Norme su spazi vettoriali complessi

Consideriamo uno spazio vettoriale complesso, V, di dimensione finita e su cui sia stata fissata una base. Senza perdita di generalità, possiamo quindi supporre di lavorare con lo spazio vettoriale $\mathbb{C}^n$ con la base canonica $\mathcal{E} = \{e_1, \dots, e_n\}$. Ogni spazio vettoriale complesso è anche uno spazio vettoriale reale (di dimensione doppia); in particolare, identificheremo $\mathbb{R}^n$ con il sottospazio vettoriale reale, formato dai vettori $x_1e_1 + \cdots + x_ne_n$ per cui $x_1, \dots, x_n$ appartengono ad $\mathbb{R}$ ed indichiamo con Π l'insieme

$$\Pi = \{ \, x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^n \mid x_i \geq 0, \; i = 1, \dots, n \, \}$$

e con Π^0 il sottoinsieme in cui $x_i > 0$ per ogni $i = 1, \dots, n$. Sui vettori di Π (o, più in generale, sui vettori di $\mathbb{R}^n$) si può definire una relazione di ordine parziale, ponendo $x_1e_1 + \cdots + x_ne_n \leq y_1e_1 + \cdots + y_ne_n$ se, e solo se, $x_i \leq y_i$ per ogni $i = 1, \dots, n$. Possiamo estendere il valore assoluto complesso ad un'applicazione $| \; | : \mathbb{C}^n \to \Pi$ che manda il vettore $v = z_1e_1 + \cdots + z_ne_n \in \mathbb{C}^n$ sul vettore $|v| = |z_1|e_1 + \cdots + |z_n|e_n$. Si ha $|v + w| \leq |v| + |w|$ e $|cv| = |c||v|$, per ogni $v, w \in \mathbb{C}^n$ e per ogni $c \in \mathbb{C}$.

Sullo spazio vettoriale $\mathbb{C}^n$ si introducono delle norme per poter parlare di "distanza" tra i suoi punti ed avere una nozione di convergenza (limite).

Definizione. Una *norma* su $\mathbb{C}^n$ è una funzione $\| \; \| : \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ che gode delle seguenti proprietà:
- $\|v\| \geq 0$ e $\|v\| = 0 \Leftrightarrow v = 0$;
- $\|cv\| = |c|\|v\|$ per ogni scalare $c \in \mathbb{C}$;
- $\|v + w\| \leq \|v\| + \|w\|$.

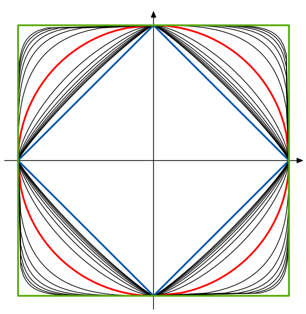
Fissata una norma, la *distanza* tra due vettori v, w di $\mathbb{C}^n$ è $\|v - w\|$ e una successione di vettori $(v_n)_{n \in \mathbb{N}}$ converge ad un vettore $v \in \mathbb{C}^n$ se, per ogni $\varepsilon > 0$ esiste un indice n_0 tale che $n \geq n_0 \Rightarrow \|v_n - v\| < \varepsilon$. Analogamente, si potrà parlare di convergenza di una serie di vettori, andando a studiare la convergenza della successione delle somme parziali dei termini della serie(†).

Fissata una norma, si definiscono le *sfere* (dette anche "palle") centrate in un punto $v_0 \in \mathbb{C}^n$ e di raggio r un numero reale positivo, come $B(v_0, r) = \{ \, v \in \mathbb{C}^n \mid \|v - v_0\| \leq r \, \}$ (si parlerà di *sfere aperte* e si scriverà $B^0(v_0, r)$ o $B(v_0, r^-)$, se si impone la disuguaglianza stretta $\|v - v_0\| < r$). Le proprietà della norma, ci dicono che $B(v_0, r) = v_0 + B(0, r)$ e che $v \in B(0, r)$ se, e solo se, $\frac{1}{r}v \in B(0, 1)$; quindi, la conoscenza della sfera unitaria $B(0, 1)$ fornisce informazioni sulla norma in questione. Diremo che due norme $\| \; \|_a$ e $\| \; \|_b$ sono *(topologicamente) equivalenti*, se esistono due costanti positive, k e K, tali che $k\|v\|_a \leq \|v\|_b \leq K\|v\|_a$ per ogni $v \in \mathbb{C}^n$. La relazione tra le due norme è ovviamente simmetrica, perché da quanto scritto si deduce $\frac{1}{K}\|v\|_b \leq \|v\|_a \leq \frac{1}{k}\|v\|_b$ per ogni $v \in \mathbb{C}^n$; ed è equivalente a dire che $B_b(0, k) \subseteq B_a(0, 1) \subseteq B_b(0, K)$ (o, analogamente, $B_a(0, 1/K) \subseteq B_b(0, 1) \subseteq B_a(0, 1/k)$), ove $B_a(0, r)$ è la sfera di raggio r nella norma $\| \; \|_a$ e $B_b(0, r)$ è la sfera di raggio r nella norma $\| \; \|_b$.

Esercizio. Si mostri che la relazione di equivalenza tra norme definita sopra è anche transitiva, ovvero che se $\| \; \|_a$ è equivalente a $\| \; \|_b$ e $\| \; \|_b$ è equivalente a $\| \; \|_c$, allora $\| \; \|_a$ è equivalente a $\| \; \|_c$. □

Esercizio. Siano $\| \; \|_a$ e $\| \; \|_b$ due norme equivalenti su $\mathbb{C}^n$. Si mostri che la successione $(v_n)_{n \in \mathbb{N}}$ converge a v per la norma $\| \; \|_a$ se, e solo se, converge a v per la norma $\| \; \|_b$. □

Vi è una famiglia di norme che usualmente si considera sullo spazio $\mathbb{C}^n$. Per ogni numero reale $p \geq 1$ (e per il simbolo ∞), per ogni vettore $v = z_1e_1 + \cdots + z_ne_n \in \mathbb{C}^n$, si pone



$$\|v\|_p = \left(\sum_{i=1}^{n} |z_i|^p \right)^{1/p} \qquad \text{e} \qquad \|v\|_\infty = \max_{i=1,\dots,n} |z_i|.$$

Dimostreremo che ciascuna di queste definizioni produce una norma, ovvero che valgono le proprietà descritte nella definizione. Osserviamo che, per ognuna delle norme date e per ogni vettore v, si ha $\||v|\|_p = \|v\|_p$, ove $|v|$ indica l'immagine di v

(†) Si può parlare di *successioni divergenti* e dire che $\lim_{n \to \infty} v_n = \infty$ se, per ogni $K > 0$ esiste un indice n_0 tale che $n \geq n_0 \Rightarrow \|v_n\| > K$. Come per le successioni e le serie numeriche, ci sono successioni che non sono né convergenti, né divergenti.

in Π. Abbiamo rappresentato sopra le intersezioni con $\mathbb{R}^2$ di alcune delle sfere unitarie in $\mathbb{C}^2$. In particolare, in azzurro compare la sfera per la norma $\|\ \|_1$; in rosso la sfera per la norma $\|\ \|_2$, che è la *norma euclidea* (o hermitiana), indotta dal consueto prodotto scalare hermitiano $\langle\ |\ \rangle$; ed in verde la sfera per la norma $\|\ \|_\infty$. Le altre si dispongono in ordine crescente, al crescere di p, nel modo descritto dal disegno. Come il disegno lascia intuire, le norme sono tutte topologicamente equivalenti. Infatti si ha $\|v\|_\infty \leq \|v\|_p \leq \sqrt[p]{n}\|v\|_\infty$, per ogni $v \in \mathbb{C}^n$. Quindi, qualunque norma si usi, non cambia il limite delle successioni in $\mathbb{C}^n$.

Le prime due proprietà delle norme sono di verifica immediata, mentre non è evidente la proprietà di subadditività, se si escludono forse la norma $\|\ \|_1$ e la norma $\|\ \|_\infty$ (esercizio!). Nel caso della norma euclidea quella proprietà si otteneva dalla disuguaglianza di Cauchy-Schwartz. Nel caso delle norme $\|\ \|_p$, la relazione $\|v+w\|_p \leq \|v\|_p + \|w\|_p$ prende anche il nome di *disuguaglianza di Minkowski* e la sua dimostrazione sarà l'oggetto del seguito di queste note.

Cominciamo osservando che, dato $p > 1$, allora esiste un unico numero $q = \frac{p}{p-1}$ tale che $\frac{1}{p} + \frac{1}{q} = 1$. La relazione tra p e q si potrebbe formalmente estendere prendendo $p = 1$ e $q = \infty$ (o $p = \infty$ e $q = 1$) ed in molti contesti ciò viene fatto. Per quanto riguarda la disuguaglianza di Minkowski, il risultato per le due norme "estreme" può essere ottenuto in modo diretto e quindi nel seguito ci limiteremo a supporre che p appartenga alla semiretta $(1, \infty)$.


Otto Ludwig Hölder (1859-1937)
Hermann Minkowski (1864-1909)

Osservazione. [disuguaglianza di Young] *Siano a e b due numeri reali non-negativi e $p, q \in (1, +\infty)$ tali che $\frac{1}{p} + \frac{1}{q} = 1$. Allora si ha $ab \leq \dfrac{a^p}{p} + \dfrac{b^q}{q}$.*

dim. Se uno dei due numeri è uguale a zero, la disuguaglianza è banalmente vera. Se a e b sono entrambi positivi, il risultato discende dal fatto che il logaritmo è una funzione crescente e concava (in ogni punto il suo grafico è al di sotto della retta tangente) sulla semiretta dei numeri reali positivi. Dati i due punti del grafico $(x, \log x)$ e $(y, \log y)$, il punto $t(x, \log x) + (1-t)(y, \log y)$ del segmento che li congiunge sta quindi al di sotto del relativo punto del grafico del logaritmo, per ogni numero reale $t \in [0, 1]$. Ovvero si ha $\log(tx + (1-t)y) \geq t \log x + (1-t) \log y$. In particolare, presi $t = \frac{1}{p}$, $x = a^p$, $y = b^q$, si ottiene da ciò

$$\log(\tfrac{1}{p}a^p + \tfrac{1}{q}b^q) \geq \tfrac{1}{p}\log(a^p) + \tfrac{1}{q}\log(b^q) = \log ab,$$

che è quanto volevamo. **CVD** □

Grazie alla precedente osservazione possiamo dimostrare la seguente generalizzazione della disuguaglianza di Cauchy-Schwartz

Proposizione. [disuguaglianza di Hölder] *Siano $p, q \in (1, +\infty)$ tali che $\frac{1}{p} + \frac{1}{q} = 1$. Presi comunque, v e w in $\mathbb{C}^n$, si ha*

$$|\langle v|w\rangle| \leq \|v\|_p\,\|w\|_q.$$

dim. Se uno dei due vettori è nullo, la disuguaglianza è banalmente vera. Supponiamo quindi che i due vettori siano non nulli e sostituiamo ai vettori dati i vettori $\frac{v}{\|v\|_p}$ e $\frac{w}{\|w\|_q}$. Supponiamo quindi $v = \sum_{i=1}^n x_i e_i$ e $w = \sum_{i=1}^n y_i e_i$ con $\|v\|_p = 1 = \|w\|_q$. Si hanno le disuguaglianze

$$|\langle v|w\rangle| = \left|\sum_{i=1}^n \bar{x}_i y_i\right| \leq \sum_{i=1}^n |x_i||y_i| \leq \sum_{i=1}^n \left(\frac{|x_i|^p}{p} + \frac{|y_i|^q}{q}\right) = \frac{{\|v\|_p}^p}{p} + \frac{{\|w\|_q}^q}{q} = \frac{1}{p} + \frac{1}{q} = 1;$$

la seconda delle quali è un'applicazione della disuguaglianza di Young. **CVD** □

Prendendo $p = q = 2$, si ottiene come caso particolare un'altra dimostrazione della disuguaglianza di Cauchy-Schwartz.

Possiamo concludere il percorso con la

Proposizione. [disuguaglianza di Minkowski] *Sia $p \in (1, +\infty)$. Presi comunque, v e w in $\mathbb{C}^n$, si ha*

$$\|v + w\|_p \leq \|v\|_p + \|w\|_p.$$

dim. Siano $v = \sum_{i=1}^n x_i e_i$ e $w = \sum_{i=1}^n y_i e_i$. Se uno dei due vettori è nullo la disuguaglianza è banalmente vera, così come se la somma $v + w = 0$. Supponiamo quindi $v + w \neq 0$ e osserviamo che si ha

$$\|v + w\|_p^p = \sum_{i=1}^n |x_i + y_i|^p \leq \sum_{i=1}^n (|x_i| + |y_i|)|x_i + y_i|^{p-1} = \langle |v| \, |z\rangle + \langle |w| \, |z\rangle,$$

ove $|v|, |w| \in \Pi$ e $z = \sum_{i=1}^n |x_i + y_i|^{p-1} e_i$.

Ricordando che per ogni vettore $u \in \mathbb{C}^n$ si ha $\|v\|_p = \||v|\|_p$, applicando la disuguaglianza di Hölder si ottiene

$$\|v + w\|_p^p \leq (\|v\|_p + \|w\|_p) \|z\|_q,$$

ove $q = \frac{p}{p-1}$ e quindi

$$\|z\|_q = \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{\frac{p-1}{p}} = \|v + w\|_p^{p-1}.$$

Dividendo entrambo i termini della disuguaglianza per $\|v + w\|_p^{p-1}$, si ottiene quanto desiderato. **CVD** □

Esercizio. Sia $\zeta : \mathbb{C}^n \to \mathbb{C}$ una forma lineare ($\zeta \in \mathbb{C}^{n*}$).

(a) Si dimostri che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $\|v - w\|_\infty < \delta \Rightarrow |\zeta(v) - \zeta(w)| < \varepsilon$.

(b) Si dimostri che quanto affermato in (a) resta vero per qualsiasi norma $\| \ \|_p$.

(c) Si concluda che ζ è una funzione (uniformemente) continua e quindi che $\lim_{k\to\infty} v_k = v_0$ per la norma $\| \ \|_p$ in $\mathbb{C}^n$ implica $\lim_{k\to\infty} \zeta \circ v_k = \zeta \circ v_0$. □

Le norme in questione (ed ora possiamo chiamarle così con proprietà di linguaggio) sono tutte equivalenti e un vettore $v_0 \in \mathbb{C}^n$ è il limite di una successione di vettori $(v_k)_{k\in\mathbb{N}}$ (rispetto ad una qualsiasi delle norme precedenti), se, e solo se, la successione converge a v_0 coordinata per coordinata, ovvero $\lim_{k\to\infty} e_i^* \circ v_k = e_i^* \circ v_0$ per $i = 1, \dots, n$; ove $\mathcal{E}^* = \{e_1^*, \dots, e_n^*\}$ è la base duale della base canonica. Inoltre, è facile verificare che la convergenza per ciascuna delle coordinate di una successione di vettori implica la convergenza nelle norme date.

Esercizio. Si verifichi che una successione di vettori $(v_k)_{k\in\mathbb{N}}$ in $\mathbb{C}^n$ converge nella norma $\| \ \|_p$ se, e solo se, soddisfa alla condizione di Cauchy: per ogni $\varepsilon > 0$ esiste un indice n_0 tale che, per $n, m \geq n_0$ si ha $\|v_n - v_m\|_p < \varepsilon$. □

Per quanto visto, può sembrare forse eccessivo considerare su $\mathbb{C}^n$ l'intera famiglia di norme tra loro equivalenti, dato che non differiscono i limiti delle successioni passando dall'una all'altra. Vi è però modo di generalizzare la situazione passando a spazi di dimensione infinita e, in questo caso, la peculiarità di ciascuna norma assume rilevanza. Diamo un rapido cenno alla questione, senza nessuna pretesa di riuscire anche solo ad introdurla in modo soddisfacente.

Lo spazio $\mathbb{C}^n$, al variare di n, può essere identificato con lo spazio vettoriale $\mathscr{F}_{I_n}(\mathbb{C})$ delle funzioni a valori complessi definite sull'insieme $I_n = \{1, 2, \dots, n\}$, identificando il vettore $v = x_1 e_1 + \cdots + x_n e_n$ con la funzione $i \mapsto x_i$, $i = 1, \dots, n$. In particolare, il vettore e_i della base canonica si identifica con la funzione $e_i : I_n \to \mathbb{C}$, definita da $e_i(j) = \delta_{ij}$ (simbolo di Kronecker), per $i = 1, \dots, n$. Al variare di n, lo spazio $\mathscr{F}_{I_n}(\mathbb{C})$ si può immergere in $\mathscr{F}_{I_{n+1}}(\mathbb{C})$, identificando le funzioni $e_1, \dots, e_n$ del primo spazio con gli omonimi elementi del secondo (che differenza c'è tra una funzione e la sua immagine?). In questo modo si può parlare dell'unione di tutti questi spazi che viene a formare lo spazio $\mathscr{F}_{\mathbb{N}}^0(\mathbb{C})$, delle successioni $(x_n)_{n\in\mathbb{N}}$ *quasi sempre nulle*, ovvero quelle per cui esiste un intero n_0 tale che $x_n = 0$ per $n > n_0$. Questo spazio è uno spazio vettoriale complesso, di dimensione infinita, che ha come base le funzioni $e_i : \mathbb{N} \to \mathbb{C}$, per $i = 1, 2, \dots$, ove $e_i(j) = \delta_{ij}$ per ogni $j \in \mathbb{N}$. Anche su questo spazio possiamo mettere le norme $\| \ \|_p$ per $p \in [1, \infty]$ che hanno le proprietà che avevano in $\mathbb{C}^n$ (dati due vettori, di questo spazio posso sempre considerarli come elementi di un $\mathbb{C}^{n_0}$ per n_0 abbastanza elevato). Quello che accade di diverso è che posso avere successioni di elementi di $\mathscr{F}_{\mathbb{N}}^0(\mathbb{C})$ che convergono ad una successione che non sta più in questo spazio (perché ha infiniti termini diversi da 0).

Esercizio. Si consideri la successione $(v_n)_{n\in\mathbb{N}}$ in $\mathscr{F}^0_{\mathbb{N}}(\mathbb{C})$, definita ponendo $v_n = \sum_{i=1}^{n} \frac{1}{i^\alpha} e_i$ per ogni numero naturale $n \geq 1$. Si studi per quali valori di α la successione converge nella norma $\| \ \|_p$ ad un elemento di $\mathscr{F}^0_{\mathbb{N}}(\mathbb{C})$ oppure ad una successione di numeri complessi al di fuori di questo spazio. $\square$

Per ogni p tra 1 ed ∞, l'insieme $\ell^p_{\mathbb{C}}$ (spazio vettoriale, in realtà) indica la "chiusura" dello spazio $\mathscr{F}^0_{\mathbb{N}}(\mathbb{C})$ rispetto alla norma $\| \ \|_p$ (per la definizione precisa rimandiamo ad un successivo corso di Analisi Matematica), ovvero alle successioni quasi sempre nulle si aggiungono quelle che si ottengono come limite nella norma fissata.

Una volta arrivati alle successioni, ovvero alle funzioni da $\mathbb{N}$ su $\mathbb{C}$, viene naturale pensare di estendere le costruzioni presentate a più generali funzioni a valori complessi, definite su intervalli di $\mathbb{R}$ o su più generali sottoinsiemi di $\mathbb{R}^n$, ... ma questo passaggio, ben più del precedente, richiede un supplemento di lavoro per rendere sufficientemente agili e generali le tecniche dell'Analisi Matematica (ad esempio, le somme che definiscono la norma, dovranno essere sostituite da integrali sull'insieme di definizione delle funzioni che si considerano e sarà rilevante "quale integrale" usare). Tutto ciò non può certo essere fatto nello spazio di queste pagine.