CAPITOLO 8

Prodotti scalari e spazi euclidei

1. Prodotto scalare

Definizione 8.1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, dove $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$. Un *prodotto scalare* è una funzione che ad ogni coppia di vettori u, v appartenenti a V associa lo scalare $\langle u, v\rangle \in \mathbb{K}$, con le seguenti proprietà:

(1) $\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$, per ogni $u_1, u_2, v \in V$ e per ogni $a, b \in \mathbb{K}$;
(2) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ per ogni $u, v \in V$ (dove $\overline{\langle v, u\rangle}$ indica il numero complesso coniugato di $\langle v, u\rangle$);
(3) per ogni $u \in V$ vale $\langle u, u\rangle \geq 0$ e $\langle u, u\rangle = 0$ vale se e solo se $u = O$.

Uno spazio vettoriale reale V munito di un prodotto scalare si dice *spazio euclideo*.

Osserviamo che, per la proprietà (2), il prodotto scalare $\langle u, u\rangle$ è sempre un numero reale, dunque ha senso la disuguaglianza che compare nella proprietà (3).

Per ogni vettore $u \in V$, scriveremo

$$||u|| = \sqrt{\langle u, u\rangle}$$

e diremo che $||u||$ è la *norma* di u.

Osservazione 8.2. Dalle proprietà (1) e (2) segue che

$$\langle u, av_1 + bv_2\rangle = \overline{a}\langle u, v_1\rangle + \overline{b}\langle u, v_2\rangle$$

Se il campo $\mathbb{K}$ è $\mathbb{R}$ ovviamente tutti i simboli di coniugazione complessa che compaiono nelle formule precedenti possono essere ignorati.

L'esempio principale di prodotto scalare è dato dal prodotto scalare standard in $\mathbb{R}^n$. Dati due vettori u, v, rappresentati, rispetto alla base standard, dalle colonne $u = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ \dots \\ a_n \end{pmatrix}$ e $v = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix}$ il prodotto scalare standard è il seguente:

$$\langle u, v\rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

Osservazione 8.3. Il prodotto scalare standard estende a tutti gli spazi $\mathbb{R}^n$ un concetto che in $\mathbb{R}^2$ e $\mathbb{R}^3$ ci è già familiare. Potete facilmente verificare che in questi casi per esempio $||u||$ coincide con quella che avete chiamato lunghezza del vettore u e che $||v-u||$ coincide con la distanza fra i vettori u e v. Inoltre $\langle u, v\rangle = 0$ se e solo se u e v sono ortogonali fra loro. Potete anche già osservare in $\mathbb{R}^2$ che vale la seguente relazione:

$$\langle u, v\rangle = ||u||||v|| \cos\theta$$

dove θ è l'angolo compreso fra i vettori u e v. Dimostreremo più in generale questa relazione fra qualche paragrafo.

Ecco altri due esempi di prodotto scalare:

Esempio 8.4. In $\mathbb{C}^n$, dati i vettori

$$u = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ \dots \\ a_n \end{pmatrix}, \quad v = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix}$$

scritti rispetto alla base standard, abbiamo il prodotto scalare

$$\langle u, v \rangle = a_1 \overline{b_1} + a_2 \overline{b_2} + \cdots + a_n \overline{b_n}$$

Esempio 8.5. Nello spazio vettoriale delle funzioni continue reali nell'intervallo $[a, b]$,[1] abbiamo il seguente prodotto scalare:

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

2. Ortogonalità

Sia V uno spazio vettoriale munito di un prodotto scalare. Estendiamo il concetto di perpendicolarità che ci è noto in $\mathbb{R}^2$ e $\mathbb{R}^3$ dando la seguente definizione:

Definizione 8.6. Due vettori $u, v \in V$ sono detti *ortogonali* se

$$\langle u, v \rangle = 0$$

Notiamo che se due vettori verificano $\langle u, v \rangle = 0$ vale anche che $\langle v, u \rangle = \overline{\langle u, v \rangle} = 0$ dunque il concetto di ortogonalià non dipende dall'ordine con cui stiamo considerando i due vettori. Inoltre dalla definizione si ricava subito che il vettore O è ortogonale a tutti i vettori di V:

$$\langle O, v \rangle = \langle 0O, v \rangle = 0\langle O, v \rangle = 0$$

Sia U un sottospazio di V. L'insieme dato dai vettori di V che sono ortogonali ad ogni vettore di U si indica con $U^\perp$:

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \; \forall \, u \in U\}$$

Si verifica facilmente che $U^\perp$ è un sottospazio vettoriale di V, chiamato il *sottospazio ortogonale* a U.

Esercizio 8.7. Fare questa verifica.

Definizione 8.8. Un insieme $\{u_1, ..., u_r\}$ di vettori si dice *ortogonale* se i suoi elementi sono vettori a due a due ortogonali fra loro. L'insieme $\{u_1, ..., u_r\}$ si dice *ortonormale* se è ortogonale e se per ogni i vale $||u_i|| = 1$.

---

[1] Questo spazio vettoriale, a differenza di tutti gli altri spazi considerati in queste dispense, non ha una base finita; si tratta dunque di uno spazio vettoriale di dimensione infinita.

Esempio 8.9. Nello spazio euclideo $\mathbb{R}^n$ con il prodotto scalare standard, i vettori della base standard ($u_1 = \begin{pmatrix} 1 \\ 0 \\ ... \\ ... \\ 0 \end{pmatrix}$ etc...) costituiscono un insieme ortonormale.

Teorema 8.10. *Sia V uno spazio vettoriale munito di un prodotto scalare. Un insieme ortonormale di vettori $\{u_1, ..., u_r\}$ è linearmente indipendente e, per ogni $v \in V$, il vettore*

$$w = v - \langle v, u_1 \rangle u_1 - \langle v, u_2 \rangle u_2 - \cdots - \langle v, u_r \rangle u_r$$

è ortogonale ad ognuno degli u_i, dunque appartiene a $Span(u_1, ..., u_r)^\perp$.

Dimostrazione. Supponiamo di avere una combinazione lineare degli u_i uguale a O:

$$a_1 u_1 + \cdots + a_r u_r = O$$

Per dimostrare la lineare indipendenza dei vettori $u_1, ..., u_r$ dobbiamo dimostrare che per ogni i vale $a_i = 0$.

Fissato un i, consideriamo il prodotto scalare di entrambi i membri dell'uguaglianza con u_i:

$$\langle a_1 u_1 + \cdots + a_r u_r, u_i \rangle = \langle O, u_i \rangle$$

ovvero

$$a_1 \langle u_1, u_i \rangle + \cdots + a_r \langle u_r, u_i \rangle = 0$$

Vista la ortonormalità dell'insieme $\{u_1, ..., u_r\}$ questo implica $a_i = 0$.

Ripetendo questa considerazione per ogni $i = 1, ..., r$ si dimostra che i vettori $u_1, ..., u_r$ sono linearmente indipendenti.

Verifichiamo adesso che w è ortogonale a u_i per ogni i:

$$\langle w, u_i \rangle = \langle v, u_i \rangle - \langle v, u_1 \rangle \langle u_1, u_i \rangle - \cdots - \langle v, u_r \rangle \langle u_r, u_i \rangle =$$

$$= \langle v, u_i \rangle - \langle v, u_1 \rangle 0 - \cdots - \langle v, u_i \rangle 1 - \cdots - \langle v, u_r \rangle 0 = 0$$

□

Il teorema appena dimostrato è l'ingrediente principale del procedimento di ortogonalizzazione di Gram-Schmidt, che permette di ottenere una base ortonormale "modificando" una base data.

Teorema 8.11 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Sia V uno spazio vettoriale munito di un prodotto scalare, e sia $v_1, ...v_n$ una qualunque base di V. Allora esiste una base $u_1, ..., u_n$ di V che è ortonormale e tale che, per ogni $i = 1, ..., n$, $u_i \in Span(v_1, ...v_i)$.*

Dimostrazione. Per prima cosa poniamo $u_1 = \frac{v_1}{||v_1||}$. L'insieme $\{u_1\}$ è ortonormale.

Come secondo passo consideriamo il vettore $v_2 - \langle v_2, u_1 \rangle u_1$: questo vettore è diverso da O per la lineare indipendenza di v_1, v_2 e per il Teorema 8.10, applicato riferendosi all'insieme ortonormale $\{u_1\}$, è ortogonale a u_1. Inoltre appartiene al sottospazio $Span(v_1, v_2)$.

È dunque un buon candidato per essere il nostro u_2; l'unico problema è che non ha norma 1. Per rimediare a questo basta moltiplicarlo per un opportuno scalare. Poniamo dunque :

$$u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{||v_2 - \langle v_2, u_1 \rangle u_1||}$$

Per le osservazioni precedenti, l'insieme $\{u_1, u_2\}$ è ortonormale.

Facciamo un ulteriore passo, per illustrare meglio il procedimento. A questo punto consideriamo il vettore $v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$. Questo vettore è diverso da O (altrimenti v_3 apparterrebbe a $Span(v_1, v_2)$ contro l'ipotesi che $v_1, v_2, ..., v_n$ è una base) e per il Teorema 8.10, applicato riferendosi all'insieme ortonormale $\{u_1, u_2\}$, è ortogonale a $Span(u_1, u_2)$.

Inoltre, visto che u_1 e u_2 appartengono a $Span(v_1, v_2)$, il vettore $v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$ appartiene a $Span(v_1, v_2, v_3)$ e ha dunque tutte le proprietà da noi richieste, eccetto quella di avere norma 1.

Come sopra, poniamo dunque:

$$u_3 = \frac{v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2}{||v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2||}$$

Una semplice dimostrazione per induzione ci mostra che possiamo proseguire questo procedimento ricorsivo fino a compiere n passi e definire u_n come

$$u_n = \frac{v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}}{||v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}||}$$

L'insieme $\{u_1, ..., u_n\}$ sarà ortonormale e tale che, per ogni i, $u_i \in Span(v_1, ..., v_i)$.

In particolare, i vettori $u_1, ..., u_n$ sono una base di V, visto che per il Teorema 8.10 sono linearmente indipendenti e $n = dim\ V$.

□

Il teorema appena dimostrato dimostra in particolare che ogni spazio V munito di un prodotto scalare ha una base ortonormale.

Esercizio 8.12. Applicare il procedimento di ortogonalizzazione alla seguente base di $\mathbb{R}^2$:

$$v_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

Esercizio 8.13. Trovare una base ortonormale di $\mathbb{R}^3$ il cui primo vettore sia $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

Osservazione 8.14 (Angolo fra due vettori in $\mathbb{R}^n$)**.** Consideriamo $V = \mathbb{R}^n$ munito del prodotto scalare standard. Siano u e v due vettori linearmente indipendenti di $\mathbb{R}^n$. Come sappiamo dal Teorema 8.10, applicato partendo dall'insieme ortonormale di vettori dato da un solo vettore, ossia $\{\frac{u}{||u||}\}$, il vettore

$$w = v - \langle v, \frac{u}{||u||} \rangle \frac{u}{||u||}$$

è ortogonale a $\frac{u}{||u||}$. Possiamo riscrivere questa relazione così

$$v = w + \langle v, \frac{u}{||u||} \rangle \frac{u}{||u||}$$

e ci accorgiamo che ci indica che nel piano $Span(u,v)$ i vettori $O, z = \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}, v$ determinano un triangolo rettangolo, con l'ipotenusa di lunghezza $||v||$ e i due cateti di lunghezza rispettivamente $||w||$ e $|\langle v, \frac{u}{||u||}\rangle| = \frac{|\langle v,u\rangle|}{||u||}$ (vedi Figura 1).

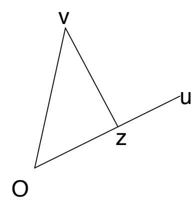

FIGURA 1. Dati due vettori $u, v \in \mathbb{R}^n$, i vettori $O, z = \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}, v$ determinano un triangolo rettangolo, visto che $w = v - \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}$ è ortogonale alla retta generata da u.

Chiamiamo θ l'angolo compreso fra l'ipotenusa e il cateto di lunghezza $\frac{|\langle v,u\rangle|}{||u||}$, ovvero l'angolo compreso fra u e v (andrebbe indicato quale dei due angoli stiamo considerando, ma siccome ne prenderemo il coseno non occorre in realtà precisare). Come sappiamo dalla geometria elementare vale

$$||v|| \cos\theta = \frac{|\langle v,u\rangle|}{||u||}$$

da cui ricaviamo la relazione

$$||u|| ||v|| \cos\theta = |\langle v,u\rangle|$$

a cui avevamo già accennato nella Osservazione 8.3. Come potete facilmente verificare, tale relazione vale anche se u e v non sono linearmente indipendenti, dunque vale per qualsiasi coppia di vettori in $\mathbb{R}^n$.

Introducendo il prodotto scalare standard, abbiamo dunque "recuperato" la trigonometria all'interno dell'algebra lineare.

3. La disuguaglianza di Cauchy-Schwarz

Dato un prodotto scalare $\langle\ ,\ \rangle$ su V, abbiamo il seguente teorema:

Teorema 8.15 (Disuguaglianza di Cauchy-Schwarz)**.** *Per ogni coppia di vettori $u, v \in V$ vale*

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

DIMOSTRAZIONE. Cominciamo osservando che se $v = O$, la disuguaglianza diventa $0 \leq 0$ dunque è vera.

Supponiamo dunque $v \neq O$, e calcoliamo, usando le proprietà del prodotto scalare, il quadrato della norma del vettore $u - \langle u,v\rangle tv$, dove $t \in \mathbb{R}$.

$$||u - \langle u,v\rangle tv||^2 = \langle u - \langle u,v\rangle tv, u - \langle u,v\rangle tv\rangle =$$
$$= \langle u,u\rangle - \overline{\langle u,v\rangle} t\langle u,v\rangle - \langle u,v\rangle t\langle v,u\rangle + \overline{\langle u,v\rangle}\langle u,v\rangle t^2\langle v,v\rangle =$$
$$= ||u||^2 - 2t(|\langle u,v\rangle|)^2 + (|\langle u,v\rangle|)^2 t^2 ||v||^2$$

Osserviamo che $||u - \langle u,v\rangle tv||^2 \geq 0$ dunque possiamo scrivere

$$||u||^2 - 2t(|\langle u,v\rangle|)^2 + (|\langle u,v\rangle|)^2 t^2 ||v||^2 \geq 0$$

Questo vale per qualsiasi $t \in \mathbb{R}$. Scegliendo adesso $t = \frac{1}{||v||^2}$ ricaviamo:

$$||u||^2 - \frac{(|\langle u, v\rangle|)^2}{||v||^2} \geq 0$$

da cui si ottiene la tesi. $\square$

Esempio 8.16. Nel caso in cui V sia $\mathbb{R}^n$ col prodotto scalare standard, la disuguaglianza di Cauchy-Schwarz si traduce così. Dati

$$u = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ \dots \\ a_n \end{pmatrix}, \quad v = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix}$$

vale

$$|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

che può essere pensata anche, svincolandosi dai vettori e dagli spazi vettoriali, come una disuguaglianza che riguarda due qualsiasi n-uple $a_1, ..., a_n$ e $b_1, ..., b_n$ di numeri reali.

Esempio 8.17. Nel caso in cui V sia lo spazio delle funzioni continue reali definite sull'intervallo $[0, 1]$, col prodotto scalare descritto nell'Esempio 8.5, la disuguaglianza di Cauchy-Schwarz si traduce nella seguente importante disuguaglianza fra integrali:

$$\left(\int_0^1 f(t)g(t)dt\right)^2 \leq \int_0^1 f(t)^2 dt \int_0^1 g(t)^2 dt$$

4. Sottospazi ortogonali

Sia V, come nei paragrafi precedenti, uno spazio vettoriale munito di prodotto scalare.

Proposizione 8.18. *Sia U un sottospazio vettoriale di V. Allora esiste una base ortonormale di U che è un sottoinsieme di una base ortonormale di V.*

Dimostrazione. Partiamo da una qualunque base $v_1, ..., v_r$ di U, ed estendiamola ad una base $v_1, ..., v_n$ di V (usando il Teorema 2.19). Ora utilizziamo il procedimento di ortogonalizzazione di Gram-Schmidt alla base $v_1, ..., v_n$: otteniamo una base ortonormale $u_1, ..., u_n$ di V, con la proprietà che i vettori $u_1, ..., u_r$ appartengono a $Span(v_1, ..., v_r)$, ossia a U. Dunque $u_1, ..., u_r$ è una base ortonormale di U, che è un sottoinsieme di una base ortonormale di V, come volevamo.

$\square$

Teorema 8.19. *Sia U un sottospazio vettoriale di V. Allora V si decompone come somma diretta di U e di $U^\perp$:*

$$V = U \oplus U^\perp$$

In particolare vale che $dim\ U^\perp = dim\ V - dim\ U$.

Dimostrazione. Osserviamo innanzitutto che $U \cap U^{\perp} = \{O\}$, dunque i due sottospazi sono in somma diretta.

Infatti se $v \in U \cap U^{\perp}$ allora $\langle v, v \rangle = 0$ (in altre parole, visto che $v \in U^{\perp}$, v deve essere ortogonale a tutti i vettori di U, fra cui se stesso). Ma per la proprietà (3) del prodotto scalare da $\langle v, v \rangle = 0$ si ricava $v = O$.

Non ci resta che dimostrare che $U + U^{\perp} = V$, ossia che ogni vettore di V può essere scritto come somma di un vettore di U e di un vettore di $U^{\perp}$. Poniamo $r = dim\ U$. Per la Proposizione 8.18 sappiamo che esiste una base ortonormale $u_1, ..., u_n$ di V tale che i primi r vettori $u_1, ..., u_r$ sono una base ortonormale di U.

Un qualunque vettore $v \in V$ può essere scritto come combinazione lineare dei vettori di questa base:

$$v = a_1 u_1 + \cdots a_r u_r + a_{r+1} u_{r+1} + \cdots + a_n u_n$$

Osserviamo a questo punto che $a_1 u_1 + \cdots a_r u_r \in U$ come osservato sopra, mentre $a_{r+1} u_{r+1} + \cdots + a_n u_n \in U^{\perp}$ per la ortonormalità della base $u_1, ..., u_n$, dunque la scrittura

$$v = (a_1 u_1 + \cdots a_r u_r) + (a_{r+1} u_{r+1} + \cdots + a_n u_n)$$

esprime v come somma di un vettore di U e di un vettore di $U^{\perp}$.

$\square$

Esercizio 8.20. Sia V uno spazio vettoriale di dimensione finita munito di un prodotto scalare e sia U un sottospazio vettoriale di V. Dimostrare che $(U^{\perp})^{\perp} = U$.

Osservazione 8.21 (Ortogonalità e sistemi lineari)**.** Consideriamo $V = \mathbb{R}^n$ munito del prodotto scalare standard. Dato un sottospazio W, con base $w_1, ..., w_r$, dove $w_1 = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ ... \\ a_n \end{pmatrix}$, $w_2 = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_n \end{pmatrix}$, etc... se pensiamo a come si scrive il prodotto scalare ci rendiamo conto i vettori di $W^{\perp}$ sono esattamente le soluzioni del sistema

$$\begin{pmatrix} a_1 & a_2 & ... & a_n \\ b_1 & b_2 & ... & b_n \\ ... & ... & ... & ... \\ ... & ... & ... & ... \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ... \\ ... \\ 0 \end{pmatrix}$$

Esercizio 8.22. Dimostrare che, dato un sottospazio U di uno spazio V munito di un prodotto scalare, e data una base ortonormale $u_1, ..., u_n$ di V tale che i primi r vettori $u_1, ..., u_r$ sono una base ortonormale di U, la mappa $T : V \to V$ data da $T(v) = v - \langle v, u_1 \rangle u_1 - \langle v, u_2 \rangle u_2 - \cdots - \langle v, u_r \rangle u_r$ è lineare, ha per immagine $U^{\perp}$ e ha nucleo uguale a U. Dimostrare inoltre che $T^2 = T$ (dunque, in base all'Esercizio 7.47 possiamo dire che T è una proiezione di V su $U^{\perp}$; in particolare T viene chiamata la *proiezione ortogonale* di V su U).

5. Esercizi

Esercizio 8.23. Sia V lo spazio vettoriale delle funzioni continue reali nell'intervallo $[-\pi, \pi]$, munito del prodotto scalare:

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$$

Verificare che il seguente sottoinsieme è ortogonale:

$$1, \cos t, cos 2t, ..., \sin t, \sin 2t, ...$$

[Attenzione, come vedete si tratta di un insieme infinito: del resto lo spazio vettoriale che stiamo considerando è, a differenza di tutti gli altri spazi considerati in queste dispense, di dimensione infinita. L'insieme proposto è molto importante in matematica, in particolare nella teoria delle *serie di Fourier*.]

Esercizio 8.24. Dato uno spazio vettoriale V munito di un prodotto scalare, dimostrare che la norma soddisfa la seguente *disuguaglianza triangolare*:

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

Esercizio 8.25. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, -2, 3, 3)$ e $(3, -5, 4, 4)$, trovare una base del sottospazio $W^\perp$.

Esercizio 8.26. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, -2, 3, 3)$ e $(3, -5, 4, 4)$, trovare una base ortonormale di W usando il procedimento di ortogonalizzazione di Gram-Schmidt.

Esercizio 8.27. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, 1, 0, 0)$ e $(0, 1, 1, 0)$, $(0, 0, 1, 1)$, trovare una base ortonormale di W usando il procedimento di ortogonalizzazione di Gram-Schmidt.

Esercizio 8.28. Si consideri $\mathbb{C}^3$ munito del prodotto scalare descritto nell'Esempio 8.4. Sia W il sottospazio generato dai vettori $(1, \imath, 1)$ e $(1+\imath, -0, 1)$. Trovare una base ortonormale per W.

CAPITOLO 9

Qualche appunto sul teorema spettrale

1. Introduzione al teorema spettrale

Cominciamo con l'enunciare, senza dimostrazione, il seguente teorema.

Teorema 9.1. *Sia $T : V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione finita sul campo $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$) e munito di prodotto scalare. Esiste allora un endomorfismo T^* univocamente determinato tale che*

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle$$

per ogni $u, v \in V$. Se si fissa in V una base ortonormale vale che la matrice $[T^*]$ di T rispetto a tale base è la trasposta della coniugata di $[T]$.

Sottolineiamo che per l'esistenza dell'endomorfismo T^* è fondamentale l'ipotesi che lo spazio vettoriale V abbia dimensione finita.

Non dimostreremo questo teorema, ma vogliamo verificarlo in un caso particolare. Consideriamo lo spazio vettoriale $\mathbb{C}^3$ con la sua base standard (che è ortonormale). Sia $T : \mathbb{C}^3 \to \mathbb{C}^3$ l'endomorfismo dato, rispetto alla base standard, dalla matrice

$$[T] = \begin{pmatrix} 2 & \imath & 3 \\ 0 & \imath & 1 \\ 1 & 2\imath & 3 \end{pmatrix}$$

Consideriamo due vettori di $\mathbb{C}^3$:

$$u = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, v = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

e calcoliamo $\langle T(u), v\rangle$. Osserviamo che:

$$T(u) = \begin{pmatrix} 2x_1 + \imath x_2 + 3x_3 \\ \imath x_2 + x_3 \\ x_1 + 2\imath x_2 + 3x_3 \end{pmatrix}$$

dunque

$$\langle T(u), v\rangle = (2x_1 + \imath x_2 + 3x_3)\overline{y}_1 + (\imath x_2 + x_3)\overline{y}_2 + (x_1 + 2\imath x_2 + 3x_3)\overline{y}_3$$

Consideriamo ora la matrice associata a T^* nella base standard. Per il Teorema 9.1 sappiamo che è uguale alla trasposta coniugata:

$$[T^*] = \begin{pmatrix} 2 & 0 & 1 \\ -\imath & -\imath & -2\imath \\ 3 & 1 & 3 \end{pmatrix}$$

Calcoliamo $\langle u, T^*(v)\rangle$. Osserviamo che:

$$T^*(v) = \begin{pmatrix} 2y_1 + y_3 \\ -\imath y_1 - \imath y_2 - 2\imath y_3 \\ 3y_1 + y_2 + 3y_3 \end{pmatrix}$$

e quindi

$$\langle u, T^*(v)\rangle = x_1\overline{(2y_1 + y_3)} + x_2\overline{(-\imath y_1 - \imath y_2 - 2\imath y_3)} + x_3\overline{(3y_1 + y_2 + 3y_3)}$$

Svolgendo i calcoli si verifica facilmente che $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, in accordo col Teorema 9.1.

Definizione 9.2. L'endomorfismo T^* si dice *aggiunto* di T. Se vale che $T = T^*$ allora T si dice *autoaggiunto*.

Osserviamo che nel caso in cui $\mathbb{K} = \mathbb{R}$, se fissiamo una base ortonormale di V e consideriamo un endomorfismo T autoaggiunto, allora la matrice $[T]$ rispetto a tale base è simmetrica, ossia, indicando con $[T]^t$ la trasposta di $[T]$, vale

$$[T] = [T]^t$$

Viceversa, se un endomorfismo è rappresentato, rispetto ad una base ortonormale, da una matrice simmetrica, allora si mostra facilmente che è autoaggiunto. Per questo motivo nel caso in cui $\mathbb{K} = \mathbb{R}$ un endomorfismo autoaggiunto si dice anche *simmetrico*.

Teorema 9.3. *Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita sul campo $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Sia λ un autovalore di T. Allora $\lambda \in \mathbb{R}$.*

Dimostrazione. Sia λ un autovalore di T e sia $v \in V - \{O\}$ un autovettore relativo a λ. Dimostreremo che $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$. Visto che $\langle v, v\rangle \neq 0$ questo implica che $\lambda = \overline{\lambda}$ ossia che $\lambda \in \mathbb{R}$.

Per dimostrare che $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$ osserviamo che, per l'ipotesi che T è autoaggiunto, vale:

$$\langle T(v), v\rangle = \langle v, T(v)\rangle$$

Ora

$$\langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle$$

mentre

$$\langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

dunque $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$ come volevamo. □

Studiamo adesso alcune proprietà degli endomorfismi simmetrici.

Teorema 9.4. *Sia $T : V \to V$ un endomorfismo simmetrico di uno spazio vettoriale V di dimensione finita su $\mathbb{K} = \mathbb{R}$. Allora sono vere le seguenti affermazioni:*

(1) *Il polinomio caratteristico $p_T(t)$ ha tutte le radici reali e si fattorizza come prodotto di fattori lineari su $\mathbb{R}$.*
(2) *L'endomorfismo T ha almeno un autovettore non nullo.*
(3) *Siano $v_1, ..., v_r$ autovettori di T relativi ad autovalori distinti $\lambda_1, ..., \lambda_r$. Allora l'insieme di vettori $\{v_1, ..., v_r\}$ è ortogonale.*

Dimostrazione. Consideriamo il polinomio caratteristico $p_T(t)$. Per il teorema fondamentale dell'algebra tale polinomio si fattorizza in $\mathbb{C}[t]$ come prodotto di fattori lineari:

$$p_T(t) = (t-\lambda_1)^{a_1}(t-\lambda_2)^{a_2}\cdots(t-\lambda_k)^{a_k}$$

Le radici $\lambda_1, ..., \lambda_k$ del polinomio sono gli autovalori di T. Dato che T è simmetrico, per il Teorema 9.3 sappiamo che tali autovalori appartengono a $\mathbb{R}$. Dunque la fattorizzazione scritta sopra è in realtà una fattorizzazione di $p_T(t)$ in $\mathbb{R}[t]$. Questo dimostra il punto (1).

Inoltre dalla fattorizzazione di $p_T(t)$ deduciamo che T ha almeno un autovalore, dunque per dimostrare il punto (2) basta prendere un autovettore relativo a questo autovalore.

Per quel che riguarda il punto (3), sia v un autovettore relativo all'autovalore λ e sia w un autovettore relativo all'autovalore μ, con $\lambda, \mu \in \mathbb{R}$ e $\lambda \neq \mu$. Per l'ipotesi che T è simmetrico vale:

$$\langle T(v), w\rangle = \langle v, T(w)\rangle$$

Ora osserviamo che

$$\langle T(v), w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle$$

e anche

$$\langle v, T(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle$$

Dalle uguaglianze scritte sopra deduciamo che

$$\lambda\langle v, w\rangle = \mu\langle v, w\rangle$$

Se fosse $\langle v, w\rangle \neq 0$ questo sarebbe assurdo. Dunque vale $\langle v, w\rangle = 0$, ossia v e w sono ortogonali.

□

Vogliamo ora dimostrare il teorema spettrale, che afferma che un endomorfismo simmetrico è diagonalizzabile, e inoltre esiste una base ortonormale che lo diagonalizza.[1] Premettiamo il seguente lemma:

Lemma 9.5. *Sia V uno spazio vettoriale sul campo $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$, di dimensione finita e munito di prodotto scalare. Sia $T : V \to V$ un endomorfismo, e sia W un sottospazio di V invariante per T, ossia tale che $T(W) \subset W$. Allora $W^\perp$ è invariante per T^*.*

Dimostrazione. Consideriamo un elemento $v \in W^\perp$. Dobbiamo dimostrare che $T^*(v) \in W^\perp$. Ora, per ogni $w \in W$ vale che $T(w) \in W$, dunque possiamo scrivere:

$$0 = \langle T(w), v\rangle = \langle w, T^*(v)\rangle$$

dunque $T^*(v)$ è ortogonale a ogni $w \in W$. Questo dimostra che $T^*(v) \in W^\perp$.

□

Teorema 9.6 (Teorema spettrale, caso reale)**.** *Sia $T : V \to V$ un endomorfismo simmetrico di uno spazio vettoriale V su $\mathbb{K} = \mathbb{R}$ di dimensione finita maggiore di 0. Allora esiste una base ortonormale di V i cui elementi sono autovettori di T.*

[1] L'aggettivo 'spettrale' proviene da alcune applicazioni, ed è collegato allo 'spettro' di una grandezza fisica. L'insieme degli autovalori è lo 'spettro' dell'endomorfismo.

Dimostrazione. Si dimostra per induzione sulla dimensione di V. Se la dimensione di V è uguale ad 1 si vede subito che l'enunciato è vero (basta prendere la base formata da un vettore non nullo di norma 1). Supponiamo dunque che $dim\ V = n > 1$ e che l'enunciato sia vero per ogni endomorfismo simmetrico di uno spazio vettoriale di dimensione $n-1$.

Per il Teorema 9.4 esiste un autovettore v_1 di T, che possiamo supporre di norma uguale a 1 (a meno di moltiplicarlo per scalare). Sia $W = Span(v_1)$ e consideriamo $W^\perp$. Come sappiamo dal Teorema 8.19 il sottospazio $W^\perp$ ha dimensione $n-1$. Inoltre sappiamo dal Lemma 9.5 che $W^\perp$ è invariante per T^*. Ma T^*, visto che T è simmetrico, coincide con T. Dunque $W^\perp$ è invariante per T e ha senso considerare l'endomorfismo T ristretto a $W^\perp$:

$$T_{|W^\perp} : W^\perp \to W^\perp$$

Osserviamo che $T_{|W^\perp}$ è simmetrico.[2] Per ipotesi induttiva sappiamo allora che esiste una base ortonormale $v_2, ..., v_n$ di $W^\perp$ composta di autovettori di $T_{|W^\perp}$. È immediato a questo punto verificare che $v_1, v_2, ..., v_n$ è una base ortonormale di V composta da autovettori di T.

□

Esempio 9.7. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Osserviamo che l'endomorfismo T è simmetrico (infatti la sua matrice rispetto alla base standard, che è una base ortonormale, è simmetrica). Dunque per il Teorema Spettrale sappiamo che T è diagonalizzabile e ci aspettiamo di trovare un base ortonormale che lo diagonalizza. Verifichiamo. Il polinomio caratteristico è

$$p_T(t) = t^2 - 2t - 3 = (t-3)(t+1)$$

dunque gli autovalori sono 3 e -1, entrambi con molteplicità algebrica 1. Questo già ci conferma che T è diagonalizzabile. Troviamo l'autospazio V_3:

$$V_3 = Ker(T - 3I) = Ker \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$$

L'autospazio V_3 ha dunque dimensione 1 (come previsto, dato che la molteplicità algebrica dell'autovalore 3 è 1) ed è generato dal vettore $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Troviamo l'autospazio V_{-1}:

$$V_{-1} = Ker(T + I) = Ker \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

[2] Per controllare in dettaglio questo fatto, osserviamo che $(T_{|W^\perp})^*$ è l'endomorfismo di $W^\perp$ che soddisfa

$$\langle T_{|W^\perp}(u_1), u_2 \rangle = \langle u_1, (T_{|W^\perp})^*(u_2) \rangle$$

per ogni $u_1, u_2 \in W^\perp$. Ma per le proprietà di T^* vale anche che

$$\langle T(u_1), u_2 \rangle = \langle u_1, T^*(u_2) \rangle$$

per ogni $u_1, u_2 \in W^\perp$, da cui, per l'unicità dell'aggiunto (vedi Teorema 9.1) segue che $(T^*)_{|W^\perp}$ e $(T_{|W^\perp})^*$ coincidono. Ma $T^* = T$ perché T è simmetrico, dunque $(T^*)_{|W^\perp} = T_{|W^\perp}$ e in conclusione $(T_{|W^\perp})^* = T_{|W^\perp}$ ossia $T_{|W^\perp}$ è simmetrico.

L'autospazio V_{-1} ha dunque dimensione 1 (come previsto) ed è generato dal vettore $w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. I due autovettori v e w sono ortogonali (in accordo col Teorema 9.4, visto che sono due autovettori di un endomorfismo simmetrico relativi a due autovalori distinti). Concludiamo dunque che la base ortornormale data dai vettori

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

è una base di autovettori per T, in accordo col teorema spettrale.

Per completezza, diamo l'enunciato generale del teorema spettrale che vale anche nel caso $\mathbb{K} = \mathbb{C}$:

Teorema 9.8 (Teorema spettrale, enunciato valido anche su $\mathbb{C}$)**.** *Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita su $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Allora esiste una base ortonormale di V i cui elementi sono autovettori di T.*

2. Endomorfismi simmetrici definiti positivi o negativi

Come abbiamo visto nel precedente paragrafo, un endomorfismo autoaggiunto è diagonalizzabile e che tutti i suoi autovalori sono reali. Per alcune applicazioni che incontrerete nel secondo semestre del corso (Analisi II) è utile sapere che segno hanno gli autovalori.

Definizione 9.9. Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita su $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Se tutti i suoi autovalori sono numeri reali ≥ 0 si dice che T è *semidefinito positivo*. Se tutti i suoi autovalori sono numeri reali ≤ 0 si dice che T è *semidefinito negativo*. In particolare, se tutti i suoi autovalori sono numeri reali > 0 si dice che T è *definito positivo*; se invece tutti i suoi autovalori sono numeri reali < 0 si dice che T è *definito negativo*.

La definizione appena data si applica in particolare nel caso reale agli endomorfismi simmetrici.

Consideriamo per esempio cosa accade nel caso in cui V ha dimensione 2. Ricordiamo innanzitutto che, dato un endomorfismo, sono ben definiti la sua traccia e il suo determinante (ossia non dipendono dalla base che si fissa e dalla particolare matrice che rappresenta l'endomorfismo in quella base: riguardate il Paragrafo 3 del Capitolo 5).

Proposizione 9.10. *Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo simmetrico. Vale che:*

(1) *T è definito positivo se e solo se il suo determinante e la sua traccia sono entrambi > 0.*
(2) *T è definito negativo se e solo se il suo determinante è > 0 e la sua traccia è < 0.*

Dimostrazione. Come sappiamo dal teorema spettrale, T è diagonalizzabile. Siano λ_1, λ_2 i suoi autovalori (eventualmente coincidenti). Fissata una base di autovettori, la matrice che rappresenta T rispetto a tale base è:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

da cui si ricava subito che il determinante di T è $\lambda_1\lambda_2$ e la traccia è $\lambda_1 + \lambda_2$. La dimostrazione a questo punto consiste in una analisi molto facile: per esempio il determinante è > 0 se solo se λ_1 e λ_2 sono diversi da 0 e hanno entrambi lo stesso segno. Se anche la traccia è > 0 allora λ_1 e λ_2 devono essere entrambi positivi, etc..(finite voi per esercizio).

$\square$

Esercizio 9.11. Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo simmetrico. Dimostrate che se il suo determinante è $= 0$ e la sua traccia è ≥ 0 allora T è semidefinito positivo e che se il suo determinante è $= 0$ e la sua traccia è < 0 allora T è semidefinito negativo. Cosa si può dire di T se il suo determinante e la sua traccia sono entrambi uguali a 0?

3. Esercizi

Esercizio 9.12. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.13. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.14. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.15. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.16. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.17. Dimostrare il Teorema 9.8.[Suggerimento: resta da dimostrare il caso $\mathbb{K} = \mathbb{C}$. Si può procedere per induzione. Si comincia prendendo un autovettore dell'endomorfismo T, e questa volta per dire che esiste basta osservare che un endomorfismo di uno spazio vettoriale complesso ha sempre un autovettore (infatti c'è sempre un autovalore perché il polinomio caratteristico ha sempre una radice complessa).]

Esercizio 9.18 (Esercizio di ripasso)**.** Verificate di saper dimostrare che, dato uno spazio vettoriale V di dimensione n su un qualunque campo $\mathbb{K}$, e dato un endomorfismo diagonalizzabile $T : V \to V$, il determinante di T è uguale al prodotto di tutti gli autovalori di T (ciascuno compare come fattore un numero di volte uguale alla sua molteplicità) e la traccia di T è uguale alla somma di tutti gli autovalori (ciascuno compare come addendo un numero di volte uguale alla sua molteplicità). [Suggerimento: prendete spunto dalla dimostrazione della Proposizione 9.10....]

Esercizio 9.19. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 6 & 18 \\ -1 & 10 \end{pmatrix}$$

Scrivere il polinomio caratteristico di T senza calcolare $det(tI - [T])$ ma solo basandosi sulla traccia e sul determinante di T.