

Metodi Matematici I

Francesco Sermi

Pisa, 16 febbraio 2025

Notazioni

Riporto all'inizio del libro le notazioni adottate all'interno di questo documento:

\mathbb{C} campo dei numeri complessi

\times prodotto cartesiano

Span insieme delle combinazioni lineari dei vettori

dim dimensione dello spazio vettoriale

Ker kernel di un'applicazione lineare

Im immagine di un'applicazione lineare

Indice

1	Prima lezione	4
2	Seconda lezione	8
2.1	Norma	9
3	Terza lezione	12
3.1	Autovalori e autovettori	15
4	Prima esercitazione	16
4.1	13/09/2024	17
5	Quarta lezione	18
5.1	Teorema spettrale	20
6	Seconda esercitazione	23

CAPITOLO 1

Prima lezione

Facciamo qualche piccolo richiamo agli spazi vettoriali finiti-dimensionali. E' doveroso riportare la definizione di spazio vettoriale.

Definizione 1.1 (spazio vettoriale). Sia dato un campo \mathbb{K} e un insieme V su cui sono definite due operazioni $+$: $V \times V \rightarrow V$ che soddisfa le proprietà di gruppo abeliano e \cdot : $\mathbb{K} \times V \rightarrow V$ che soddisfa le seguenti proprietà:

- ① $\forall x, y \in V, \forall \alpha \in \mathbb{K}, (x + y) = \alpha x + \alpha y$
- ② $\forall \alpha, \beta \in \mathbb{K}, \forall x \in V, \alpha(\beta x) = \beta(\alpha x).$

Diremo che la tupla $(V, \mathbb{K}, +, \cdot)$ è uno spazio vettoriale.

Oss.: Dalla proprietà di gruppo abeliano segue che esiste un elemento neutro rispetto all'operazione $+$ e un elemento neutro rispetto all'operazione \cdot .

Notazione: all'interno di questi appunti, spesso, faremo riferimento a V come spazio vettoriale, sottintendendo naturalmente le operazioni $+$ e \cdot definite su di esso. Inoltre, se non specificato, supporremo di lavorare con dei \mathbb{C} -spazi vettoriali, ovvero spazi vettoriali definiti sul campo \mathbb{C} .

Iniziamo considerando il più "semplice" degli spazi vettoriali, ovvero \mathbb{C}^n : tale spazio vettoriale è costituito dalle n -uple di numeri $(\alpha_1, \dots, \alpha_n)$ con $\forall i \in \{1, \dots, n\}, \alpha_i \in \mathbb{C}$. Questo può essere uno spazio vettoriale prendendo la $+$ e \cdot definite come

$$\begin{aligned}(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) &= (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n) \\ \alpha(\alpha_1, \dots, \alpha_n) &= (\alpha\alpha_1, \dots, \alpha\alpha_n).\end{aligned}$$

Definizione 1.2 (combinazione lineare). Sia $\{x_1, \dots, x_n\} \subset V$ un insieme di vettori appartenenti a V \mathbb{K} -spazio vettoriale. Diremo che

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

è una combinazione lineare dei vettori $x_1, \dots, x_n \forall \alpha_1, \dots, \alpha_n$. L'insieme di tutte le possibili combinazioni lineari di x_1, \dots, x_n è detto $\text{Span}(\{x_1, \dots, x_n\})$.

Proposizione 1.1. $\text{Span}(\{x_1, \dots, x_n\})$ è un sottospazio vettoriale di V .

Introduciamo adesso il concetto di vettori linearmente indipendenti

Definizione 1.3 (vettori linearmente indipendenti). Sia $\{x_1, \dots, x_n\} \subset V$ un insieme di vettori. Diremo che essi sono linearmente indipendenti se l'equazione

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 0$$

ha come unica soluzione quella banale, ovvero $\forall i \in \{1, \dots, n\}, \alpha_i = 0$.

Oss.: Se l'equazione non ha soluzione banale, allora uno qualunque di quei vettori è in funzione degli altri: infatti, se supponiamo che $\alpha_i \neq 0$ con $i \in \{1, \dots, n\}$ allora

$$\begin{aligned} \alpha_i x_i &= -(\alpha_1 x_1 + \dots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \dots + \alpha_n x_n) \implies \\ \implies x_i &= -\frac{1}{\alpha_i} (\alpha_1 x_1 + \dots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \dots + \alpha_n x_n). \end{aligned}$$

Definizione 1.4 (dimensione di uno spazio vettoriale). Sia V uno spazio vettoriale. Diremo che $\dim(V) = n$ è il numero massimo di vettori linearmente indipendenti che si possono trovare in V e diremo che esso è la dimensione dello spazio.

Naturalmente dobbiamo distinguere vari casi:

- ① se \exists massimo numero di vettori linearmente indipendenti, allora $\exists n \in \mathbb{N} : \dim(V) = n$;
- ② se \nexists massimo numero di vettori linearmente indipendenti, allora $\forall n \in \mathbb{N}, \exists x_1, \dots, x_n$ vettori linearmente indipendenti. In tal caso diremo che $\dim(V) = \infty$.

Definizione 1.5 (base di uno spazio vettoriale). Sia V uno spazio vettoriale. Diremo che un insieme $\{x_1, \dots, x_m\} \subset V$ di vettori linearmente indipendenti è una base di V se

$$\forall v \in V, \exists \alpha_1, \dots, \alpha_m \in \mathbb{C} : v = \alpha_1 x_1 + \dots + \alpha_m x_m. \quad (\star)$$

Oss.: Naturalmente la (\star) implica banalmente che $\text{Span}(\{x_1, \dots, x_m\}) = V$, ovvero i vettori x_1, \dots, x_n "generano" tutto lo spazio.

Supponendo di aver mostrato che ogni spazio vettoriale ammette una base (che supponiamo già assodato), mostriamo la dimensione dello spazio coincide con il numero di vettori che costituiscono una base.

Lemma 1.0.1 (dello scambio). Sia V uno spazio vettoriale e $\mathcal{B} = \{x_1, \dots, x_m\}$ un insieme di vettoriali linearmente indipendenti e sia $0 \neq w \in \text{Span}(\mathcal{B})$. Allora sappiamo che $\exists \alpha_1, \dots, \alpha_m \in \mathbb{K}$ tali che

$$w = \alpha_1 x_1 + \dots + \alpha_m x_m.$$

Allora per qualche $i \in \{1, \dots, m\}, \alpha_i \neq 0$, dunque preso $\mathcal{B}' = (\mathcal{B} \setminus \{x_i\}) \cup \{w\}$ avremo che

- ① \mathcal{B}' è un insieme di vettori linearmente indipendenti;
- ② $\text{Span}(\mathcal{B}') = \text{Span}(\mathcal{B})$.

Corollario 1.0.1. Se $A = \{u_1, \dots, u_k\}$ linearmente indipendente e $B = \{w_1, \dots, w_h\}$ linearmente indipendente con $B \subset \text{Span}(A)$ allora $\exists B' \subset A$ con $\#B = \#B' = h$ tale che

- ① $A' = (A \setminus B') \cup B$ è ancora linearmente indipendente;
- ② $\text{Span}(A') = \text{Span}(A)$

Teorema 1.1. Sia V uno spazio vettoriale tale che $\dim(V) = n$ e sia $\mathcal{B} = \{x_1, \dots, x_m\} \subset V$ una base di V . Allora $m = n$.

Dimostrazione. dalla definizione di dimensione, abbiamo chiaramente che $m \leq n$. Supponiamo per assurdo che $m \neq n$, ovvero $m < n$: in tal caso possiamo trovare $C = \{x_1, \dots, x_n\}$ vettori linearmente indipendenti e $\mathcal{B} = \{y_1, \dots, y_m\}$ base di V . In tal caso preso $x_i \in C$ allora

$$\exists \alpha_1, \dots, \alpha_m \in \mathbb{C} : \alpha_1 x_1 + \dots + \alpha_m x_m = x_i.$$

Sappiamo che $\exists \alpha_i \neq 0$ per qualche $i \in \{1, \dots, m\}$ (se così non fosse allora $x_i = 0$ e questo contraddice l'ipotesi che i vettori $\{x_1, \dots, x_n\}$ siano linearmente indipendenti). Supponiamo, senza perdita di generalità, che $\alpha_i \neq 0$, dunque

$$y_i = -\frac{1}{\alpha_i}(\alpha_1 y_1 + \dots + \alpha_{i-1} y_{i-1} + \alpha_{i+1} y_{i+1} + \dots + \alpha_m y_m - x_i).$$

Possiamo dunque sostituire a y_i il vettore x_i e, per il lemma precedente, sappiamo che $\{y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_m\}$ è ancora una base di V . Ma allora, sostituendo i vettori di $\{x_1, \dots, x_m\}$ all'interno della base \mathcal{B} , abbiamo un assurdo siccome questo implica che $\forall k > m, \exists \alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{C}$ tali che

$$x_k = \alpha_1 y_1 + \dots + \alpha_m y_m.$$

il che contraddice l'ipotesi che l'insieme x_1, \dots, x_n sia un insieme linearmente indipendente di vettori. \square

Fissando una base $\mathcal{B} = \{x_1, \dots, x_n\}$, possiamo definire un isomorfismo non canonico tra V e \mathbb{C}^n .

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & \mathbb{C}^n \\ \psi & & \psi \\ x & \longleftrightarrow & (\alpha_1, \dots, \alpha_n) \\ \parallel & & \\ \alpha_1 x_1 + \dots + \alpha_n x_n & & \end{array}$$

Figura 1.1: Diagramma commutativo tra V e \mathbb{C}^n : fissando una base, è possibile studiare un qualunque spazio vettoriale supponendo di trovarci in \mathbb{C}^n .

Proposizione 1.2. Sia V uno spazio vettoriale e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V , allora possiamo definire $\varphi : V \mapsto \mathbb{C}^n$ tale che, preso $V \ni x = \alpha_1 v_1 + \dots + \alpha_n v_n$, $x \xrightarrow{\varphi} (\alpha_1, \dots, \alpha_n)$

Dimostrazione. Dalle formule della dimensione di un'applicazione lineare (che verranno dimostrate più avanti) segue banalmente che $\dim(V) = \dim(\text{Im})(\varphi)$ siccome $\text{Ker}(\varphi) = \{0\}$ chiaramente. \square

Definizione 1.6 (applicazione lineare). Siano V, W spazi vettoriali sul medesimo campo \mathbb{K} . Allora diremo che $A : V \rightarrow W$ è un'applicazione lineare se

$$A(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 A(v_1) + \alpha_2 A(v_2) \forall \alpha_1, \alpha_2 \in \mathbb{K}, \forall v_1, v_2 \in V$$

Oss.: Per ogni applicazione lineare $A(0) = 0$. Ciò segue dal fatto che $A(0) = A(0) + A(0) \implies A(0) = 0$.

Nel caso finito-dimensionale non abbiamo "problemi" nel dominio di A , siccome essa è lineare su tutto A : nel caso infinito-dimensionale saremo costretti a restringerci a dei sottospazi. Ora andiamo naturalmente a dare nuovamente la definizione del kernel e dell'immagine di un'applicazione lineare

Definizione 1.7 (kernel). Siano V, W due spazi vettoriali sul medesimo campo e sia $A : V \rightarrow W$ un'applicazione lineare. Allora definiamo il kernel di A come

$$\text{Ker}(A) = \{x \in V : A(x) = 0\} \subset V.$$

Definizione 1.8 (immagine). Siano V, W due spazi vettoriali sul medesimo campo e sia $A : V \rightarrow W$ un'applicazione lineare. Allora definiamo l'immagine di A come

$$\text{Im}(A) = \{y \in W | \exists x \in V : A(x) = y\} \subset W.$$

Proposizione 1.3. $\text{Ker}(A)$ e $\text{Im}(A)$ è un sottospazio vettoriale.

Definizione 1.9 (applicazione suriettiva). Siano V, W due spazi vettoriali sul medesimo campo \mathbb{K} . Allora A è suriettivo $\iff \text{Im}(A) = W$.

Proposizione 1.4. Siano V, W due spazi vettoriali sul medesimo campo \mathbb{K} . Allora A è iniettivo $\iff \text{Ker}(A) = \{0\}$.

Dimostrazione.

\Rightarrow : se A è iniettivo chiaramente segue che $\text{Ker}(A) = \{0\}$.

\Leftarrow : se $\text{Ker}(A) = \{0\}$ allora avremo che se per assurdo $\exists v, w \in V : A(v) = A(w)$ con $v \neq w$ allora $A(v - w) = 0 \implies v - w \in \text{Ker}(A) \implies v = w$. \square

Teorema 1.2 (della dimensione). Siano V, W due spazi vettoriali sul medesimo campo, V a dimensione finita e sia $A : V \rightarrow W$ un'applicazione lineare. Allora avremo che

$$\dim(V) = \dim(\text{Ker}(A)) + \dim(\text{Im}(A)).$$

Oss.: Se $A : V \rightarrow V$ allora è facile vedere che se $\dim(\text{Ker})(A) = 0 \iff \text{Ker}(A) = \{0\} \iff A$ è iniettiva, allora $\dim(V) = \dim(\text{Im}(V))$, ovvero A è suriettivo.

Dimostrazione. Sappiamo che $\text{Ker}(A)$ è un sottospazio vettoriale. Supponiamo allora che $B = \{v_1, \dots, v_r\}$ sia una base di tale sottospazio, allora sappiamo che è possibile completare tale base di $\text{Ker}(A)$ ad una base di V . Sia allora $B' = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ una base di V : per mostrare la tesi è sufficiente mostrare che i vettori $f(v_{r+1}), \dots, f(v_n)$ generano l'immagine di A . Sappiamo, per definizione, che la base dell'immagine di f è data da $\{f(v_1), \dots, f(v_n)\}$ ma sappiamo che i vettori $f(v_1), \dots, f(v_r)$ sono nulli, quindi l'immagine è generata dai vettori $f(v_{r+1}), \dots, f(v_n)$. Verifichiamo adesso che l'immagine di A ha dimensione $n - r$, permettendoci di concludere: per fare ciò osserviamo che l'equazione

$$\begin{aligned} \alpha_{r+1}f(v_{r+1}) + \dots + \alpha_n f(v_n) = 0 &\implies f(\alpha_{r+1}v_{r+1} + \dots + \alpha_n v_n) = 0 \implies \\ &\implies \alpha_{r+1}v_{r+1} + \dots + \alpha_n v_n \in \text{Ker}(A). \end{aligned}$$

Ma allora questo implica che quest'ultimo vettore è una combinazione lineare dei vettori v_1, \dots, v_r che formano una base di $\text{Ker}(A)$, dunque

$$\alpha_{r+1}v_{r+1} + \dots + \alpha_n v_n = \alpha_1 v_1 + \dots + \alpha_r v_r \implies -\alpha_1 v_1 - \dots - \alpha_r v_r + \alpha_{r+1}v_{r+1} + \dots + \alpha_n v_n = 0.$$

Essendo i vettori v_1, \dots, v_n linearmente indipendenti, si conclude che $\forall i \in \{1, \dots, n\}, \alpha_i = 0$ dunque $f(v_{r+1}), \dots, f(v_n)$ sono linearmente indipendenti, permettendoci dunque di affermare che $\dim(\text{Im})(A) = n - r \implies \dim(\text{Im})(A) + \dim(\text{Ker})(A) = n - r + r = n = \dim V$ \square

CAPITOLO 2

Seconda lezione

La scorsa lezione siamo arrivati alla formula della dimensione. Ora, come fatto prima, vogliamo vedere se è possibile studiare l'azione di un'applicazione lineare fra V e W di dimensione, rispettivamente, $\dim(V) = n$ e $\dim(W) = m$ tramite gli isomorfismi fra $V \longleftrightarrow \mathbb{C}^n$ e $W \longleftrightarrow \mathbb{C}^m$.

$$\begin{array}{ccc}
 x \in V & \xrightarrow{A} & W \ni y \\
 \updownarrow \varphi_n & & \updownarrow \varphi_m \\
 (\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n & \xrightarrow{[A]} & \mathbb{C}^m \ni (\mu_1, \dots, \mu_m)
 \end{array}
 \quad \text{dove } [A] = \underbrace{\begin{pmatrix} a_{ij} \end{pmatrix}}_{n \text{ colonne}} \} m \text{ righe}$$

Figura 2.1: Diagramma non commutativo fra V , W e gli spazi vettoriali ad esso isomorfi \mathbb{C}^n e \mathbb{C}^m : possiamo studiare l'azione di un'applicazione lineare A tramite la matrice associata $[A]$.

Ricordiamo che un'applicazione lineare è caratterizzata completamente dalla conoscenza dei $w_i = A(v_i)$ nelle coordinate della base \mathcal{B}' di W : infatti posto $A(e_i) = \sum_{j=1}^m a_{ji} f_j$ (siccome la matrice associata altro non è che la matrice che ha, nelle colonne, le coordinate del nuovo vettore ottenuto applicandoci A si ottiene leggendo la colonna i -esima e "scorrendo" lungo le righe) allora avremo che, preso $x \in V$, possiamo concludere che

$$x = \sum_{i=1}^n \alpha_i e_i \implies A(x) = \sum_{i=1}^n \alpha_i A(e_i) = \sum_{i=1}^n \alpha_i \sum_{j=1}^m a_{ji} f_j = \sum_{j=1}^m f_j \sum_{i=1}^n a_{ji} \alpha_i = \sum_{j=1}^m \mu_j f_j$$

dove abbiamo posto $\mu_i = \sum_{j=1}^m a_{ji} \alpha_i$.

Un'altra cosa molto importante che accade a dimensione finita è il fatto che lo spazio delle applicazioni da uno spazio vettoriale V ad uno spazio vettoriale W è uno spazio vettoriale: possiamo anche facilmente mostrare, sempre tramite l'isomorfismo fra $V \longleftrightarrow \mathbb{C}^n$ e $W \longleftrightarrow \mathbb{C}^m$, che le dimensioni di tale spazio è proprio pari al prodotto della dimensione dei due spazi, ovvero $\dim(V)\dim(W)$.

Definizione 2.1 (spazio delle applicazioni lineari fra V e W). Siano V e W due spazi vettoriali a dimensione finita. Indicheremo con $\mathcal{L}(V, W)$ lo spazio delle applicazioni lineari da V a W . Se consideriamo lo spazio degli operatori di V (ovvero le applicazioni lineari da V a V) lo indicheremo con $\mathcal{L}(V)$.

Consideriamo la restrizione al sottospazio degli operatori invertibili in V : allora sappiamo che $\exists A^{-1} \in \mathcal{L}(V) : A \circ A^{-1} = A^{-1} \circ A = \text{Id}$ dove $\circ : \mathcal{L}(V) \times \mathcal{L}(V) \rightarrow \mathcal{L}(V)$ è l'operazione di composizione fra operatori, che spesso indicheremo con "prodotto" per comodità (infatti, la composizione di due applicazioni lineari corrisponde al prodotto fra le due matrici associate, facendo attenzione all'ordine).

Definizione 2.2 (algebra degli operatori). Sia V uno spazio vettoriale a dimensione finita, allora le operazioni $+$: $\mathcal{L}(V) \times \mathcal{L}(V) \rightarrow \mathcal{L}(V)$ e \circ : $\mathcal{L}(V) \times \mathcal{L}(V) \rightarrow \mathcal{L}(V)$ formano un'algebra su $\mathcal{L}(V)$.

2.1 Norma

Per schiarire le idee, iniziamo con un'esempio di norma: pensando $C = \mathbb{C}$ come spazio vettoriale di dimensione 1, allora sappiamo che ad ogni $z \in C$ è possibile definire

$$\|z\| = |z| = \sqrt{z^*z} = \sqrt{x^2 + y^2} \quad \text{dove } z = x + iy \text{ con } x, y \in \mathbb{R},$$

dove $\|\cdot\|$ è chiamata, in \mathbb{C} , modulo.

La nozione di *norma*, dunque, altro non è che la generalizzazione della nozione di modulo e, dunque, del concetto di lunghezza ad ogni spazio vettoriale.

Ora può risultare più semplice capire la definizione formale

Definizione 2.3 (norma). Sia V un \mathbb{C} -spazio vettoriale (di dimensione finita o infinita). Chiamiamo norma un'applicazione $\|\cdot\| : V \rightarrow \mathbb{R}^{\geq 0}$ tale che $x \mapsto \|x\| \in \mathbb{R}^{\geq 0}$ che soddisfa le seguenti proprietà

- ① $\|x\| \geq 0, \|x\| = 0 \iff x = 0$;
- ② $\|\alpha x\| = |\alpha| \|x\| \quad \forall \alpha \in \mathbb{C}, \forall x \in V$;
- ③ (disuguaglianza triangolare) $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$.

Diremo che la coppia $(V, \|\cdot\|)$ è uno spazio normato.

Innanzitutto vediamo che, dalla ②, siamo in grado di mostrare che la norma è fissata a meno di una costante: infatti, preso un vettore $0 \neq e_1 \in V$, abbiamo che

$$x \in V \implies x = \alpha e_1 \implies \|x\| = \|\alpha e_1\| \stackrel{②}{=} |\alpha| \|e_1\| \in \mathbb{R}^{\geq 0}.$$

Già con $\dim(V) = 2$ esistono una moltitudine di norme diverse. Un esempio sono sicuramente

- $\|z\|_{\infty} = \max_{i=1,2} |\alpha_i|$;
- $\|z\|_1 = |\alpha_1| + |\alpha_2|$.

dove abbiamo fissato una base di V $\{e_1, e_2\}$ e $\alpha_1 e_1 + \alpha_2 e_2 = x \in V \iff z = \varphi(x) = (\alpha_1, \alpha_2) \in \mathbb{C}^2$, ovvero abbiamo considerato l'isomorfismo fra $V \longleftrightarrow \mathbb{C}^2$.

In generale, esistono delle strutture che sono ancora più generali degli spazi normati, ovvero gli spazi metrici, ovvero degli insiemi in cui è possibile definire una funzione, detta *distanza*, che quantifica la distanza fra gli elementi (non è detto però che essa rappresenti una "lunghezza" nel senso comune)

Definizione 2.4 (spazio metrico). Sia V un insieme su cui è definita una funzione detta *distanza* $d : V \times V \rightarrow \mathbb{R}^{\geq 0}$ tale che:

- ① $d(x, y) = 0 \iff x = y \quad \forall x, y \in V$;
- ② $d(x, y) = d(y, x) \quad \forall x, y \in V$;
- ③ (disuguaglianza triangolare) $\forall v, w, z \in V, d(v, z) \leq d(v, w) + d(w, z)$.

Oss.: Possiamo munire un insieme sempre della cosiddetta *distanza discreta*, definita come

$$d(x, y) = \begin{cases} 1 & \text{se } x \neq y \\ 0 & \text{se } x = y \end{cases}$$

e verificare che essa possiede tutte le proprietà richiesta dalla definizione

Proposizione 2.1. Sia $(V, \|\cdot\|)$ uno spazio normato. Allora $d(\cdot_1, \cdot_2) = \|\cdot_2 - \cdot_1\|$ è una distanza

Dimostrazione. La dimostrazione verte nel dimostrare che la distanza definita sopra soddisfa tutte le proprietà richieste: notiamo che

$$d(x, y) = \|y - x\| = 0 \iff y - x = 0 \implies x = y$$

dove si è usato la proprietà ① della norma. Successivamente osserviamo che

$$\|x - y\| = \|(-1)(y - x)\| = |-1| \|y - x\| = \|y - x\| \implies d(x, y) = d(y, x) \quad \forall x, y \in V.$$

Infine presi $x, y, z \in V$

$$d(x, z) = \|x - z\| = \|x - y + y - z\| \leq d(x, y) + d(y, z)$$

dove si è usato la proprietà ③ della norma (disuguaglianza triangolare). □

Possiamo introdurre ancora un'altra struttura, ovvero quella dello spazio vettoriale munito di prodotto scalare definito positivo. **Nota bene:** nel nostro caso andremo a dare una definizione "leggermente" diversa di prodotto scalare, siccome richiediamo che esso sia definito positivo, ovvero che $\forall v \in V, (v, v) > 0$. La definizione più generale di prodotto scalare non richiede ciò, tuttavia, siccome siamo interessati non solo ad avere un prodotto scalare ma siamo interessati anche alla norma indotta dal prodotto scalare, segue necessariamente che gli unici prodotti scalari "utili" siano proprio quelli definiti positivi (o quelli definiti negativi, con qualche piccolo cambiamento di segno) siccome sono quelli per cui la norma può avere un senso geometrico.

Definizione 2.5 (prodotto scalare). Sia V uno spazio vettoriale sul campo \mathbb{C} , allora chiamiamo prodotto scalare un'applicazione $(\cdot, \cdot) : V \times V \rightarrow \mathbb{C}$ che possiede le seguenti proprietà:

- ① $\forall x \in V, (x, x) \in \mathbb{R}^{\geq 0}, (x, x) = 0 \iff x = 0$;
- ② $\forall x, y \in V, (x, y) = (y, x)^*$;
- ③ $\forall x, y \in V, \forall \alpha \in \mathbb{C}, (x, \alpha y) = \alpha(x, y)$;
- ④ $\forall x, y, z \in V, (x, y + z) = (x, y) + (x, z)$.

Oss.: Imporre che $(x, x) \in \mathbb{R}$ è ridondante, siccome dalla ② abbiamo che $(x, x) = (x, x)^* \implies (x, x) \in \mathbb{R}^{\geq 0}$.

Oss.: Osserviamo che (\cdot, \cdot) è lineare in entrambi le componenti rispetto a $+$: $V \times V \rightarrow V$: infatti presi $x, y, z \in V$ avremo che $(x + y, z) = (z, x + y)^* = ((z, x) + (z, y))^* = (z, x)^* + (z, y)^* = (x, z) + (y, z)$.

Presi $z, w \in \mathbb{C}^n$ con $z = (z_1, \dots, z_n)$ e $w = (w_1, \dots, w_n)$ possiamo introdurre il seguente prodotto scalare

$$(z, w) = \sum_{i=1}^n z_i^* w_i$$

e possiamo facilmente mostrare che presenta tutte le buone proprietà richieste dalla definizione di prodotto scalare, siccome

$$\textcircled{1} \quad (z, z) = \sum_{i=1}^n z_i^* z_i = \sum_{i=1}^n |z_i|^2 \geq 0 \text{ e } (z, z) = 0 \iff \sum_{i=1}^n |z_i|^2 = 0 \iff z = 0;$$

$$\textcircled{2} \quad (z, w) = \sum_{i=1}^n z_i^* w_i = \sum_{i=1}^n (w_i^* z_i)^* = \left(\sum_{i=1}^n w_i^* z_i \right)^* = (w, z)^* \text{ dove si è usato il fatto che } \forall n \in \mathbb{N}, \forall z_1, \dots, z_n \in \mathbb{C}, (z_1 + \dots + z_n)^* = z_1^* + \dots + z_n^*;$$

$$\textcircled{3} \quad (z, w_1 + w_2) = \sum_{i=1}^n z_i^* (w_1 + w_2)_i = \sum_{i=1}^n z_i^* w_{1i} + \sum_{i=1}^n z_i^* w_{2i} = (z, w_1) + (z, w_2);$$

$$\textcircled{4} \quad (z, \alpha w) = \sum_{i=1}^n z_i^* (\alpha w)_i = \alpha \sum_{i=1}^n z_i^* w_i = \alpha (z, w).$$

CAPITOLO 3

Terza lezione

Per adesso fissiamo $V = \mathbb{C}^n$. Consideriamo $z = (z_1, \dots, z_n)$ tale che $\forall i, z_i \in \mathbb{C}$ e consideriamo il prodotto scalare definito alla fine della scorsa lezione

$$(z, w) = \sum_{i=1}^n z_i^* w_i.$$

Vogliamo adesso mostrare che $\|\cdot\| = \sqrt{(\cdot, \cdot)}$ è una norma. Prima di procedere, tuttavia, è sicuramente conveniente stabilire dei nomi per le norme precedentemente individuate:

① $\|z\|_\infty = \max_i |z_i|$ viene detta **norma infinito** (o norma del max);

② $\|z\|_2 = \sqrt{\sum_{i=1}^n |z_i|^2}$ viene detta norma standard;

③ $\|z\|_1 = \sqrt{\sum_{i=1}^n |z_i|}$

Mostriamo che $\|z\|_2 = \sqrt{(z, z)}$ è una norma: per fare ciò abbiamo bisogno di mostrare la seguente disuguaglianza

Proposizione 3.1 (disuguaglianza di Schwarz).

$$\forall z, w \in \mathbb{C}^n, |(z, w)| \leq \|z\| \|w\|$$

Dimostrazione. Osserviamo che presi $z, w \in \mathbb{C}^n$ abbiamo che $(\|w\|^2 z - (w, z)w, \|w\|^2 z - (w, z)w) \geq$

0, pertanto si deve avere che

$$\begin{aligned}
\sum_{i=1}^n ||w||^2 z_i - (w, z) w_i|^2 &\geq 0 \iff \sum_{i=1}^n (||w||^2 z_i - (w, z) w_i)^* (||w||^2 z_i - (w, z) w_i) \geq 0 \iff \\
\sum_{i=1}^n (||w||^4 |z_i|^2 - ||w||^2 z_i^* w_i (w, z) - ||w||^2 (w, z)^* w_i^* z_i + (w, z)^* (w, z) |w_i|^2) &\geq 0 \iff \\
\sum_{i=1}^n ||w||^4 |z_i|^2 - \sum_{i=1}^n ||w||^2 z_i^* w_i (w, z) - \sum_{i=1}^n ||w||^2 w_i^* z_i (w, z)^* + \sum_{i=1}^n (w, z)^* (w, z) |w_i|^2 &\geq 0 \iff \\
||w||^4 |z|^2 - ||w||^2 (z, w) (w, z) - ||w||^2 \overline{(w, z)} (w, z)^* + ||w||^2 (w, z)^* \overline{(w, z)} &\geq 0 \iff \\
||w||^2 (||w||^2 |z|^2 - (w, z)^* (w, z)) &\geq 0 \iff ||w||^2 |z|^2 - |(w, z)|^2 \geq 0 \implies \\
\implies |(w, z)| &\leq ||z|| ||w||,
\end{aligned}$$

dove abbiamo potuto eseguire la radice senza problemi perché entrambi sono dei numeri maggiori o uguali a 0. \square

Proposizione 3.2. $|| \cdot ||_2$ è una norma

Dimostrazione. La dimostrazione verte chiaramente nel mostrare che $|| \cdot ||_2$ soddisfa le proprietà richieste dalla definizione di norma.

Per mostrare la ① abbiamo chiaramente che $||z||_2 = \sqrt{\sum_i^n |z_i|^2} = 0 \iff z_i = 0 \forall i$. Per quanto riguarda la ② abbiamo che $||\alpha z||_2 = |\alpha| ||z||$, infatti $||\alpha z|| = \sqrt{(\alpha z, \alpha z)} = \sqrt{\alpha^* \alpha (z, z)} = |\alpha| \sqrt{(z, z)} = |\alpha| ||z||$.

Infine, per mostrare la ③, si ha che $||z + w||^2 = (z + w, z + w) = (z, z) + (z, w) + (w, z) + (w, w) = ||z||^2 + ||w||^2 + (z, w) + (z, w)^* = ||z||^2 + ||w||^2 + 2\operatorname{Re}((z, w))$ ma osserviamo che $\operatorname{Re}((z, w)) \leq |(z, w)| \leq ||z|| ||w||$ per la disuguaglianza di Schwarz: questo ci permette di concludere, siccome $||z + w||^2 = ||z||^2 + 2\operatorname{Re}((z, w)) + ||w||^2 \leq ||z||^2 + 2||z|| ||w|| + ||w||^2 = (||z|| + ||w||)^2 \implies ||z + w|| \leq ||z|| + ||w||$ dove abbiamo potuto prendere la radice senza problemi siccome $|| \cdot ||$ è sempre positiva. \square

Definizione 3.1 (vettori ortogonali). Sia V uno spazio vettoriale sul campo \mathbb{C} munito di un prodotto scalare. Sia $\dim(V) = n < \infty$, allora $z, w \in V$ si dicono *ortogonali* se $(z, w) = 0$.

Definizione 3.2 (base ortogonale). Sia V uno spazio vettoriale sul campo \mathbb{C} munito di prodotto scalare (\cdot, \cdot) con $\dim(V) = n$. Diremo che $\{e_i\}_{i \in \{1, \dots, n\}} \subset V$ è una base ortogonale di V se $\{e_i\}_{i \in \{1, \dots, n\}}$ è una base di V e se $(e_i, e_j) = \delta_{ij} \forall i \neq j \in \{1, \dots, n\}$.

Oss.: Il teorema di Lagrange ci assicura sempre di poter trovare una base ortogonale. Il teorema di Sylvester ci assicura, invece, di poter trovare una base ortonormale.

Possiamo sempre ortogonalizzare una base tramite il procedimento di Grand-Schmidt: sia $\{v_i\}_{i \in \{1, \dots, n\}}$ una base di V , allora posto $e_1 = \frac{v_1}{||v_1||}$ dobbiamo togliere a v_2 la componente "parallela" a e_1 affinché siano ortogonali, pertanto poniamo $e_2 = \frac{v_2 - (e_1, v_2)e_1}{||v_2 - (e_1, v_2)e_1||}$ e possiamo vedere che

$$(e_1, e_2) = (e_1, \frac{v_2}{||v_2 - (e_1, v_2)e_1||}) - (e_1, e_2)(e_1, \frac{e_1}{||v_2 - (e_1, v_2)e_1||}) = 0.$$

Possiamo ripetere il medesimo procedimento per tutti gli altri vettori: prendendo v_3 dobbiamo toglierli la componente parallela sia a e_1 che a e_2 ; per v_4 la componente parallela a e_1, e_2 e e_3 e

così via. Procedendo per induzione, abbiamo che l' i -esimo vettore è ortogonalizzato da

$$e_i = \frac{v_i - \sum_{j=1}^{i-1} (v_i, e_j) e_j}{\|v_i - \sum_{j=1}^{i-1} (v_i, e_j) e_j\|}$$

Presa $\mathcal{B} = \{e_i\}$ una base ortonormale possiamo sempre considerare il solito isomorfismo fra $V \longleftrightarrow \mathbb{C}$ avremo che

$$(z, w) = \left(\sum_{i=1}^n z_i e_i, \sum_{j=1}^n w_j e_j \right) = \sum_{i=1}^n \sum_{j=1}^n z_i^* w_j (e_i, e_j) = \sum_{i=1}^n \sum_{j=1}^n z_i^* w_j \delta_{ij} = \sum_{i=1}^n z_i^* w_i.$$

Oss.: Possiamo capire, dalla forma in cui si presenta in base ortonormale il prodotto scalare, che il prototipo perfetto di prodotto scalare è quello usuale che avevamo definito in \mathbb{C}^n in base canonica a cui, possibilmente, vorremmo ricondurci sempre trovando sempre una base ortonormale.

Inoltre, in una base ortonormale $\{e_1, \dots, e_n\}$, è possibile ricavare le coordinate di un vettore in maniera abbastanza agevole: infatti, sfruttando l'ortogonalità fra i vettori della base, si vede che, preso $z \in V$ con coordinate (rispetto alla base ortogonale) (z_1, \dots, z_n) , allora

$$(e_i, z) = (e_i, \sum_{j=1}^n z_j e_j) = \sum_{j=1}^n z_j (e_i, e_j) = \sum_{j=1}^n z_j \delta_{ij} = z_i.$$

Possiamo anche far vedere facilmente che, preso $A : V \rightarrow V$, avremo che, prendendo la base ortonormale $\{e_i\}_{i \in \{1, \dots, n\}}$, la seguente relazione per i termini che compongono la matrice associata

$$\begin{aligned} w = A(z) &= \sum_{i=1}^n z_i A(e_i) = \sum_{i=1}^n w_j e_j \text{ dove } w_j = (e_j, w) = (e_j, w) = (e_j, \sum_{i=1}^n z_i A(e_i)) = \\ &= (e_j, \sum_{i=1}^n z_i \sum_{l=1}^n (Ae_i, e_l) e_l) = \sum_{i=1}^n z_i (e_j, \sum_{l=1}^n (Ae_i, e_l) e_l) = \sum_{i=1}^n \sum_{l=1}^n z_i [(Ae_i, e_l)(e_j, e_l)] = \\ &= \sum_{i=1}^n z_i (Ae_i, e_j) \end{aligned}$$

Nel caso di spazio vettoriale munito di prodotto scalare è comodo definire l'operatore aggiunto

Definizione 3.3 (operatore autoaggiunto). Sia V un \mathbb{C} -spazio vettoriale, allora preso $T : V \rightarrow V$ definiamo l'operatore aggiunto come quell'operatore $T^\dagger : V \rightarrow V$ tale che

$$(T^\dagger w, v) = (w, Tv) \quad \forall v, w \in V.$$

Oss.: In $\dim(V) = 1$ l'aggiunto è l'operatore complesso coniugato.

Osserviamo che, in base ortogonale, $[T]_{ij} = (e_i, Te_j)$ ma $[T^\dagger]_{ij} = (e_i, T^\dagger e_j) = (T^\dagger e_j, e_i)^* = (e_j, Te_i)^* = [T^*]_{ji}$, dunque vale (solo in base ortogonale!) che $T^\dagger = (T^t)^*$. Possiamo inoltre facilmente mostrare che $(T^\dagger)^\dagger = T$: infatti, preso $(e_i, Te_j) = (T^\dagger e_i, e_j) = (e_j, T^\dagger e_i)^* = ((T^\dagger)^\dagger e_j, e_i)^* = (e_i, (T^\dagger)^\dagger e_j)$ e ciò è possibile se e solo se $(T^\dagger)^\dagger = T$. Infine, possiamo mostrare che

Proposizione 3.3.

$$\begin{aligned} (T_1 T_2)^\dagger &= T_2^\dagger T_1^\dagger \\ (\alpha_1 T_1 + \alpha_2 T_2)^\dagger &= \alpha_1^* T_1^* + \alpha_2^* T_2^* \end{aligned}$$

Definizione 3.4 (operatore autoaggiunto). Sia V uno \mathbb{C} -spazio vettoriale e sia $T : V \rightarrow V$ un operatore. Diremo che esso è autoaggiunto se $T^\dagger = T$.

Per quanto visto prima, sappiamo che $[T]_{ij} = [T^\dagger]_{ji} \implies [T^\dagger] = [T]$.

3.1 Autovalori e autovettori

Definizione 3.5. Sia V un \mathbb{C} -spazio vettoriale e sia $A : V \rightarrow V$ un operatore. Allora diremo che $\lambda \in \mathbb{C}$ è un autovalore se $\exists v \in V : A(v) = \lambda v$.

Proposizione 3.4. Dato $\lambda \in \mathbb{C}$ l'insieme dei possibili autovettori è un sottospazio vettoriale $V_\lambda \subset V$

$$\lambda \in \mathbb{C} \longrightarrow V_\lambda \text{ autospazio.}$$

Sia $\lambda \in \mathbb{C}$ un autovalore e sia V_λ un autospazio con $\dim(V_\lambda) \geq 1$. Allora $\forall v \in V_\lambda, A(v) = \lambda v$ e presa $(\underbrace{v_1, \dots, v_m}_{\text{base di } V_\lambda}, \underbrace{v_{m+1}, \dots, v_n}_{\text{base di } V \setminus V_\lambda})$ avremo che

$$[A] = \begin{pmatrix} \lambda \text{Id}_{m \times m} & *_{(n-m) \times (n-m)} \\ 0_{(n-m) \times (n-m)} & *_{(n-m) \times (n-m)} \end{pmatrix}$$

CAPITOLO 4

Prima esercitazione

A lezione abbiamo visto che, presi $x, y \in \mathcal{H}$, con \mathcal{H} spazio di Hilbert allora

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

Possiamo dare una dimostrazione più formale di tale disuguaglianza nella seguente maniera

Dimostrazione. sappiamo che $(\cdot, \cdot) \geq 0 \forall z \in \mathcal{H}$. Considerando allora $\|x + \alpha y\|^2 \geq 0 \forall \alpha \in \mathbb{C}$. Scrivendo α in forma esponenziale avremo che $\exists t \in [0, +\infty), \theta \in [0, 2\pi] : \alpha = te^{i\theta}$, pertanto

$$\begin{aligned} p(t) &= (x + te^{i\theta}y, x + te^{i\theta}y) = \|x\|^2 + (te^{i\theta})(x, y) + (te^{i\theta}(x, y))^* + \|y\|^2 = \\ &= \|x\|^2 + 2\operatorname{Re}(e^{i\theta}(x, y)) + t^2\|y\|^2 \implies \Delta^2 = \operatorname{Re}(e^{i\theta}(x, y))^2 - \|x\|^2\|y\|^2 \leq 0 \end{aligned}$$

ma siccome questa relazione è valida $\forall \theta \in [0, 2\pi]$ allora possiamo prendere un $\bar{\theta}$ per cui $e^{i\bar{\theta}}(x, y) \in \mathbb{R}^{\geq 0}$. In tal caso avremo che

$$\operatorname{Re}(e^{i\bar{\theta}}(x, y)) = e^{i\bar{\theta}}(x, y) = |(x, y)|.$$

Questo ci permette di concludere allora che

$$|(x, y)|^2 \leq \|x\|^2 \|y\|^2 \implies |(x, y)| \leq \|x\| \|y\|.$$

□

Adesso andiamo ad introdurre le p -norme

Definizione 4.1. $x \in \mathcal{H}, \|x\|_p = (\sum_k |x_k|^p)^{\frac{1}{p}}, p \geq 1$.

Proposizione 4.1 (disuguaglianza di Holder).

Siano $p, q \geq 1 : \frac{1}{p} + \frac{1}{q} = 1$. Allora

$$|\sum_i x_i y_i| \leq (\sum_i |x_i|^p)^{\frac{1}{p}} (\sum_i |y_i|^q)^{\frac{1}{q}}$$

Dimostrazione. Per dimostrare questo fatto bisogna sfruttare la convessità: ricordiamo che una funzione $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ se, presi due punti $x, y \in D$ allora vale la seguente relazione

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \quad \forall t \in [0, 1] \quad (\star)$$

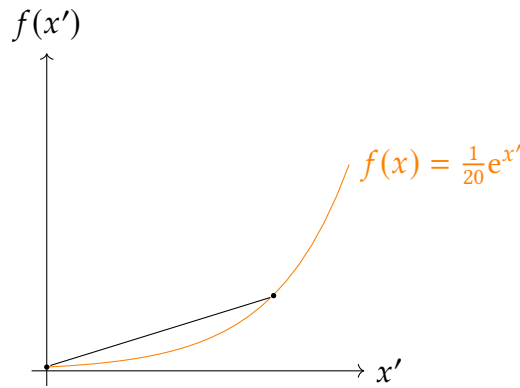


Figura 4.1: Esempio di funzione convessa: la relazione (★) ci dice che il segmento che congiunge i due punti $(x, f(x))$ e $(y, f(y))$ sta sempre sopra la funzione nell'intervallo $[x, y]$

Si osserva che $\forall t \in [0, 1]$

$$\begin{aligned} \ln \sum_k |x_k y_k| &\leq \ln \sum_k |x_k| |y_k| \leq \ln \sum_k e^{\frac{1}{p} \ln |x_k|^p + \frac{1}{q} \ln |y_k|^q} \leq \ln \frac{1}{p} \sum_k e^{\ln |x_k|^p} + \frac{1}{q} \sum_k e^{\ln |y_k|^q} \implies \\ \implies \sum_k |x_k y_k| &\leq \left(\sum_k |x_k|^p \right)^{\frac{1}{p}} \left(\sum_k |y_k|^q \right)^{\frac{1}{q}} \end{aligned}$$

□

4.1 13/09/2024

Sia $\{e_k\}_{k=1}^N$ è base ortonormale ($\iff (e_p, e_q) = \delta_{pq}$), preso $N_n = \text{Span}(\{e_k\}_{k=1}^N)$ e preso $Te_k = \frac{1}{\sqrt{k}} \sum_{p=1}^k e_p$. Mostrare che V_n è T -invariante ($\iff TV_n \subseteq V_n$).

Dimostrazione. Preso un generico vettore $x \in V_n$ allora

$$x = \alpha_1 e_1 + \dots + \alpha_N e_N \implies Tx = \sum_{k=1}^N \alpha_k T(e_k) = \sum_{k=1}^N \frac{1}{\sqrt{k}} \sum_{l=1}^k e_l = \sum_{k=1}^N \sum_{l=1}^k \frac{x_k}{\sqrt{k}} e_l.$$

Posto $\beta_k = \sum_{k=1}^N \frac{x_k}{\sqrt{k}} \implies Tx = \sum_{k=1}^N \sum_{l=1}^k \beta_k e_l$, dunque è pur sempre combinazione degli elementi della base di V_n . Altrimenti potevamo osservare che

$$[T|_{V_n}]_{pq} = (e_p, Te_q) = (e_p, \frac{1}{\sqrt{q}} \sum_{l=1}^q e_l) = \frac{1}{\sqrt{q}} \sum_{l=1}^q (e_p, e_l) = \frac{1}{\sqrt{q}} \delta_{pl}.$$

Se $p \geq q \implies (e_p, Te_q) = 0$, se $p \leq q \implies (e_p, Te_q) = \frac{1}{\sqrt{q}}$ dunque la matrice $[T|_{V_n}]$ è una matrice triangolare superiore, pertanto ogni vettore della base verrà rimandato in V_n .

□

CAPITOLO 5

Quarta lezione

Consideriamo $A : V \rightarrow V$ e sia $V_\lambda \subset V$ autospazio. Preso $v_\lambda \in V_\lambda$ allora $Av_\lambda = \lambda v_\lambda$ fissiamo una base del tipo $\underbrace{v_1, \dots, v_m}_{\text{base di } V_\lambda}, \underbrace{v_{m+1}, \dots, v_n}_{\text{resto di } V}$ e scriviamo A in forma matriciale

$$[A] = \begin{pmatrix} \lambda \text{Id}_{m \times m} & *_{(m) \times (n-m)} \\ 0_{(n-m) \times (m)} & *_{(n-m) \times (n-m)} \end{pmatrix}$$

Supponiamo che siano $\lambda_1, \dots, \lambda_k$ autovalori $\implies \exists V_{\lambda_1}, \dots, V_{\lambda_k} \subset V$: chiaramente $\dim(V) \geq \sum_k \dim(V_{\lambda_k})$ dunque

$$[A] = \begin{bmatrix} \lambda_1 \text{Id} & 0 & 0 & \dots & * \\ 0 & \lambda_2 \text{Id} & 0 & \dots & * \\ \vdots & \vdots & \ddots & \ddots & * \\ 0 & 0 & 0 & \lambda_k \text{Id} & * \\ 0 & 0 & 0 & \dots & * \end{bmatrix}.$$

Se $\dim(V) = \sum_k \dim(V_{\lambda_k})$ si dice che A è completamente diagonalizzabile o diagonalizzabile.

Osserviamo che se $T = T^*$ e se lavoriamo in un \mathbb{C} -spazio vettoriale munito di prodotto scalare allora $(v_\lambda, Tv_\lambda) = \lambda(v_\lambda, v_\lambda) = (T^*v_\lambda, v_\lambda) = (\lambda^*v_\lambda, v_\lambda) = \lambda^*(v_\lambda, v_\lambda) \implies (\lambda - \lambda^*)(v_\lambda, v_\lambda) = 0 \implies \lambda = \lambda^*$. Abbiamo quindi mostrato la seguente proposizione

Proposizione 5.1. Sia $A : V \rightarrow V$ dove V \mathbb{C} -spazio vettoriale e A è un operatore autoaggiunto. Allora gli autovalori di A sono tutti reali.

Dato $V_\lambda \subset V$ se scegliamo una base ortonormale come $\underbrace{e_1, \dots, e_m}_{\text{base di } V_\lambda}, \underbrace{e_{m+1}, \dots, e_n}_{\text{resto di } V}$ allora

$$[T] = \begin{bmatrix} \lambda \text{Id}_{m \times m} & 0_{m \times (n-m)} \\ 0_{(n-m) \times m} & [T]_{|V_\lambda^\perp} \end{bmatrix} \quad \text{dove } [T] = [T]^{t*} \text{ e si ha che } [T]_{|V_\lambda^\perp} = [T]_{|V_\lambda^\perp}^\dagger$$

Proposizione 5.2. Se $(V, (\cdot, \cdot))$ è uno spazio munito di prodotto scalare definito positivo allora $W \subset V$ è un sottospazio vettoriale allora

$$V = W \oplus W^\perp$$

Dimostrazione. Dalla formula delle dimensioni dell'ortogonale abbiamo che

$$\dim(V) - \dim(W \cap V^\perp) = \dim(W) + \dim(W^\perp)$$

ma osserviamo che $V^\perp = \{0\} \implies \dim(W^\perp) = \dim(V) - \dim(W)$. A questo punto si osserva che $W \cap W^\perp = \{0\}$ (se per assurdo $\exists v \neq 0 \in W \cap W^\perp \implies (v, w) = 0 \forall w \in W \implies (v, v) = 0 \implies$ assurdo! Il prodotto scalare è definito positivo). Dunque, avremo che W e W^\perp possono essere messi in somma diretta e generano V siccome la somma delle loro dimensioni è coincide con quella di V . \square

Proposizione 5.3. Se $\alpha_1 \neq \alpha_2 \in \mathbb{R}$ sono due distinti autovalori dell'operatore autoaggiunto $T : V \rightarrow V$ con V \mathbb{C} -spazio vettoriale allora $V_{\lambda_1} \perp V_{\lambda_2}$.

Dimostrazione. Se $\alpha_1 \neq \alpha_2 \in \mathbb{C}$ allora, presi $v_1 \in V_{\alpha_1}$ e $v_2 \in V_{\alpha_2}$, si ha che $(v_2, Tv_1) = (v_2, \lambda_1 v_1) = \lambda_1(v_2, v_1) = (Tv_2, v_1) = \lambda_2(v_2, v_1) \implies (\lambda_1 - \lambda_2)(v_2, v_1) = 0 \implies (v_2, v_1) = 0$. Dall'arbitrarietà dei due vettori appartenenti ai due distinti sottospazi segue che i due sottospazi sono ortogonali. \square

Consideriamo sempre $A : V \rightarrow V$ con $\{v_1, \dots, v_n\}$ base di V e, usando sempre la solita strategia dell'isomorfismo fra basi, possiamo rappresentare l'azione di A sui vettori tramite la sua matrice associata, dunque se $v = \sum_i \alpha_i v_i$ e $Av = \sum_i \mu_i v_i$ allora $\mu_i = \sum_j [A]_{ij} \alpha_j$ dunque $v \xrightarrow{A} Av$ e $\{\alpha_i\} \xrightarrow{[A]} \{\sum_j [A]_{ij} \alpha_j\}$.

Una cosa che possiamo fare è quello di cambiare la base in partenza e in arrivo: nel caso degli operatori completamente diagonalizzabili è sicuramente conveniente tenere come base in partenza e in arrivo la base costituita da autovettori, tuttavia questa operazione di cambio di base si può pensare di fare anche per operatori del tipo $A : V \rightarrow W$ dove $V \neq W$. Presa $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{w_1, \dots, w_n\}$ due distinte basi di V , allora possiamo definire N come matrice di cambio base tale che $w_i = \sum_j N_{ji} v_j$, dunque preso $x \in V$ allora

$$x = \sum_i \alpha_i w_i = \sum_i \alpha_i \sum_j N_{ji} v_j = \sum_j \sum_i N_{ji} \alpha_i v_j = \sum_j \mu_j v_j \quad \text{dove } \mu_j = \sum_i N_{ji} \alpha_i$$

Osserviamo, posto che $v_i = \sum_j N_{ji}^{-1} w_j$, si ha che la matrice di A nella nuova base \mathcal{B}' è chiaramente pari a $[A]' = [N]^{-1}[A][N]$: questo perché

$$\begin{aligned} Aw_i &= A \left(\sum_j N_{ji} v_j \right) = \sum_j N_{ji} Av_j = \sum_j \sum_s N_{ji} A_{sj} v_s = \sum_j \sum_s \sum_r N_{ji} A_{sj} N_{rs}^{-1} w_r = \\ &= \sum_r \sum_s \sum_j N_{rs}^{-1} A_{sj} N_{ji} w_r = \\ &= \sum_r (N^{-1} A N)_{ri} w_r \end{aligned}$$

Osserviamo adesso che la condizione affinché esista un autovalore λ è che l'equazione $Av = \lambda v$ abbia soluzione: in base \mathcal{B} possiamo scrivere che

$$\begin{aligned} Av = \lambda v &\iff Av - \lambda v = 0 \iff Av - \lambda \text{Id} v = 0 \iff \sum_j (A_{ij} - \lambda \delta_{ij}) v_j = 0 \iff \\ &\iff \det([A] - \lambda \text{Id}) = 0. \end{aligned}$$

Abbiamo quindi mostrato il seguente teorema

Teorema 5.1. Sia $A : V \rightarrow V$ con V \mathbb{C} -spazio vettoriale a dimensione finita, allora $\lambda \in \mathbb{C}$ è autovalore se e solo se

$$\det([A] - \lambda \text{Id}) = 0.$$

E' un'equazione di grado n che ha $0 < k \leq n$ radici complesse. Vale la seguente proposizione

Proposizione 5.4. $\det([A] - \lambda \text{Id})$ è indipendente dalla base scelta

Dimostrazione. Osserviamo che se riscriviamo la matrice associata ad A in una qualunque altra base allora $[A]' = [N^{-1}][A][N]$

$$\det([A]' - \lambda \text{Id}) = \det([N^{-1}][A][N] - \lambda[N^{-1}][N]) = \det[N^{-1}](\det([A] - \lambda \text{Id})[N]) = \det([A] - \lambda \text{Id})$$

dove si è usato il teorema di Binet nell'ultimo passaggio. Si conclude che se $\lambda \in \mathbb{C}$ è radice di $\det([A] - \lambda \text{Id})$ allora è radice di $\det([A]' - \lambda \text{Id})$, quindi gli autovalori non dipendono dalla base scelta. \square

Definizione 5.1 (polinomio caratteristico). Sia V un \mathbb{C} -spazio vettoriale a dimensione finita e sia $A : V \rightarrow V$ un operatore, allora diremo che

$$p_A(\lambda) = \det([A] - \lambda \text{Id})$$

è il polinomio caratteristico.

Oss.: Lavorando in \mathbb{C} (che ricordiamo essere algebricamente chiuso) allora

$$p_A(\lambda) = \prod_{i=1}^k (\lambda - \bar{\lambda}_i)^{n_i}$$

dove $0 < k \leq n$ è il numero di radici complesse, $\bar{\lambda}_i$ è l' i -esima radice che il nostro polinomio possiede e n_i è la molteplicità algebrica dell' i -esima radice.

Osserviamo che l'esistenza di una radice λ implica che $\dim V_\lambda \geq 1$ ma è un risultato ben noto che se $m_a(\lambda)$ è la molteplicità algebrica di λ allora deve risultare che

$$m_a(\lambda_i) = \dim(V_{\lambda_i})$$

Esempio 5.1 (matrice dove $m_a(\lambda) \neq \dim(V_\lambda)$).

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \implies p_A(\lambda) = (1 - \lambda)^2$$

5.1 Teorema spettrale

Teorema 5.2 (spettrale per gli operatori autoaggiunti). Sia $(V, (\cdot, \cdot))$ uno spazio vettoriale munito di prodotto scalare con $\dim(V) < \infty$. Se $T : V \rightarrow V$ è un operatore autoaggiunto allora

- ① T è diagonalizzabile;
- ② T possiede autovalori reali

Dimostrazione. La ② è già stata mostrata. Per quanto riguarda la ① procediamo per induzione sulla dimensione di V : per $n = 1$ è banalmente vero; mostriamo che $n \implies n + 1$ sia v un autovettore di T , allora sappiamo che

$$V = \text{Span}(v) \oplus (\text{Span}(v))^\perp$$

pertanto possiamo considerare la base $\mathcal{B} = \{v, \underbrace{w_1, \dots, w_n}_{\text{base di } (\text{Span}(v))^\perp}\}$. Osserviamo allora che preso $w \in (\text{Span}(v))^\perp$ allora

$$(Tv, Tw) = (\lambda v, Tw) = 0$$

dunque l'unica possibilità è che $Tw \in (\text{Span}(v))^\perp \implies (\text{Span}(v))^\perp$ è T -invariante e chiaramente $\forall u \in (\text{Span}(v))^\perp, T|_{(\text{Span}(v))^\perp}(u) = T(u)$, dunque $T|_{(\text{Span}(v))^\perp}$ è sempre autoaggiunto: possiamo dunque applicare l'ipotesi induttiva e concludere che è diagonalizzabile \square

Possiamo generalizzare il seguente risultato anche ai cosiddetti operatori normali

Definizione 5.2 (operatore normale). Sia $T : V \rightarrow V$ con V \mathbb{C} -spazio vettoriale. Diremo che T è normale se T commuta con la sua aggiunta, ovvero se

$$TT^\dagger = T^\dagger T$$

Possiamo facilmente vedere che gli operatori autoaggiunti e gli operatori unitari rientrano in questa categoria: infatti osserviamo che se T è autoaggiunto allora

$$T^\dagger T = T^2 = TT^\dagger$$

mentre per gli operatori unitari

$$T^\dagger T = T^{-1}T = \text{Id} = TT^{-1} = TT^\dagger.$$

Osserviamo che vale la seguente proposizione

Lemma 5.2.1. Se $T : V \rightarrow V$ è un operatore normale allora λ è un autovalore di $T \iff \bar{\lambda}$ è un autovalore di T^\dagger . Inoltre

$$V_\lambda(T) = V_{\bar{\lambda}}(T)$$

Dimostrazione. E' sufficiente mostrare che $\forall v \in V, T(v) = \lambda v \iff T^\dagger(v) = \bar{\lambda}T(v)$. Osserviamo che se mostriamo che $T(v) = \lambda v \implies T^\dagger(v) = \bar{\lambda}v$ allora segue che se $T^\dagger(v) = \lambda'v \implies (T^\dagger)^\dagger(v) = T(v) = \bar{\lambda}'v \implies T^\dagger(v) = \bar{\lambda}v$: è dunque sufficiente mostrare una sola delle due implicazioni, per esempio la \implies . Si osserva

$$T^\dagger(v) = \bar{\lambda}v \iff (T^\dagger(v) - \bar{\lambda}v, T^\dagger(v) - \bar{\lambda}v) = 0.$$

Dunque è sufficiente mostrare che vale la seconda proposizione per dimostrare questo lemma: siccome T è normale, T e T^\dagger commutano e dunque preso $v \in V_\lambda(T)$

$$TT^\dagger(v) = T^\dagger T(v) = \lambda T^\dagger(v) \implies T^\dagger(v) - \bar{\lambda}v \in V_\lambda(T),$$

d'altra parte abbiamo che, preso $w \in V_\lambda(T)$, la seguente relazione

$$(T^\dagger(v), w) = (v, Tw) = (v, \lambda w) = \lambda(v, w) = (\bar{\lambda}v, w) \implies (T^\dagger(v) - \bar{\lambda}v, w) = 0$$

e prendendo $w = T^\dagger(v) - \bar{\lambda}v$ (che sappiamo appartenere a $V_\lambda(T)$ siccome abbiamo visto che T e T^\dagger preservano i rispettivi autospazi) avremo che

$$(T^\dagger(v) - \bar{\lambda}v, T^\dagger(v) - \bar{\lambda}v) = 0 \iff T^\dagger(v) = \bar{\lambda}v$$

\square

Possiamo generalizzare la dimostrazione fatta per gli operatori autoaggiunti anche per gli operatori normali

Teorema 5.3 (teorema spettrale per gli operatori normali in dimensione finita). Sia V un \mathbb{C} –spazio vettoriale a dimensione finita munito di prodotto scalare (\cdot, \cdot) , sia \mathcal{B} una base ortonormale e sia $T : V \rightarrow V$ un operatore normale. Allora $\exists \mathcal{B}'$ base di autovettori ortonormali rispetto al prodotto scalare

Dimostrazione. Preso $v \in V$ autovettore di T allora possiamo decomporre V come

$$V = \text{Span}(v) \oplus (\text{Span}(v))^\perp$$

e osservare che $\text{Span}(v)$ è T –invariante $\iff T^\dagger$ –invariante. Ne segue allora che anche $(\text{Span}(v))^\perp$ è T –invariante e T^\dagger –invariante: infatti preso $z \in (\text{Span}(v))^\perp$ chiaramente $(v, z) = 0$ e

$$(v, T(z)) = (T^\dagger(v), z) = (\bar{\lambda}v, z) = \lambda(v, z) = 0 \implies T(z) \in (\text{Span}(v))^\perp$$

e similmente si mostra che $T^\dagger(w) \in (\text{Span}(v))^\perp$, dunque $(\text{Span}(v))^\perp$ è T –invariante e T^\dagger –invariante. Chiaramente $(T|_{(\text{Span}(v))^\perp})^\dagger = T^\dagger|_{(\text{Span}(v))^\perp}$ (dove il primo aggiunto va inteso rispetto a $(\cdot, \cdot)_{(\text{Span}(v))^\perp}$), da cui segue che $T|_{(\text{Span}(v))^\perp}$ è un operatore normale su $(\text{Span}(v))^\perp$ e procedendo per induzione si ottiene la tesi. \square

CAPITOLO 6

Seconda esercitazione

Riprendiamo l'esercizio della scorsa volta.