Lezione: Spazi vettoriali complessi

$\boxed{\text{Numeri complessi:}}$ $\mathbb{C} = \mathbb{R}^2$ come un spazio vettoriale reale, dotato della seguente operazione di moltiplicazione vettoriale:

$$(a,b)(c,d) = (ac - bd, ad + bc) \qquad \text{(E1)}$$

Meno formalmente, se scriviamo $1 = (1,0), \quad i = (0,1)$ allora $\mathbb{C} = \{a + b\,i \mid a,\, b \in \mathbb{R}\}$ dove $a + ib = a1 + ib = (a,b)$ e

$$\begin{cases} (a+ib)(c+id) = (ac - bd) + i(ad + bc) \\ (a+ib) + (c+id) = (a+c) + i(b+d) \end{cases} \qquad \text{(E2)}$$

Proprietà: $z_1,\, z_2,\, z_3 \in \mathbb{C}, \quad 0 = (0,0)$

1. $z_1 + z_2 = z_2 + z_1$
2. $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$
3. $z_1 z_2 = z_2 z_1$
4. $(z_1 z_2) z_3 = z_1 (z_2 z_3)$
5. $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$
6. $z_1 + 0 = z_1, \quad (z_1)(1) = z_1$
7. $\forall z_1\ \exists!\ -z_1$ tale che $z_1 + (-z_1) = 0$
8. $\forall z_1 \neq 0\ \exists!\ (z_1)^{-1}$ tale che $z_1 (z_1)^{-1} = 1$

Proprietà 7: Chiaramente $-(a+ib) = (-a) + i(-b)$
Proprietà 8: Si trova che,

$$\frac{1}{a+ib} = \frac{a - ib}{a^2 + b^2}, \quad \text{se } (a,b) \neq (0,0) \qquad \text{(E3)}$$

Coniugazione complessa: $\overline{a+ib} = a - ib$ (E4)

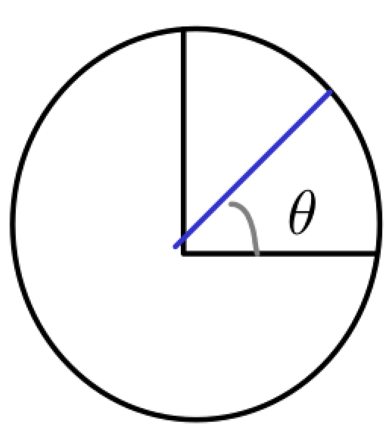
Valore assoluto (anche detto norma): $|a+ib| = \sqrt{a^2 + b^2}$ (E5)

Proprietà:

9. $|\overline{a+ib}| = |a+ib|$
10. $|z|^2 = z\bar{z}$
11. $|z_1 + z_2| \leq |z_1| + |z_2|$
12. $|z_1 \pm z_2| \geq ||z_1| - |z_2||$

Nota: Proprietà 10 implica $z = \dfrac{\bar{z}}{|z|^2}, \quad z \neq 0$

Forma Polare: Se $|x+iy| = 1$ allora $x + iy = \cos(\theta) + i\sin(\theta)$. La quantità θ è ben definita a meno di multipli interi di 2π.



$(x,y) = (\cos(\theta), \sin(\theta))$

$x + iy = \cos(\theta) + i\sin(\theta)$

In generale, se $|z| = r \neq 0$ allora $z/|z|$ ha valore assoluto 1. Quindi $z/|z| = \cos(\theta) + i\sin(\theta)$.
Insomma:

$$z = r(\cos(\theta) + i\sin(\theta)) \qquad \text{(E6)}$$

Chiamiamo (E6) la forma polare di z.

Esempio Chiave:

$$\begin{aligned}(\cos(\alpha)+i\sin(\alpha))(\cos(\beta)+i\sin(\beta)) &= (\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta))+i(\sin(\alpha)\cos(\beta)+\sin(\beta)\cos(\alpha) \\ &= \cos(\alpha+\beta)+i\sin(\alpha+\beta)\end{aligned}$$

Quindi,

$$\left.\begin{aligned} z_1 &= r_1(\cos(\theta_1)+i\sin(\theta_1)) \\ z_2 &= r_2(\cos(\theta_2)+i\sin(\theta_2)) \end{aligned}\right\} \implies z_1 z_2 = r_1 r_2(\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)) \qquad \text{(E7)}$$

Formula di de Moivre: $(\cos(\theta)+i\sin(\theta))^n = \cos(n\theta)+i\sin(n\theta)$

Esempio: Dato un intero positivo n le soluzioni dell'equazione $z^n = 1$ sono chiamate le radici n-essime dell'unità. Chiarmente

$$z^n = 1 \implies |z|^n = 1 \implies |z| = 1 \implies z = \cos(\theta)+i\sin(\theta)$$

Allora

$$z^n = \cos(n\theta)+i\sin(n\theta) = 1 \implies \cos(n\theta) = 1, \quad \sin(n\theta) = 0$$

Quindi $n\theta = 2\pi m$ per qualche intero m. Poiché $z^n = 1$ ha n soluzioni, vediamo che

$$z = \cos(2\pi k/n)+i\sin(2\pi k/n), \qquad k = 0,\ldots,n-1 \qquad \text{(E8)}$$

è la lista completa delle radici n-essime dell'unità.

Esempio: Le radici terze dell'unità sono $z = \cos(0)+i\sin(0) = 1$ e

$$z = \cos(2\pi/3)+i\sin(2\pi i/3) = \frac{-1}{2}+i\frac{\sqrt{3}}{2}, \qquad z = \cos(4\pi/3)+i\sin(4\pi/3) = \frac{-1}{2}-i\frac{\sqrt{3}}{2}$$

Nota che

$$z^n = 1 \implies (\bar{z})^n = \overline{z^n} = \bar{1} = 1$$

(Ricorda: Le radici complesse di un polinomio con coefficienti reali vengono in coppie coniugate)

Formula di Eulero: Per ogni numero reale θ, $e^{i\theta} = \cos(\theta)+i\sin(\theta)$. In particolare, l' equazione (E7) diventa:

$$z_1 = r_1 e^{i\theta_1}, \quad z_2 = r_2 e^{i\theta_2} \implies z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$$

e l'equazione (E8) è $z = e^{2\pi i k/n}, \quad k = 0,\ldots,n-1$.

Polinomi:

Formalmente, un polinomio con coefficienti complessi è una mappa $f:\{0\}\cup\mathbb{N}\to\mathbb{C}$ tale che $\{n \mid f(n)\neq 0\}$ è un insieme finito.

L'addizione polinomiale è definita dalla regola: $(f+g)(n)=f(n)+g(n)$.

La moltiplicazione polinomiale è definita dalla regola:

$$(f*g)(n)=\sum_{k=0}^{n} f(k)g(n-k)$$

La mappa f corrisponde a un polinomio $p_f(z)$ nel senso usuale della regola

$$p_f(z)=\sum_{\{k:f(k)\neq 0\}} f(k)z^k, \qquad p_f(z)+p_g(z)=p_{f+g}(z), \qquad p_f(z)p_g(z)=p_{f*g}(z)$$

Il grado $\deg(f)$ di $f\neq 0$ è il massimo dei valori di n tale che $f(n)\neq 0$.

<u>Teorema fondamental dell'algebra</u>: Sia f un polinomio non costante con coefficienti reali o complessi di grado $n\geq 1$. Allora, f ha esattamente n radici sui numeri complessi, quando contate con molteplicità.

In altre parole, sia

$$f(x)=f_n x^n+\cdots+f_0, \quad f_n\neq 0 \implies f(x)=f_n(x-r_1)\cdots(x-r_n) \tag{E9}$$

Se il fattore $(x-\lambda)$ appare k volte nell'equazione (E9), diciamo che λ è una radice di f con molteplicità k. Allora, λ è una radice di f con molteplicità k se e solo se

$$f(x)=(x-\lambda)^k g(x), \qquad g(\lambda)\neq 0$$

Polinomio caratteristico:

Sia A una matrice $n\times n$ con numeri reali o complessi come voci. Allora, il polinomio caratteristico di A è

$$p_A(t)=\det(tI_n-A) \tag{E10}$$

<u>Esempio</u>:

$$A=\begin{pmatrix}2 & 3\\ 1 & 2\end{pmatrix} \implies p_A(t)=\det\begin{pmatrix}t-2 & -3\\ -1 & t-2\end{pmatrix}=t^2-4t+1$$

<u>Lemma</u>: Se A e B sono matrici simili allora $p_A(t)=p_B(t)$.
<u>Dimostrazione</u>:

$$A=C^{-1}BC \implies tI_n-A=tI_n-C^{-1}BC=C^{-1}(tI_n-B)C$$
$$\therefore\ p_A(t)=\det(tI_n-A)=\det(C^{-1}(tI_n-B)C)=\det(CC^{-1}(tI_n-B))=p_B(t)$$

Esempio: $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$

$$\implies p_A(t) = t^2 - 4t + 1, \qquad p_B(t) = t^2 - 4t - 1$$

Dunque, A e B non sono matrici simili.

Si potrebbe sperare che viceversa, se due matrici hanno lo stesso polinomio caratteristico, allora sono simili. L'esempio seguente dimostra che questo non è vero:

Esempio:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \implies p_A(t) = p_B(t) = t^2$$

Se A e B sono simili allora esiste una matrice C tale che $B = C^{-1}AC$. Tuttavia, poiché A è la matrice identità, $C^{-1}AC = A$. Quindi $B = A$, che è falso.

In somma:
(i) Se A e B sono simili, allora A e B hanno lo stesso polinomio caratteristico.
In particolare, se due matrici hanno polinomi caratteristici diversi non sono simili
(ii) Il viceversa (matrici con lo stesso polinomio caratteristico sono simili) non è vero.

Polinomio caratteristico di una mappa lineare: Sia U uno spazio vettoriale a dimensione finita e $L : U \to U$ sia una mappa lineare. Sia A (risp. A') la matrice di L rispetto ad una scelta di base B (risp. B') di U. Allora, A e A' sono matrici simili, quindi $p_A(t) = p_{A'}(t)$.

Si definisce il polinomio caratteristico di L essere il polinomio caratteristico di A che rappresenta L dopo la scelta di una base per U.

Esempio: $U = P_1[x], \quad L(p) = x\dfrac{dp}{dx} + p(x)$

$$B = \{1, x\} \implies A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \text{ è la matrice di L}$$

$$p_L(t) = p_A(t) = (t-1)(t-2)$$

Teorema di Hamilton-Cayley: Sia A una matrice quadrata con polinomio caratteristico $p(t)$. Allora $p(A) = 0$.
Dimostrazione: Questo può essere dimostrato usando idee simili allo sviluppo di Laplace.

Per chiarire:

$$p(t) = p_n t^n + \cdots + p_1 t + p_0 \implies p(A) = p_n A^n + \cdots + p_1 A + p_0 A^0, \qquad A^0 = I$$

Esempio:

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \implies p_A(t) = t^2 - 4t + 1$$

$$A^2 = \begin{pmatrix} 7 & 12 \\ 4 & 7 \end{pmatrix} \implies A^2 - 4A + I = \begin{pmatrix} 7 & 12 \\ 4 & 7 \end{pmatrix} - \begin{pmatrix} 8 & 12 \\ 4 & 8 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Sia $A=(a_{ij})$ una matrice $n \times n$ (voci reali o complesse). Per $i=1,\dots,n$ sia

$$R_i = \sum_{j \neq i} |a_{ij}|$$

e sia

$$D_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i\}$$

Ciascuno di questi dischi è chiamato disco di Gershgorin di A.

Il seguente risultato fornisce la posizione approssimativa delle radici del polinomio caratteristico:

<u>Teorema del cerchio di Gershgorin</u>: Sia A una matrice quadrata con polinomio caratteristico p. Allora, ogni radice di p è contenuta in almeno uno dei dischi di Gershgorin di A.

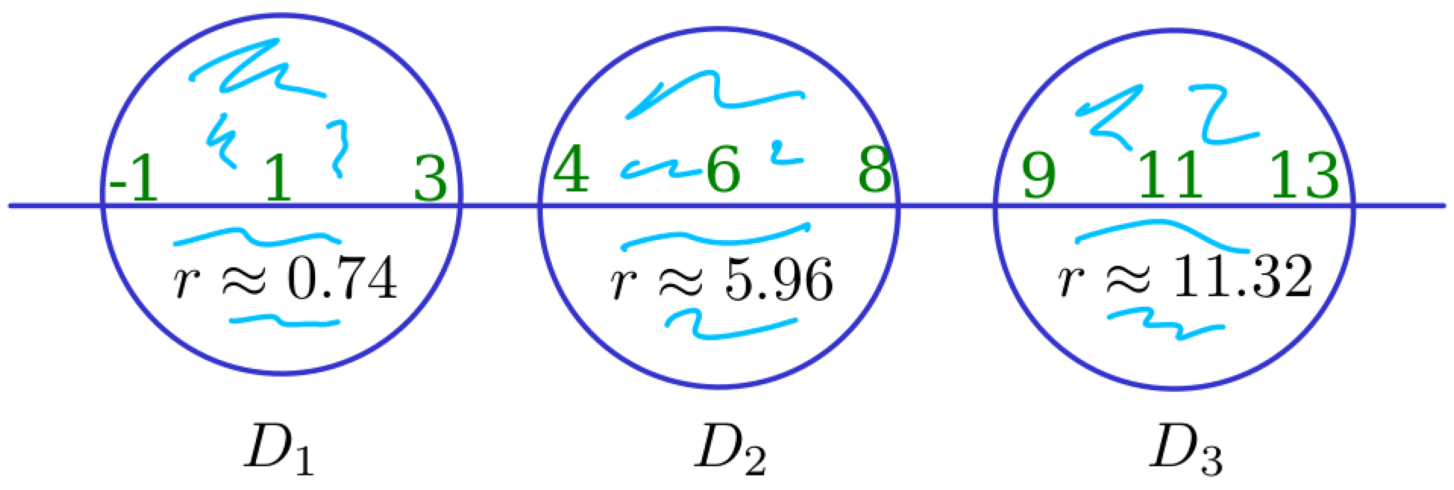
Questo teorema può essere migliorato come segue:

<u>Teorema</u>: Supponiamo che S sia l'unione di k dischi di Gershgorin di A e che S' sia l'unione di $n-k$ dischi di Gershgorin di A. Se S e S' sono disgiunti allora S contiene k radici di p_A e S' contiene $n-k$ radici di p_A.

In particolare, un disco Gershgorin che è disgiunto da ogni altro disco Gershgorin contiene una radice del polinomio caratteristico.

<u>Esempio:</u> Verificare che il polinomio caratteristico della seguente matrice ha tre radici distinte

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 6 & 1 \\ 1 & 1 & 11 \end{pmatrix} \implies R_1 = R_2 = R_3 = 2$$



Ogni disco è disgiunto dal resto, e quindi contiene esattamente una radice del polinomio caratteristico.

Spazi vettoriali complessi:

In analogia con gli spazi vettoriali reali (gli oggetti che abbiamo studiato finora), uno spazio vettoriale complesso è un insieme V dotato di due operazioni:

(i) Addizione vettoriale

$$V \times V \to V, \qquad (u,v) \mapsto u+v$$

(ii) Moltiplicazione per scalari complessi.

$$\mathbb{C} \times V \to V, \qquad (c,v) \mapsto c \cdot v \quad \text{(di solito scritto) } cv$$

Queste due operazioni devono soddisfare gli 8 assiomi dello spazio vettoriale elencati a pagina 5 della lezione 5, dove gli scalari reali sono sostituiti con scalari complessi.

Esempi di base di spazi vettoriali complessi:

(1) $\mathbb{C}^n$, $(z_1, \dots, z_n) + (w_1, \dots, w_n) = (z_1 + w_1, \dots, z_n + w_n)$

$c(z_1, \dots, z_n) = (cz_1, \dots, cz_n)$

$\text{Base} = \{e_1, \dots, e_n\}, \quad e_1 = (1, 0, \dots, 0), \ \dots, \ e_n = (0, \dots, 0, 1)$

(2) Polinomi con coefficienti complessi $\mathbb{C}[z]$, $\text{Base} = \{1, z, z^2, \dots\}$

(3) $P_n[z]$ = Polinomi con coefficienti complessi e grado minore o uguale a n (incluso lo zero)

$\text{Base} = \{1, z, \dots, z^n\}$

(4) $M_{n \times m}$ = matrici $n \times m$ con coefficienti complessi. Addizione e moltiplicazione scalare componente per componente, proprio come $\mathbb{C}^n$.

$\text{Base} = \{E_{ij} \mid 1 \le i \le n, \ 1 \le j \le m\}$

(Ogni voce di E_{ij} è zero tranne il posto (i,j) che è 1.)

(5) S è un insieme. $\mathbb{C}^S = \{\text{funzioni } f : S \to \mathbb{C}\}$

Se S è un insieme finito, allora le funzioni indicatrici

$$\chi_s(t) = \begin{cases} 1, & t = s \\ 0, & t \ne s \end{cases}$$

formano una base $\{\chi_s \mid s \in S\}$ di $\mathbb{C}^S$.

Ogni parte del corso che abbiamo discusso finora e che utilizza solo gli assiomi di uno spazio vettoriale si applica agli spazi vettoriali complessi.

Esempi: Sottospazi, span, indipendenza lineare, base, dimensione, eliminazione gaussiana, gli algoritmi per trovare una base per lo spazio riga, immagine e kernel, il teorema del rango, prodotti di spazi vettoriali, la formula di Grassmann, la somma di sottospazi e l'algoritmo di Zassehaus.

Cosa non è valido (senza modifiche): Tutto ciò che coinvolge il prodotto scalare.

Ingenuamente,

$$((1, i), (1, i)) = 1^2 + i^2 = 0$$

Quindi, tutto ciò che abbiamo fatto (ad esempio la disuguaglianza di Cauchy-Schwarz) che richiede $(v, v) \ge 0$ e $(v, v) = 0$ se e solo $v = 0$ fallisce per gli spazi vettoriali complessi.

Per risolvere questo problema, introduciamo la nozione di prodotto hermitiano.
(pagina seguente)

L'esempio base di un prodotto hermitiano è $\mathbb{C}^n$ dotato della mappa

$$\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \qquad \langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = z_1 \bar{w}_1 + \cdots + z_n \bar{w}_n$$

Con questa scelta, si ha:

(A1) $$\begin{aligned} \langle u+v, w \rangle &= \langle (u_1, \ldots, u_n) + (v_1, \ldots, v_n), (w_1, \ldots, w_n) \rangle \\ &= \langle (u_1 + v_1, \ldots, u_n + v_n), (w_1, \ldots, w_n) \rangle \\ &= (u_1 + v_1)\bar{w}_1 + \cdots + (u_n + v_n)\bar{w}_n \\ &= (u_1 \bar{w}_1 + \cdots + u_n \bar{w}_n) + (v_1 \bar{w}_1 + \cdots + v_n \bar{w}_n) \\ &= \langle u, w \rangle + \langle v, w \rangle \end{aligned}$$

$$\begin{aligned} \langle \lambda u, w \rangle &= \langle \lambda (u_1, \ldots, u_n), (w_1, \ldots, w) \rangle \\ &= \lambda u_1 \bar{w}_1 + \cdots + \lambda u_n \bar{w}_n = \lambda \langle u, w \rangle \end{aligned}$$

(A2) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ a causa di:

$$\begin{aligned} \overline{\langle v, u \rangle} &= \overline{v_1 \bar{u}_1 + \cdots + v_n \bar{u}_n} \\ &= \bar{v}_1 u_1 + \cdots + \bar{v}_n u_n = \langle u, v \rangle \end{aligned}$$

Nota: Ciò implica che $\langle u, \lambda v \rangle = \overline{\langle \lambda v, u \rangle} = \overline{\lambda \langle v, u \rangle} = \bar{\lambda} \overline{\langle v, u \rangle} = \bar{\lambda} \langle u, v \rangle$

(A3) $\langle z, z \rangle = \langle (z_1, \ldots, z_n), (z_1, \ldots, z_n) \rangle = |z_1|^2 + \cdots + |z_n|^2 \geq 0$

$\langle z, z \rangle = 0 \iff z = 0$

Adottiamo queste tre condizioni come assiomi di un prodotto hermitiano.

Norm. Se V è uno spazio vettoriale complesso dotato di un prodotto hermitiano definiamo:

$$|v| = \sqrt{\langle v, v \rangle} \geq 0$$

Teorema (Disuguaglianza di Cauchy-Schwarz): Se V è uno spazio vettoriale complesso dotato di un prodotto hermitiano allora:

$$|\langle u, v \rangle| \leq |u||v|$$

(e l'uguaglianza implica che un vettore è un multiplo scalare dell'altro)

Teorema (Disuguaglianza del Triangolo): Sia V uno spazio vettoriale complesso dotato di un prodotto hermitiano. Allora,

$$|u + v| \leq |u| + |v|$$

Per via di questi due teoremi:

(a) Possiamo definire gli angoli usando la solita formula

$$\langle u, v \rangle = |u||v| \cos\theta, \qquad (u \neq 0, \ v \neq 0)$$

(b) La funzione norma soddisfa le seguenti condizioni:

(N1) $|v| \geq 0, \qquad |v| = 0 \iff v = 0$

(N2) $|\lambda v| = |\lambda||v|, \qquad \lambda \in \mathbb{C}, \ v \in V$

(N3) $|u + v| \leq |u| + |v|$

Le condizioni (N1)-(N3) sono gli assiomi di una norma su uno spazio vettoriale complesso. È spesso più facile costruire una norma che un prodotto hermitiano

Esempio:

$$v \in \mathbb{C}^n, \quad \|(v_1, \dots, v_n)\| = \max\{|v_1|, \dots, |v_n|\}$$

(N1): $\|v\| = \max\{|v_1|, \dots, |v_n|\} \geq 0, \quad \max\{|v_1|, \dots, |v_n|\} = 0 \implies v = 0$

(N2): $\|\lambda v\| = \max\{|\lambda v_1|, \dots, |\lambda v_n|\} = \lambda \max\{|v_1|, \dots, |v_n|\} = |\lambda| \|v\|$

(N3): $\|u + v\| = \max\{|u_1 + v_1|, \dots, |u_n + v_n|\} \leq \max\{|u_1|, \dots, |u_n|\} + \max\{|v_1|, \dots, |v_n|\}$
$= \|u\| + \|v\|$

Esempio: $v \in \mathbb{C}^n, \qquad \|(v_1, \dots, v_n)\| = |v_1| + \cdots + |v_n|$

Trasposta Coniugata: Sia A una matrice nxm allora A^* è la matrice mxn

$$A^* = (\bar{A})^t, \qquad (a_{ij})^* = (\alpha_{ij}), \quad \alpha_{ij} = \bar{a}_{ji}$$

Lemma: Rispetto al prodotto hermitiano standard su $\mathbb{C}^n$,

$$\langle (z_1, \dots, z_n), (w_1, \dots, w_n) \rangle = z_1 \bar{w}_1 + \cdots + z_n \bar{w}_n$$

Abbiamo

$$\langle Au, v \rangle = \langle v, A^* u \rangle$$