{f}|g)_{L^2} = (2\pi)^n (f|\mathcal{F}^{-1}\{g\})_{L^2}$$

Da ciò, è possibile dimostrare facilmente la **formula di Plancherel**, estensione della formula di Parseval: data $f \in \mathcal{S}(\mathbb{R}^n)$, si ha:

$$\left\| \hat{f} \right\|_{L^2}^2 = (\mathcal{F}f | \mathcal{F}f) = (f | \overline{\mathcal{F}f}) = (2\pi)^n (f | \mathcal{F}^{-1} \mathcal{F}f) = (2\pi)^n (f | f) = (2\pi)^n \|f\|_{L^2}^2$$

ossia:

$$\left\| \hat{f} \right\|_{L^2}^2 = (2\pi)^n \|f\|_{L^2}^2$$

che, prendendo la radice, diventa:

$$\left\| \hat{f} \right\|_{L^2} = (2\pi)^{\frac{n}{2}} \|f\|_{L^2}, \forall f \in S(\mathbb{R}^n)$$

Questa formula non ci permette ancora di dire che a ogni funzione in L^2 può essere applicata la trasformata di Fourier, bensì ci dice che, quando calcoliamo la trasformata di Fourier di una funzione di Schwartz, la sua norma sarà uguale, a meno della moltiplicazione per una costante nota; tuttavia, questa formula suggerisce anche che è possibile **estendere** la trasformata di Fourier dall'insieme delle funzioni di Schwartz a quello delle funzioni in L^2 , mantenendo questa eguaglianza, applicando il **teorema di estensione continua**. Infatti, la trasformata \mathcal{F} è un'applicazione **lineare, continua**, come applicazione da $S(\mathbb{R}^n)$ in $S(\mathbb{R}^n)$, dove $S(\mathbb{R}^n)$ è visto come sottoinsieme di $L^2(\mathbb{R}^n)$, e tutto ciò in $L^2(\mathbb{R}^n)$. Dal momento che $S(\mathbb{R}^n)$ è un insieme denso in $L^2(\mathbb{R}^n)$ (non lo abbiamo mai visto, ma si può dimostrare), è possibile estendere la trasformata di Fourier da $L^2(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$, e la formula di Plancherel sarà sempre valida.

In L^2 , dunque, la trasformata di Fourier non è definita come integrale, ma come estensione continua della trasformata integrale definita su S , quest'ultimo visto come sottoinsieme di L^2 . Non solo: siccome la trasformata inversa ha la stessa forma (circa) della trasformata diretta, anche \mathcal{F}^{-1} manda da L^2 in L^2 ; dunque, \mathcal{F} è un isomorfismo di L^2 , e la formula di Plancherel mostra anzi che, a meno del $(2\pi)^{\frac{n}{2}}$, \mathcal{F} è un'isometria suriettiva.

Cenni a un legame tra trasformata e serie di Fourier

La trasformata di Fourier e serie di Fourier possono essere in qualche maniera legate tra loro, se analizzate nell'ottica giusta. La trasformata di Fourier può essere definita, in modo molto più astratto, su un gruppo mediano topologico

compatto; si avrebbe una definizione simile, anche se, al posto dell'esponenziale, si usa un **carattere** del gruppo, ossia un omomorfismo tra il gruppo G e \mathbb{C} .

Con questa definizione, a seconda del fatto che il gruppo sia S^1 , ossia la circonferenza della struttura di gruppo, la trasformata di Fourier coincide con la serie di Fourier; se il gruppo G è \mathbb{R} , invece, si ottiene la trasformata appena definita.

In questo framework, dunque, la stessa struttura può essere vista in due modi diversi, a seconda del gruppo che si sceglie; al gruppo G poi viene associato un *gruppo duale*, \hat{G} , dove:

- nel caso di $G = S^1$, ossia la circonferenza, si ha che $\hat{G} = \mathbb{Z}$;
- nel caso di $G = \mathbb{R}$, si ha che $\hat{G} = \mathbb{R}$.

x varia su G ; ξ varia in \hat{G} ; questo ricorda il fatto che, dunque, intuitivamente, nel caso di $G = S^1$, si ha un comportamento *periodico*, dal momento che ci si muove sempre nella circonferenza; in uscita, si ottengono numeri razionali, ossia delle *righe*. Se invece x varia su \mathbb{R} , la trasformata avrà una variabile ξ che varia anch'essa su \mathbb{R} , dal momento che *non si ha periodicità*.

Spesso, la trasformata viene indicata, intuitivamente, come una versione della serie in cui integrale e sommatoria si scambiano, a causa del fatto che il periodo tende a diventare *infinito* (si ha *aperiodicità* vista come limite di un periodo infinito), e dunque le varie righe della serie tendono ad avvicinarsi, fino a ottenere la continuità; in verità, questo tipo di ragionamento non è formalmente corretto, dal momento che, per passare da serie a trasformata e viceversa, sarebbe necessario ricorrere a questo tipo di astrazione, invece che cercare di saltare direttamente dall'una all'altra.

Calcolo della trasformata di Fourier della funzione gaussiana

Si consideri la funzione:

$$f(x) = e^{-a|x|^2}, \quad a > 0$$

Verrà ora dimostrato che:

$$\hat{f}(\xi) = \left(\frac{\pi}{a}\right)^{\frac{n}{2}} e^{-\frac{|\xi|^2}{4a}}, \quad x \in \mathbb{R}^n$$

Si osservi che, per $a = \frac{1}{2}$, f è l'autofunzione della trasformata di Fourier, dal momento che trasforma f in sé stessa. Tra le varie proprietà della trasformata di Fourier ve ne è una riguardante un principio di indeterminazione tra la funzione $f(x)$ e la sua trasformata $\hat{f}(\xi)$; questo principio afferma che:

$$\|xf\|_{L^2} \left\| \xi \hat{f} \right\|_{L^2} = \|xf\|_{L^2} \|f'\|_{L^2} \geq \frac{1}{4\pi}$$

Se la funzione f è una gaussiana, il \geq diventa un $=$, dal momento che la gaussiana è la funzione per cui l'indeterminazione è minimizzata.

Si vuole ora calcolare la trasformata di Fourier di una gaussiana, con argomento $x \in \mathbb{R}^n$; vediamo che:

$$\hat{f}(\xi) = \int e^{-jx\xi} e^{-|x|^2} dx = \int e^{-jx_1\xi_1} e^{-x_1^2} dx_1 \int e^{-jx_2\xi_2} e^{-x_2^2} dx_2 \dots \int e^{-jx_n\xi_n} e^{-x_n^2} dx_n$$

Dal momento che tutti questi integrali sono nella stessa forma; essendo una produttoria di integrali, è sufficiente essere in grado di calcolare l'integrale in una singola variabile; si consideri dunque il caso $n = 1$, $a = 1$. Si ha:

$$\hat{f}(\xi) = \int e^{-jx\xi} e^{-x^2} dx$$

Un trucco che spesso si utilizza, quando si deve calcolare un integrale in una certa variabile (x) contenente un altro parametro (ξ), è basato sul derivare la funzione rispetto al parametro (ξ). Dunque:

$$\frac{d}{d\xi} \hat{f}(\xi) = \int [-jx] e^{-x^2} e^{-jx\xi} dx$$

Si moltiplica e divide per due, e così si osserva il seguente fatto:

$$= \frac{j}{2} \int e^{-jx\xi} \frac{d}{dx} (e^{-x^2}) dx$$

si può a questo punto applicare il precedente risultato "tecnico", ossia l'integrazione per parti su \mathbb{R} (quello senza i termini valutati agli estremi), e così ottenere:

$$= -\frac{j}{2} \int \left[\frac{d}{dx} e^{-j\xi x} \right] e^{-x^2} dx = -\frac{j}{2} \int (-j\xi) e^{-j\xi x} e^{-x^2} dx = -\frac{\xi}{2} \hat{f}(\xi)$$

Riassumendo, abbiamo ottenuto la seguente relazione:

$$\frac{d}{d\xi} \hat{f}(\xi) = -\frac{\xi}{2} \hat{f}(\xi)$$

Questa è un'equazione differenziale, nella forma $y'(x) = a(x)y(x)$. Per risolverla, è sufficiente calcolare una primitiva $A(x)$ di $a(x)$:

$$A(x) = \int a(x) dx = -\frac{\xi^2}{4} + K$$

da cui:

$$\hat{f}(\xi) = ce^{-\frac{\xi^2}{4}}$$

si deve ancora calcolare $c = e^K$; per fare ciò, possiamo sfruttare un'informazione nota: valutare la trasformata di Fourier in $\xi = 0$; infatti, per $\xi = 0$, si ha che:

$$\hat{f}(0) = c = \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

(la dimostrazione di questo risultato è nota). Dunque:

$$c = \sqrt{\pi}$$

La tecnica utilizzata in questo ambito è generale: quando si ha una situazione di questo genere, una buona idea è provare a derivare rispetto al parametro (non di integrazione), al fine di ottenere qualcosa di questo genere.

Note aggiuntive sulla convoluzione

Esiste un legame tra il prodotto in un dominio (x) e la convoluzione nel dominio reciproco (ξ), e viceversa. Date due funzioni di Schwartz, $f, g \in S(\mathbb{R}^n)$, valgono infatti le seguenti proprietà:

1. $\mathcal{F}\{f * g\} = \hat{f} \hat{g}$
2. $\mathcal{F}\{fg\} = \frac{1}{(2\pi)^n} \hat{f} * \hat{g}$

Le dimostrazioni di queste due proprietà sono simili; si vuole dunque dimostrare la prima:

$$(f * g)(x) = \int f(y)g(x - y)dy$$

si vuole ora calcolare la trasformata di ciò:

$$\mathcal{F}f * g = \int e^{-jx\xi} \int f(y)g(x - y)dydx$$

a questo punto, sommiamo e sottraiamo all'esponente di $e^{-j\xi x}$ un y , come segue:

$$e^{-jx\xi} = e^{-j(x-y)\xi} e^{-jy\xi}$$

da cui, l'integrale diviene:

$$= \iint e^{-jy\xi} f(y) e^{-j(x-y)\xi} g(x - y) dy dx$$

a questo punto, si ha che $f(y)g(x - y) \in S(\mathbb{R}^n)$; questo, è il prodotto di due funzioni di Schwartz, dunque questo sarà ancora una funzione di Schwartz; questo, ci permette di applicare il teorema di Fubini-Tonelli, ottenendo:

$$= \int e^{-jy\xi} f(y) \int e^{-j(x-y)\xi} g(x - y) dx dy$$

Nell'integrale più interno, y è solo un parametro, ma l'operatore di integrazione non agisce su di esso; di conseguenza, è possibile effettuare il seguente cambio di variabili:

$$x - y = z$$

da qui, si può dimostrare che $dx = dz$; infatti:

$$dx = \left| \det \frac{\partial \varphi}{\partial z} \right| dz$$

dove il modulo è il modulo del determinante della matrice jacobiana; in questo caso, si ha che:

$$x_1 = y_1 + z_1$$

$$x_2 = y_2 + z_2$$

e così via; in questo caso, la jacobiana è l'identità (si deriva $\frac{\partial x_i}{\partial z_i}$); di conseguenza, $dx = dz$, anche in più variabili; tutto ciò per dire che questo è uguale a:

$$e^{-iy\xi} f(y) \left[\int e^{-jz\xi} g(z) dz \right] dy = \hat{g}(\xi) \int e^{-iy\xi} f(y) dy = \hat{g}(\xi) \hat{f}(\xi)$$

La dimostrazione è così ultimata.

Dalla precedente formula, è possibile trarre la seguente conclusione: l'antitrasformata di Fourier del prodotto è uguale alla convoluzione delle antitrasformate; infatti:

$$\mathcal{F}^{-1} \{fg\} = \mathcal{F}^{-1} \{f\} * \mathcal{F}^{-1} \{g\}$$

Infatti, se si prende l'antitrasformata di $\mathcal{F} \{f * g\}$, si ha:

$$\mathcal{F}^{-1} \{\mathcal{F} \{f * g\}\} = \mathcal{F} \{\hat{f}\hat{g}\}$$

ma dunque:

$$f * g = \mathcal{F}^{-1} \{\hat{f}\hat{g}\}$$

A questo punto: si definiscano $f = \mathcal{F}^{-1} \{u\}$, $g = \mathcal{F}^{-1} \{v\}$; si ha dunque:

$$\hat{f} = u \quad \hat{g} = v$$

quindi, questa formula dice che:

$$\mathcal{F}^{-1} \{u\} \mathcal{F}^{-1} \{v\} = \mathcal{F}^{-1} \{uv\}$$

Un'osservazione: queste formule sono state introdotte a partire da funzioni f, g di Schwartz; in verità, è possibile fare le stesse considerazioni per $f, g \in L^1(\mathbb{R}^n)$; infatti, si può dimostrare facilmente che $f * g \in L^1(\mathbb{R}^n)$; in altre parole, L^1 è un'algebra di convoluzione, senza l'unità; questo accade dal momento che la δ non è una funzione, bensì una distribuzione, di conseguenza l'elemento identità non esiste.

Esempio applicativo: equazione del calore

Si vuole a questo punto usare quanto appena studiato per motivare l'ultimo passaggio per ricavare la soluzione dell'equazione del calore.

$$\begin{cases} \partial_t u - \Delta u = 0, & (t, x) \in [0, +\infty) \times \mathbb{R}^n \\ u(0, x) = u_0(x) \end{cases}$$

Si era visto che $u_0 \in S(\mathbb{R}^n)$; dunque:

$$\hat{u}(t, \xi) = \int e^{-ix\xi} u(t, x) dx$$

Si arriva a ottenere, coi vari passaggi precedentemente eseguiti:

$$\hat{u}(t, \xi) = \hat{u}_0(\xi) e^{-t|\xi|^2}$$

ora, dobbiamo “tornare indietro”; tuttavia, abbiamo una formula per farlo, dal momento che abbiamo visto che:

$$u(t, x) = \mathcal{F}^{-1} \left\{ \hat{u}_0(\xi) e^{t|\xi|^2} \right\} = u_0(x) * \mathcal{F}^{-1} \left\{ e^{-t|\xi|^2} \right\}$$

rimane solo dunque da calcolare questa antitrasformata; ma noi sappiamo che:

$$\mathcal{F} \left\{ e^{-a|x|^2} \right\} = \left(\frac{\pi}{a} \right)^{\frac{n}{2}} e^{-\frac{|\xi|^2}{4a}}$$

dunque: usando $a = \frac{1}{4t}$, si può scrivere:

$$\mathcal{F}^{-1} \left\{ e^{-t|\xi|^2} \right\} = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int e^{-\frac{|x-y|}{4t}} u_0(y) dy$$

Osservazione: l'antitrasformata deve avere senso! Questo fatto è però verificato, dal momento che si può dimostrare che, con queste condizioni, l'equazione del calore è **ben posta**: questo, grazie ai dati iniziali scelti.

7.4.4 Trasformata di Fourier di distribuzioni

L'idea è la solita: estendere i concetti prima visti sulle funzioni, studiando l'azione della trasformata di Fourier su delle funzioni test. L'idea di partenza dunque è: data $f \in S(\mathbb{R}^n)$ $\varphi \in C_c^{(\infty)}(\mathbb{R}^n)$ (che è un sottoinsieme di $S(\mathbb{R}^n)$), si ha che:

$$\mathcal{F}\{f\} = \hat{f}$$

certamente è Schwartz, ma si può pure pensare come distribuzione! Dunque, ha senso vedere come essa agisce su una funzione test; l'idea è:

$$\int \hat{f}\varphi d\xi = \int f\hat{\varphi}d\xi$$

ossia, applicando le solite formule e i cambi di variabile (formule precedentemente dimostrate con $f(x)$ e $g(x)$ negli integrali doppi), si vuole scaricare l'operazione di trasformazione secondo Fourier sulla funzione test.

Ora, considerando $u \in \mathcal{D}'(\mathbb{R}^n)$, $\hat{u}(\varphi)$ è:

$$\hat{u}(\varphi) = u(\hat{\varphi})$$

in verità, **ciò è sbagliato**: questo non funziona, dal momento che, se φ è una funzione test, essa è a supporto compatto; di conseguenza, la sua trasformata di Fourier dovrà essere, per la proprietà del principio di indeterminazione, a supporto non limitato; di conseguenza, $\hat{\varphi} \notin C_c^{(\infty)}$; la trasformata di φ , in altre parole, non è più una funzione test.

La teoria delle distribuzioni "classica" non è compatibile, per questo motivo, con la trasformata di Fourier; tuttavia, l'idea di Schwartz, a questo punto, è stata quella di cambiare lo spazio delle funzioni test: considerare $\varphi \in S(\mathbb{R}^n)$; in questo caso, dunque, la u non sarà più una distribuzione, ossia in \mathcal{D}' , bensì sarà solo in $S'(\mathbb{R}^n)$: sarà una **distribuzione temperata**. Se dunque prendiamo $u \in S'(\mathbb{R}^n)$, $\varphi \in S(\mathbb{R}^n)$, abbiamo:

$$\hat{u}(\varphi) = u(\hat{\varphi})$$

E, $\hat{u} \in \mathcal{D}'$: anche essa è una distribuzione temperata! Infatti, se si considera $\varphi_i \rightarrow \varphi$ in $S(\mathbb{R}^n)$, si ha:

$$\hat{u}(\varphi_i) \rightarrow u(\hat{\varphi})$$

Questo si può dimostrare, così:

$$\hat{u}(\varphi_i) = u(\hat{\varphi}_i) \rightarrow u(\hat{\varphi}) = \hat{u}(\varphi)$$

si ha che $\hat{\varphi}_i \rightarrow \hat{\varphi}$, nello spazio di Schwartz, dal momento che \mathcal{F} è una applicazione da S in S , continua:

$$\mathcal{F} : S(\mathbb{R}^n) \rightarrow S(\mathbb{R}^n)$$

Esempio

Si consideri il seguente esempio: $\hat{\delta}$, dove $\delta \in S'(\mathbb{R}^n)$ (la delta di Dirac è una distribuzione temperata). Vediamo come agisce su una funzione di Schwartz:

$$\hat{\delta}(\varphi) = \delta(\hat{\varphi}) = \hat{\varphi}(0) = \int e^{-j\xi x} \varphi(x) dx \Big|_{\xi=0} = \int 1 \times \varphi(x) dx = \langle 1, \varphi \rangle$$

ossia, è l'azione della distribuzione "1" su φ .

7.4.5 Antitrasformata di Fourier di distribuzioni

Si vuole a questo punto introdurre l'antitrasformata di Fourier di distribuzioni; la validità della formula verrà anche dimostrata, al fine di poi dimostrarla anche nel caso di funzioni, in maniera elegante.

Dat $u \in S'(\mathbb{R}^n)$, si ha:

$$\mathcal{F}^{-1} \{u\}(\varphi) = u(\mathcal{F}^{-1} \{\varphi\}), \varphi \in S(\mathbb{R}^n)$$

Osservando che:

$$\mathcal{F}^{-1} \{\varphi\} = (2\pi)^{-n} \int e^{jx\xi} \varphi(\xi) d\xi, \in S'(\mathbb{R}^n)$$

Ora, vogliamo dimostrare che la \mathcal{F}^{-1} così definita è veramente l'inversa della distribuzione; non solo: vogliamo vedere che:

$$\mathcal{F}^{-1} : S'(\mathbb{R}^n) \rightarrow S'(\mathbb{R}^n)$$

e che:

$$\mathcal{F} : S'(\mathbb{R}^n) \rightarrow S'(\mathbb{R}^n)$$

non solo:

$$\mathcal{F}^{-1} : S(\mathbb{R}^n) \rightarrow S(\mathbb{R}^n)$$

e che:

$$\mathcal{F} : S(\mathbb{R}^n) \rightarrow S(\mathbb{R}^n)$$

Un'osservazione: pensando nel senso classico alla trasformata di Fourier, verrebbe da dire che, dal momento che l'integrabilità è richiesta, che tutti questi discorsi valgano solo per funzioni fortemente decrescenti; in verità, nel senso distribuzionale, tutte queste considerazioni hanno senso in un contesto molto più ampio, ossia quello delle funzioni a **crescita lenta**, come per esempio i polinomi; questo fatto, intuitivamente, è dovuto al fatto che l'esponenziale complesso, grazie alle proprie oscillazioni, riesce in qualche modo a compensare la crescita della funzione, mettendo a posto le cose.

Al fine di introdurre la dimostrazione, è necessario calcolare la trasformata della funzione identicamente uguale a 1:

$$f(x) = 1 \forall x$$

Si considera una successione $f_m(x)$ di funzioni Schwartz che tendano a 1:

$$f_m(x) = e^{-\frac{|x|^2}{m}}$$

Si ha che $f_m(x) \rightarrow 1$, in $S'(\mathbb{R}^n)$. Si vuole sfruttare la seguente idea: come la successione di funzioni f_m converge a 1, così la trasformata della successione di funzioni tenderà alla trasformata di 1 ! In questo modo, dunque, si sta dimostrando questa trasformata per mezzo di un'approssimazione mediante identità approssimata, per poi passare al limite; si ha che:

$$\mathcal{F} \{f_m(x)\} = e^{-\frac{|x|^2}{m}} = (\pi m)^{\frac{n}{2}} e^{-m \frac{|\xi|^2}{4}} \rightarrow \mathcal{F} \{1\}$$

A questo punto, al fine di ottenere a partire da ciò un'identità approssimata (si vedano i dettagli nella parte precedente dedicata ad esse degli appunti), si moltiplica e divide per $(2\pi)^n$:

$$(2\pi)^n \left(\frac{m}{4\pi} \right)^{\frac{n}{2}} e^{-m \frac{|\xi|^2}{4}}$$

si può verificare facilmente che questa è un'identità approssimata. Si può dunque studiare l'azione su funzioni test di queste identità, e il limite sarà l'azione sul limite della successione; questo, per funzioni test, ossia per funzioni appartenenti a $C_c^{(\infty)}(\mathbb{R}^n)$; la stessa cosa, tuttavia, vale utilizzando funzioni test appartenenti a $S(\mathbb{R}^n)$; si può dunque dimostrare che:

$$\mathcal{F}\{1\} = (2\pi)^n \delta$$

Teorema di inversione

A questo punto si può enunciare il teorema di inversione, descrivendo sommariamente le prime conseguenze; il teorema di inversione afferma che valgono le seguenti formule:

1. La prima formula riguarda le funzioni di Schwartz, e afferma che:

$$\mathcal{F}^{-1}\{\mathcal{F}\{\varphi\}\} = \varphi, \forall \varphi \in S(\mathbb{R}^n)$$

2. La seconda formula scambia semplicemente l'ordine di applicazione della trasformata di Fourier:

$$\mathcal{F}\{\mathcal{F}^{-1}\{\varphi\}\} = \varphi, \forall \varphi \in S(\mathbb{R}^n)$$

3. La terza e la quarta formula sono l'analogo delle prime due, però per quanto concerne le distribuzioni temperate; nel dettaglio, la prima di queste afferma che:

$$\mathcal{F}^{-1}\{\mathcal{F}\{u\}\} = u, \forall u \in S'(\mathbb{R}^n)$$

4. L'ultima formula rappresenta l'analogo della seconda, per le distribuzioni:

$$\mathcal{F}\{\mathcal{F}^{-1}\{u\}\} = u, \forall u \in S'(\mathbb{R}^n)$$

Tra breve alcune dimostrazioni fondamentali saranno introdotte; prima di ciò, tuttavia, si vuole proporre la seguente osservazione: dimostrare questo teorema significa dimostrare che la trasformata di Fourier è un isomorfismo non solo di S in sé stesso, ma anche di S' in sé stesso; questa è una cosa che è stata precedentemente detta, al fine di motivare la definizione degli spazi di Schwartz e delle distribuzioni temperate:

$$\begin{aligned} \mathcal{F} &: S \rightarrow S \\ \mathcal{F} &: S' \rightarrow S' \end{aligned}$$

Questo ci dice che, sia per S' , sia per S , \mathcal{F} è una biiezione continua: un isomorfismo; volendo, si ha qualcosa in più: S' , infatti, è anche un isomorfismo da L^2 in L^2 .

A questo punto, si vuole dimostrare la prima proprietà; per fare ciò, scriviamo la definizione di antitrasformata di Fourier, al fine dunque di verificarla:

$$\mathcal{F}^{-1} \{ \mathcal{F} \{ \varphi(x) \} \} = (2\pi)^{-n} \int e^{jx\xi} \hat{\varphi}(\xi) d\xi$$

A questo punto, al fine di procedere, si utilizza una proprietà della trasformata di Fourier (che non è stata precedentemente introdotta): la modulata della trasformata, ossia la moltiplicazione per $e^{ja\xi}$. Questa proprietà afferma che:

$$e^{ja\xi} \hat{f}(\xi) = e^{ja\xi} \int e^{-jx\xi} f(x) dx = \int e^{-j(x-a)\xi} f(x) dx$$

A questo punto, si ridefinisca $x \leftrightarrow x - a$, ossia $x \rightarrow x + a$ si ottiene:

$$= \int e^{-jx\xi} f(x + a) dx$$

Questa non è altri che l'azione della distribuzione "1" sulla funzione f , traslata di a . Si è ottenuto:

$$e^{ja\xi} \hat{f}(\xi) = \int e^{-jx\xi} f(x + a) dx$$

Se al posto di f mettiamo φ , è possibile leggere quanto segue, riscrivendo l'integrale di partenza usando questa relazione:

$$(2\pi)^{-n} \int e^{jx\xi} \hat{\varphi}(\xi) d\xi = (2\pi)^{-n} \iint e^{-jx\xi} \varphi(\xi + x) d\xi dx$$

ossia, qui è stato usato $a = x$; questo si può vedere come segue:

$$= (2\pi)^{-n} \int 1 \times \mathcal{F} \{ \varphi(\cdot, x) \} dx$$

la variabile "·" sarebbe semplicemente la variabile ξ , che è scomparsa dopo l'integrazione; rispetto all'integrale più interno, x è solo un parametro fissato; l'esponenziale è sparito nel calcolo dell'integrale più interno, calcolato come trasformata di Fourier rispetto a ξ . Per ispezione si può dedurre che questa

non è altri che l'azione della distribuzione "1", sulla trasformata di queste funzioni φ , traslate di x :

$$= (2\pi)^{-n} \langle 1, \mathcal{F} \{ \varphi(\cdot + x) \} \rangle$$

Quindi, poi, è possibile scaricare la trasformata di φ sulla distribuzione:

$$\begin{aligned} (2\pi)^{-n} \langle 1, \mathcal{F} \{ \varphi(\cdot + x) \} \rangle &= (2\pi)^{-n} \langle \hat{1}, \varphi(\cdot, x) \rangle = (2\pi)^{-n} \langle (2\pi)^n \delta, \varphi(\cdot + x) \rangle \\ &= \langle \delta, \varphi(\cdot + x) \rangle = \varphi(x) \end{aligned}$$

Spieghiamo alcuni passaggi con calma: nel secondo passaggio è stata usata la trasformata di 1, che è stata dimostrata precedentemente; quindi, fatte le semplificazioni del caso, ci si è ritrovati con il calcolo dell'azione di una δ , su una $\varphi(\cdot + x)$; δ agisce sulla "variabile della distribuzione", ossia \cdot , essendo x solo un parametro; di conseguenza, δ valuta \cdot in 0, ottenendo $\varphi(x)$ alla fine.

La dimostrazione della proprietà 2 è analoga alla prima, dunque non la si propone. Si vuole dire qualcosa in più per quanto riguarda il punto 3 (il punto 4 sarà analogo al punto 3): per dimostrare questo punto, si vuole verificare che:

$$\mathcal{F}^{-1} \{ \mathcal{F} \{ u(\varphi) \} \} = u(\varphi), \forall \varphi \in S(\mathbb{R}^n)$$

per fare ciò, si vuole vedere che:

$$\mathcal{F}^{-1} \{ \mathcal{F} \{ u(\varphi) \} \} = \mathcal{F} \{ u(\mathcal{F}^{-1} \{ \varphi \}) \} = u(\mathcal{F} \{ \mathcal{F}^{-1} \{ \varphi \} \})$$

questo, ricordando come si applica la trasformata di Fourier (o, indifferentemente, l'antitrasformata), a una funzione test. Tuttavia, per quanto concerne le funzioni di Schwartz, abbiamo appena dimostrato che:

$$\mathcal{F}^{-1} \{ \mathcal{F} \{ \varphi \} \} = u(\mathcal{F} \{ \mathcal{F}^{-1} \{ \varphi \} \}) = \varphi$$

ma dunque:

$$u(\mathcal{F} \{ \mathcal{F}^{-1} \{ \varphi \} \}) = u(\varphi)$$

e la dimostrazione è completata.

Cenni al caso di funzioni in L^p

Quanto è stato detto è vero in $S(\mathbb{R}^n)$; si può dire qualcosa di simile per $L^p(\mathbb{R}^n)$? La risposta è sì, dal momento che:

$$L^p(\mathbb{R}^n) \subset S(\mathbb{R}^n)$$

Tuttavia, è possibile dire qualcosa in più: abbiamo visto che la trasformata di Fourier è un'applicazione lineare continua tra i seguenti spazi:

$$\mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^\infty(\mathbb{R}^n)$$

e

$$\mathcal{F} : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$$

A questo punto ci si potrebbe chiedere: per $p \in (1, 2)$, cosa capita? Ossia, per gli spazi in un qualche senso intermedi³. Dal momento che \mathcal{F} agisce come visto, ed è un'applicazione lineare continua, allora per il teorema di interpolazione di Riesz-Thorin, allora \mathcal{F} definisce un operatore lineare continuo tra una scala di spazi intermedi (intermedi riferendoci appunto all'indice) tra $L^1(\mathbb{R}^n)$ e $L^2(\mathbb{R}^n)$ come “spazi dominio”, e $L^1(\mathbb{R}^n)$ e $L^\infty(\mathbb{R}^n)$ come “spazi codominio”. Questo teorema garantisce che:

$$\mathcal{F} : L^p(\mathbb{R}^n) \rightarrow L^{p'}(\mathbb{R}^n), \quad \frac{1}{p} + \frac{1}{p'} = 1, \quad 1 \leq p \leq 2$$

Non solo: questo teorema permette di affermare che:

$$\hat{f}_{L^{p'}} \leq (2\pi)^{\frac{n}{p'}} \|f\|_{L^p}$$

Se invece $p > 2$; la trasformata di $f \in L^p(\mathbb{R}^n)$, in generale, è una distribuzione, e non più una funzione.

Qualche parola in più su questo teorema: in generale esso dice che, se:

$$\begin{cases} A : L^{p_1}(\mathbb{R}^n) \rightarrow L^{q_1}(\mathbb{R}^n) \\ A : L^{p_2}(\mathbb{R}^n) \rightarrow L^{q_2}(\mathbb{R}^n) \end{cases}$$

³è improprio dire intermedi nel senso insiemistico, dal momento che non si può dire nulla riguardo al fatto che una funzione che appartiene a uno spazio L^p appartenga a un $L^{p'}$, $p' > p$ o $p' < p$, dunque “intermedi” in senso figurato, guardando la grandezza dell'indice: $1 < 2$, anche se ciò poi in senso insiemistico ciò non vuol dire nulla: $L^1(\mathbb{R}^n) \not\subset L^2(\mathbb{R}^n)$, e $L^2(\mathbb{R}^n) \not\subset L^1(\mathbb{R}^n)$

Se A è un operatore lineare continuo, allora, l'operatore:

$$A : L^{p_\vartheta} \rightarrow L^{q_\vartheta}$$

dove:

$$p_\vartheta = \frac{1 - \vartheta}{p_1} + \frac{\vartheta}{p_2}$$

$$q_\vartheta = \frac{1 - \vartheta}{q_1} + \frac{\vartheta}{q_2}$$

È ancora un operatore lineare continuo. Questo risultato è legato alla convessità: infatti, l'insieme dei p_ϑ e q_ϑ è un insieme **convesso**.

7.4.6 Proprietà aggiuntive della trasformata di Fourier di distribuzioni

Si vuole a questo punto aggiungere ancora qualcosa riguardo alle trasformate di Fourier di distribuzioni, e alle principali proprietà. Nel dettaglio, si vogliono discutere le due proprietà legate a derivazione e moltiplicazione per la variabile, anche nel caso di trasformate di Fourier applicate alle distribuzioni.

Data $u \in S'(\mathbb{R}^n)$, data:

$$\partial_i = \frac{\partial}{\partial x_i}$$

allora:

1. La prima proprietà è:

$$\partial_i \hat{u} = \mathcal{F} \{j^{-1} x_i u\}$$

2. La seconda proprietà è:

$$\xi_i \hat{u} = j^{-1} \mathcal{F} \{\partial_i u\}$$

Verifichiamo la prima proprietà (la seconda si verifica in modo molto simile):

$$\partial_i \hat{u}(\varphi) = -\hat{u}(\partial_i \varphi)$$

nulla di nuovo: si scarica la derivata sulla funzione test, e si cambia il segno, come visto precedentemente; a questo punto, si può anche scaricare la trasformata di Fourier:

$$= -u(\mathcal{F}\{\partial_i\varphi\})$$

a questo punto, è possibile applicare la proprietà della trasformata di Fourier su φ , ottenendo:

$$= -u(j\xi_i\hat{\varphi}) = -j\xi_i u(\hat{\varphi})$$

Ora, è possibile applicare la trasformata di Fourier a tutto ciò, invece che alla sola funzione test (definizione di azione della trasformata di Fourier su funzioni test, ma “al contrario”):

$$= \mathcal{F}\{-j\xi_i u(\varphi)\}$$

volendo, poi, è possibile fare uno scambio di variabili: chiamare ξ_i come x_i , e scrivere:

$$\mathcal{F}\{-j\xi_i u(\varphi)\} \iff \mathcal{F}\{-jx_i u(\varphi)\}$$

7.4.7 Esempio pratico - Cenni ai teoremi di struttura

Si considera a questo punto un esempio pratico di applicazione: il caso $n = 1$, ossia \mathbb{R} ; si consideri $f(x_i) = x_i$, $f \in \mathbb{R}$. Calcoliamo la trasformata di Fourier di ciò:

$$\mathcal{F}\{x_i\}$$

questo si può vedere come segue:

$$\mathcal{F}\{x_i\} = \mathcal{F}\{x_i \times 1\} = j\partial_i(\mathcal{F}\{1\}) = j(2\pi)^n \partial_i \delta = j(2\pi)^n \delta'$$

Calcolando la trasformata di Fourier di un monomio di primo grado, è stata ottenuta la derivata prima della delta di Dirac, centrata in 0; con monomi di secondo grado, è possibile trovare la derivata seconda della delta, e così via. Questo suggerisce, per linearità, che, se si calcola la trasformata di Fourier di un polinomio di grado m , si ottiene proprio:

$$\mathcal{F} \left\{ \sum_{k=1}^m a_k x^k \right\} = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha \delta$$

ossia, si ottiene una combinazione lineare delle derivate delle delta di Dirac.

Esiste un teorema, detto **teorema di struttura per le distribuzioni a supporto in un punto**, che afferma che **ogni distribuzione** che abbia supporto in un punto (come, per esempio, la delta di Dirac), si può scrivere come una combinazione lineare finita di derivate della δ in quel punto; qua negli esempi si sta sempre considerando ciò nel punto $\{0\}$, ma ovviamente la cosa è valida per qualsiasi punto.

Esiste un teorema analogo, del quale si vuole solamente proporre un rapido cenno: ogni distribuzione a supporto compatto si può scrivere come una combinazione delle derivate di funzioni continue: questo è il teorema di struttura per distribuzioni a supporto compatto.

7.5 Spazi di Sobolev

Si vuole a questo punto concludere la trattazione proponendo una rapida introduzione agli spazi di Sobolev, al fine di motivarne la definizione e di proporre alcuni risultati fondamentali.

Uno degli obiettivi principali dell'Analisi Matematica è quello di studiare la regolarità delle funzioni; nell'Analisi in una variabile, per esempio, si può partire dallo studio di funzioni continue, per poi ottenerne di più regolari; la regolarità è legata alla derivabilità: quanto più una funzione è regolare, tanto più essa è derivabile; in altre parole, a seconda di quante derivate sono continue, la funzione sarà più "liscia", meno "frastagliata".

Al fine di classificare una funzione in termini della sua regolarità, viene definita una certa scala di spazi di funzioni: $C^{(k)}(\mathbb{R})$; per esempio, le funzioni continue sono quelle nel caso $k = 0$: $C^{(0)}$; le funzioni derivabili una volta sono in $C^{(1)}$; le funzioni derivabili k volte sono in $C^{(k)}$. L'operatore di derivazione, applicato a una funzione, si può vedere come un'applicazione:

$$\frac{d}{dx} : C^{(k)} \rightarrow C^{(k-1)}$$

quindi, esso riduce la regolarità della funzione di un ordine.

Questo, riguarda il caso scalare, ossia quando siamo in \mathbb{R} ; ciò permette di stabilire una teoria solida per quanto riguarda le equazioni differenziali ordinarie (ODE), basandosi su questi spazi $C^{(k)}$.

Questo tipo di classificazione, per le PDE (equazioni alle derivate parziali), non è accettabile; si consideri l'equazione alle derivate parziali

$$\Delta u = f$$

in \mathbb{R}^n . Se fossimo su \mathbb{R} , avremmo la certezza che, con $f \in C^{(0)}(\mathbb{R})$, si ha $u \in C^{(2)}(\mathbb{R})$; questa certezza crolla in \mathbb{R}^n , dal momento che non è detto che, in questo caso più generale, se $f \in C^{(0)}(\mathbb{R}^n)$, allora $u \in C^{(2)}(\mathbb{R}^n)$.

Ci si pone dunque il seguente problema: introdurre una classe di spazi con i quali si possa misurare, nel caso multidimensionale (\mathbb{R}^n), la regolarità delle funzioni; in altre parole, per esempio, Δ deve divenire, per questi spazi, una biiezione. Vedremo tra breve che questi spazi non son altro che gli spazi di Sobolev. Il motivo per cui essi vengono definiti in questo punto della trattazione, riguarda la trasformata di Fourier: uno dei modi per definire questi spazi è proprio attraverso la trasformata di Fourier delle funzioni che vi appartengono. Si definisce il seguente spazio di Sobolev:

$$H^s(\mathbb{R}^n) = \left\{ u \in S'(\mathbb{R}^n) : \left(\int (1 + |\xi|^2)^s |\hat{f}(\xi)|^2 d\xi \right)^{\frac{1}{2}} < \infty \right\}$$

dove s , come k nel caso delle funzioni unidimensionali, misura la regolarità della funzione: maggiore è s , più regolare è la funzione. Il termine:

$$\left(\int (1 + |\xi|^2)^s |\hat{f}(\xi)|^2 d\xi \right)^{\frac{1}{2}} \triangleq \|f\|_{H^s}$$

è detto **norma Sobolev**.

$\hat{f}(\xi)$ è una distribuzione temperata; di conseguenza, essa si può identificare come una funzione misurabile; questa, dovrà soddisfare, per appartenere allo spazio H^s , la condizione appena scritta. Si può attribuire intuitivamente un significato al termine:

$$(1 + |\xi|^2)^s$$

per s grande, il grafico di questa funzione in ξ si stringe, variando più rapidamente, in maniera più brusca; di conseguenza, più s è grande, più la

trasformata deve essere piccola, regolare, per “compensare” la crescita di questo termine; in altre parole, la \hat{f} , al crescere di s , dovrà essere molto piccola all’infinito, al fine di compensare la crescita di $(1 + |\xi|^2)^s$. Per la formula:

$$\mathcal{F}\{\partial_i f\} = jx_i f$$

più la s è grande, più la trasformata decade all’infinito, decrescendo più rapidamente, e dunque più è regolare.

Si vuole a questo punto spiegare dove sta il vantaggio nella definizione di questi spazi. Si definisca, mediante le seguenti **parentesi giapponesi**, un generico operatore:

$$\langle \xi \rangle = (1 + |\xi|^2)^{\frac{1}{2}}$$

dato un operatore $\langle D \rangle$ definito come:

$$\langle D \rangle^2 = 1 - \Delta = 1 - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \dots - \frac{\partial^2}{\partial x_n^2}$$

allora, è possibile dimostrare che questo $\langle D \rangle^2$ è un isomorfismo tra H^s e H^{s-2} ! Esattamente quello che si voleva! In altre parole:

$$\langle D \rangle^2 : H^s(\mathbb{R}^n) \rightarrow H^{s-2}(\mathbb{R}^n)$$

questo significa che, nel senso degli spazi di Sobolev, l’applicazione di questo operatore “abbassa la regolarità di due ordini”: esattamente quello che volevamo, per operatori definiti su \mathbb{R}^n ! Essendo un isomorfismo, essendo biiettivo, tutto ciò suggerisce che questi spazi sono assolutamente idonei al fine di studiare la regolarità su \mathbb{R}^n .

7.5.1 Casi particolari di spazi di Sobolev

Si vuole a questo punto studiare qualche caso particolare di spazio di Sobolev, al fine di trarre qualche conclusione aggiuntiva.

Il primo caso che può essere di interesse è $s = 0$, ossia $H^0(\mathbb{R}^n)$; in questo caso, si può vedere che:

$$H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$$

infatti, la definizione ricade nella semplice convergenza della trasformata di Fourier e, essendo la trasformata un isomorfismo di $L^2(\mathbb{R}^n)$, è sufficiente verificare la convergenza di:

$$\int_{\mathbb{R}^n} (1 + |\xi|^2)^0 |\hat{f}|^2 = \int_{\mathbb{R}^n} |\hat{f}|^2$$

e perché ciò sia soddisfatto, è sufficiente che la trasformata di Fourier esista.

Abbiamo già visto che più s è grande, più la condizione diventa “esigente”, nel senso che le funzioni appartenenti allo spazio di Sobolev saranno più regolari; si può tuttavia dire qualcosa in più, nel caso di $s = k \in \mathbb{N}$; in questo caso, infatti, si può dimostrare che la norma di Sobolev, H^k , è **equivalente** a:

$$\|u\|_{H^k} \equiv \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^2}$$

dove questa derivata è in senso **distribuzionale**.

Un altro caso di interesse è quello di $s < 0$; è possibile attribuire un significato a questi spazi dal momento che si può dimostrare che:

$$(H^s(\mathbb{R}^n))' = H^{-s}(\mathbb{R}^n)$$

ossia, il duale di $H^s(\mathbb{R}^n)$ altri non è che $H^{-s}(\mathbb{R}^n)$; questo è il **teorema di identificazione del duale di $H^s(\mathbb{R}^n)$** , dove per “duale” si intende “spazio di distribuzioni”.

Tutti questi spazi sono su \mathbb{R}^n ; verrebbe a questo punto spontaneo chiedersi se è possibile limitare l'indagine a un aperto $\Omega \subset \mathbb{R}^n$. Questo si può fare, ma non come è stato fatto finora; la definizione mediante trasformata di Fourier non è applicabile, dal momento che l'ambito in cui le trasformate di Fourier sono definite è quello delle distribuzioni temperate, di conseguenza il comportamento all'infinito deve essere tenuto in conto; non avendo senso definire la trasformata di Fourier su aperti, questa definizione non può essere fatta per questa via.

È possibile definire gli spazi di Sobolev su aperti Ω passando per la teoria dell'interpolazione, sviluppata a partire dal teorema di Riesz-Thorin precedentemente citato; facendo ciò nel caso di $s = k$, $k \in \mathbb{N}$, è possibile ottenere:

$$H^k(\Omega), k \in \mathbb{N} = \left\{ u \in L^2(\Omega) : \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^2(\Omega)} < \infty \right\}$$

dove, ancora, le derivate sono fatte in senso distribuzionale. Si definisce analogamente una norma Sobolev sull'aperto come:

$$\|u\|_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \|\partial^\alpha u\|_{L^2(\Omega)}$$

con la teoria dell'interpolazione, si può trattare anche il caso di s non intero.

Note sugli spazi comunemente usati nelle PDE

Degli spazi di Sobolev, solo un piccolo sottoinsieme è di solito utile nello studio della teoria delle equazioni alle derivate parziali. Un primo spazio molto usato è $H^1(\Omega)$; un secondo spazio spesso utilizzato è $H_0^1(\Omega)$; questo si definisce come:

$$H_0^1(\Omega) = \overline{C_c^{(\infty)}(\Omega)} \text{ in } H_1(\Omega)$$

ossia, $H_0^1(\Omega)$ è la chiusura di $C_c^{(\infty)}$, in H_1 ; in parole povere, questo è l'insieme delle funzioni in H_1 che si annullano al bordo $\partial\Omega$ dell'aperto.

Questo tipo di spazi è molto utile quando si studia un problema ai limiti di Dirichlet, per esempio del tipo:

$$\begin{cases} \Delta u = f & \text{in } \Omega \\ u = 0 & \text{su } \partial\Omega \end{cases}$$

in questo caso, la soluzione va cercata tra l'insieme delle funzioni che si annullano su $\partial\Omega$; da qui, l'utilità di questo spazio di funzioni. Si può dimostrare che il duale di questo spazio $H_0^1(\Omega)$ è:

$$(H_0^1(\Omega))' = H^{-1}(\Omega)$$

di conseguenza, la soluzione del problema va ricercata in $H^1(\Omega)$; poi, applicando due volte l'operatore di derivazione (in altre parole, applicando il Δ), quello che si ottiene è una "discesa" in H^{-1} : la regolarità è diminuita di due ordini. Questo conferma nuovamente quanto detto precedentemente.

Ambientando il lemma di Lax-Milgram in questi spazi, è possibile verificare l'esistenza e l'unicità di queste soluzioni, e dunque determinare se il problema è ben posto.