CAPITOLO 1

Spazi vettoriali euclidei ed hermitiani

Per svariati motivi di natura matematica, fisica e scientifica, è interessante studiare la geometria degli spazi $\mathbb{R}^n$ e $\mathbb{C}^n$ dove, oltre alle usuali strutture lineari, consideriamo anche le strutture metriche indotte da quelli che gli anglofoni chiamano i "dot products"

$$\mathbb{R}^n \times \mathbb{R}^n \xrightarrow{\cdot} \mathbb{R}, \qquad x \cdot y = x^T y = \sum_i x_i y_i,$$

$$\mathbb{C}^n \times \mathbb{C}^n \xrightarrow{\cdot} \mathbb{C}, \qquad x \cdot y = x^T \overline{y} = \sum_i x_i \overline{y_i},$$

e che noi chiameremo *prodotto scalare canonico* (nel caso reale) e *prodotto hermitiano canonico* (nel caso complesso).

Il primo problema che dobbiamo affrontare è che tali prodotti non sono invarianti per cambio di coordinate; questa considerazione ci conduce in maniera naturale alle nozioni di spazio vettoriale euclideo e spazio vettoriale hermitiano, che introdurremo nelle prossime sezioni.

1.1. Spazi vettoriali euclidei

Iniziamo introducendo una classe particolare di applicazioni bilineari (Definizione **??**) a valori reali.

Definizione 1.1.1. Sia V uno spazio vettoriale reale. Un'applicazione

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}, \qquad (u, v) \mapsto \langle u, v \rangle,$$

si dice una **forma bilineare simmetrica** se è bilineare, ossia separatamente lineare in ogni variabile, e $\langle u, v \rangle = \langle v, u \rangle$ (simmetria) per ogni $u, v \in V$.

Definizione 1.1.2. Siano V uno spazio vettoriale reale. Una forma bilineare simmetrica $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ si dice **definita positiva** se $\langle v, v \rangle > 0$ per ogni vettore non nullo $0 \neq v \in V$. Le forme bilineari simmetriche definite positive sono anche dette **prodotti scalari** (non canonici).

Ad esempio, il dot product $\langle x, x \rangle = x^T x$ è un prodotto scalare su $\mathbb{R}^n$: infatti per ogni vettore x si ha $\langle x, x \rangle = \sum_{i=1}^n x_i^2$ e, siccome siamo sui numeri reali, si ha $\langle x, x \rangle \geq 0$ e vale $\langle x, x \rangle = 0$ se e solo se $x = 0$.

Definizione 1.1.3. Uno **spazio vettoriale euclideo** è uno spazio vettoriale reale di dimensione finita corredato di un prodotto scalare.

Abbiamo visto che $\mathbb{R}^n$ con il prodotto scalare canonico è uno spazio vettoriale euclideo. Più in generale, per ogni n-upla di numeri reali positivi $a_1, \ldots, a_n > 0$, lo spazio $\mathbb{R}^n$ equipaggiato con il prodotto scalare $\langle x, y \rangle = \sum_{i=1}^n a_i x_i y_i$ è uno spazio vettoriale euclideo.

Convenzione. Da ora in poi, salvo avviso contrario, intenderemo $\mathbb{R}^n$ come spazio vettoriale euclideo dotato del prodotto scalare canonico $x \cdot y = x^T y$. Inoltre, riprendendo la notazione introdotta nella dimostrazione del Corollario **??**, per ogni $x \in \mathbb{R}^n$ denotiamo $\|x\| = \sqrt{x \cdot x} = \sqrt{\sum_i x_i^2}$.

Esempio 1.1.4. Ogni sottospazio di uno spazio vettoriale euclideo, dotato della naturale restrizione del prodotto scalare, è ancora uno spazio vettoriale euclideo. Più in generale, se $(V, \langle\cdot,\cdot\rangle)$ è uno spazio vettoriale euclideo e $f\colon U \to V$ è un'applicazione lineare iniettiva, allora $\langle u_1, u_2\rangle = \langle f(u_1), f(u_2)\rangle$ è un prodotto scalare su U.

Esempio 1.1.5. Sia V uno spazio vettoriale reale di dimensione n e sia $x\colon V \to \mathbb{R}^n$ il sistema di coordinate associato ad una base $v_1, \ldots, v_n \in V$. Allora l'applicazione

$$\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}, \qquad \langle u, v\rangle = x(u) \cdot x(v),$$

è un prodotto scalare e la coppia $(V, \langle\cdot,\cdot\rangle)$ è uno spazio vettoriale euclideo. Si noti che $\langle v_i, v_i\rangle = 1$ per ogni i e $\langle v_i, v_j\rangle = 0$ per ogni $i \neq j$.

Diremo che due vettori v, w in uno spazio vettoriale euclideo sono **ortogonali** o **perpendicolari**, e scriveremo $v \perp w$, se $\langle v, w\rangle = 0$. Per capire il senso di tale definizione basta osservare che in $\mathbb{R}^2$ tale nozione coincide con quella usuale di perpendicolarità (Figura 1.1).

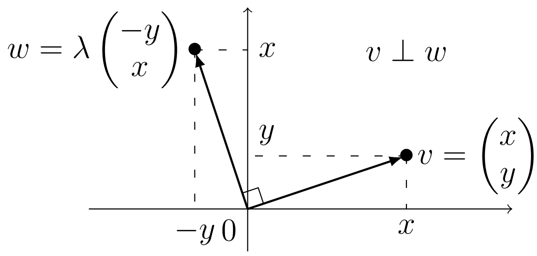

Figura 1.1. In $\mathbb{R}^2$, $w \perp \begin{pmatrix} x \\ y \end{pmatrix}$ se e solo se w è un multiplo di $\begin{pmatrix} -y \\ x \end{pmatrix}$.

Teorema 1.1.6 (disuguaglianza di Cauchy–Schwarz). *Per ogni coppia di vettori v, w in uno spazio vettoriale euclideo si ha*

$$\langle v, w\rangle^2 \leq \langle v, v\rangle\langle w, w\rangle,$$

con l'uguaglianza vera se e solo se v, w sono linearmente dipendenti.

Dimostrazione. Il risultato è banale se almeno uno tra v e w è il vettore nullo, non è quindi restrittivo supporre v e w entrambi diversi da 0. Se v, w sono linearmente dipendenti esiste uno scalare $a \in \mathbb{R}$ tale che $w = av$; in tal caso $\langle v, w\rangle = a\langle v, v\rangle$, $\langle w, w\rangle = a^2\langle v, v\rangle$, da cui segue l'uguaglianza $\langle v, w\rangle^2 = \langle v, v\rangle\langle w, w\rangle$.

Se v, w sono linearmente indipendenti allora $\langle w, w\rangle > 0$ ed il vettore $r = \langle w, w\rangle v - \langle v, w\rangle w$ è diverso da 0. Quindi $\langle r, r\rangle > 0$ e si ha

$$0 < \frac{\langle r, r\rangle}{\langle w, w\rangle} = \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2.$$

□

Dimostriamo adesso che tutti gli spazi vettoriali euclidei si ottengono come nell'Esempio 1.1.5 per un opportuno sistema di coordinate; di conseguenza, la geometria di un qualunque spazio vettoriale euclideo si riconduce mediante un'opportuna scelta della base a quella di $\mathbb{R}^n$ con il prodotto scalare canonico. Il punto chiave è che negli spazi vettoriali euclidei il prodotto scalare consente di stratificare la collezione delle basi in varie caste, dove la più elevata in gerarchia è quella dalle basi ortonormali.

Definizione 1.1.7. Una successione di vettori $u_1, \ldots, u_n$ in uno spazio vettoriale euclideo si dice:

(1) **ortonormale**, se $\langle u_i, u_i\rangle = 1$ per ogni i e $\langle u_i, u_j\rangle = 0$ per ogni $i \neq j$;
(2) **ortogonale** se $\langle u_i, u_i\rangle \neq 0$ per ogni i e $\langle u_i, u_j\rangle = 0$ per ogni $i \neq j$.

Una base che è anche una successione ortonormale (risp.: ortogonale) si dice una **base ortonormale** (risp.: **base ortogonale**).

Il termine ortonormale è la forma contratta di *normale ortogonale*; le condizioni $\langle u_i, u_i\rangle = 1$ sono quelle di normalità, mentre le condizioni $\langle u_i, u_j\rangle = 0$ per $i \neq j$ sono quelle di ortogonalità.

Ad esempio: la base canonica di $\mathbb{R}^n$ è una base ortonormale rispetto al prodotto scalare canonico; i vettori

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad u_3 = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}$$

sono una base ortonormale di $\mathbb{R}^3$ rispetto al prodotto scalare canonico.

LEMMA 1.1.8. *Ogni successione ortogonale è formata da vettori linearmente indipendenti.*

DIMOSTRAZIONE. Si abbia $a_1u_1 + \cdots + a_nu_n = 0$ con $u_1, \ldots, u_n$ successione ortogonale. Allora per ogni indice i si ha

$$0 = \langle 0, u_i\rangle = \langle \sum_j a_ju_j, u_i\rangle = \sum_j a_j\langle u_j, u_i\rangle = a_i\langle u_i, u_i\rangle,$$

e siccome $\langle u_i, u_i\rangle \neq 0$ ne consegue $a_i = 0$. □

LEMMA 1.1.9. *Una base $v_1, \ldots, v_n$ di uno spazio vettoriale euclideo V è ortonormale se e solo se per ogni $v \in V$ vale $v = \sum_i\langle v, v_i\rangle v_i$.*

DIMOSTRAZIONE. Se $v_1, \ldots, v_n \in V$ è una base ortonormale e $v \in V$ si può scrivere $v = \sum_i a_iv_i$ con $a_i \in \mathbb{R}$. Ma allora $\langle v, v_j\rangle = \sum_i a_i\langle v_i, v_j\rangle = a_j$ per ogni j.

Viceversa, se $v = \sum_i\langle v, v_i\rangle v_i$ per ogni v, in particolare si ha $v_j = \sum_i\langle v_j, v_i\rangle v_i$ per ogni j, e siccome i vettori v_i sono linearmente indipendenti, questo è possibile solo se la base $v_1, \ldots, v_n$ è ortonormale. □

Segue dal Lemma 1.1.9 che in uno spazio vettoriale euclideo le coordinate di un vettore v rispetto ad una base ortonormale $u_1, \ldots, u_n$ sono gli n numeri reali $\langle v, u_1\rangle, \ldots, \langle v, u_n\rangle$.

TEOREMA 1.1.10 (di ortogonalizzazione). *Sia $v_1, \ldots, v_n$ una successione di vettori linearmente indipendenti in uno spazio vettoriale euclideo $(V, \langle,\rangle)$. Allora esiste un'unica successione ortogonale $w_1, \ldots, w_n$ in V tale che*

$$\mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(w_1, \ldots, w_i) \quad e \quad \langle w_i, v_i\rangle = \langle w_i, w_i\rangle > 0$$

per ogni $i = 1, \ldots, n$.

DIMOSTRAZIONE. (*Esistenza.*) Dimostriamo per induzione su n che la successione $w_1, \ldots, w_n \in V$, definita dalle formule ricorsive,

$$w_1 = v_1, \qquad w_k = v_k - \sum_{i=1}^{k-1}\frac{\langle v_k, w_i\rangle}{\langle w_i, w_i\rangle}w_i,$$

soddisfa le condizioni richieste. Per $n = 1$ tutto segue dal fatto che $w_1 = v_1 \neq 0$. Possiamo supporre per induzione che $w_1, \ldots, w_{n-1}$ sia una successione ortogonale tale che

$$\mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(w_1, \ldots, w_i) \quad \text{e} \quad \langle w_i, v_i\rangle = \langle w_i, w_i\rangle > 0$$

per ogni $i = 1, \ldots, n-1$. Allora è ben definito il vettore

$$w_n = v_n - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} w_i$$

ed è immediato osservare che $\mathrm{Span}(w_1, \ldots, w_n) = \mathrm{Span}(v_1, \ldots, v_n)$. Siccome $v_n \notin \mathrm{Span}(v_1, \ldots, v_{n-1})$ si ha $w_n \notin \mathrm{Span}(w_1, \ldots, w_{n-1})$; a maggior ragione $w_n \neq 0$ e $\langle w_n, w_n \rangle > 0$. Per ogni $k < n$ si ha

$$\langle w_n, w_k \rangle = \langle v_n, w_k \rangle - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_k \rangle = \langle v_n, w_k \rangle - \frac{\langle v_n, w_k \rangle}{\langle w_k, w_k \rangle} \langle w_k, w_k \rangle = 0$$

e questo prova che la successione $w_1, \ldots, w_n$ è ortogonale. Per finire,

$$\langle w_n, w_n \rangle = \langle v_n, w_n \rangle - \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_n \rangle = \langle v_n, w_n \rangle.$$

(*Unicità.*) Data un'altra successione ortogonale $z_1, \ldots, z_n$ con le proprietà richieste, possiamo supporre per induzione che $z_i = w_i$ per ogni $i = 1, \ldots, n-1$. Siano $a_1, \ldots, a_n \in \mathbb{R}$ tali che $z_n = a_n v_n - \sum_{i=1}^{n-1} a_i w_i$, allora $\langle z_n, z_n \rangle = a_n \langle z_n, v_n \rangle$ da cui segue $a_n = 1$. Inoltre, per ogni $k < n$ si ha

$$0 = \langle z_n, z_k \rangle = \langle z_n, w_k \rangle = \langle v_n, w_k \rangle - a_k \langle w_k, z_k \rangle$$

da cui segue $a_k = \langle v_n, w_k \rangle / \langle w_k, w_k \rangle$.

□

Teorema 1.1.11 (di ortonormalizzazione). *Sia $v_1, \ldots, v_n$ una successione di vettori linearmente indipendenti in uno spazio vettoriale euclideo $(V, \langle,\rangle)$. Allora esiste un'unica successione ortonormale $u_1, \ldots, u_n$ in V tale che*

$$\mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(u_1, \ldots, u_i) \quad e \quad \langle u_i, v_i \rangle > 0$$

per ogni $i = 1, \ldots, n$. Inoltre, se per qualche $k \leq n$ la sottosuccessione $v_1, \ldots, v_k$ è ortonormale, allora $u_i = v_i$ per ogni $i = 1, \ldots, k$.

Dimostrazione. Per provare l'esistenza basta prendere una successione ortogonale $w_1, \ldots, w_n$ come nel Teorema 1.1.10 e definire $u_i = w_i / \sqrt{\langle w_i, w_i \rangle}$ per ogni i.

Viceversa, se $u_1, \ldots, u_n$ è una successione ortonormale con le proprietà richieste, allora la successione $w_i = \langle v_i, u_i \rangle u_i$, $i = 1, \ldots, n$, soddisfa le condizioni del Teorema 1.1.10.

Dunque, se $x_1, \ldots, x_n \in V$ è una successione ortonormale con le stesse caratteristiche, allora $\langle v_i, u_i \rangle u_i = \langle v_i, x_i \rangle x_i$ per ogni indice i e quindi esistono $a_1, \ldots, a_n > 0$ tali che $x_i = a_i u_i$ per ogni i, e dalle condizioni di normalità $1 = \langle x_i, x_i \rangle = a_i^2 \langle u_i, u_i \rangle = a_i^2$ segue $a_i = 1$ per ogni i. L'ultima affermazione del teorema segue immediatamente dall'unicità. □

Riepilogando, nelle precedenti procedure (ortogonalizzazione e ortonormalizzazione), le successioni ortogonale ed ortonormale sono state costruite in maniera esplicita e ricorsiva usando il cosiddetto *processo di Gram–Schmidt*, da ora in poi abbreviato, con licenza eufonica, in POGS:

$$\text{POGS}: \quad \begin{array}{c}\text{vettori}\\ \text{indipendenti}\\ v_1, \ldots, v_n\end{array} \rightsquigarrow \begin{array}{c}\text{successione}\\ \text{ortogonale}\\ w_1, \ldots, w_n\end{array} \rightsquigarrow \begin{array}{c}\text{successione}\\ \text{ortonormale}\\ u_1, \ldots, u_n\end{array}$$

$$
\begin{aligned}
&w_1 = v_1, \qquad u_1 = \frac{w_1}{\sqrt{\langle w_1, w_1\rangle}} = \frac{v_1}{\sqrt{\langle v_1, v_1\rangle}},\\
&w_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, w_i\rangle}{\langle w_i, w_i\rangle} w_i = v_k - \sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i,\\
&u_k = \frac{w_k}{\sqrt{\langle w_k, w_k\rangle}}, \qquad w_k = \langle v_k, u_k\rangle u_k, \qquad k = 1, \dots, n.
\end{aligned}
\tag{1.1.1}
$$

ESEMPIO 1.1.12. Vogliamo trovare una base ortonormale u_1, u_2 del piano $V \subseteq \mathbb{R}^3$ di equazione $x_1 + x_2 - x_3 = 0$. Prendiamo una base qualunque di V, ad esempio $v_1 = (1, 0, 1)^T$, $v_2 = (0, 1, 1)^T$. Applicando il POGS si ottiene

$$
w_1 = v_1, \qquad u_1 = \frac{w_1}{\sqrt{\langle w_1, w_1\rangle}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},
$$

$$
w_2 = v_2 - \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix}, \qquad u_2 = \frac{w_2}{\sqrt{\langle w_2, w_2\rangle}} = \frac{\sqrt{2}}{\sqrt{3}} \begin{pmatrix} -1/2 \\ 1 \\ 1/2 \end{pmatrix}.
$$

COROLLARIO 1.1.13. *Ogni spazio vettoriale euclideo $(V, \langle,\rangle)$ possiede basi ortonormali ed esistono isomorfismi lineari $f \colon V \to \mathbb{R}^n$ tali che*

$$
\langle v_1, v_2\rangle = f(v_1) \cdot f(v_2), \qquad \textit{per ogni } v_1, v_2 \in V.
$$

DIMOSTRAZIONE. Per l'esistenza delle basi ortonormali basta applicare il POGS ad una qualunque base. Sia $f \colon V \to \mathbb{R}^n$ il sistema di coordinate associato ad una base ortonormale $v_1, \dots, v_n$, ossia

$$
f(a_1 v_1 + \cdots + a_n v_n) = (a_1, \dots, a_n)^T \qquad \text{per ogni } a_1, \dots, a_n \in \mathbb{R}.
$$

Per ogni coppia di vettori $u = \sum_i a_i v_i$ e $w = \sum_j b_j v_j$ si ha

$$
\langle u, w\rangle = \sum_{i,j} a_i b_j \langle v_i, v_j\rangle = \sum_i a_i b_i = f(u) \cdot f(w).
$$

□

COROLLARIO 1.1.14. *In uno spazio vettoriale euclideo ogni successione ortonormale si estende ad una base ortonormale.*

DIMOSTRAZIONE. Basta estendere ad una base qualunque e poi POGS. □

Esercizi.

1. Dati due vettori v, w in uno spazio vettoriale euclideo, provare che $v \perp w$ se e solo se $\langle v + w, v + w\rangle = \langle v, v\rangle + \langle w, w\rangle$.

2. Trovare una base ortonormale del sottospazio $V \subseteq \mathbb{R}^4$ di equazione $x_1 - x_2 = x_3 - x_4 = 0$.

3. Trovare una base v_1, v_2 per il sottospazio $U \subseteq \mathbb{R}^3$ definito dall'equazione $x - y + 3z = 0$ tale che $v_1^T v_2 = 0$ e $(0, 0, 1)v_1 = 1$.

4 (Identità di Parseval). Sia $u_1, \dots, u_n$ una base ortonormale di uno spazio vettoriale euclideo V. Provare che per ogni $v, w \in V$ vale

$$
\langle v, w\rangle = \sum_{i=1}^{n} \langle v, u_i\rangle \langle u_i, w\rangle.
$$

5. Siano $u_1, \dots, u_n$, $v_1, \dots, v_n$ e $w_1, \dots, w_n$ tre basi ortonormali di uno spazio vettoriale euclideo. Dimostrare che

$$t_i = \sum_j \langle u_i, v_j \rangle w_j, \qquad i = 1, \dots, n,$$

è ancora una base ortonormale.

6. Trovare una base ortonormale dello spazio delle matrici reali 2×2 a traccia nulla rispetto al prodotto scalare $\langle A, B \rangle = \mathrm{Tr}(A^T B)$.

1.2. Proiezioni e riflessioni ortogonali

Dato un qualsiasi sottoinsieme S di uno spazio vettoriale euclideo V denotiamo con $S^\perp \subseteq V$ l'insieme dei vettori ortogonali a tutti gli elementi di S:

$$S^\perp = \{v \in V \mid \langle v, s \rangle = 0 \text{ per ogni } s \in S\}.$$

È evidente che se $R \subseteq S$ allora $S^\perp \subseteq R^\perp$ e si verifica facilmente che $S^\perp$ è un sottospazio vettoriale; più precisamente, $S^\perp$ è l'intersezione di tutti i nuclei delle applicazioni lineari

$$\langle \cdot, w \rangle \colon V \to \mathbb{R}, \qquad v \mapsto \langle v, s \rangle, \qquad s \in S.$$

Per i nostri obiettivi è fondamentale osservare che se S è un sottoinsieme finito di V e $W = \mathrm{Span}(S)$, allora $S^\perp = W^\perp$: infatti, siccome $S \subseteq W$ vale $W^\perp \subseteq S^\perp$. Viceversa, se $S = \{s_1, \dots, s_h\}$, allora per per ogni $v \in S^\perp$ ed ogni $w = \sum t_i s_i \in W$ si ha $\langle v, w \rangle = \sum_i t_i \langle v, s_i \rangle = 0$ e quindi $v \in W^\perp$.

Definizione 1.2.1. Dato un sottospazio vettoriale W di uno spazio vettoriale euclideo, il sottospazio $W^\perp$ viene chiamato l'**ortogonale** di W.

Lemma 1.2.2. *Per ogni sottospazio W di uno spazio vettoriale euclideo V si ha $V = W \oplus W^\perp$ e $(W^\perp)^\perp = W$. In particolare,* $\dim W + \dim W^\perp = \dim V$.

Dimostrazione. Siano $n = \dim V$, $m = \dim W$ e prendiamo una qualunque base $v_1, \dots, v_n$ di V tale che $W = \mathrm{Span}(v_1, \dots, v_m)$. Applicando l'ortonormalizzazione a tale base troviamo una base ortonormale $u_1, \dots, u_n$ di V tale che $W = \mathrm{Span}(u_1, \dots, u_m)$. Per dimostrare che $V = W \oplus W^\perp$ basta quindi provare che $W^\perp = \mathrm{Span}(u_{m+1}, \dots, u_n)$. Sia $v = \sum_{i=1}^n a_i u_i$ un generico vettore di V, abbiamo già osservato che $v \in W^\perp$ se e solo se $\langle v, u_i \rangle = 0$ per ogni $i = 1, \dots, m$ e, siccome $\langle v, u_i \rangle = a_i$ si ha $v \in W^\perp$ se e solo se $a_1 = \cdots = a_m = 0$, ossia se e solo se $v \in \mathrm{Span}(u_{m+1}, \dots, u_n)$.

Dalla simmetria del prodotto scalare segue subito che $W \subseteq (W^\perp)^\perp$ ed i due sottospazi hanno la stessa dimensione. $\square$

Fissato un sottospazio $W \subseteq V$, ogni vettore $v \in V$ si scrive in modo unico come

$$v = v_W + v_{W^\perp}, \qquad \text{con} \qquad v_W \in W,\ v_{W^\perp} \in W^\perp. \tag{1.2.1}$$

Possiamo quindi definire, tra le altre, le applicazioni lineari

$$P_W, R_W \colon V \to V, \qquad P_W(v) = v_W, \quad R_W(v) = v_W - v_{W^\perp} = v - 2v_{W^\perp},$$

chiamate rispettivamente **proiezione ortogonale** e **riflessione ortogonale** rispetto ad W. Notiamo che P_W è la composizione della proiezione sul primo fattore $V = W \oplus W^\perp \to W$ con il morfismo di inclusione $W \to V$, e si può scrivere $R_W = 2P_W - I$.

Osserviamo che $P_W(v) = v$ se e solo se $v \in W$ e $P_W(v) = 0$ se e solo se $v \in W^\perp$. Quindi $V = W \oplus W^\perp$ coincide con la decomposizione di V in autospazi per entrambi gli operatori P_W; più precisamente, W è l'autospazio relativo all'autovalore 1 per P_W, R_W, mentre e $W^\perp$ è l'autospazio relativo all'autovalore 0 per P_W e -1 per R_W. In particolare, $\det(R_W) = (-1)^{\dim W^\perp}$.

Il calcolo esplicito della proiezione ortogonale P_W e dei suoi parenti stretti

$$P_{W^\perp} = I - P_W, \qquad R_W = 2P_W - I, \qquad R_{W^\perp} = I - 2P_W,$$

risulta particolarmente facile quando conosciamo una base ortogonale di W. Infatti, per quanto già visto nella dimostrazione del Lemma 1.2.2, se $w_1, \dots, w_k$ è una base ortogonale di W, allora:

$$(1.2.2) \qquad P_W(v) = \sum_{i=1}^k \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i, \qquad P_{W^\perp}(v) = v - \sum_{i=1}^k \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i,$$

$$(1.2.3) \qquad R_W(v) = -v + 2\sum_{i=1}^k \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i, \qquad R_{W^\perp}(v) = v - 2\sum_{i=1}^k \frac{\langle v, w_i\rangle}{\langle w_i, w_i\rangle} w_i.$$

Se invece $w_1, \dots, w_k$ è una base qualunque di W, che non vogliamo ortogonalizzare, possiamo calcolare $P_W(v)$ scrivendo $P_W(v) = a_1w_1 + \cdots + a_kw_k$ e risolvendo il sistema di k equazioni lineari $\langle v - P_W(v), w_i\rangle = 0$, $i = 1, \dots, k$, nelle incognite $a_1, \dots a_k$; il fatto che tale sistema abbia soluzione unica è sostanzialmente equivalente all'Esercizio 11.

Esercizi.

7. Sia $U \subseteq \mathbb{R}^4$ il piano generato dai vettori $(1, 1, 0, 0)^T$ e $(1, 2, 3, 4)^T$. Trovare un piano $V \subseteq \mathbb{R}^4$ tale che $P_UP_V = P_VP_U = 0$, dove P_U e P_V sono le proiezioni ortogonali su U e V rispettivamente.

8. Scrivere le matrici che rappresentano nella base canonica le proiezioni ortogonali di $\mathbb{R}^3$ sui piani di equazione

$$2x + y - 3z = 0, \qquad x + 2y - z = 0, \qquad x + y + z = 0.$$

9. Determinare l'immagine del vettore $v = (-2, 0, 1)^T$ nella riflessione ortogonale $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ rispetto alla retta generata dal vettore $u = (1, 2, 1)^T$.

10. Sia $f = R_VR_W\colon \mathbb{R}^3 \to \mathbb{R}^3$ la composizione delle riflessioni ortogonali rispetto ai piani $V = \{x + y = 0\}$ e $W = \{y + z = 0\}$. Trovare la matrice che rappresenta f nella base canonica.

11. Sia $v_1, \dots, v_n$ una successione di vettori in uno spazio vettoriale euclideo V. Dimostrare che la matrice simmetrica $A \in M_{n,n}(\mathbb{R})$ di coefficienti $a_{ij} = \langle v_i, v_j\rangle$ è invertibile se e solo se i vettori v_i sono linearmente indipendenti.

12 (☕). Siano $u_1, \dots, u_n$ e $w_1, \dots, w_m$ due successioni di vettori in uno spazio vettoriale euclideo V. Dimostrare che il rango della matrice $n \times m$ di coefficienti $\langle u_i, w_j\rangle$ è uguale al rango della restrizione a $\operatorname{Span}(u_1, \dots, u_n)$ della proiezione ortogonale su $\operatorname{Span}(w_1, \dots, w_m)$.

1.3. Il teorema spettrale reale

Come preannunciato, in questo capitolo useremo spesso il termine *operatore lineare* come sinonimo di applicazione lineare.

Lemma 1.3.1. *Siano U spazio vettoriale reale non nullo di dimensione finita e $f\colon U \to U$ un operatore lineare. Esistono allora due vettori $u, v \in U$ e due numeri $a, b \in \mathbb{R}$ tali che*

$$u \neq 0, \quad f(u) = au - bv, \quad f(v) = bu + av.$$

Dimostrazione. Non è restrittivo supporre $U = \mathbb{R}^n$ e di conseguenza $f = L_A$ per una matrice $A \in M_{n,n}(\mathbb{R})$. Possiamo pensare A come una matrice a coefficienti complessi e quindi estendere nel modo canonico L_A ad un'applicazione lineare $L_A \colon \mathbb{C}^n \to \mathbb{C}^n$: se scriviamo i vettori di $\mathbb{C}^n$ nella forma $x + iy$, con $x, y \in \mathbb{R}^n$ ed i unità immaginaria, allora si ha $L_A(x + iy) = L_A(x) + iL_A(y)$.

Sia $u + iv \in \mathbb{C}^n$ un autovettore per L_A con autovalore $a + ib \in \mathbb{C}$; a meno di moltiplicare l'autovettore per i non è restrittivo supporre $u \neq 0$. Dunque

$$L_A(u) + iL_A(v) = (a + ib)(u + iv) = (au - bv) + i(av + bu),$$

che equivale a $L_A(u) = au - bv$ e $L_A(v) = bu + av$. □

Definizione 1.3.2. Un operatore lineare $f \colon V \to V$ di uno spazio vettoriale euclideo si dice **simmetrico**, o **autoaggiunto**, se $\langle f(u), v\rangle = \langle u, f(v)\rangle$ per ogni $u, v \in V$.

Una possibile motivazione del termine simmetrico è contenuta nella seguente proposizione.

Proposizione 1.3.3. *Per un operatore lineare $f \colon V \to V$ di uno spazio vettoriale euclideo sono equivalenti:*

(1) *f è simmetrico;*
(2) *f si rappresenta con una matrice simmetrica rispetto ad ogni base ortonormale;*
(3) *esiste una base ortonormale nella quale f si rappresenta con una matrice simmetrica.*

In particolare, se $\dim V = n$, allora gli operatori simmetrici formano un sottospazio vettoriale di $\mathrm{Hom}_{\mathbb{R}}(V, V)$ di dimensione $n(n + 1)/2$.

Dimostrazione. Sia $u_1, \ldots, u_n$ una qualunque base ortonormale di V e sia $A = (a_{ij})$ la matrice che rappresenta f in tale base, ossia $f(u_i) = \sum_j u_j a_{ji}$ per ogni i. Se f è simmetrico, allora per ogni $i, j = 1, \ldots, n$ si ha

$$a_{ji} = \langle f(u_i), u_j\rangle = \langle u_i, f(u_j)\rangle = a_{ij}.$$

Viceversa, se A è simmetrica, allora per ogni $x = \sum_i x_i u_i$ ed ogni $y = \sum_i y_i u_i$ in V si ha

$$\langle f(x), y\rangle = \sum_{i,j} x_i y_j \langle f(u_i), u_j\rangle = \sum_{i,j,k} x_i y_j a_{ki}\langle u_k, u_j\rangle = \sum_{i,j} x_i y_j a_{ji},$$
$$\langle x, f(y)\rangle = \sum_{i,j} x_i y_j \langle u_i, f(u_j)\rangle = \sum_{i,j,k} x_i y_j a_{kj}\langle u_i, u_k\rangle = \sum_{i,j} x_i y_j a_{ij},$$

e dalla simmetria di A segue l'uguaglianza delle due espressioni. □

Teorema 1.3.4 (teorema spettrale reale). *Sia $f \colon V \to V$ un operatore lineare di uno spazio vettoriale euclideo. Allora f è simmetrico se e solo se esiste una base ortonormale di V formata da autovettori per f.*

Dimostrazione. Dimostriamo prima l'implicazione più semplice; in una base ortonormale di autovettori l'operatore f si rappresenta con una matrice diagonale e quindi f è simmetrico per la Proposizione 1.3.3.

Supponiamo adesso che f sia simmetrico e dimostriamo, come passo intermedio, che ogni sottospazio f-invariante $0 \neq U \subseteq V$ possiede autovettori. Per il Lemma 1.3.1 esistono $u, v \in U$, con $u \neq 0$, e due numeri reali $a, b \in \mathbb{R}$ tali che $f(u) = au - bv$, $f(v) = bu + av$. Siccome f è simmetrico si ha

$$\begin{aligned} 0 &= \langle u, f(v)\rangle - \langle f(u), v\rangle \\ &= (a\langle u, v\rangle + b\langle u, u\rangle) - (a\langle u, v\rangle - b\langle v, v\rangle) = b(\langle u, u\rangle + \langle v, v\rangle). \end{aligned}$$

Siccome $u \neq 0$ si ha $\langle u,u\rangle + \langle v,v\rangle > 0$, quindi $b = 0$ ed u è un autovettore per f. Per il Teorema **??** l'endomorfismo f è triangolabile ed esiste una base $u_1,\dots,u_n$ tale che $f(\mathrm{Span}(u_1,\dots,u_i)) \subseteq \mathrm{Span}(u_1,\dots,u_i)$ per ogni $i = 1,\dots,n$; per il Teorema 1.1.11 possiamo supporre che $u_1,\dots,u_n$ sia una base ortonormale e dimostriamo che questo implica necessariamente che gli u_i sono tutti autovettori. Supponiamo di aver dimostrato per un certo $1 \leq m \leq n$ che $u_1,\dots,u_{m-1}$ sono autovettori. Se $f(u_m) = \sum_{i=1}^{m} a_i u_i$ si ottiene

$$a_i = \langle f(u_m), u_i\rangle = \langle u_m, f(u_i)\rangle = 0 \text{ per ogni } i < m.$$

□

Nel caso di $\mathbb{R}^n$ con il prodotto scalare canonico, il teorema spettrale è del tutto equivalente al seguente risultato, che estende e migliora il Corollario **??**.

Corollario 1.3.5. *Una matrice $A \in M_{n,n}(\mathbb{R})$ è simmetrica se e solo se possiede una base di autovettori ortonormale rispetto al prodotto scalare canonico.*

Esempio 1.3.6. Tutte le proiezioni e le riflessioni ortogonali sono operatori simmetrici.

Sia W un sottospazio di $\mathbb{R}^n$ e proviamo che $\langle P_W(u), v\rangle = \langle P_W(u), P_W(v)\rangle$ per ogni $u, v \in V$. Segue dalla definizione di proiezione che $P_W(u) \in W$ e $v - P_W(v) \in W^\perp$. Quindi $\langle P_W(u), v - P_W(v)\rangle = 0$ e

$$\begin{aligned}\langle P_W(u), v\rangle &= \langle P_W(u), P_W(v) + (v - P_W(v))\rangle \\ &= \langle P_W(u), P_W(v)\rangle + \langle P_W(u), v - P_W(v)\rangle = \langle P_W(u), P_W(v)\rangle.\end{aligned}$$

Scambiando u con v si ottiene $\langle P_W(u), P_W(v)\rangle = \langle u, P_W(v)\rangle$ e, di conseguenza, la simmetria di P_W. Per dimostrare che R_W è un operatore simmetrico basta osservare che $R_W = 2P_W - I$ e ricordare che gli operatori simmetrici formano un sottospazio vettoriale.

Esercizi.

13. Sia $A = (a_{ij}) \in M_{n,n}(\mathbb{R})$ e si assuma che esistano $b_1,\dots,b_n > 0$ tali che $b_i a_{ij} = b_j a_{ji}$ per ogni i, j. Dimostrare che A è diagonalizzabile.

14. Calcolare una base ortonormale di autovettori per la matrice simmetrica reale

$$\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

15. Sia $f\colon V \to V$ un operatore simmetrico di uno spazio vettoriale euclideo e siano $u, v \in V$ autovettori corrispondenti ad autovalori distinti. Provare che $u \perp v$.

16 (☕). Sia $A \in M_{n,n}(\mathbb{R})$ simmetrica con tutti gli autovalori distinti. Provare che per ogni matrice antisimmetrica $E \in M_{n,n}(\mathbb{R})$ esiste unica una matrice simmetrica $B \in M_{n,n}(\mathbb{R})$ tale che $E = AB - BA$ e $\mathrm{Tr}(A^i B) = 0$ per ogni $i = 0,\dots,n-1$.

1.4. Criterio di Sylvester e regola dei segni di Cartesio

In questa sezione affronteremo il problema di come fare a capire se una forma bilineare simmetrica reale è, oppure non è, definita positiva, e quindi a saper distinguere gli spazi vettoriali euclidei dalle patacche.

Definizione 1.4.1. Diremo che una matrice simmetrica reale $A \in M_{n,n}(\mathbb{R})$ è **definita positiva** se $x^T Ax > 0$ per ogni vettore $x \in \mathbb{R}^n$ diverso da 0.

Notiamo che in una matrice simmetrica definita positiva $A = (a_{ij})$ i coefficienti sulla diagonale principale sono tutti positivi: infatti si ha $a_{ii} = e_i^T Ae_i$ dove, come al solito $e_1, \dots, e_n$ indica la base canonica di $\mathbb{R}^n$. Il viceversa è generalmente falso per matrici di ordine $n \geq 2$. Ad esempio la matrice $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ non è definita positiva in quanto $(1, -1) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -2$.

Proposizione 1.4.2. *Siano $v_1, \dots, v_n$ una base di uno spazio vettoriale reale V e $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ una forma bilineare simmetrica. Allora $\langle \cdot, \cdot \rangle$ è definitia positiva, ossia è un prodotto scalare, se e solo se la matrice simmetrica $A \in M_{n,n}(\mathbb{R})$ di coefficienti $a_{ij} = \langle v_i, v_j \rangle$ è definita positiva nel senso della Definizione 1.4.1.*

Dimostrazione. Sia $L_{\mathbf{v}} : \mathbb{R}^n \to V$, $L_{\mathbf{v}}(x) = \sum_i x_i v_i$, l'isomorfismo lineare definito dalla base $v_1, \dots, v_n$. Chiaramente la forma $\langle \cdot, \cdot \rangle$ è definita positiva se e solo se $\langle L_{\mathbf{v}}(x), L_{\mathbf{v}}(x) \rangle > 0$ per ogni $0 \neq x \in \mathbb{R}^n$. Per concludere basta osservare che

$$\langle L_{\mathbf{v}}(x), L_{\mathbf{v}}(x) \rangle = \sum_{i,j} x_i x_j \langle v_i, v_j \rangle = \sum_i x_i \sum_j a_{ij} x_j = x^T Ax.$$

□

Teorema 1.4.3. *Una matrice simmetrica reale è definita positiva se e solo se i suoi autovalori sono tutti positivi.*

Dimostrazione. Sia $A \in M_{n,n}(\mathbb{R})$ simmetrica, per il teorema spettrale esiste una base $u_1, \dots, u_n$ di $\mathbb{R}^n$ ortonormale rispetto al prodotto scalare canonico tale che $Au_i = \lambda_i u_i$, $\lambda_i \in \mathbb{R}$, per ogni indice i. Se A è definita positiva, allora per ogni autovalore λ_i si ha $0 < u_i^T Au_i = \lambda_i u_i^T u_i = \lambda_i$. Viceversa, se $\lambda_i > 0$ per ogni i, allora per ogni vettore $0 \neq x = \sum a_i u_i \in \mathbb{R}^n$ si ha

$$x^T Ax = \sum_{i,j} a_i a_j u_i^T Au_j = \sum_{i,j} a_i a_j \lambda_j u_i^T u_j = \sum_i a_i^2 \lambda_i > 0.$$

□

Il criterio di Sylvester. Per ogni matrice $A \in M_{n,n}(\mathbb{R})$ e per ogni $k = 1, \dots, n$ denotiamo con $A[k] \in M_{k,k}(\mathbb{R})$ il minore principale formato dai coefficienti contenuti nelle prime k righe e k colonne. Ad esempio, se

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

allora

$$A[1] = (1), \qquad A[2] = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \qquad A[3] = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Alcuni autori chiamano le sottomatrici $A[k]$ i *minori principali di nord–ovest.*

Teorema 1.4.4 (Criterio di Sylvester). *Per una matrice simmetrica reale* $A \in M_{n,n}(\mathbb{R})$ *le seguenti condizioni sono equivalenti:*

(1) A *è definita positiva;*
(2) $\det(A[k]) > 0$ *per ogni* $k = 1, \dots, n$.

Dimostrazione. Supponiamo che A sia definita positiva; per il Teorema 1.4.3 tutti gli autovalori di A sono positivi e quindi $\det(A) = \det(A[n]) > 0$. Per ogni $1 \le k \le n$ sia $i\colon \mathbb{R}^k \to \mathbb{R}^n$ l'applicazione $i(x_1, \dots, x_k)^T = (x_1, \dots, x_k, 0, \dots, 0)^T$; in particolare il sottospazio $i(\mathbb{R}^k)$ è quello dei vettori con le ultime $n-k$ coordinate uguali a 0. Notiamo che $i(x)^T A i(x) = x^T A[k] x$ per ogni $x \in \mathbb{R}^k$. Dunque se A è definita positiva, anche le matrici $A[k]$ sono definite positive e di conseguenza i loro determinanti sono tutti positivi.

Supponiamo adesso $\det(A[k]) > 0$ per ogni k e proviamo che gli autovalori di A sono tutti positivi; per induzione su n possiamo assumere già dimostrato che il minore principale $A[n-1]$ sia definito positivo. Siano adesso $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ gli autovalori di A contati con molteplicità e scegliamo una base ortonormale $u_1, \dots, u_n$ tale che $Au_i = \lambda_i u_i$ per ogni i. Supponiamo per assurdo $\lambda_1 \le 0$; siccome per ipotesi $\det(A) = \det(A[n]) = \lambda_1 \lambda_2 \cdots \lambda_n > 0$ deve essere necessariamente $\lambda_1 \le \lambda_2 < 0$, ossia λ_1 e λ_2 entrambi negativi. Per la formula di Grassmann $i(\mathbb{R}^{n-1}) \cap \operatorname{Span}(u_1, u_2) \neq 0$ ed esistono $x \in \mathbb{R}^{n-1}$, $x \neq 0$, $a, b \in \mathbb{R}$ tali che $i(x) = au_1 + bu_2$. La contraddizione segue dalle disuguaglianze

$$\begin{aligned} 0 < x^T A[n-1] x &= (au_1 + bu_2)^T A (au_1 + bu_2) \\ &= (au_1 + bu_2)^T (a\lambda_1 u_1 + b\lambda_2 u_2) = a^2\lambda_1 + b^2\lambda_2 \le 0. \end{aligned}$$

□

Il criterio di Sylvester si presta ad un approccio algoritmico per stabilire se una matrice simmetrica reale A è definita positiva. Infatti, se ad una riga aggiungiamo dei multipli scalari delle righe precedenti, i determinanti dei minori $A[k]$ non cambiano. Di conseguenza si può procedere nel modo seguente:

(1) si guarda il coefficiente a_{11}; se $a_{11} \le 0$ allora la matrice non è definita positiva ed il processo si ferma;
(2) se $a_{11} > 0$, aggiungendo alle righe $2, 3, \dots$ opportuni multipli della prima riga si annullano i coefficienti a_{i1} per ogni $i > 1$;
(3) nella matrice ottenuta si guarda il coefficiente a_{22}; se è ≤ 0 allora la matrice A non è definita positiva;
(4) se $a_{22} > 0$, aggiungendo alle righe $3, 4, \dots$ opportuni multipli della seconda riga si annullano i coefficienti a_{i2} per ogni $i > 2$;
(5) e così via, mutatis mutandis, con le righe dalla tre in poi.

Se il processo non si interrompe prematuramente e si arriva alla fine ad una matrice triangolare con tutti i coefficienti sulla diagonale principale positivi, allora la matrice di partenza è definita positiva.

Osservazione 1.4.5. Il Teorema 1.4.4 contiene solo la prima parte del criterio di Sylvester. La seconda parte, che dimostreremo nel Corollario 2.6.12, afferma che se i determinanti delle sottomatrici $A[k]$ sono tutti diversi da 0, allora il numero di autovalori negativi della matrice A è uguale al numero dei cambiamenti di segno della successione di $n+1$ numeri reali

$$1,\ \det(A[1]),\ \det(A[2]),\ \dots,\ \det(A[n]).$$

La regola dei segni di Cartesio. Sappiamo che ogni polinomio a coefficienti reali di grado n possiede al più n radici reali. In questa sezione illustreremo un semplice criterio per determinare quante radici reali positive e quante radici reali negative può avere al massimo un determinato polinomio a coefficienti reali.

Definizione 1.4.6. Per ogni $p(t) \in \mathbb{R}[t]$ denotiamo con $s(p(t))$ il numero dei cambiamenti di segno nella successione ordinata dei coefficienti non nulli di $p(t)$.

Ad esempio, la successione dei coefficienti non nulli di $t^9 + 2t^6 - 3t^4 + 4t^3 - 5$ è $1, 2, -3, 4, -5$ ed il numero dei cambiamenti di segno è 3; similmente

$$s(t^8) = 0, \quad s(t^3 + t^2 - t + 1) = 2, \quad s(t^4 + t^2 + 2) = 0.$$

È chiaro che $s(p(t))$ non cambia se ordiniamo coefficienti non nulli da quello di grado più basso a quello di grado più alto; inoltre, il numero dei cambiamenti di segno non cambia se moltiplichiamo il polinomio per un qualunque monomio at^k con $a \neq 0$.

Lemma 1.4.7. *Sia $p(t) \in \mathbb{R}[t]$ un polinomio non nullo di grado n. Allora $s(p(t)) + s(p(-t)) \leq n$.*

Dimostrazione. Il risultato è chiaro se $p(t) = at^n$ per qualche $0 \neq a \in \mathbb{R}$. Non è quindi restrittivo supporre $p(t) = at^n + q(t)$ dove $q(t) = bt^m + \cdots$ è un polinomio di grado $m < n$, e $a, b \neq 0$. Allora $s(p(t)) \leq s(q(t)) + 1$ e l'uguaglianza vale se e solo se a, b hanno diverso segno; similmente $s(p(-t)) \leq s(q(-t)) + 1$ e l'uguaglianza vale se e solo se $(-1)^n a, (-1)^m b$ hanno diverso segno. Per induzione sul grado possiamo supporre $s(q(t)) + s(q(-t)) \leq m$. Se $m \leq n - 2$ abbiamo finito; se invece $m = n - 1$ allora a, b hanno lo stesso segno se e solo se $(-1)^n a, (-1)^{n-1} b$ hanno segno diverso e quindi $s(p(t)) + s(p(-t)) = s(q(t)) + s(q(-t)) + 1$. □

Teorema 1.4.8. *Sia $p(t)$ un polinomio di grado positivo a coefficienti reali. Allora il numero di radici reali positive di $p(t)$, contate con molteplicità, è minore od uguale al numero $s(p(t))$ dei cambiamenti di segno.*

Ad esempio, il polinomio $p(t) = t^4 - t^2 - 1$ ha un solo cambiamento di segno, e dunque possiede al più una radice reale positiva.

Dimostrazione. Per dimostrare il teorema è sufficiente dimostrare la formula

$$s((t-c)p(t)) > s(p(t)), \qquad \text{per ogni } 0 \neq p(t) \in \mathbb{R}[t],\ c \in \mathbb{R},\ c > 0. \tag{1.4.1}$$

Infatti se $c_1, \ldots, c_k$ sono le radici reali positive, contate con molteplicità, di $p(t)$, è possibile scrivere

$$p(t) = (t - c_1)(t - c_2) \cdots (t - c_k) q(t)$$

e quindi $s(p(t)) \geq k + s(q(t)) \geq k$.

Per dimostrare la formula (1.4.1) supponiamo il polinomio $p(t)$ di grado n, con $t = 0$ radice di molteplicità $m \geq 0$, e scriviamo

$$p(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_m t^m, \qquad a_m, a_n \neq 0.$$

Denotiamo $s = s(p(t))$ e consideriamo la successione strettamente decrescente $n = i_0 > i_1 > \cdots > i_s \geq m$ definita ricorsivamente come

$$i_0 = n, \qquad i_{k+1} = \max\{j \mid j < i_k,\ a_j a_{i_k} < 0\}, \quad 0 \leq k < s.$$

Scrivendo $(t-c)p(t) = b_{n+1}t^{n+1} + \cdots + b_m t^m$, per concludere la dimostrazione è sufficiente dimostrare che

$$b_{n+1} = b_{i_0+1},\ b_{i_1+1}, \ldots,\ b_{i_s+1},\ b_m,$$

è una successione di numeri reali a segni alterni. Dato che, per ogni $k = 0, \ldots, s$, il coefficiente a_{i_k+1} è nullo oppure di segno opposto a a_{i_k}, ne segue che $b_{i_k+1} = a_{i_k} - c a_{i_k+1}$

ha lo stesso segno di a_{i_k}. Infine, siccome a_{i_s} ha lo stesso segno di a_m ne segue che b_{i_s+1} e $b_m = -ca_m$ hanno segni opposti. □

Corollario 1.4.9 (Regola dei segni di Cartesio). *Sia $p(t)$ un polinomio di grado positivo a coefficienti reali. Se tutte le radici di $p(t)$ sono reali, allora il numero di radici positive, contate con molteplicità, è uguale al numero dei cambiamenti di segno della successione ordinata dei coefficienti non nulli di $p(t)$.*

Dimostrazione. Se $t = 0$ è una radice di molteplicità α, dividendo il polinomio per t^α il numero dei cambi di segno resta invariato; non è quindi restrittivo supporre $p(0) \neq 0$ e quindi che tutte le radici siano reali e non nulle.

Sia n il grado di $p(t)$ e denotiamo con $c_1, \dots, c_k$ le radici positive e $-d_{k+1}, \dots, -d_n$ quelle negative, contate con molteplicità. Le radici del polinomio $p(-t)$ sono pertanto $-c_1, \dots, -c_k, d_{k+1}, \dots, d_n$ e quindi per il Teorema 1.4.8 si hanno le disuguaglianze $s(p(t)) \geq k$, $s(p(-t)) \geq n - k$.

Per il Lemma 1.4.7 si ha $s(p(t)) + s(p(-t)) \leq n$ e quindi l'unica possibilità è che $s(p(t)) = k$, $s(p(-t)) = n - k$.

Notiamo incidentalmente che se ν è la molteplicità della radice $t = 0$, allora $s(p(t)) + s(p(-t)) + \nu < \deg(p(t))$ è una condizione sufficiente per l'esistenza di radici complesse non reali; l'esempio $p(t) = t^2 + t + 1$ mostra che tale condizione non è necessaria. □

Tornando all'algebra lineare, se sappiamo che una matrice $A \in M_{n,n}(\mathbb{R})$ è triangolabile, o meglio ancora diagonalizzabile, e $p_A(t)$ è il suo polinomio caratteristico, allora:

(1) la molteplicità algebrica dell'autovalore 0 è la più piccola potenza di t che compare in $p_A(t)$ con coefficiente non nullo;
(2) la somma delle molteplicità algebriche degli autovalori positivi è uguale al numero dei cambiamenti di segno della successione ordinata dei coefficienti non nulli di $p_A(t)$;
(3) la somma delle molteplicità algebriche degli autovalori negativi è la differenza tra n e la somma dei due numeri precedenti.

Ad esempio, se sappiamo che il polinomio caratteristico di una matrice triangolabile reale è $t^4 - t^3 - 7t^2 + t + 6$, allora tale matrice ha zero autovalori nulli, due autovalori positivi e due autovalori negativi, tutti contati con molteplicità.

Corollario 1.4.10. *Una matrice simmetrica reale è definita positiva se e solo se il coefficiente di t^i nel polinomio caratteristico è positivo per ogni i pari, e negativo per ogni i dispari.*

Esercizi.

17. Siano $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$ gli autovalori, ordinati in ordine crescente, di una matrice simmetrica reale $A = (a_{ij}) \in M_{n,n}(\mathbb{R})$. Provare che per ogni vettore $x \in \mathbb{R}^n$ valgono le disuguaglianze

$$\lambda_1 \|x\|^2 \leq x^T A x \leq \lambda_k \|x\|^2.$$

Dedurre che $\lambda_1 \leq a_{ii} \leq \lambda_k$ per ogni $i = 1, \dots, n$.

18. Dedurre dal criterio di Sylvester che

$$\begin{vmatrix} 1 & 1 & a \\ 1 & 2 & b \\ a & b & -3 \end{vmatrix} \leq 0$$

per ogni $a, b \in \mathbb{R}$.

19 (Decomposizione LU). Sia $A \in M_{n,n}(\mathbb{K})$, non necessariamente simmetrica, che abbia tutti i minori principali di nord–ovest $A[k]$ invertibili. Sia $e_1, \ldots, e_n$ la base canonica di $\mathbb{K}^n$ e denotiamo con $\pi_i \colon \mathbb{K}^n \to \mathbb{K}^i$ la proiezione sulle prime coordinate. Provare che per ogni $i = 0, \ldots, n-1$ esiste unico un vettore $v_i \in \operatorname{Span}(e_1, \ldots, e_i)$ tale che $\pi_i L_A(e_{i+1} + v_i) = 0$. Dedurre che A si fattorizza in modo unico come $A = LU$, con L triangolare inferiore ed U triangolare superiore con tutti i coefficienti sulla diagonale principale uguali a 1.

20. Siano $A, B \in M_{2,2}(\mathbb{R})$ matrici simmetriche definite positive. Mostrare con un esempio che la matrice simmetrica $AB + BA$ non è in generale definita positiva. (Suggerimento: prendere A diagonale con autovalori molto diversi tra loro.)

21. Mostrare con un esempio che il Corollario 1.4.9 è in generale falso se il polinomio $p(t)$ possiede alcune radici complesse non reali.

22. Sia $p(t) = \sum_{i=0}^{n} a_i t^i$ un polinomio di grado n con tutte le radici reali non nulle. Provare che se $a_i = 0$ allora a_{i-1} e a_{i+1} sono $\neq 0$ ed hanno segni opposti. (Suggerimento: guardare alla dimostrazione del Corollario 1.4.9.)

23. Mostrare che:
(1) l'equazione $x^{1901} = 10 + \sum_{i=1}^{935} x^{2i}$ possiede una sola soluzione reale;
(2) il polinomio $t^7 + t^6 + t^4 + t^2$ possiede esattamente 3 radici reali contate con molteplicità.

24. Sia $p(t)$ un polinomio avente tutte le radici reali e siano $a < b \in \mathbb{R}$ tali che $p(a), p(b) \neq 0$. Determinare una formula per il numero di radici comprese tra a e b.

25. Provare che, nella dimostrazione del Teorema 1.4.8 la differenza $s((t-c)p(t)) - s(p(t))$ è dispari. Trovare un esempio con $p(t)$ di grado 2 e $s((t-c)p(t)) - s(p(t)) = 3$.

26 (Ⓐ). Provare che la matrice simmetrica reale di coefficienti

$$a_{ij} = \frac{1}{i+j-1} = \int_0^1 t^{i-1} t^{j-1}\, dt, \qquad 1 \leq i, j \leq n.$$

è definita positiva, per ogni $n > 0$. Suggerimento:

$$x^T A y = \int_0^1 \left(\sum_{i=1}^{n} x_i t^{i-1} \right) \left(\sum_{j=1}^{n} y_j t^{j-1} \right) dt.$$

27. Sia $t > 1$ un numero reale. Dimostrare che la matrice simmetrica di coefficienti $a_{ij} = t^{(i-1)(j-1)}$, $i, j = 1, \ldots, n$, è definita positiva.

28. Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica definita positiva. Mostrare che se A^2 è un multiplo scalare dell'identità, allora anche A è un multiplo scalare dell'identità.

29. Sia $A \in M_{n,n}(\mathbb{R})$ simmetrica e definita positiva. Si dimostri che per ogni $\varepsilon \in \mathbb{R}$ di valore assoluto sufficientemente piccolo, la matrice $A + \varepsilon I$ è definita positiva.

1.5. Matrici ortogonali ed isometrie

Dato che il prodotto scalare canonico in $\mathbb{R}^n$ è sostanzialmente un prodotto righe per colonne, si vede facilmente che le m colonne di una matrice $A \in M_{n,m}(\mathbb{R})$ formano una successione ortonormale in $\mathbb{R}^n$ se e solo se $A^T A$ è la matrice identità di ordine m.

Definizione 1.5.1. Una matrice $E \in M_{n,n}(\mathbb{R})$ si dice **ortogonale** se le colonne di E sono una base ortonormale o, equivalentemente, se $E^T E = I$, ossia $E^{-1} = E^T$.

Per ogni $n \geq 1$ si denota con

$$O_n(\mathbb{R}) = \{E \in M_{n,n}(\mathbb{R}) \mid E^T E = I\}$$

l'insieme di tutte le matrici ortogonali di ordine n. Si osserva che:

(1) le matrici identità sono ortogonali;
(2) se $E, F \in O_n(\mathbb{R})$, allora anche $EF \in O_n(\mathbb{R})$; infatti si ha $(EF)^T(EF) = F^T E^T EF = F^T F = I$;
(3) se $E \in O_n(\mathbb{R})$ allora $\det(E) = \pm 1$ ed $E^{-1} = E^T \in O_n(\mathbb{R})$; infatti $\det(E^T) = \det(E)$ e quindi $1 = \det(E^T E) = \det(E)^2$. Siccome le operazioni di inversa e trasposta commutano tra loro, si ha $(E^T)^{-1} = (E^{-1})^T = (E^T)^T = E$ e questo implica che $E^T \in O_n(\mathbb{R})$.

Possiamo usare le matrici ortogonali per dare un enunciato equivalente al Corollario 1.3.5.

Corollario 1.5.2. *Una matrice $A \in M_{n,n}(\mathbb{R})$ è simmetrica se e solo se è diagonalizzabile con matrici ortogonali, ossia se e solo se esiste $E \in O_n(\mathbb{R})$ tale che $E^{-1}AE$ sia diagonale.*

Dimostrazione. Se per qualche $E \in O_n(\mathbb{R})$ la matrice $D = E^{-1}AE$ è diagonale, allora $A = EDE^{-1} = EDE^T = ED^T E^T = (EDE^T)^T = A^T$. Viceversa, se A è simmetrica, per il teorema spettrale esiste $u_1, \ldots, u_n \in \mathbb{R}^n$ base ortonormale di autovettori per A. Se indichiamo con $E \in M_{n,n}(\mathbb{R})$ la matrice che ha $u_1, \ldots, u_n$ come vettori colonna, allora E è una matrice ortogonale e la matrice $E^{-1}AE$ è diagonale. □

Definizione 1.5.3. Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali euclidei. Diremo che f è una **isometria**, o **ortogonale**, se è bigettiva e preserva i prodotti scalari, ossia se $\langle f(v_1), f(v_2)\rangle = \langle v_1, v_2\rangle$ per ogni $v_1, v_2 \in V$.

Ad esempio, se $u_1, \ldots, u_n$ è una qualunque base ortonormale di uno spazio vettoriale euclideo V, allora l'applicazione $L_{\mathbf{u}}\colon \mathbb{R}^n \to V$, $L_{\mathbf{u}}(x) = \sum_i x_i u_i$, è una isometria. È chiaro che composizione ed inverse di isometrie sono ancora isometrie.

Teorema 1.5.4. *Per un'applicazione lineare $f\colon V \to W$ tra spazi vettoriali euclidei le seguenti condizioni sono equivalenti:*

(1) *f è una isometria;*
(2) *f trasforma ogni base ortonormale di V in una base ortonormale di W;*
(3) *esiste una base ortonormale di V trasformata da f in una base ortonormale di W.*

Dimostrazione. Se $f\colon V \to W$ è una isometria e $v_1, \ldots, v_n$ è una base ortonormale di V, allora dalle uguaglianze $\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle$ per ogni i, j, segue che $f(v_1), \ldots, f(v_n)$ è una successione ortonormale. Per ipotesi f è lineare bigettiva e quindi trasforma basi in basi.

Viceversa se per una qualunque base ortonormale $v_1, \ldots, v_n$ di V accade che $f(v_1), \ldots, f(v_n)$ è una base ortonormale di W, allora f è un isomorfismo lineare dato che per ogni $x = \sum x_i v_i, y = \sum y_i v_i \in V$ si ha:

$$\langle x, y\rangle = \sum_{i,j} x_i y_j \langle v_i, v_j\rangle = \sum_i x_i y_i,$$

$$\langle f(x), f(y)\rangle = \sum_{i,j} x_i y_j \langle f(v_i), f(v_j)\rangle = \sum_i x_i y_i.$$

□

Per definizione, una matrice $E \in M_{n,n}(\mathbb{R})$ è ortogonale se e solo se trasforma la base canonica in una base ortonormale di $\mathbb{R}^n$. Allora, una conseguenza immediata del Teorema 1.5.4 è che l'applicazione lineare $L_E \colon \mathbb{R}^n \to \mathbb{R}^n$ è una isometria se e solo se $E \in O_n(\mathbb{R})$.

LEMMA 1.5.5. *Siano $v_1, \ldots, v_n$ e $u_1, \ldots, u_n$ due basi di uno spazio vettoriale Euclideo e sia $B \in M_{n,n}(\mathbb{R})$ la corrispondente matrice di cambio di base:*

$$(v_1, \ldots, v_n) = (u_1, \ldots, u_n)B.$$

Se due delle seguenti condizioni sono vere, allora è vera anche la terza:

(1) *la base $u_1, \ldots, u_n$ è ortonormale;*
(2) *la base $v_1, \ldots, v_n$ è ortonormale;*
(3) *la matrice B è ortogonale.*

DIMOSTRAZIONE. Come tutte le matrici di cambiamento di base, B è invertibile. Dato che B è ortogonale se e solo se B^{-1} è ortogonale, basta dimostare che se vale (1) allora le condizioni (2) e (3) sono equivalenti. Si ha

$$\langle v_i, v_j \rangle = \sum_{h,k} b_{hi} b_{kj} \langle u_h, u_k \rangle = \sum_h b_{hi} b_{hj},$$

e la sommatoria più a destra non è altro che il coefficiente (i, j) del prodotto $B^T B$. Quindi $B^T B = I$ se e solo se la base $v_1, \ldots, v_n$ è ortonormale. □

Per ogni $0 \neq u \in V$ denotiamo $S_u = R_{u^\perp}$ la riflessione ortogonale rispetto all'iperpiano $u^\perp$; dunque, per (1.2.3) si ha:

$$S_u(x) = x - 2\frac{\langle x, u \rangle}{\langle u, u \rangle} u.$$

Si noti che $S_u = S_{\lambda u}$ per ogni numero reale $\lambda \neq 0$ e quindi è sempre possibile supporre $\langle u, u \rangle = 1$.

Il prossimo teorema implica in particolare che ogni isometria di uno spazio vettoriale euclideo è composizione di riflessioni rispetto ad iperpiani.

TEOREMA 1.5.6. *Sia $f \colon V \to V$ una isometria di uno spazio vettoriale euclideo e denotiamo con* $\operatorname{Fix}(f) = \{v \in V \mid f(v) = v\} = \operatorname{Ker}(f - I)$ *il sottospazio vettoriale dei punti fissi di f. Esiste allora una base $u_1, \ldots, u_k$ di $\operatorname{Fix}(f)^\perp$ tale che $f = S_{u_1} \cdots S_{u_k}$. In particolare,* $\det(f) = (-1)^{\operatorname{rg}(f-I)}$.

DIMOSTRAZIONE. Dimostriamo prima per induzione su $k = \dim \operatorname{Fix}(f)^\perp$ che possiamo scrivere $f = S_{u_1} \cdots S_{u_r}$ con $r \leq k$ e $u_i \in \operatorname{Fix}(f)^\perp$ per ogni i.

Se $k = 0$ vuol dire che $\operatorname{Fix}(f) = V$, ossia $f = I$. Supponiamo $k > 0$ e sia $v \in \operatorname{Fix}(f)^\perp$ un qualunque vettore non nullo. Allora $f(v) \in \operatorname{Fix}(f)^\perp$ dato che per ogni $u \in \operatorname{Fix}(f)$ si ha $0 = \langle u, v \rangle = \langle f(u), f(v) \rangle = \langle u, f(v) \rangle$.

Siccome $v \notin \operatorname{Fix}(f)$, il vettore $u_1 = f(v) - v \in \operatorname{Fix}(f)^\perp$ è diverso da 0 e possiamo considerare l'isometria $g = S_{u_1} f$. Se $x \in \operatorname{Fix}(f)$, allora $\langle u_1, x \rangle = 0$, $g(x) = S_{u_1}(x) = x$ e questo prova $\operatorname{Fix}(f) \subseteq \operatorname{Fix}(g)$. Inoltre

$$g(v) = S_{u_1}(f(v)) = f(v) - 2\frac{\langle f(v), f(v) - v \rangle}{\langle f(v) - v, f(v) - v \rangle}(f(v) - v).$$

Usando di nuovo il fatto che f è una isometria si ottiene

$$\begin{aligned} \langle f(v) - v, f(v) - v \rangle &= \langle f(v), f(v) \rangle - 2\langle f(v), v \rangle + \langle v, v \rangle \\ &= 2\langle f(v), f(v) \rangle - 2\langle f(v), v \rangle \end{aligned}$$

e quindi $g(v) = f(v) - (f(v) - v) = v$. Ma allora $v \in \mathrm{Fix}(g)$, $\mathrm{Fix}(f)$ è un sottospazio proprio di $\mathrm{Fix}(g)$ e quindi $\mathrm{Fix}(g)^\perp$ è un sottospazio proprio di $\mathrm{Fix}(f)^\perp$.

Per induzione si ha $g = S_{u_2} \cdots S_{u_r}$ con $r-1 \le \dim \mathrm{Fix}(g)^\perp$ e $u_2, \ldots, u_r \in \mathrm{Fix}(g)^\perp \subseteq \mathrm{Fix}(f)^\perp$. In conclusione abbiamo dimostrato che $f = S_{u_1} \cdots S_{u_r}$ con $r \le \dim \mathrm{Fix}(f)^\perp$ e $u_i \in \mathrm{Fix}(f)^\perp$ per ogni i. Sia $W = \mathrm{Span}(u_1, \ldots, u_r)$, allora $W^\perp \subseteq \mathrm{Fix}(S_{u_i})$ per ogni i e quindi $W^\perp \subseteq \mathrm{Fix}(f)$, $W = W^{\perp\perp} \supseteq \mathrm{Fix}(f)^\perp$ e questo è possibile solo se $r = k$ e $W = \mathrm{Fix}(f)^\perp$. □

Esempio 1.5.7. Sul piano euclideo $\mathbb{R}^2$ esistono molti esempi non banali di isometrie lineari.

(1) **Rotazioni**: per ogni numero reale α, la rotazione attorno all'origine in $\mathbb{R}^2$ di angolo α è data dalla matrice ortogonale

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

(2) **Riflessioni**: dato $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos(\beta) \\ \sin(\beta) \end{pmatrix}$ = vettore unitario di $\mathbb{R}^2$ (unitario significa $\|u\|^2 = u_1^2 + u_2^2 = 1$), la riflessione associata è data dalla matrice

$$S_u = \begin{pmatrix} u_2^2 - u_1^2 & -2u_1u_2 \\ -2u_1u_2 & u_1^2 - u_2^2 \end{pmatrix} = \begin{pmatrix} -\cos(2\beta) & -\sin(2\beta) \\ -\sin(2\beta) & \cos(2\beta) \end{pmatrix}.$$

Proposizione 1.5.8. *Ogni isometria di $\mathbb{R}^2$ in sé è una rotazione oppure una riflessione: il primo caso accade se e solo se il determinante è $+1$, il secondo se e solo se il determinante è -1.*

Dimostrazione. È immediato osservare che $\det(R_\alpha) = 1$ e $\det(S_u) = -1$. Viceversa, sia $A \in O_2(\mathbb{R})$ una matrice ortogonale, ossia $AA^T = I$, siccome $|A| = |A^T|$, per il teorema di Binet si ha $|A|^2 = 1$ e quindi $|A| = \pm 1$.

Se $|A| = 1$, poiché i vettori colonna di A sono una base ortonormale si può scrivere

$$\begin{pmatrix} \cos(\alpha) & \sin(\beta) \\ \sin(\alpha) & \cos(\beta) \end{pmatrix},$$

dove $\alpha, \beta \in \mathbb{R}$ devono soddisfare le condizioni

$$1 = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) = \cos(\alpha + \beta),$$
$$0 = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) = \sin(\alpha + \beta),$$

da cui segue $\alpha + \beta = 2k\pi$, e quindi $\cos(\beta) = \cos(\alpha)$, $\sin(\beta) = -\sin(\alpha)$.

Se $|A| = -1$ la matrice possiede due autovalori distinti, uno negativo λ ed uno positivo μ. Dato che A è ortogonale ne segue che $\lambda^2 = \mu^2 = 1$, $\lambda = -1$, $\mu = +1$. Se u è un autovettore relativo a λ, allora $Au = -u$ e AS_u è l'identità, in quanto isometria di determinante 1 e con punti fissi diversi dal vettore nullo. □

Esercizi.

30. Provare che $O_n(\mathbb{R})$ e $SO_n(\mathbb{R}) = \{E \in O_n(\mathbb{R}) \mid \det(E) = 1\}$ sono gruppi di matrici, vedi Osservazione **??**.

31 (Decomposizione QR). Sia $A \in M_{n,n}(\mathbb{R})$ invertibile. Provare che esiste unica una decomposizione $A = QR$ con Q matrice ortogonale e R matrice triangolare superiore con elementi sulla diagonale principale positivi. (Sugg.: POGS alle colonne di A).

32. Un'applicazione lineare $f\colon V \to W$ tra spazi vettoriali euclidei si dice una **similitudine** se è bigettiva e preserva le perpendicolarità, ossia se $f(v_1) \perp f(v_2)$ ogniqualvolta $v_1 \perp v_2$. Dimostrare che ogni similitudine è un multiplo scalare di una isometria, ossia che ogni similitudine è composizione di una isometria ed una omotetia. (Suggerimento: fissata una base ortonormale $u_1, \ldots, u_n$, utilizzare le perpendicolarità $u_i \perp u_j$ e $(u_i - u_j) \perp (u_i + u_j)$ pwer ogni $i \neq j$.)

33. Sia $E \in O_n(\mathbb{R})$ con n dispari. Dimostrare che esiste $v \neq 0 \in V$ tale che $E^2 v = v$.

34. Per ogni vettore non nullo $u \in \mathbb{R}^n$ indichiamo con S_u la riflessione ortogonale rispetto all'iperpiano perpendicolare a u. Dimostrare che:
   (1) se u, v, w sono vettori linearmente indipendenti, allora la composizione $S_v S_w$ non è l'identità e la composizione $S_u S_v S_w$ non è una riflessione rispetto ad un iperpiano;
   (2) per ogni isometria $f\colon \mathbb{R}^n \to \mathbb{R}^n$ vale $f S_u f^{-1} = S_{f(u)}$;
   (3) siano $u, v, w, z \in \mathbb{R}^3$ vettori non nulli, allora esistono $a, b \in \mathbb{R}^3$ tali che $S_u S_v S_w S_z = S_a S_b$.

35. Dimostrare che un'applicazione lineare $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ è:
   (1) una proiezione ortogonale se e solo se $A^2 = A = A^T$;
   (2) una riflessione ortogonale se e solo se $A^2 = I = AA^T$.

36. Sia $A \in M_{n,n}(\mathbb{R})$ antisimmetrica. Dimostrare che:
   (1) $I - A^2$ è simmetrica e definita positiva;
   (2) $I - A$ è invertibile;
   (3) $(I + A)(I - A)^{-1} = (I - A)^{-1}(I + A)$;
   (4) la matrice $E = (I - A)^{-1}(I + A) = (I - A)^{-2}(I - A^2)$ è ortogonale, con determinante $+1$ e tutti gli autovalori diversi da -1.

Viceversa, provare che se E è una matrice ortogonale con $\det(E + I) \neq 0$ (e quindi con $\det(E) = 1$, vedi Esercizio **??**), allora la matrice $A = (E-I)(E+I)^{-1}$ è antisimmetrica.

L'applicazione $A \mapsto E$, dalle matrici antisimmetriche reali alle matrici ortogonali con autovalori $\neq -1$, viene detta *trasformazione di Cayley*.

37. Provare che, in uno spazio vettoriale euclideo, le riflessioni ortogonali sono le uniche isometrie triangolabili.

1.6. Regressione lineare e valori singolari

Diremo che una matrice simmetrica reale è **semidefinita positiva** se tutti i suoi autovalori sono non negativi; lo stesso argomento usato nella dimostrazione del Teorema 1.4.3, con $\leq$ al posto di $<$, mostra che una matrice simmetrica $A \in M_{n,n}(\mathbb{R})$ è semidefinita positiva se e solo se $x^T A x \geq 0$ per ogni $x \in \mathbb{R}^n$.

Se $A \in M_{n,n}(\mathbb{R})$ è simmetrica e semidefinita positiva, allora per ogni $B \in M_{n,m}(\mathbb{R})$ la matrice $B^T A B \in M_{m,m}(\mathbb{R})$ è ancora simmetrica e semidefinita positiva; infatti, per ogni $y \in \mathbb{R}^m$ si ha $y^T B^T A B y = (By)^T A (By) \geq 0$. In particolare, se $E \in M_{n,n}(\mathbb{R})$ è ortogonale, allora $E^T = E^{-1}$ e quindi $E^{-1} A E$ è simmetrica semidefinita positiva.

Lemma 1.6.1. *Per ogni $A \in M_{n,m}(\mathbb{R})$ la matrici simmetriche $A^T A \in M_{m,m}(\mathbb{R})$ e $AA^T \in M_{n,n}(\mathbb{R})$ sono semidefinite positive ed hanno lo stesso rango di A.*

Dimostrazione. Per ogni $x \in \mathbb{R}^m$, si ha $x A^T A x = (Ax)^T (Ax) \geq 0$ e questo prova che $A^T A$ è semidefinita positiva. Per dimostrare che $\operatorname{rg}(A) = \operatorname{rg}(A^T A)$ proviamo che $Ax = 0$ se e solo se $A^T A x = 0$. Se $Ax = 0$ a maggior ragione $A^T A x = 0$; viceversa, se $A^TAx = 0$ allora $(Ax)^T(Ax) = x^T(A^TAx) = 0$ e questo è possibile solo se $Ax = 0$. Considerando A^T al posto di A, troviamo che AA^T è definita positiva dello stesso rango di A^T, che già sappiamo uguale al rango di A. □

Teorema 1.6.2. *Sia $A \in M_{n,m}(\mathbb{R})$, allora per ogni vettore $b \in \mathbb{R}^n$ l'equazione*

$$A^TAx = A^Tb, \qquad x \in \mathbb{R}^m, \tag{1.6.1}$$

possiede soluzioni, che sono tutti e soli i punti di minimo assoluto della funzione

$$f\colon \mathbb{R}^m \to \mathbb{R}, \qquad f(x) = \|Ax - b\|^2.$$

Dimostrazione. Consideriamo le due applicazioni lineari:

$$L_{A^T}\colon \mathbb{R}^n \to \mathbb{R}^m, \qquad L_{A^TA}\colon \mathbb{R}^m \to \mathbb{R}^m\,.$$

Siccome $L_{A^TA} = L_{A^T}L_A$, l'immagine di L_{A^TA} è contenuta nell'immagine di L_{A^T}; per il Lemma 1.6.1 le matrici A^T e A^TA hanno lo stesso rango e questo prova che le applicazioni L_{A^TA} e L_{A^T} hanno la stessa immagine. In particolare, per ogni $b \in \mathbb{R}^m$ il vettore A^Tb appartiene all'immagine di L_{A^TA} e quindi l'equazione $A^TAx = A^Tb$ è risolubile. Per concludere resta da provare che vale $A^TAx = A^Tb$ se e solo se $\|Ay-b\|^2 \geq \|Ax - b\|^2$ per ogni $y \in \mathbb{R}^m$.

Dati $x, y \in \mathbb{R}^m$, denotando $z = A^T(Ax - b) \in \mathbb{R}^m$ si ha

$$(Ay - Ax)^T(Ax - b) = (y - x)^TA^T(Ax - b) = (y - x)^Tz$$

e quindi

$$\begin{aligned}\|Ay - b\|^2 &= \|(Ay - Ax) + (Ax - b)\|^2\\ &= \|Ay - Ax\|^2 + 2(Ay - Ax)^T(Ax - b) + \|Ax - b\|^2\\ &= \|Ay - Ax\|^2 + 2(y - x)^Tz + \|Ax - b\|^2.\end{aligned}$$

Se $z = 0$ allora

$$|Ay - b\|^2 = \|Ay - Ax\|^2 + \|Ax - b\|^2 \geq \|Ax - b\|^2 \text{ per ogni } y.$$

Se invece $z \neq 0$, prendendo $y = x + tz$ con $t \in \mathbb{R}$ si ottiene

$$\|Ay - b\|^2 = t^2\|Az\|^2 + 2t\|z\|^2 + \|Ax - b\|^2$$

e quindi $\|Ax - b\|^2 > \|A(x + tz) - b\|^2$ per ogni t negativo e di valore assoluto sufficientemente piccolo. □

Corollario 1.6.3 (Formule di regressione lineare).
1) Per ogni matrice $A \in M_{n,m}(\mathbb{R})$ ed ogni vettore $b \in \mathbb{R}^n$ vale l'uguaglianza

$$A^TAx = A^Tb, \qquad x \in \mathbb{R}^m,$$

se e solo se Ax è la proiezione ortogonale di b sull'immagine di L_A.
2) Per ogni matrice $B \in M_{k,n}(\mathbb{R})$ ed ogni vettore $c \in \mathbb{R}^n$ vale l'uguaglianza

$$BB^Tx = Bc, \qquad x \in \mathbb{R}^k,$$

se e solo se $c - B^Tx$ è la proiezione ortogonale di c sul nucleo di L_B.

Dimostrazione. Sappiamo, per il Teorema 1.6.2 che entrambe le equazioni $A^TAx = A^Tb$ e $BB^Tx = Bc$ nell'incognita x ammettono soluzioni.

Vale $A^TAx - A^Tb = 0$ se e solo se se per ogni $v \in \mathbb{R}^m$ vale

$$0 = v^T(A^TAx - A^Tb) = (Av)^T(Ax - b),$$

ossia se e solo se Ax è la proiezione ortogonale di b sull'immagine di L_A.

Vale $BB^Tx = Bc$ se e solo se $c - B^Tx$ appartiene al nucleo di L_B; in aggiunta, se $BB^Tx = Bc$ allora per ogni $y \in \operatorname{Ker} L_B$ si ha

$$y^T(c - (c - B^Tx)) = y^T(B^Tx) = (By)^Tx = 0$$

e questo significa che $c - B^Tx$ è la proiezione ortogonale di c su $\operatorname{Ker} L_B$. $\square$

Possiamo interpretare i due punti del Corollario 1.6.3 come metodi costruttivi per calcolare le proiezioni ortogonali su un sottospazio $H \subseteq \mathbb{R}^n$; il primo per i sottospazi definiti in forma parametrica, il secondo per quelli in forma cartesiana.

Ricordiamo che avere un sottospazio $H \subseteq \mathbb{R}^n$ di dimensione m descritto in forma parametrica significa avere una matrice $A \in M_{n,m}(\mathbb{R})$ tale che $H = \{Ax \mid x \in \mathbb{R}^m\}$. In tal caso A ha rango massimo m e quindi la matrice $A^TA \in M_{n,n}(\mathbb{R})$ è invertibile. Per il Corollario 1.6.3 la matrice che rappresenta la proiezione ortogonale su H è quindi $A(A^TA)^{-1}A^T$.

Ricordiamo anche che avere un sottospazio $H \subseteq \mathbb{R}^n$ di dimensione $n - m$ descritto in forma cartesiana significa avere una matrice $B \in M_{m,n}(\mathbb{R})$ di rango m tale che $H = \{x \in \mathbb{R}^n \mid Bx = 0\}$. In questo caso è la matrice $BB^T \in M_{m,m}$ ad essere invertibile Per il Corollario 1.6.3 la matrice che rappresenta la proiezione ortogonale su H è quindi $I - B^T(BB^T)^{-1}B$.

Esempio 1.6.4. Vogliamo calcolare la proiezione ortogonale del vettore $v = (1, 1, 1)^T \in \mathbb{R}^3$ sul sottospazio H di equazione $x_1 + 2x_2 + 3x_3 = 0$. Scrivendo l'equazione del sottospazio come $Bx = 0$, dove $B = (1, 2, 3) \in M_{1,3}(\mathbb{R})$, abbiamo visto che $p_H(v) = v - B^Ty$, dove $BB^Ty = Bv$. Siccome $BB^T = 14$ e $Bv = 6$ si ha $y = 3/7$ e quindi

$$p_H(v) = v - \frac{3}{7}B^T = \left(\frac{4}{7}, \frac{1}{7}, \frac{-2}{7}\right)^T.$$

Lemma 1.6.5. *Siano $A \in M_{n,m}(\mathbb{R})$ e $u_1, \dots, u_m \in \mathbb{R}^m$ una base ortonormale di autovettori per la matrice simmetrica $A^TA \in M_{m,m}(\mathbb{R})$. Allora i vettori $Au_1, \dots, Au_m$ sono tra loro perpendicolari.*

Dimostrazione. Denotiamo con $\lambda_i \in \mathbb{R}$ l'autovalore di A^TA corrispondente all'autovettore u_i. Per ogni $i \neq j$ si ha

$$(Au_i) \cdot (Au_j) = (Au_i)^T(Au_j) = u_i^TA^TAu_j = \lambda_j u_i^Tu_j = 0.$$

Giova sottolineare che il lemma si applica ad ogni base ortonormale di autovettori, che esiste per il teorema spettrale, ma che non è unica in generale. $\square$

Teorema 1.6.6 (decomposizione a valori singolari). *Sia $A \in M_{n,m}(\mathbb{R})$ di rango r. Allora esistono due basi ortonormali $u_1, \dots, u_m \in \mathbb{R}^m$, $v_1, \dots, v_n \in \mathbb{R}^n$ ed una successione monotona $\lambda_1 \geq \dots \geq \lambda_r > 0$ di numeri reali positivi tali che*

$$Au_i = \begin{cases} \lambda_i v_i & \text{se } i \leq r \\ 0 & \text{se } i > r. \end{cases}$$

Inoltre, la successione $\lambda_1, \dots, \lambda_r$ è unica e $\lambda_1^2, \dots, \lambda_r^2$ sono gli autovalori positivi, contati con molteplicità, della matrice A^TA.

Dimostrazione. Sia $\mu_1 \geq \mu_2 \geq \dots \geq \mu_m \geq 0$ la successione, ordinata in senso decrescente, degli autovalori della matrice simmetrica semidefinita positiva $A^TA \in M_{m,m}(\mathbb{R})$ e sia $u_1, \dots, u_m \in \mathbb{R}^m$ una base ortonormale tale che $A^TAu_i = \mu_i u_i$ per ogni i. Siccome le matrici A^TA ed A hanno lo stesso nucleo e stesso rango, si ha $\mu_i > 0$ per $i \leq r$, $\mu_i = 0$ per $i > r$ e $\operatorname{Ker}(A) = \operatorname{Span}(u_{r+1}, \cdots, u_m)$.

Dunque, l'immagine di L_A, che ha dimensione r, è generata dai vettori $Au_1, \dots, Au_r$ che pertanto sono linearmente indipendenti e tra loro perpendicolari per il Lemma 1.6.5.

Se poniamo $\lambda_i = \|Au_i\|$, i vettori

$$v_1 = \frac{Au_1}{\lambda_1}, \quad \cdots \quad v_r = \frac{Au_r}{\lambda_r},$$

sono una successione ortonormale che può essere estesa ad una base ortonormale $v_1, \dots, v_n$ di $\mathbb{R}^n$. Si ha dunque

$$Au_i = \begin{cases} \lambda_i v_i & \text{se } i \leq r \\ 0 & \text{se } i > r. \end{cases}$$

Per ogni $i = 1, \dots, r$ abbiamo

$$\lambda_i^2 = \|Au_i\|^2 = (Au_i)^T(Au_i) = u_i^T A^T A u_i = \mu_i \|u_i\|^2 = \mu_i$$

e questo prova sia le disuguaglianze $\lambda_i \geq \lambda_j$ per ogni $i \leq j$, sia l'unicità della successione $\lambda_1, \dots, \lambda_r$. □

DEFINIZIONE 1.6.7. I numeri $\lambda_1, \dots, \lambda_r$ introdotti nel Teorema 1.6.6 sono detti i **valori singolari** della matrice A. Equivalentemente, i valori singolari di A sono le radici quadrate degli autovalori positivi della matrice $A^T A$.

ESEMPIO 1.6.8. I valori singolari delle matrici

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad A^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix},$$

sono $\sqrt{3}$ e 1, dato che gli autovalori di $A^T A$ sono $3, 1, 0$ e gli autovalori di AA^T sono $3, 1$.

PROPOSIZIONE 1.6.9. *Una matrice $A \in M_{n,m}(\mathbb{R})$ e la sua trasposta $A^T \in M_{m,n}(\mathbb{R})$ hanno gli stessi valori singolari. Conseguentemente, la matrici $A^T A$ e AA^T hanno gli stessi autovalori positivi.*

DIMOSTRAZIONE. Nella stesse notazioni del Teorema 1.6.6, grazie all'unicità dei valori singolari, basta dimostrare che

$$A^T v_i = \begin{cases} \lambda_i u_i & \text{se } i \leq r \\ 0 & \text{se } i > r. \end{cases}$$

Per il Lemma 1.1.9, applicato al prodotto scalare canonico, si ha

$$A^T v_i = \sum_{j=1}^{n} (A^T v_i \cdot u_j) u_j = \sum_{j=1}^{n} (v_i \cdot Au_j) u_j = \sum_{j=1}^{r} (v_i \cdot \lambda_j v_j) u_j$$

e tutto segue dal fatto che la base $v_1, \dots, v_m$ è ortonormale. □

Esercizi.

38. Determinare le proiezioni ortogonali del vettore $(1, 1, 1, 1)^T \in \mathbb{R}^4$ sugli iperpiani H, K di equazioni $x_1 + x_2 + x_3 + x_4 = 0$, $x_1 - x_2 - x_3 + x_4 = 0$ e sulla loro intersezione.

39 (♡). Dire se esiste una matrice $A \in M_{4,4}(\mathbb{R})$ take che

$$A^T A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 2 & 0 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

40. Abbiamo visto che ogni matrice reale B ha lo stesso rango di $B^T B$. Trovare una matrice 2×2 a coefficienti complessi per cui la precedente proprietà non vale.

41. Data la matrice

$$X = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix},$$

dire se L_X è una proiezione ortogonale.

42. Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica semidefinita positiva. Dimostrare che:
(1) $\det(I + A) \geq 1$;
(2) se $x^T A x = 0$, allora $Ax = 0$;
(3) se $\mathrm{rg}(A) \leq 1$, allora esiste $B \in M_{n,1}(\mathbb{R})$ tale che $A = BB^T$.

43. Sia A una matrice antisimmetrica reale. Dimostrare che $-A^2$ è una matrice simmetrica semidefinita positiva.

44. Se gli autovalori di una matrice simmetrica $A \in M_{3,3}(\mathbb{R})$ sono $1, 0, -2$, quali sono i suoi valori singolari?

45 (radici quadrate). Siano $A \in M_{n,n}(\mathbb{R})$ simmetrica semidefinita positiva ed r un intero positivo. Dimostrare che:
(1) esiste una matrice P simmetrica e semidefinita positiva tale che $P^r = A$ (usare il teorema spettrale per ridursi al caso diagonale);
(2) ogni base di autovettori per A è anche una base di autovettori di P (dal fatto che P è diagonalizzabile con tutti gli autovalori non negativi, dedurre che per ogni $\lambda \geq 0$ il nucleo di $P - \lambda I$ è uguale al nucleo di $A - \lambda^r I$);
(3) la matrice P è unica.

In particolare esiste unica una matrice P tale che $P^2 = A$.

46 (☕). Sia $A \in M_{n,n}(\mathbb{R})$ simmetrica definita positiva. Si dimostri che vi è una matrice triangolare superiore B tale che $A = B^T B$.

47 (decomposizione polare). Sia $A \in M_{n,n}(\mathbb{R})$ invertibile. Dimostrare che esistono uniche una matrice P simmetrica definita positiva ed una matrice E ortogonale tali che $A = EP$. (Per l'esistenza usare la decomposizione a valori singolari, per l'unicità osservare che $P^2 = A^T A$.)

48 (☕). Siano $A, B \in M_{n,n}(\mathbb{R})$ matrici simmetriche semidefinite positive. Dimostrare che la matrice $AB + I$ è invertibile e che ogni autovalore di AB è non negativo.

1.7. Spazi vettoriali hermitiani

Le stesse considerazioni che hanno portato alla nozione di spazio vettoriale euclideo, trasportate nel caso complesso, portano alla definizione di spazio vettoriale hermitiano.

DEFINIZIONE 1.7.1. Una **forma hermitiana** su uno spazio vettoriale complesso V è un'applicazione

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$$

che soddisfa le seguenti condizioni:
(1) $\langle \cdot, \cdot \rangle$ è lineare sulla prima variabile, e cioè, per ogni $u \in V$ l'applicazione $\langle \cdot, u \rangle : V \to \mathbb{C}$, $v \mapsto \langle v, u \rangle$, è lineare;
(2) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ per ogni $u, v \in V$.

Segue dalla definizione che ogni forma hermitiana $\langle\cdot,\cdot\rangle$ è $\mathbb{R}$-lineare ma non $\mathbb{C}$-lineare sulla seconda variabile. Infatti, per ogni $u,v\in V$ ed ogni $a\in\mathbb{C}$ si ha

$$\langle u,av\rangle=\overline{\langle av,u\rangle}=\overline{a\langle v,u\rangle}=\overline{a}\,\overline{\langle v,u\rangle}=\overline{a}\langle u,v\rangle.$$

Inoltre, per ogni $v\in V$ si ha $\langle v,v\rangle=\overline{\langle v,v\rangle}$ e quindi $\langle v,v\rangle\in\mathbb{R}$.

Definizione 1.7.2. Un **prodotto hermitiano** su uno spazio vettoriale complesso V è una forma hermitiana

$$\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{C}$$

tale che $\langle v,v\rangle>0$ per ogni $0\neq v\in V$.

Ad esempio, il prodotto hermitiano canonico $\langle x,y\rangle=x^T\overline{y}=\sum_i x_i\overline{y_i}$ su $\mathbb{C}^n$ è un prodotto hermitiano.

Esempio 1.7.3 (Matrici hermitiane). Ricordiamo dalla Definizione **??** che una matrice $H\in M_{n,n}(\mathbb{C})$ si dice hermitiana se $H^T=\overline{H}$, dove $\overline{H}$ è la matrice ottenuta coniugando tutti i coefficienti di H. Separando le parti reali e immaginarie dei coefficienti si può scrivere $H=A+iB$, con $A,B\in M_{n,n}(\mathbb{R})$. Allora $H^T=A^T+iB^T$, $\overline{H}=A-iB$ e quindi H è hermitiana se e solo se A è simmetrica e B è antisimmetrica.

Data una matrice $H=(h_{ij})\in M_{n,n}(\mathbb{C})$, l'applicazione

$$\langle\cdot,\cdot\rangle\colon\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C},\qquad\langle x,y\rangle=x^TH\overline{y}=\sum_{i,j}x_ih_{ij}\overline{y_j}$$

è una forma hermitiana se e solo se H è hermitiana. Se H è hermitiana e $\langle v,v\rangle=v^TH\overline{v}>0$ per ogni $v\neq 0$ diremo che H è *definita positiva*.

Definizione 1.7.4. Uno **spazio vettoriale hermitiano** è uno spazio vettoriale complesso di dimensione finita corredato di un prodotto hermitiano.

In letteratura sono frequenti i termini **spazio vettoriale unitario** e **spazio vettoriale metrico complesso** con lo stesso significato di spazio vettoriale hermitiano.

La nozione di successione ortonormale si estende naturalmente al caso hermitiano, cambiando però nome:

Definizione 1.7.5. Una successione di vettori $u_1,\ldots,u_n$ in uno spazio vettoriale hermitiano si dice un **successione unitaria** se $\langle u_i,u_i\rangle=1$ per ogni i e $\langle u_i,u_j\rangle=0$ per ogni $i\neq j$. Una base che è anche una successione unitaria si dice una **base unitaria**.

Ad esempio, la base canonica di $\mathbb{C}^n$ è unitaria rispetto al prodotto hermitiano canonico. I vettori

$$u_1=\frac{1}{\sqrt{2}}\begin{pmatrix}i\\1\end{pmatrix},\quad u_2=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$$

sono una base unitaria di $\mathbb{C}^2$ rispetto al prodotto hermitiano canonico.

Ogni successione unitaria è formata da vettori linearmente indipendenti; la dimostrazione è identica al caso ortonormale e viene lasciata come esercizio per il lettore.

Proposizione 1.7.6 ((dis)uguaglianza di Bessel). *Sia $u_1,\ldots,u_n$ una successione unitaria in uno spazio vettoriale hermitiano V. Allora per ogni $v\in V$ si ha*

$$\langle v,v\rangle\geq\sum_{i=1}^n|\langle v,u_i\rangle|^2$$

e vale l'uguaglianza se e solo se $v\in\mathrm{Span}(u_1,\ldots,u_n)$.

Dimostrazione. Consideriamo il vettore $w = v - \sum_{i=1}^{n}\langle v, u_i\rangle u_i$. Allora

$$\langle u_k, w\rangle = \langle u_k, v\rangle - \sum_{i=1}^{n}\overline{\langle v, u_i\rangle}\langle u_k, u_i\rangle = \langle u_k, v\rangle - \overline{\langle v, u_k\rangle} = 0$$

per ogni k e quindi

$$0 \leq \langle w, w\rangle = \langle v, w\rangle = \langle v, v\rangle - \sum_{i=1}^{n}|\langle v, u_i\rangle|^2,$$

con l'uguaglianza che vale se e solo se $w = 0$. Se $w = 0$ è chiaro che $v \in \mathrm{Span}(u_1, \ldots, u_n)$. Viceversa, se $v \in \mathrm{Span}(u_1, \ldots, u_n)$, diciamo $v = \sum a_j u_j$, allora $\langle v, u_i\rangle = a_i$ per ogni i e

$$\langle v, v\rangle = \sum_{i,j} a_i\overline{a_j}\langle u_i, u_j\rangle = \sum_i |a_i|^2 = \sum_{i=1}^{n}|\langle v, u_i\rangle|^2.$$

$\square$

Anche il POGS vale, con lievissime modifiche, negli spazi vettoriali hermitiani.

Teorema 1.7.7. *Sia $v_1, \ldots, v_n$ una successione di vettori linearmente indipendenti in uno spazio vettoriale hermitiano $(V, \langle\cdot,\cdot\rangle)$. Allora esiste un'unica successione unitaria $u_1, \ldots, u_n$ in V tale che*

(1) $\mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(u_1, \ldots, u_i)$*;*
(2) $\langle u_i, v_i\rangle$ *è un numero reale positivo;*

per ogni $i = 1, \ldots, n$. Se $v_1, \ldots, v_k$ è una successione unitaria per qualche $k \leq n$, allora $u_i = v_i$ per ogni $i = 1, \ldots, k$.

Dimostrazione. Anche qui la dimostrazione è quasi identica al caso euclideo e ne diamo solo un accenno. Per l'esistenza, si pone

$$u_1 = \frac{v_1}{\sqrt{\langle v_1, v_1\rangle}},$$

e poi per ricorrenza

$$u_i = \frac{w_i}{\sqrt{\langle w_i, w_i\rangle}}, \qquad \text{dove} \qquad w_i = v_i - \sum_{j=1}^{i-1}\langle v_i, u_j\rangle u_j, \quad i = 1, \ldots, n. \tag{1.7.1}$$

Segnaliamo che, a differenza del caso euclideo, qui occorre fare attenzione all'ordine dei vettori nella Formula (1.7.1), dato che in generale $\langle v_i, u_j\rangle \neq \langle u_j, v_i\rangle$.

Per l'unicità, se $x_1, \ldots, x_n \in V$ è una successione unitaria con le stesse caratteristiche, dimostriamo per induzione su n che $u_i = x_i$ per ogni i. Per l'ipotesi induttiva $x_i = u_i$ per ogni $i < n$ e $x_n \in \mathrm{Span}(u_1, \ldots, u_n)$, diciamo $x_n = \sum_i a_i u_i$.

Ma allora $0 = \langle x_n, x_j\rangle = \langle x_n, u_j\rangle = a_j$ per ogni $j < n$ e quindi $x_n = a_n u_n$. Dalla condizione $0 < \langle x_n, v_n\rangle = a_n\langle u_n, v_n\rangle$ segue che a_n è un numero reale positivo, e dalla condizione $1 = \langle x_n, x_n\rangle = |a_n|^2\langle u_n, u_n\rangle$ segue $a_n = 1$, ossia $x_n = u_n$. $\square$

Corollario 1.7.8. *In uno spazio vettoriale hermitiano ogni successione unitaria si estende ad una base unitaria.*

Su $\mathbb{C}^n$, dotato del prodotto hermitiano canonico, la condizione che i vettori colonna di una matrice $U \in M_{n,n}(\mathbb{C})$ formino una base unitaria può essere scritta nella forma $U^T\overline{U} = I$, o equivalentemente, $U^{-1} = \overline{U}^T$.

Definizione 1.7.9. Una matrice $U \in M_{n,n}(\mathbb{C})$ si dice **unitaria** se è invertibile e se $U^{-1} = \overline{U}^T$. Si denota con $U_n(\mathbb{C}) \subseteq M_{n,n}(\mathbb{C})$ il sottoinsieme delle matrici unitarie.

Equivalentemente, una matrice è unitaria se trasforma i vettori della base canonica in una base unitaria per il prodotto hermitiano canonico. Un'altra caratterizzazione delle matrici unitarie è stata data nella Sezione **??**.

PROPOSIZIONE 1.7.10. *Gli autovalori ed i coefficienti del polinomio caratteristico di una matrice hermitiana sono numeri reali. Gli autovalori ed il determinante di una matrice unitaria sono numeri complessi di modulo* 1.

DIMOSTRAZIONE. Sia H una matrice hermitiana, allora

$$\det(H - tI) = \det(H^T - tI) = \det(\overline{H - tI}) = \overline{\det(H - tI)},$$

e quindi tutti i coefficienti del polinomio caratteristico sono reali. Se λ è un autovalore con autovettore $v \in \mathbb{C}^n$ si ha

$$\lambda v^T\overline{v} = (Hv)^T\overline{v} = v^T\overline{H}\overline{v} = v^T\overline{(Hv)} = \overline{\lambda} v^T\overline{v},$$

da cui segue $\lambda = \overline{\lambda}$.

Se U è unitaria, allora $\det(U)^{-1} = \det(U^{-1}) = \det(\overline{U}^T) = \det(\overline{U}) = \overline{\det(U)}$ da cui segue che $\det(U)$ ha modulo 1. Se λ è un autovalore con autovettore $v \in \mathbb{C}^n$ si ha

$$\overline{\lambda} v^T\overline{v} = v^T\overline{Uv} = v^T\overline{U}\overline{v} = (\overline{U}v)^T\overline{v} = (U^{-1}v)^T\overline{v} = \lambda^{-1}v^T\overline{v}$$

da cui segue $\lambda^{-1} = \overline{\lambda}$. □

Esercizi.

49. Determinare per quali valori di $t \in \mathbb{R}$ la matrice hermitiana

$$\begin{pmatrix} 1 & it \\ -it & 1 \end{pmatrix} \in M_{2,2}(\mathbb{C})$$

è definita positiva.

50. Trovare una base unitaria del sottospazio $H \subseteq \mathbb{C}^3$ di equazione

$$x - iy + z = 0, \qquad i = \sqrt{-1}.$$

51. Siano V uno spazio vettoriale hermitiano e $W \subseteq V$ un sottospazio vettoriale. Definiamo l'ortogonale di W come

$$W^\perp = \{v \in V \mid \langle w, v\rangle = 0 \quad \forall\, w \in W\}.$$

Dimostrare che $W^\perp$ è un sottospazio vettoriale e che $W \oplus W^\perp = V$.

52. Sia $H = A + iB \in M_{n,n}(\mathbb{C})$ con $A, B \in M_{n,n}(\mathbb{R})$. Provare che H è hermitiana definita positiva se e solo se la matrice a blocchi

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in M_{2n,2n}(\mathbb{R})$$

è simmetrica e definita positiva.

53. Siano $A, B \in M_{n,n}(\mathbb{C})$ matrici hermitiane. Provare che $\mathrm{Tr}(AB) \in \mathbb{R}$ e mostrare con un esempio che i coefficienti sulla diagonale principale di AB non sono, in generale, numeri reali.

54. Dimostrare che ogni matrice $L \in M_{n,n}(\mathbb{C})$ si decompone in modo unico nella forma $L = H + iK$ con H, K hermitiane.

55 (disuguaglianza di Cauchy–Schwarz). Dimostrare che per ogni coppia di vettori v, w in uno spazio vettoriale hermitiano si ha

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle,$$

ed il segno di uguaglianza vale se e solo se v, w sono linearmente dipendenti.

56. Provare che $U_n(\mathbb{C})$ e $SU_n(\mathbb{C}) = \{E \in U_n(\mathbb{C}) \mid \det(E) = 1\}$ sono gruppi di matrici, vedi Osservazione **??**.

1.8. Il teorema spettrale complesso

Prima di enunciare il teorema spettrale complesso abbiamo bisogno di introdurre la nozione di operatore aggiunto.

Teorema 1.8.1. *Sia $f\colon V \to W$ un'applicazione lineare tra spazi vettoriali hermitiani. Esiste allora un'unica applicazione lineare $f^*\colon W \to V$ tale che $\langle f(v), w\rangle = \langle v, f^*(w)\rangle$ per ogni $v \in V$ ed ogni $w \in W$.*

Dimostrazione. Fissiamo una base unitaria $v_1, \dots, v_n$ di V, una base unitaria $w_1, \dots, w_m$ di W e sia $A = (a_{ij})$ la matrice che rappresenta f in tali basi, ossia

$$f(v_i) = \sum_j a_{ji} w_j.$$

Dimostriamo prima l'unicità di f^*; se $B = (b_{ij})$ è la matrice che rappresenta f^* nelle basi suddette, ossia

$$f^*(w_j) = \sum_i b_{ij} v_i,$$

si hanno le relazioni

$$\langle f(v_i), w_j\rangle = a_{ji}, \quad \langle v_i, f^*(w_j)\rangle = \overline{b_{ij}}, \quad i, j = 1, \dots, n,$$

da cui segue necessariamente $b_{ij} = \overline{a_{ji}}$ per ogni i, j, ossia $B = \overline{A}^T$.

Per la prova dell'esistenza basta provare che la formula appena trovata per f^* è consistente con le proprietà richieste. Per ogni $v = \sum_i x_i v_i \in V$ ed ogni $w = \sum_j y_j w_j$ si ha

$$\langle f(v), w\rangle = \sum_{i,j} x_i \overline{y_j} \langle f(v_i), w_j\rangle = \sum_{i,j} a_{ji} x_i \overline{y_j},$$
$$\langle v, f^*(w)\rangle = \sum_{i,j} x_i \overline{y_j} \langle v_i, f^*(w_j)\rangle = \sum_{i,j} \overline{b_{ij}} x_i \overline{y_j},$$

e l'uguaglianza tra le due espressioni segue immediatamente da $B = \overline{A}^T$. □

Definizione 1.8.2. Nelle notazioni del Teorema 1.8.1, l'applicazione f^* viene detta **operatore aggiunto** di f.

Nel caso in cui $f\colon V \to V$ è un endomorfismo di uno spazio vettoriale hermitiano, se A è la matrice che rappresenta f in una base unitaria, abbiamo visto nella dimostrazione del Teorema 1.8.1 che f^* è rappresentato, nella stessa base, dalla matrice trasposta coniugata $\overline{A}^T$.

Esempio 1.8.3. Sia $f\colon V \to V$ un endomorfismo di uno spazio vettoriale hermitiano. Allora:

1) Dato un numero complesso λ, si ha $(f - \lambda I)^* = f^* - \overline{\lambda} I$.

2) Il doppio aggiunto f^{**} è uguale all'operatore di partenza f; in particolare $\langle f^*(u), v\rangle = \langle u, f(v)\rangle$. Questo è chiaro sia guardando le matrici che rappresentano f, f^* in una base

unitaria, sia utilizzando l'unicità dell'operatore aggiunto, dato che per ogni $v, w \in V$ si ha

$$\langle u, f^{**}(v)\rangle = \langle f^*(u), v\rangle = \overline{\langle v, f^*(u)\rangle} = \overline{\langle f(v), u\rangle} = \langle u, f(v)\rangle.$$

Definizione 1.8.4. Un endomorfismo $f\colon V \to V$ di uno spazio vettoriale hermitiano si dice un **operatore normale** se commuta con il suo aggiunto, ossia se $ff^* = f^*f$.

I due principali esempi di operatori normali $f\colon V \to V$ sono:

(1) gli operatori, detti **autoaggiunti** o **hermitiani**, tali che $f = f^*$; sono quelli che in una base unitaria sono rappresentati da una matrice hermitiana;
(2) gli operatori, detti **unitari** o **isometrie**, tali che $f^{-1} = f^*$; sono quelli che in una base unitaria sono rappresentati da una matrice unitaria.

Lemma 1.8.5. *Sia $f\colon V \to V$ un operatore normale. Allora f ed f^* hanno gli stessi autovettori. Inoltre, se $f(u) = \lambda u$ con $\lambda \in \mathbb{C}$, allora $f^*(u) = \overline{\lambda} u$.*

Dimostrazione. Sia $u \in V$ un autovettore per f, con autovalore λ e consideriamo l'endomorfismo $g = f - \lambda I$. Abbiamo visto che $g^* = f^* - \overline{\lambda} I$ e siccome

$$gg^* - g^*g = (f - \lambda I)(f^* - \overline{\lambda} I) - (f^* - \overline{\lambda} I)(f - \lambda I) = 0$$

anche g è un operatore normale. Dato che $g(u) = 0$ si ha

$$0 = \langle g^*g(u), u\rangle = \langle gg^*(u), u\rangle = \langle g^*(u), g^*(u)\rangle$$

e questo è possibile solo se $g^*(u) = 0$, ossia se $f^*(u) = \overline{\lambda} u$. □

Teorema 1.8.6 (teorema spettrale complesso). *Sia $f\colon V \to V$ un operatore normale in uno spazio hermitiano. Allora esiste una base unitaria di autovettori per f.*

Come potete osservare i teoremi spettrali reale e complesso si assomigliano molto, ed infatti sarebbe possibile, molti lo fanno, inglobarli in un unico teorema. Tuttavia, data la notevole varietà di dimostrazioni, tutte istruttive ed interessanti, non solo abbiamo separato i teoremi ma forniremo anche delle prove basate su diversi approcci.

Dimostrazione. Diamo due diverse dimostrazioni del teorema: la prima utilizza il Lemma 1.8.5, la seconda la (dis)uguaglianza di Bessel.

Prima dimostrazione. Dato che sui complessi ogni endomorfismo possiede autovettori, per il Lemma 1.8.5 (o se preferite per l'Esercizio **??**) gli operatori f, f^* possiedono un autovettore comune, ossia esistono $0 \neq u_1 \in V$ e $\lambda \in \mathbb{C}$ tali che $f(u_1) = \lambda u_1$ e $f^*(u_1) = \overline{\lambda} u_1$. A meno di moltiplicare u per una costante possiamo supporre $\langle u_1, u_1\rangle = 1$. Consideriamo adesso l'iperpiano ortogonale a u_1:

$$W = u_1^\perp = \{v \in V \mid \langle v, u_1\rangle = 0\}.$$

Siccome $u_1 \notin W$ si ha $V = \operatorname{Span}(u_1) \oplus W$; dimostriamo che W è un sottospazio invariante per f ed f^*. Per ogni $v \in W$ si ha

$$\langle f(v), u_1\rangle = \langle v, f^*(u_1)\rangle = \lambda\langle v, u_1\rangle = 0,$$

$$\langle f^*(v), u_1\rangle = \overline{\langle u_1, f^*(v)\rangle} = \overline{\langle f(u_1), v\rangle} = \overline{\lambda\langle u_1, v\rangle} = \overline{\lambda}\langle v, u_1\rangle = 0.$$

È chiaro che in tali condizioni la restrizione di f^* a W coincide con l'aggiunto della restrizione di f a W. Per induzione sulla dimensione esiste una base unitaria $u_2, \ldots, u_n$ di W formata da autovettori per f ed f^*. In conclusione $u_1, u_2, \ldots, u_n$ è la base cercata.

Seconda dimostrazione. Sappiamo che f è triangolabile e, come nella dimostrazione del teorema spettrale reale, esiste una base unitaria $u_1, \ldots, u_n$ di V tale che $f(u_j) \in \mathrm{Span}(u_1, \ldots, u_j)$ per ogni j. In particolare, per la (dis)uguaglianza di Bessel si ha

$$\begin{aligned}\langle f^*(u_j), f^*(u_j)\rangle &= \langle ff^*(u_j), u_j\rangle = \langle f^*f(u_j), u_j\rangle = \langle f(u_j), f(u_j)\rangle \\ &= \sum_{i=1}^{j} |\langle f(u_j), u_i\rangle|^2 = \sum_{i=1}^{k} |\langle f(u_j), u_i\rangle|^2 = \sum_{i=1}^{k} |\langle f^*(u_i), u_j\rangle|^2\end{aligned}$$

per ogni $1 \leq j \leq k \leq n$ e quindi

$$\sum_{i=1}^{k}\langle f^*(u_i), f^*(u_i)\rangle = \sum_{j=1}^{k}\langle f^*(u_j), f^*(u_j)\rangle = \sum_{i,j=1}^{k} |\langle f^*(u_i), u_j\rangle|^2 \tag{1.8.1}$$

per ogni $k = 1, \ldots, n$. Ancora per la (dis)uguaglianza di Bessel si hanno le disuguaglianze

$$\langle f^*(u_i), f^*(u_i)\rangle \geq \sum_{j=1}^{k} |\langle f^*(u_i), u_j\rangle|^2 \tag{1.8.2}$$

per ogni $1 \leq i \leq k \leq n$, che confrontate con (1.8.1) si rivelano uguaglianze, ossia

$$\langle f^*(u_i), f^*(u_i)\rangle = \sum_{j=1}^{k} |\langle f^*(u_k), u_j\rangle|^2$$

per ogni $i \leq k$. Ancora Bessel ci dice quindi che $f^*(u_i) \in \mathrm{Span}(u_1, \ldots, u_k)$ per ogni $i \leq k$; in particolare $f^*(u_k) \in \mathrm{Span}(u_1, \ldots, u_k)$ per ogni k e quindi f, f^* si rappresentano entrambe con matrici triangolari superiori nella base unitaria $u_1, \ldots, u_n$. Per concludere, basta ricordare che se una matrice A rappresenta f in una base unitaria, allora nella stessa base la matrice $\overline{A}^T$ rappresenta f^*. Però le matrici $A, \overline{A}^T$ sono entrambe triangolari superiori se e solo se sono entrambe diagonali. □

Corollario 1.8.7. *Sia $H \in M_{n,n}(\mathbb{C})$ una matrice hermitiana. Allora esiste una matrice unitaria U tale che $U^{-1}HU$ è diagonale reale.*

Dimostrazione. Si prenda come U la matrice cha ha come vettori colonna una base unitaria di autovettori per H. Allora $U^{-1}HU$ è diagonale e già sappiamo che gli autovalori di H sono numeri reali. □

Corollario 1.8.8. *Sia $E \in U_n(\mathbb{C})$ una matrice unitaria. Allora esiste una matrice unitaria U tale che $U^{-1}EU$ è diagonale con i coefficienti principali di modulo* 1.

Dimostrazione. Si prenda come U la matrice cha ha come vettori colonna una base unitaria di autovettori per E. Allora $U^{-1}EU$ è diagonale e già sappiamo che gli autovalori di E hanno modulo 1. □

Esercizi.

57. Trovare una base unitaria di autovettori per la matrice hermitiana

$$\begin{pmatrix} 1 & 6i \\ -6i & 4 \end{pmatrix}.$$

58. Siano H, K matrici hermitiane non nulle dello stesso ordine. Provare che:
(1) $|\det(I + iK)| > 1$;
(2) (☕) se H è definita positiva, allora $|\det(H + iK)| > \det(H)$.

59. Provare che le matrici unitarie di ordine 2 e determinante 1 sono tutte e sole quelle del tipo

$$\begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}$$

con $a, b \in \mathbb{C}$ tali che $|a|^2 + |b|^2 = 1$.

60. Provare che una matrice diagonale è unitaria se e solo se tutti i coefficienti principali hanno modulo 1.

61. Sia U una matrice unitaria. Dimostrare che esiste una matrice unitaria S tale che $S^2 = U$.

62 (Matrici di Pauli). Per definizione, le matrici di Pauli sono:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(1) Provare che le matrici di Pauli sono una base ortonormale del sottospazio (reale) $V \subseteq M_{2,2}(\mathbb{C})$ delle matrici hermitiane a traccia nulla, dotato del prodotto scalare
$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}, \qquad \langle A, B \rangle = \frac{1}{2}\operatorname{Tr}(AB).$$
(2) Per ogni matrice unitaria $U \in U_2(\mathbb{C})$, provare che l'applicazione $\phi_U \colon V \to V$, $\phi_U(A) = UAU^{-1}$ è una isometria.
(3) Provare che ϕ_U è l'identità se e solo se U è un multiplo scalare dell'identità.
(4) Provare che per ogni coppia di matrici unitarie $U, S \in U_2(\mathbb{C})$ si ha $\phi_{US} = \phi_U \phi_S$ e dedurre dall'Esercizio 61 che le isometrie ϕ_U hanno tutte determinante $+1$.

63. Descrivere tutte le matrici $U \in M_{2,2}(\mathbb{C})$ che sono al tempo stesso unitarie ed hermitiane (come ad esempio le matrici di Pauli).

Note. Dal punto di vista computazionale, la decomposizione QR (Esercizio 31) viene spesso preferita a Gauss–Jordan nel calcolo dell'inverse di matrici reali: appurato che $\det(A) \neq 0$, si applica il POGS alle colonne di A e si ottiene $Q = AS$ con Q matrice ortogonale ed S triangolare, e quindi $A^{-1} = SQ^T$. Argomenti quali la regressione lineare e le decomposizioni LU, QR, polare ed a valori singolari, tipicamente relegati ai testi di metodi numerici e ottimizzazione, stanno recentemente trovando spazio nei testi base di algebra lineare per effetto del loro fondamentale utilizzo in intelligenza artificiale.

L'aggettivo *hermitiano* deriva da Charles Hermite (1822–1901) e pertanto si può anche scrivere “spazio Hermitiano” e “matrice Hermitiana” senza violare alcuna regola grammaticale.

CAPITOLO 2

Forme bilineari e quadratiche

In questo capitolo tratteremo la teoria delle forme bilineari e quadratiche. Presenteremo i concetti in ordine di generalità decrescente del campo base e degli spazi vettoriali coinvolti: inizieremo con spazi vettoriali definiti su campi arbitrari, per poi passare ai campi dei numeri reali e complessi; parallelamente, inizieremo con spazi vettoriali qualunque, per poi passare a quelli numerici.

2.1. Nozioni base

Anche se svilupperemo in dettaglio la teoria solamente in dimensione finita e caratteristica $\neq 2$, è comunque interessante dare alcune definizioni generali, senza alcuna restrizione sugli spazi vettoriali coinvolti.

Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Ricordiamo che un'applicazione

$$\varphi\colon V \times V \to \mathbb{K}$$

si dice una **forma bilineare** se per ogni vettore $v \in V$ le due applicazioni

$$\varphi(-, v)\colon V \to \mathbb{K}, \qquad w \mapsto \varphi(w, v),$$
$$\varphi(v, -)\colon V \to \mathbb{K}, \qquad w \mapsto \varphi(v, w),$$

sono lineari.

Esempio 2.1.1. Se $f, g\colon V \to \mathbb{K}$ sono due applicazioni lineari, allora le applicazioni

$$\varphi_1,\ \varphi_2\colon V \times V \to \mathbb{K},$$
$$\varphi_1(x, y) = f(x)g(y) + g(x)f(y), \quad \varphi_2(x, y) = f(x)g(y) - g(x)f(y),$$

sono bilineari. Più in generale, per ogni intero $n > 0$ ed ogni successione di $2n$ applicazioni lineari $f_1, \dots, f_n, g_1, \dots, g_n\colon V \to \mathbb{K}$, l'applicazione

$$\varphi\colon V \times V \to \mathbb{K}, \qquad \varphi(x, y) = \sum_{i=1}^{n} f_i(x)g_i(y),$$

è bilineare.

La seguente definizione è una ripetizione, si spera utile, di concetti già introdotti.

Definizione 2.1.2. Una forma bilineare $\varphi\colon V \times V \to \mathbb{K}$ si dice:
(1) **simmetrica** se $\varphi(x, y) = \varphi(y, x)$ per ogni $x, y \in V$;
(2) **alternante** se $\varphi(x, x) = 0$ per ogni $x \in V$.

Ad esempio, il prodotto scalare canonico è una forma bilineare simmetrica su $\mathbb{R}^n$, mentre il determinante di matrici 2×2 è una forma bilineare alternante sullo spazio dei vettori colonna $\mathbb{K}^2$. Nell''Esempio 2.1.1, la forma bilineare φ_1 è simmetrica, mentre φ_2 è alternante; usando il prodotto esterno (Sezione **??**) si può scrivere $\varphi_2 = f \wedge g$.

Gli stessi argomenti usati nello studio del determinante mostrano che ogni forma alternante φ è antisimmetrica, ossia $\varphi(x, y) = -\varphi(y, x)$ per ogni x, y.

Abbiamo visto nella Sezione **??** che le forme bilineari su uno spazio V formano uno spazio vettoriale ed è immediato osservare che se φ, ψ sono simmetriche (risp.:

alternanti), allora anche ogni loro combinazione lineare $a\varphi + b\psi$ è simmetrica (risp.: alternante).

Definizione 2.1.3. Sia V uno spazio vettoriale su di un campo $\mathbb{K}$. Un'applicazione $\Phi\colon V \to \mathbb{K}$ si dice una **forma quadratica** se:

(1) $\Phi(tx) = t^2\Phi(x)$ per ogni $t \in \mathbb{K}$ ed ogni $x \in V$;
(2) l'applicazione
$$\varphi\colon V \times V \to \mathbb{K}, \qquad \varphi(x, y) = \Phi(x + y) - \Phi(x) - \Phi(y),$$
è una forma bilineare.

Esempio 2.1.4. Se $f, g\colon V \to \mathbb{K}$ sono due applicazioni lineari, allora l'applicazione
$$\Phi\colon V \to \mathbb{K}, \qquad \Phi(x) = f(x)g(x),$$
è una forma quadratica. Infatti, per quanto visto nell'Esempio 2.1.1 l'applicazione
$$(x, y) \mapsto \Phi(x + y) - \Phi(x) - \Phi(x) = f(x)g(y) + f(y)g(x)$$
è una forma bilineare simmetrica.

Alla stessa maniera delle forme bilineari, anche le forme quadratiche possono essere sommate e moltiplicate per scalare. In particolare, ogni espressione del tipo $\sum_{i,j=1}^{n} a_{ij}x_ix_j$, con $a_{ij} \in \mathbb{K}$, definisce una forma quadratica su $\mathbb{K}^n$.

Definizione 2.1.5. Sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare. Chiameremo l'applicazione
$$\Phi\colon V \to \mathbb{K}, \qquad \Phi(x) = \varphi(x, x), \tag{2.1.1}$$
forma quadratica associata, o **indotta**, da φ.

Per mostrare che la precedente definizione abbia senso, verifichiamo che Φ è effettivamente una forma quadratica: si ha $\Phi(tx) = \varphi(tx, tx) = t^2\varphi(x, x) = t^2\Phi(x)$, e
$$\begin{aligned}(x, y) \mapsto \Phi(x + y) &- \Phi(x) - \Phi(y)\\ &= \varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y) = \varphi(x, y) + \varphi(y, x)\end{aligned}$$
è una forma bilineare. È chiaro che due forme bilineari inducono la stessa forma quadratica se e solo se la loro differenza è alternante.

Osservazione 2.1.6 (⚽). In dimensione finita non è difficile dimostrare che ogni forma quadratica è indotta da una forma bilineare (Esercizio 74) e quindi che esiste un isomorfismo naturale tra lo spazio vettoriale delle forme quadratiche ed il quoziente dello spazio vettoriale delle forme bilineari per il sottospazio delle forme alternanti.

Definizione 2.1.7 (pull-back di forme). Siano $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare ed $f\colon W \to V$ un'applicazione lineare. La forma bilineare
$$f^*\varphi\colon W \times W \to \mathbb{K}, \qquad f^*\varphi(x, y) = \varphi(f(x), f(y)),$$
viene detta il **pull-back** di φ tramite f. Similmente, se $\Phi\colon V \to \mathbb{K}$ è una forma quadratica, il suo pull-back tramite f è la forma quadratica
$$f^*\Phi\colon W \to \mathbb{K}, \qquad f^*\Phi(x) = \Phi(f(x)).$$

La Definizione 2.1.7 è ben posta: segue immediatamente dalle definizioni e dalla linearità di f che $f^*\varphi$ è una forma bilineare e $f^*\Phi$ è una forma quadratica. Inoltre, se Φ è indotta da φ, allora $f^*\Phi$ è indotta da $f^*\varphi$.

Esercizi.

64 (legge del parallelogramma). Data una forma quadratica $\Phi\colon V \to \mathbb{K}$, provare che:

$$\Phi(x+y) + \Phi(x-y) = 2\Phi(x) + 2\Phi(y), \quad \text{per ogni } x, y \in V.$$

65. Provare che se $\mathbb{K} = \mathbb{F}_2$ allora ogni funzionale lineare è anche una forma quadratica ed esistono forme quadratiche che non sono lineari. Viceversa, se $\mathbb{K} \neq \mathbb{F}_2$, allora l'unico funzionale lineare che è anche una forma quadratica è quello nullo.

66. Dimostrare o confutare le seguenti affermazioni:

(1) sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare. Allora il sottoinsieme
$$\{v \in V \mid \varphi(v, u) = 0 \text{ per ogni } u \in V\}$$
è un sottospazio vettoriale;

(2) siano $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare e $u_1, u_1 \in V$ vettori linearmente indipendenti. Allora il sottoinsieme
$$\{v \in V \mid \varphi(v, u_1)\varphi(v, u_1) = 0\}$$
è un sottospazio vettoriale;

(3) sia $\Phi\colon V \to \mathbb{K}$ una forma quadratica. Allora il sottoinsieme
$$\{v \in V \mid \Phi(v) = 0\}$$
è un sottospazio vettoriale.

67. Siano V uno spazio vettoriale e $f, g\colon V \to \mathbb{K}$ applicazioni lineari. Provare che la forma bilineare
$$\varphi\colon V \times V \to \mathbb{K}, \qquad \varphi(x, y) = f(x)g(y),$$
è simmetrica se e solo se $\operatorname{Ker} f \cup \operatorname{Ker} g$ è un sottospazio vettoriale di V; equivalentemente, φ è simmetrica se e solo se f e g sono linearmente dipendenti (nota: se V ha dimensione infinita occorre usare i risultati della Sezione **??**).

68 (Ⓐ). Sia V lo spazio vettoriale reale delle funzioni continue sull'intervallo $[0, 1]$ dimostrare che l'applicazione
$$\varphi\colon V \times V \to \mathbb{R}, \qquad \varphi(f, g) = \int_0^1 f(t)g(t)dt,$$
è una forma bilineare.

2.2. Rango e nucleo

Siano V uno spazio vettoriale di dimensione finita n e $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare; ogni base $v_1, \ldots, v_n$ di V permette di associare alla forma φ la matrice
$$B \in M_{n,n}(\mathbb{K}), \qquad B = (b_{ij}), \qquad b_{ij} = \varphi(v_i, v_j),$$
i cui coefficienti sono i valori di φ calcolati nelle n^2 coppie di elementi della base; chiameremo B *matrice associata alla forma bilineare φ nella base $v_1, \ldots, v_n$*. Nel prossimo lemma proveremo che la precedente costruzione è reversibile, ossia che la matrice B determina univocamente la forma φ, e possiamo quindi dire che φ è la *forma bilineare associata alla matrice B nella base $v_1, \ldots, v_n$*.

Lemma 2.2.1. *Sia $v_1, \ldots, v_n$ una base fissata di uno spazio vettoriale V. Per ogni matrice $B = (b_{ij}) \in M_{n,n}(\mathbb{K})$ esiste un'unica forma bilineare $\varphi\colon V \times V \to \mathbb{K}$ tale che $\varphi(v_i, v_j) = b_{ij}$ per ogni i, j.*

Dimostrazione. *Unicità.* Se φ è una forma bilineare su V, data una qualunque coppia di vettori $u = \sum x_i v_i$ e $w = \sum y_i v_i$, $x_i, y_i \in \mathbb{K}$, per bilinearità si ha

$$\varphi(u,w) = \varphi(\sum x_i v_i, \sum y_j v_j) = \sum_{i,j} x_i y_j\, \varphi(v_i, v_j) = \sum_{i,j} x_i y_j b_{ij} \tag{2.2.1}$$

ossia $\varphi(u,w) = x^T B y$, dove $x, y \in \mathbb{K}^n$ sono i vettori delle coordinate di u, v rispetto alla base $v_1, \ldots, v_n$ e $x^T B y$ denota l'usuale prodotto righe per colonne. In particolare φ è univocamente definita dagli scalari $b_{ij} = \varphi(v_i, v_j)$.

Esistenza. Basta definire φ usando la formula

$$\varphi(u,w) = \sum_{i,j} b_{ij} x_i y_j = x^T B y, \qquad u = \sum x_i v_i,\ w = \sum y_i v_i,$$

trovata nella dimostrazione dell'unicità. □

Il Lemma 2.2.1 applicato agli spazi vettoriali numerici ed alle loro basi canoniche ci dice in particolare che ogni applicazione bilineare $\varphi\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ si scrive in maniera unica come

$$\varphi(x,y) = x^T B y, \qquad B \in M_{n,n}(\mathbb{K}), \quad x, y \in \mathbb{K}^n.$$

La matrice associata ad una forma bilineare dipende fortemente dalla scelta della base: cambiando base anche la matrice B cambia. Sia $\varepsilon_1, \ldots, \varepsilon_n$ un'altra base di V, e sia $A = (a_{ij}) \in M_{n,n}(\mathbb{K})$ la corrispondente matrice di cambiamento di base:

$$(\varepsilon_1, \ldots, \varepsilon_n) = (v_1, \ldots, v_n)A \iff \varepsilon_i = \sum_j a_{ji} v_j.$$

Sia $C = (c_{ij})$ la matrice associata a φ nella base ε_i, allora

$$c_{ij} = \varphi(\varepsilon_i, \varepsilon_j) = \varphi(\sum_h a_{hi} v_h, \sum_k a_{kj} v_k) = \sum_{h,k} a_{hi} b_{hk} a_{kj}$$

e quindi $C = A^T B A$. Lo stesso conto dimostra anche che se partiamo da una qualunque matrice invertibile $A \in \mathrm{GL}_n(\mathbb{K})$, allora la matrice $A^T B A$ rappresenta la forma blineare φ nella base $(v_1, \ldots, v_n)A$.

Definizione 2.2.2. Due matrici $B, C \in M_{n,n}(\mathbb{K})$ si dicono **congruenti** se esiste una matrice invertibile $A \in \mathrm{GL}_n(\mathbb{K})$ tale che $C = A^T B A$.

Anche se a prima vista la nozione di congruenza appare quasi identica alla similitudine, le due relazioni hanno significati geometrici e comportamenti molto diversi. Infatti:

- delle matrici sono congruenti se e solo se rappresentano la medesima forma bilineare, in altrettante basi;
- delle matrici sono simili se e solo se rappresentano lo stesso endomorfismo, in altrettanti basi.

Ad esempio, per ogni scalare a tale che $a^2 \neq 0, 1$ le due matrici

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

sono congruenti ma non sono simili, mentre le due matrici

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

sono simili ma non congruenti (esercizio: perché?). Inoltre, la congruenza non è invariante per estensione degli scalari: le matrici $(1), (2) \in M_{1,1}(\mathbb{Q})$ sono congruenti su $\mathbb{R}$ ma non su $\mathbb{Q}$; le matrici $(1), (-1) \in M_{1,1}(\mathbb{R})$ sono congruenti su $\mathbb{C}$ ma non su $\mathbb{R}$.

Matrici congruenti hanno lo stesso rango: infatti se $C = A^T BA$ per qualche matrice invertibile A, allora anche A^T è invertibile e sappiamo che il rango è invariante per moltiplicazione, sia a destra che a sinistra, con matrici invertibili. Questo prova in particolare che la prossima definizione è ben posta.

Definizione 2.2.3. Il **rango** $\mathrm{rg}(\varphi)$, di una forma bilineare φ su uno spazio vettoriale di dimensione finita V, è il rango della matrice che rappresenta φ in una (qualunque) base di V.

Ad esempio, il rango della forma bilineare $(x, y) \mapsto x^T By$ su $\mathbb{K}^n$ è uguale al rango della matrice $B \in M_{n,n}(\mathbb{K})$ (basta prendere come base quella canonica).

Poiché una matrice e la sua trasposta hanno lo stesso rango, segue immediatamente che una forma bilineare φ ha il medesimo rango della forma bilineare φ^T ottenuta scambiando l'ordine delle variabili: $\varphi^T(u, v) = \varphi(v, u)$. Il prossimo risultato è l'equivalente bilineare del teorema del rango **??**.

Teorema 2.2.4. *Sia φ una forma bilineare di rango r su uno spazio vettoriale V di dimensione finita n. Allora:*

(1) *il sottospazio vettoriale* (nucleo destro)
$$\mathrm{RKer}(\varphi) = \{v \in V \mid \varphi(u, v) = 0 \text{ per ogni } u \in V\}$$
ha dimensione $n - r$;

(2) *il sottospazio vettoriale* (nucleo sinistro)
$$\mathrm{LKer}(\varphi) = \{v \in V \mid \varphi(v, u) = 0 \text{ per ogni } u \in V\}$$
ha dimensione $n - r$.

Dimostrazione. Siccome $\mathrm{LKer}(\varphi) = \mathrm{RKer}(\varphi^T)$ e le forme φ, φ^T hanno lo stesso rango, basta dimostrare l'enunciato sul nucleo destro. Fissiamo una base di V, ossia un isomorfismo di spazi vettoriali $L_{\mathbf{v}} \colon \mathbb{K}^n \to V$ e sia B la matrice associata a φ in tale base. Per costruzione $\varphi(L_{\mathbf{v}}(x), L_{\mathbf{v}}(y)) = x^T By$ e
$$L_{\mathbf{v}}^{-1}(\mathrm{RKer}(\varphi)) = \{y \in \mathbb{K}^n \mid x^T By = 0 \text{ per ogni } x \in \mathbb{K}^n\}.$$
Se, come al solito, denotiano con $e_1, \ldots, e_n \in \mathbb{K}^n$ la base canonica, allora $e_i^T By$ coincide con la i-esima coordinata di By; dunque $e_i^T By = 0$ per ogni i se e solo se $By = 0$ e, a maggior ragione,
$$L_{\mathbf{v}}^{-1}(\mathrm{RKer}(\varphi)) = \{y \in \mathbb{K}^n \mid By = 0\} = \mathrm{Ker}\, L_B.$$
Basta adesso applicare il teorema del rango all'applicazione lineare L_B per concludere la dimostrazione. □

Le stesse considerazioni fatte all'inizio della Sezione 1.2 sui sottospazi ortogonali ci portano a dire che per ogni base $v_1, \ldots, v_n$ di V si ha:
$$\mathrm{RKer}(\varphi) = \{u \in V \mid \varphi(v_i, u) = 0 \text{ per ogni } i = 1, \ldots, n\};$$
$$\mathrm{LKer}(\varphi) = \{u \in V \mid \varphi(u, v_i) = 0 \text{ per ogni } i = 1, \ldots, n\}.$$

Definizione 2.2.5. Una forma bilineare si dice **non degenere** se ha rango massimo, ossia se è rappresentata da una matrice invertibile.

Ad esempio, la forma $x^T By$ su $\mathbb{K}^n$ è non degenere se e solo se $\det(B) \neq 0$. Segue dal Teorema 2.2.4 che per una forma bilineare φ le seguenti condizioni sono equivalenti:

(1) φ è non degenere;
(2) $\mathrm{LKer}(\varphi) = 0$, ossia $\varphi(u, v) = 0$ per ogni v se e solo se $u = 0$;
(3) $\mathrm{RKer}(\varphi) = 0$, ossia $\varphi(u, v) = 0$ per ogni u se e solo se $v = 0$.

In letteratura si usa spesso le seguente terminologia:

(1) un **prodotto interno** è una forma bilineare simmetrica non degenere;
(2) una **forma simplettica** è una forma bilineare alternante non degenere.

Vediamo due esempi tipici di prodotti interni:

ESEMPIO 2.2.6. Il *prodotto interno canonico* sugli spazi vettoriali numerici $\mathbb{K}^n$ è definito come l'applicazione bilineare simmetrica

$$\mathbb{K}^n \times \mathbb{K}^n \xrightarrow{\cdot} \mathbb{K}, \qquad x \cdot y = x^T y = \sum_{i=1}^{n} x_i y_i.$$

La matrice associata nella base canonica è l'identità ed è quindi non degenere.

Negli spazi $\mathbb{R}^n$ il prodotto interno canonico coincide con il prodotto scalare canonico; su $\mathbb{C}^n$ il prodotto interno canonico è ben diverso dal prodotto hermitiano canonico.

ESEMPIO 2.2.7. Nella teoria delle forme bilineari simmetriche gioca un ruolo importante il cosiddetto *piano iperbolico* su di un campo $\mathbb{K}$, definito come lo spazio vettoriale $\mathbb{K}^2$ dotato della forma bilineare simmetrica non degenere

$$U(x, y) = x_1 y_2 + x_2 y_1 = x^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} y.$$

Esercizi.

69. Sia B la matrice associata ad una forma bilineare φ rispetto ad una base. Dimostrare che φ è una forma simmetrica (risp.: alternante) se e solo se B è una matrice simmetrica (risp.: alternante).

70. Sia $U \in M_{2,2}(\mathbb{C})$ fissata e si consideri la forma bilineare alternante

$$\varphi\colon M_{2,2}(\mathbb{C}) \times M_{2,2}(\mathbb{C}) \to \mathbb{C}, \qquad \varphi(A, B) = \operatorname{Tr}(UAB - UBA).$$

Dimostrare che $\varphi = 0$ se U è un multiplo scalare della matrice identità e che il rango di φ è uguale a 2 in tutti gli altri casi. (Suggerimento: provare che non è restrittivo supporre che la matrice U abbia la seconda riga nulla).

71 (☕, ♡). Siano V uno spazio vettoriale di dimensione finita e $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare con la proprietà che

$$\varphi(x, y) = 0 \quad \text{se e solo se} \quad \varphi(y, x) = 0.$$

Dimostrare che φ è simmetrica oppure alternante.

72 (♡). Siano V uno spazio vettoriale di dimensione finita, $\varphi\colon V \times V \to \mathbb{R}$ una forma bilineare e $f\colon V \to V$ lineare tale che $f^*\varphi = \varphi$. Dimostrare che:

(1) se φ è non degenere allora f è iniettiva;
(2) il rango di f è maggiore od uguale al rango di φ.

73. Siano φ, ψ forme bilineari sul medesimo spazio vettoriale V di dimensione finita. Dimostrare che se φ è non degenere esiste $f\colon V \to V$ lineare tale che $\psi(x, y) = \varphi(x, f(y))$. (Sugg.: se B è una matrice invertibile, allora $x^T A y = x^T B B^{-1} A y$.)

74 (♡). Sia $e_1, \ldots, e_n$ una base fissata di uno spazio vettoriale V. Dimostrare che per ogni forma quadratica $\Phi\colon V \to \mathbb{K}$ vi è un'unica forma bilineare $\varphi\colon V \times V \to \mathbb{K}$ tale che $\Phi(x) = \varphi(x, x)$ per ogni x e $\varphi(e_i, e_j) = 0$ per ogni $i > j$.

75. Siano W uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$ con spazio duale $W^\vee = \mathrm{Hom}_{\mathbb{K}}(W, \mathbb{K})$. Si considerino le seguenti forme bilineari sullo spazio $V = W \times W^\vee$:

$$g, \omega\colon V \times V \to \mathbb{K},$$

$$\begin{aligned} g((w_1, \varphi_1), (w_2, \varphi_2)) &= \varphi_1(w_2) + \varphi_2(w_1) \\ \omega((w_1, \varphi_1), (w_2, \varphi_2)) &= \varphi_1(w_2) - \varphi_2(w_1) \end{aligned} \qquad w_1, w_2 \in W,\ \varphi_1, \varphi_2 \in W^\vee.$$

Provare che g è un prodotto interno e ω una forma simplettica.

76. Sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare su uno spazio vettoriale di dimensione finita. Dimostrare che, per un intero fissato r, le seguenti condizioni sono equivalenti:

(1) φ ha rango r;
(2) l'applicazione lineare $h\colon V \to V^\vee$, $h(u)(v) = \varphi(u, v)$, ha rango r;
(3) esistono due succesioni $f_1, \ldots, f_r \in V^\vee$ e $g_1, \ldots, g_r \in V^\vee$, ciascuna formata da r funzionali lineari linearmente indipendenti tali che
$$\varphi(u, v) = \sum_{i=1}^{r} f_i(u) g_i(v);$$
(4) r è il minimo intero per cui esistono due succesioni $f_1, \ldots, f_r \in V^\vee$ e $g_1, \ldots, g_r \in V^\vee$ tali che
$$\varphi(u, v) = \sum_{i=1}^{r} f_i(u) g_i(v).$$

2.3. Forme bilineari simmetriche

Da questo punto, e fino alla fine del capitolo, tutti gli spazi vettoriali saranno assunti, salvo avviso contrario, di *dimensione finita su campi di caratteristica diversa da 2.*

Iniziamo osservando che se $2 \neq 0$ in $\mathbb{K}$, allora una forma bilineare è contemporaneamente simmetrica ed alternante se e solo se è identicamente nulla: infatti, se $\varphi(x, y) = \varphi(y, x)$ e $\varphi(x, x) = 0$ per ogni x, y si ha

$$2\varphi(x, y) = \varphi(x, y) + \varphi(y, x) = \varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y) = 0.$$

Inoltre, ogni forma bilineare φ si può scrivere in modo unico come somma di una forma simmetrica e di una alternante:

$$\varphi(x, y) = \frac{\varphi(x, y) + \varphi(y, x)}{2} + \frac{\varphi(x, y) - \varphi(y, x)}{2}.$$

Definizione 2.3.1. Data una forma quadratica $\Phi\colon V \to \mathbb{K}$, con $\mathbb{K}$ di caratteristica $\neq 2$, la forma bilineare simmetrica

$$\varphi\colon V \times V \to \mathbb{K}, \qquad \varphi(x, y) = \frac{1}{2}(\Phi(x + y) - \Phi(x) - \Phi(y)),$$

si dice **forma polare** di Φ.

Se φ è la polare di una forma quadratica Φ, allora Φ è indotta da φ, ossia $\Phi(x) = \varphi(x, x)$ per ogni x; infatti

$$\Phi(x) = \frac{1}{2}(4\Phi(x) - \Phi(x) - \Phi(x)) = \frac{1}{2}(\Phi(x + x) - \Phi(x) - \Phi(x)) = \varphi(x, x).$$

Dato che in caratteristica $\neq 2$ l'unica forma bilineare contemporaneamente simmetrica ed alternante è quella nulla, ne segue che la polare di una forma quadratica Φ è l'unica forma bilineare simmetrica φ tale che $\Phi(x) = \varphi(x, x)$.

Riepilogando, abbiamo provato che esiste un isomorfismo naturale tra lo spazio vettoriale delle forme quadratiche $\Phi\colon V \to \mathbb{K}$ e lo spazio vettoriale delle forme bilineari simmetriche $\varphi\colon V \times V \to \mathbb{K}$. Tale isomorfismo ed il suo inverso sono definiti dalle formule

$$\Phi(x) = \varphi(x, x), \qquad \varphi(x, y) = \frac{1}{2}(\Phi(x + y) - \Phi(x) - \Phi(y)). \tag{2.3.1}$$

Deduciamo in particolare che la restrizione di Φ ad un sottospazio vettoriale $W \subseteq V$ è identicamente nulla se e solo se la restrizione di φ a $W \times W$ è identicamente nulla.

Nel linguaggio algebrico la prima formula in (2.3.1) viene detta formula di *restituzione*, e la seconda formula di *polarizzazione*. Noi utilizzeremo un linguaggio meno tecnico e diremo semplicemente che φ e Φ sono associate se sono legate l'una all'altra dalle relazioni (2.3.1). Inoltre, useremo solitamente la stessa lettera greca per indicare due forme associate: minuscola per la bilineare, maiuscola per la quadratica.

Abbiamo visto che in un sistema di coordinate $x_1, \ldots, x_n$, ogni forma bilineare simmetrica φ si scrive come

$$\varphi(x, y) = \sum_{ij} b_{ij} x_i y_j$$

con la matrice (b_{ij}) simmetrica ed univocamente determinata. Come conseguenza, ogni forma quadratica Φ si esprime come un polinomio omogeneo di secondo grado nelle variabili $x_1, \ldots, x_n$. Lasciamo per esercizio al lettore la semplice verifica del fatto che se

$$\Phi(x) = \sum_{i \le j} a_{ij} x_i x_j$$

allora la forma polare è uguale a

$$\varphi(x, y) = \sum_i a_{ii} x_i y_i + \frac{1}{2} \sum_{i<j} a_{ij}(x_i y_j + x_j y_i).$$

Ad esempio la forma polare e la matrice associate alla forma quadratica $\Phi(x) = x_1^2 - x_1 x_2 - 2x_2^2$ su $\mathbb{K}^2$ sono rispettivamente:

$$\varphi(x, y) = x_1 y_1 - \frac{1}{2}(x_1 y_2 + x_2 y_1) - 2x_2 y_2, \qquad \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & -2 \end{pmatrix}.$$

Nelle forme simmetriche i nuclei destro e sinistro coincidono e quindi possiamo parlare di nucleo senza ulteriori specifiche.

Definizione 2.3.2. Il **nucleo** di una forma bilineare simmetrica φ è il sottospazio vettoriale

$$\operatorname{Ker} \varphi = \{v \in V \mid \varphi(v, w) = 0 \ \forall w \in V\} = \{v \in V \mid \varphi(w, v) = 0 \ \forall w \in V\}$$

e l'**indice di nullità** di φ è la dimensione di $\operatorname{Ker} \varphi$. Il nucleo $\operatorname{Ker} \Phi$ ed il rango $\operatorname{rg}(\Phi)$ di una forma quadratica Φ si definiscono rispettivamente come il nucleo ed il rango della forma bilineare associata.

Per il Teorema 2.2.4 vale la formula $\operatorname{rg}(\varphi) = \dim V - \dim \operatorname{Ker} \varphi$; dunque φ è non degenere se e solo se per ogni $v \neq 0$ esiste $w \in V$ tale che $\varphi(v, w) \neq 0$.

Osservazione 2.3.3. Bisogna fare attenzione alla reale possibilità che la restrizione di una forma non degenere ad un sottospazio vettoriale proprio possa essere degenere. Ad esempio il prodotto interno canonico $x^T \cdot y$ è non degenere su $\mathbb{C}^2$, ma la sua restrizione al sottospazio $\{x \in \mathbb{C}^2 \mid x_2 = ix_1\}$ è identicamente nulla e quindi degenere.

Definizione 2.3.4. Un vettore $x \in V$ si dice **isotropo** rispetto ad una forma quadratica $\Phi\colon V \to \mathbb{K}$ se $\Phi(x) = 0$.

In generale, i vettori isotropi non formano un sottospazio vettoriale. Ad esempio se $\Phi(x) = f(x)g(x)$ con $f, g \colon V \to \mathbb{K}$ lineari, allora i vettori isotropi di Φ sono dati dall'unione del nucleo di f e del nucleo di g. Ogni vettore appartenente al nucleo di una forma quadratica è isotropo, ma il viceversa è generalmente falso.

Definizione 2.3.5. Sia φ una forma bilineare simmetrica sullo spazio vettoriale V; due vettori $x, y \in V$ si dicono **φ-ortogonali** se $\varphi(x, y) = 0$. Una base $v_1, \ldots, v_n$ di V si dice φ-ortogonale se $\varphi(v_i, v_j) = 0$ per ogni $i \neq j$.

Quando la forma bilineare è chiara dal contesto parleremo più semplicemente di ortogonalità anziché di φ-ortogonalità. In una base φ-ortogonale $v_1, \ldots, v_n$, per due vettori generici $x = \sum_i x_i v_i$ e $y = \sum_i y_i v_i$ si ha

$$\varphi(x, y) = \sum_{i,j} \varphi(v_i, v_j) x_i y_j = \sum_{i=1}^{n} \lambda_i \, x_i y_i, \qquad \lambda_i = \varphi(v_i, v_i) \in \mathbb{K},$$

e la forma quadratica associata diventa

$$\Phi(x) = \sum_i \lambda_i x_i^2, \qquad \lambda_i = \Phi(v_i) \in \mathbb{K}.$$

Dato che la matrice $\varphi(v_i, v_j)$ è diagonale abbiamo dimostrato il seguente utile criterio.

Proposizione 2.3.6. *Il rango di una forma bilineare simmetrica φ è uguale al numero di indici i tali che $\varphi(v_i, v_i) \neq 0$, calcolati rispetto ad una qualunque base φ-ortogonale $v_1, \ldots, v_n$.*

I possibili dubbi sull'esistenza di basi φ-ortogonale li cancelliamo immediatamente grazie al prossimo risultato fondamentale.

Teorema 2.3.7. *Ogni forma bilineare simmetrica φ su di uno spazio vettoriale di dimensione finita su di un campo di caratteristica $\neq 2$ ammette basi φ-ortogonali.*

Dimostrazione. Diamo due distinte dimostrazioni; la prima, breve ed elegante, mostrerà solamente l'esistenza mentre la seconda darà anche un metodo di costruzione della base che potrà essere usato negli esercizi.

Prima dimostrazione. Dimostriamo il teorema per induzione su $n = \dim V$; se $n = 1$ qualsiasi vettore non nullo è una base φ-ortogonale. Supponiamo $n > 1$; se φ è identicamente nulla ogni base è φ-ortogonale; se $\varphi \neq 0$ allora anche la forma quadratica associata è diversa da 0 ed esiste un vettore $v_1 \in V$ tale che $\varphi(v_1, v_1) \neq 0$. Denotiamo $W = \{w \in V \mid \varphi(v_1, w) = 0\}$ e poiché W è il nucleo dell'applicazione lineare $\varphi(v_1, -) \colon V \to \mathbb{K}$ si ha $\dim W \geq n - 1$; siccome $v_1 \notin W$ si ha $\dim W = n - 1$ e $V = \mathbb{K}v_1 \oplus W$. Per induzione esiste una base $v_2, \ldots, v_n \in W$ che è φ-ortogonale. È allora chiaro che $v_1, v_2, \ldots, v_n$ è una base φ-ortogonale di V.

Seconda dimostrazione. Sia $n = \dim V$; diremo che una base $e_1, \ldots, e_n$ di V è *k-ortogonale* se $\varphi(e_i, e_j) = 0$ per ogni $i < k$ e per ogni $j \neq i$. Notiamo che per stabilire se una base $e_1, \ldots, e_n$ è k-ortogonale basta verificare che $\varphi(e_i, e_j) = 0$ per ogni $i < k$ e per ogni $j > i$: infatti se $j < i$ allora $j < k$, $i > j$ e quindi $\varphi(e_i, e_j) = \varphi(e_j, e_i) = 0$.

Ogni base è 1-ortogonale, mentre le basi n-ortogonali sono esattamente le basi φ-ortogonali. Illustriamo adesso un algoritmo in 4 passi che permette, a partire da una base k-ortogonale, con $1 \leq k < n$, di costruire esplicitamente una nuova base $(k + 1)$-ortogonale.

Passo 1). Sia $e_1, \ldots, e_n$ una base k-ortogonale, con $k < n$. Se $\varphi(e_k, e_j) = 0$ per ogni $j > k$, ossia se la base è già $(k+1)$-ortogonale, andate al passo 4. Altrimenti, sia $l \leq n$

il minimo indice tale che $\varphi(e_k, e_l) \neq 0$: per la k-ortogonalità della base si ha $l \geq k$. Se $k = l$ andate al passo 3; se $l > k$ andate passo 2.

Passo 2). Se siete arrivati qui è perché $\varphi(e_k, e_k) = 0$ e $\varphi(e_k, e_l) \neq 0$ per qualche $l > k$. Dalla formula $\varphi(e_k + e_l, e_k + e_l) - \varphi(e_k - e_l, e_k - e_l) = 4\varphi(e_k, e_l) \neq 0$ segue che i due addendi al primo membro non possono essere entrambi nulli e possiamo trovare $a \in \{1, -1\}$ tale che $\varphi(e_k + ae_l, e_k + ae_l) \neq 0$. Sostituite e_k con $e_k + ae_l$, lasciate invariati gli altri elementi della base e andate al passo 3.

Passo 3). Se siete arrivati qui è perché avete una base $e_1, \ldots, e_n$ k-ortogonale tale che $\varphi(e_k, e_k) \neq 0$. Poniamo $v_i = e_i$ per ogni $i \leq k$, mentre se $i > k$ poniamo

$$v_i = e_i - \frac{\varphi(e_i, e_k)}{\varphi(e_k, e_k)} e_k.$$

Si verifica facilmente che $v_1, \ldots, v_n$ è una base $(k+1)$-ortogonale: adesso andate al passo 4.

Passo 4). Se $k+1 < n$ si aumenta k di una unità e si torna al passo 1). Se $k+1 = n$ abbiamo trovato una base n-ortogonale che, per definizione, è φ-ortogonale. □

La relazione di congruenza tra matrici si estende immediatamente alle forme bilineari e quadratiche.

Definizione 2.3.8. Due forme bilineari simmetriche $\varphi, \psi \colon V \times V \to \mathbb{K}$ si dicono **congruenti** se esiste un'applicazione lineare invertibile $f \colon V \to V$ tale che $\varphi = f^*\psi$, ossia

$$\varphi(x, y) = \psi(f(x), f(y)), \qquad \text{per ogni } x, y \in V.$$

Similmente, diremo che due forme quadratiche $\Phi, \Psi \colon V \to \mathbb{K}$ sono **congruenti** se esiste un'applicazione lineare invertibile $f \colon V \to V$ tale che $\Phi = f^*\Psi$, ossia $\Phi(x) = \Psi(f(x))$ per ogni $x \in V$.

Lemma 2.3.9.

(1) *Due forme bilineari simmetriche sono congruenti se e solo se le forme quadratiche associate sono congruenti.*
(2) *La congruenza è una relazione di equivalenza sullo spazio delle forme quadratiche.*

Dimostrazione. Siano Φ, Ψ le forme quadratiche associate a due forme bilineari simmetriche φ, ψ. Se φ, ψ sono congruenti esiste $f \colon V \to V$ lineare invertibile tale che $\varphi(x, y) = \psi(f(x), f(y))$ per ogni x, y. In particolare, per ogni $x \in V$ vale $\Phi(x) = \varphi(x, x) = \psi(f(x), f(x)) = \Psi(f(x))$ e quindi anche Φ, Ψ sono congruenti.

Viceversa se Φ, Ψ sono congruenti, diciamo $\Phi(x) = \Psi(f(x))$, allora per ogni $x, y \in V$ vale

$$\begin{aligned}\varphi(x, y) &= \frac{1}{2}(\Phi(x+y) - \Phi(x) - \Phi(y)) = \frac{1}{2}(\Psi(f(x+y)) - \Psi(f(x)) - \Psi(f(y)))\\ &= \frac{1}{2}(\Psi(f(x) + f(y)) - \Psi(f(x)) - \Psi(f(y))) = \psi(f(x), f(y))\end{aligned}$$

e quindi φ, ψ sono congruenti.

Se $f, g \colon V \to V$ sono due applicazioni lineari, allora

$$(fg)^*\Phi(x) = \Phi(fg(x)) = f^*\Phi(g(x)) = g^* f^*\Phi(x)$$

da cui segue che $(fg)^*\Phi = g^* f^*\Phi$. La congruenza di forme quadratiche è una relazione di equivalenza dato che soddisfa le proprietà:

(1) *riflessiva*, poiché $I^*\Phi = \Phi$, dove I è l'identità su V;
(2) *simmetrica*, poiché se $\Phi = f^*\Psi$ allora $(f^{-1})^*\Phi = (f^{-1})^* f^*\Psi = (ff^{-1})^*\Psi = \Psi$;

(3) *transitiva*, poiché se $\Phi = f^*\Psi$ e $\Psi = g^*\Gamma$ allora $\Phi = f^*g^*\Gamma = (gf)^*\Gamma$.

□

Segue immediatamente dal Lemma 2.3.9 che la congruenza è una relazione di equivalenza anche sullo spazio delle forme bilineari simmetriche.

TEOREMA 2.3.10. *Per due forme bilineari simmetriche φ, ψ in uno spazio vettoriale V di dimensione finita n, le seguenti condizioni sono equivalenti:*

(1) *φ e ψ sono congruenti;*
(2) *per ogni base $e_1, \ldots, e_n$ di V esiste un'altra base $\epsilon_1, \ldots, \epsilon_n$ di V tale che $\varphi(e_i, e_j) = \psi(\epsilon_i, \epsilon_j)$ per ogni i, j;*
(3) *esistono due basi $e_1, \ldots, e_n$ e $\epsilon_1, \ldots, \epsilon_n$ di V tali che $\varphi(e_i, e_j) = \psi(\epsilon_i, \epsilon_j)$ per ogni i, j.*

DIMOSTRAZIONE. *(1)*⇒*(2)* Se $\varphi = f^*\psi$ per un endomorfismo invertibile f, e $e_1, \ldots, e_n$ è una base qualunque di V, allora $\epsilon_1 = f(e_1), \ldots, \epsilon_n = f(e_n)$ è ancora una base di V e vale $\varphi(e_i, e_j) = \psi(\epsilon_i, \epsilon_j)$ per ogni i, j.

L'implicazione *(2)*⇒*(3)* non è altro che il teorema di esistenza delle basi.

(3)⇒*(1)* Siano $e_1, \ldots, e_n$ e $\epsilon_1, \ldots, \epsilon_n$ due basi di V tali che $\varphi(e_i, e_j) = \psi(\epsilon_i, \epsilon_j)$ per ogni i, j, e denotiamo con f l'endomorfismo invertibile tale che $f(e_i) = \epsilon_i$ per ogni i; per concludere basta dimostrare che $\varphi = f^*\psi$. Dati due vettori $x = \sum_i x_i e_i$ e $y = \sum_i y_i e_i$ in V si ha

$$f^*\psi(x, y) = \psi(f(x), f(y)) = \sum_{i,j} x_i y_j \psi(f(e_i), f(e_j)) = \sum_{i,j} x_i y_j \psi(\epsilon_i, \epsilon_j)$$

che è uguale a

$$\varphi(x, y) = \sum_{i,j} x_i y_j \varphi(e_i, e_j).$$

□

DEFINIZIONE 2.3.11. Sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare simmetrica. Per ogni sottospazio vettoriale $W \subseteq V$ definiamo il suo **φ-ortogonale** come

$$W^{\perp\varphi} = \{v \in V \mid \varphi(v, w) = 0 \quad \text{per ogni} \quad w \in W\}.$$

In particolare, $0^{\perp\varphi} = \operatorname{Ker}(\varphi)$, $W \cap W^{\perp\varphi}$ coincide con il nucleo delle restrizione di φ a W e $W \subseteq (W^{\perp\varphi})^{\perp\varphi}$. Un sottospazio vettoriale $W \subseteq V$ si dice:

(1) **anisotropo** se $W \cap W^{\perp\varphi} = 0$, ossia se la restrizione di φ a W è non degenere;
(2) **isotropo** se $W \cap W^{\perp\varphi} \neq 0$, ossia se la restrizione di φ a W è degenere;
(3) **totalmente isotropo** se $W \subseteq W^{\perp\varphi}$, ossia se $\varphi(x, y) = 0$ per ogni $x, y \in W$.

È utile osservare che se $w_1, \ldots, w_m$ è una base di W, allora $v \in W^{\perp\varphi}$ se e solo se $\varphi(w_i, v) = 0$ per ogni $i = 1, \ldots, m$. Una implicazione è ovvia; viceversa, se $\varphi(w_i, v) = 0$ per ogni i, allora per ogni $w = \sum_i a_i w_i \in W$ si ha $\varphi(w, v) = \sum_i a_i \varphi(w_i, v) = 0$.

LEMMA 2.3.12. *Sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare simmetrica non degenere. Per ogni sottospazio vettoriale $W \subseteq V$ valgono le formule*

$$\dim W^{\perp\varphi} + \dim W = \dim V, \qquad (W^{\perp\varphi})^{\perp\varphi} = W.$$

DIMOSTRAZIONE. Sia $w_1, \ldots, w_m$ una base di W, allora $v \in W^{\perp\varphi}$ se e solo se $\varphi(w_i, v) = 0$ per ogni $i = 1, \ldots, m$; dunque $W^{\perp\varphi}$ coincide con il nucleo dell'applicazione lineare

$$V \to \mathbb{K}^m, \qquad v \mapsto \begin{pmatrix} \varphi(w_1, v) \\ \vdots \\ \varphi(w_m, v) \end{pmatrix},$$

e questo prova che $\dim W^{\perp\varphi} \geq \dim V - m = \dim V - \dim W$. Lo stesso argomento applicato ad un sottospazio complementare H di W in V, ossia tale che $H \oplus W = V$, prova che $\dim H^{\perp\varphi} \geq \dim V - \dim H = \dim W$.

Ogni elemento $x \in W^{\perp\varphi} \cap H^{\perp\varphi}$ appartiene al nucleo di φ: infatti per ogni vettore $v \in V$ possiamo scrivere $v = w + h$ con $w \in W$ e $h \in H$ ed allora si ha $\varphi(x, v) = \varphi(x, w) + \varphi(x, h) = 0$. Siccome φ è non degenere si ha $W^{\perp\varphi} \cap H^{\perp\varphi} = 0$ e quindi $\dim V \geq \dim W^{\perp\varphi} + \dim H^{\perp\varphi}$ per la formula di Grassmann; ma questo è possibile solo se $\dim W^{\perp\varphi} = \dim V - \dim W$ e $\dim H^{\perp\varphi} = \dim W$. La medesima uguaglianza dimensionale applicata a $W^{\perp\varphi}$ ci dice che $\dim(W^{\perp\varphi})^{\perp\varphi} = \dim W$ e siccome $W \subseteq (W^{\perp\varphi})^{\perp\varphi}$ per ovvii motivi si ha $W = (W^{\perp\varphi})^{\perp\varphi}$. □

Il lettore deve fare attenzione al fatto che, a differenza di quanto accade negli spazi vettoriali euclidei, in generale $W + W^{\perp\varphi} \neq V$; per la precisione, vale $W + W^{\perp\varphi} = V$ se e solo se il sottospazio W è anisotropo.

Esercizi.

77. Determinare tutti i vettori isotropi del piano iperbolico (Esempio 2.2.7) e del prodotto interno canonico su $\mathbb{C}^2$.

78. Data una forma quadratica $\Phi\colon V \to \mathbb{K}$, dimostrare che

$$\operatorname{Ker}\Phi = \{x \in V \mid \Phi(x + v) = \Phi(v) \quad \text{per ogni } v \in V\}.$$

79. Dimostrare che una forma bilineare simmetrica $\varphi\colon V \times V \to \mathbb{K}$ è non degenere se e solo se per ogni successione $u_1, \ldots, u_m$ di vettori linearmente indipendenti di V, l'applicazione

$$V \to \mathbb{K}^m, \qquad v \mapsto \begin{pmatrix} \varphi(u_1, v) \\ \vdots \\ \varphi(u_m, v) \end{pmatrix},$$

è surgettiva.

80. Sia W un sottospazio totalmente isotropo rispetto ad una forma bilineare simmetrica non degenere su V. Provare che $\dim V \geq 2 \dim W$.

81. Sia $0 \neq x$ un vettore isotropo rispetto ad una forma bilineare simmetrica non degenere φ su V. Provare che esiste un altro vettore isotropo $0 \neq y \in V$ tale che

$$\varphi(x, y) = 1, \qquad V = \operatorname{Span}(x, y) \oplus \operatorname{Span}(x, y)^{\perp\varphi}.$$

82. Sia $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare simmetrica. Dimostrare che per ogni sottospazio vettoriale $W \subseteq V$ si ha $\dim W^{\perp\varphi} = \dim V - \dim W + \dim(W \cap \operatorname{Ker}\varphi)$.

83. Sia $B = (b_{ij})$ una matrice $n \times n$ simmetrica a coefficienti in un campo $\mathbb{K}$ di caratteristica $\neq 2$ e definiamo un'applicazione $\varphi\colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ mediante la formula

$$\varphi(x, y) = \begin{vmatrix} B & y \\ x^T & 0 \end{vmatrix} = \det \begin{pmatrix} b_{11} & \ldots & b_{1n} & y_1 \\ \vdots & \ddots & \vdots & \vdots \\ b_{n1} & \ldots & b_{nn} & y_n \\ x_1 & \ldots & x_n & 0 \end{pmatrix}.$$

Dimostrare che φ è bilineare simmetrica. Determinare una relazione tra l'aggiunta classica $\operatorname{adj}(B)$ e la matrice associata alla forma φ, nonché una relazione tra il rango di B ed il rango di φ.

2.4. Eliminazione di Gauss simmetrica

Abbiamo visto nel Lemma 2.2.1 che per ogni forma bilineare simmetrica φ su $\mathbb{K}^n$, vi è un'unica matrice simmetrica $B = (b_{ij}) \in M_{n,n}(\mathbb{K})$ tale che $\varphi(x,y) = x^T By$, determinata dalla regola $b_{ij} = \varphi(e_i, e_j)$. Questo ci permette di trasportare la relazione di congruenza dalle forme bilineari simmetriche alle matrici simmetriche. Due forme bilineari simmetriche $\varphi(x,y) = x^T By$ e $\psi(x,y) = x^T Cy$ sono per definizione congruenti se esiste una matrice invertibile A tale che

$$x^T By = \varphi(x,y) = \psi(L_A(x), L_A(y)) = (Ax)^T C(Ay) = x^T(A^T CA)y$$

per ogni $x, y \in \mathbb{K}^n$. Ne deduciamo che *le matrici B, C sono congruenti se e solo se esiste $A \in \mathrm{GL}_n(\mathbb{K})$ tale che $B = A^T CA$*, in totale accordo con la precedente Definizione 2.2.2.

Teorema 2.4.1. *Su di un qualunque campo di caratteristica $\neq 2$, ogni matrice simmetrica è congruente ad una matrice diagonale.*

Dimostrazione. Sia A una matrice simmetrica, per il Teorema 2.3.7 esiste una base $v_1, \ldots, v_n$ di $\mathbb{K}^n$ che è ortogonale per la forma bilineare $x^T Ay$. Se C è la matrice (invertibile) le cui colonne sono $v_1, \ldots, v_n$, i coefficienti della matrice $C^T AC$ coincidono con i valori $v_i^T Av_j$, che si annullano per ogni $i \neq j$. □

Tra le conseguenze della riduzione a scala vi è la prova che due matrici, dello stesso ordine, hanno lo stesso nucleo se e solo se sono equivalenti per operazioni sulle righe. Possiamo dimostrare un risultato analogo per la relazione di congruenza sulle matrici simmetriche e che potremmo riassumere, in maniera breve, come: *due matrici quadrate, dello stesso ordine, sono congruenti se e solo se sono equivalenti per operazioni simmetriche su righe e colonne.* Cerchiamo di spiegare meglio la situazione, senza fare una trattazione esaustiva, ma dando solamente alcune indicazioni utili per lo svolgimento di alcuni esercizi delle prossime sezioni.

Abbiamo visto che due matrici simmetriche sono congruenti se e solo se rappresentano la stessa forma bilineare $\varphi\colon V \times V \to \mathbb{K}$ rispetto a due basi dello spazio vettoriale V. Sia $e_1, \ldots, e_n$ una base di V sulla quale agiamo con una successione finita delle delle seguenti operazioni elementari:

(1) moltiplicare un vettore della base per uno scalare non nullo;
(2) aggiungere ad un vettore della base un multiplo scalare di un altro vettore della base;
(3) scambiare tra di loro due vettori della base.

Analizziamo come cambia la matrice $A = (a_{ij}) = (\varphi(e_i, e_j))$ in ciascuno dei tre casi:

(1) se moltiplichiamo e_h per $\lambda \neq 0$, ossia se consideriamo la nuova base

$$\varepsilon_h = \lambda e_h, \qquad \varepsilon_i = e_i \quad \text{per } i \neq h,$$

allora la matrice $B = (b_{ij}) = (\varphi(\varepsilon_i, \varepsilon_j))$ che rappresenta φ nella nuova base soddisfa le uguaglianze

$$\begin{cases} b_{hh} = \lambda^2 a_{hh} & \\ b_{hi} = \lambda a_{hi}, \quad \lambda b_{ih} = \lambda a_{ih} & \text{per } i \neq h, \\ b_{ij} = a_{ij} & \text{per } i, j \neq h, \end{cases}$$

e dunque B si ottiene da A moltiplicando prima la h-esima riga per λ e poi moltiplicando per λ la colonna h-esima della matrice così ottenuta.

(2) cambiamo adesso la base aggiungendo a e_h un multiplo scalare di e_k, con $h \neq k$, ossia consideriamo una nuova base $\varepsilon_1, \ldots, \varepsilon_n$ dove

$$\varepsilon_h = e_h + \lambda e_k, \qquad \varepsilon_i = e_i \quad \text{per } i \neq h.$$

La matrice $B = (b_{ij}) = (\varphi(\varepsilon_i, \varepsilon_j))$ che rappresenta φ nella nuova base soddisfa le uguaglianze

$$\begin{cases} b_{hh} = a_{hh} + \lambda a_{hk} + \lambda a_{kh} + \lambda^2 a_{kk} & \\ b_{hi} = a_{hi} + \lambda a_{ki}, \quad b_{ih} = a_{ih} + \lambda a_{ik} & \text{per } i \neq h, \\ b_{ij} = a_{ij} & \text{per } i, j \neq h. \end{cases}$$

Dunque la matrice B si ottiene da A mediante due operazioni elementari tra loro simmetriche: nella prima si somma alla riga h la riga k moltiplicata per λ; nella seconda si somma alla colonna h la colonna k *della nuova matrice* moltiplicata per λ.

(3) se scambiamo e_h con e_k, la nuova matrice si ottiene scambiando prima le righe h e k di A e poi scambiando le colonne h e k della matrice così ottenuta.

Vediamo un esempio numerico in cui le precedenti tre operazioni elementari simmetriche sono eseguite sulla matrice

$$(\varphi(e_i, e_j)) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Come nella Sezione **??** useremo le abbreviazioni $R_i \leftrightarrow R_j$, aR_i e $R_i + aR_j$ per indicare le operazioni elementari sulle righe, e $C_i \leftrightarrow C_j$, aC_i e $C_i + aC_j$ per quelle corrispondenti sulle colonne.

Se moltiplichiamo il secondo vettore della base per 2 otteniamo

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{2R_2} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{2C_2} \begin{pmatrix} 1 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Se aggiungiamo al primo vettore della base il triplo del secondo otteniamo

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{R_1+3R_2} \begin{pmatrix} 4 & 4 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{C_1+3C_2} \begin{pmatrix} 16 & 4 & 3 \\ 4 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}.$$

Se scambiamo tra di loro il primo ed il secondo vettore della base otteniamo

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Con il termine di **eliminazione di Gauss simmetrica** intenderemo la trasformazione in forma diagonale di una matrice simmetrica mediante una successione finita di operazioni elementari sulle rigonne (righe-colonne) descritte sopra ai punti (1),(2), (3).

Ribadiamo che, in questa procedura, ogni operazione elementare sulle righe deve essere immediatamente seguita dalla medesima operazione elementare sulle colonne.

Esempio 2.4.2. Un caso dove la seconda rigonna viene moltiplicata per 2:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow{2R_2} \begin{pmatrix} 0 & 1 & 1 \\ 2 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow{2C_2} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 4 & 4 \\ 1 & 4 & 3 \end{pmatrix}.$$

Esempio 2.4.3. Cerchiamo una matrice diagonale congruente a

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & -3 \end{pmatrix}.$$

Possiamo applicare l'eliminazione di Gauss simmetrica con l'obiettivo di annullare tutti i coefficienti esterni alla diagonale principale. Come primo passo annulliamo il coefficiente di riga 2 e colonna 1, sottraendo alla seconda rigonna (riga-colonna) la prima rigonna:

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & -3 \end{pmatrix} \xrightarrow{R_2 - R_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & -3 \end{pmatrix} \xrightarrow{C_2 - C_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix}.$$

Come secondo passo sottraiamo alla terza rigonna la prima rigonna:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix} \xrightarrow{R_3 - R_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix} \xrightarrow{C_3 - C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

Esempio 2.4.4. Eseguiamo l'eliminazione di Gauss simmetrica della matrice

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

Iniziamo scambiando tra loro prima e seconda rigonna:

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix}.$$

Poi togliamo alla seconda rigonna la prima ed alla terza il doppio della prima:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix}.$$

Si noti che, per i fissato, tutte le operazioni del tipo $R_i + aR_h, C_i + bC_k$ commutano tra loro. Per finire togliamo la seconda rigonna alla terza:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Il fatto che sia sempre possibile, su campi di caratteristica diversa da 2, trasformare una matrice simmetrica in una matrice diagonale mediante operazioni sulle rigonne è del tutto equivalente al Teorema 2.4.1. Un eventuale approccio algoritmico dovrebbe quindi seguire la seconda dimostrazione del Teorema 2.3.7 e risulterebbe alquanto complicato. Alternativamente, possiamo, guidati dall'esperienza maturata nella riduzione a scala tradizionale, procedere per tentativi, ben sapendo che andrà spesso bene ma anche che ci si può infilare in un vicolo cieco; in tal caso, niente panico, ritentiamo e saremo più fortunati.

Esercizi.

84. Provare che ogni matrice simmetrica a scala è diagonale.

85. Dimostrare che su ogni campo di caratteristica 2 la matrice

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

non è congruente ad alcuna matrice diagonale.

86. Mediante operazioni elementari sulle rigonne, mettere in forma diagonale le seguenti matrici a coefficienti razionali:

$$\begin{pmatrix}1&1&1\\1&2&0\\1&0&-3\end{pmatrix},\quad \begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix},\quad \begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix},\quad \begin{pmatrix}1&2&3\\2&4&5\\3&5&6\end{pmatrix}.$$

2.5. Applicazioni ortogonali e riflessioni

In tutta la sezione denoteremo φ una forma bilineare simmetrica non degenere su di uno spazio vettoriale V di dimensione finita su di un campo $\mathbb{K}$ di caratteristica diversa da 2.

Definizione 2.5.1. Un'applicazione lineare $f\colon V\to V$ si dice **ortogonale** rispetto ad una forma bilineare simmetrica non degenere $\varphi\colon V\times V\to\mathbb{K}$, o più semplicemente φ-ortogonale, se

$$f^*\varphi=\varphi \quad\Leftrightarrow\quad \varphi(f(x),f(y))=\varphi(x,y) \quad \text{per ogni} \quad x,y\in V.$$

Denoteremo con $O(V,\varphi)\subseteq \operatorname{Hom}_{\mathbb{K}}(V,V)$ l'insieme delle applicazioni φ-ortogonali.

Dall'ipotesi che φ sia non degenere segue che ogni applicazione φ-ortogonale è invertibile: infatti se $f(x)=0$, allora per ogni y vale

$$\varphi(x,y)=\varphi(f(x),f(y))=0$$

e quindi $x\in\operatorname{Ker}\varphi=0$. È immediato osservare che $O(V,\varphi)$ è un gruppo di matrici: infatti è non vuoto (contiene l'identità) e se $f,g\in O(V,\varphi)$, allora anche $f^{-1},fg\in O(V,\varphi)$. In linea di principio possiamo considerare la precedente definizione anche quando φ è degenere; in tal caso occorre però imporre la condizione, non più automatica, che ogni applicazione φ-ortogonale sia invertibile.

Esempio 2.5.2. Sia $W\subseteq V$ un sottospazio anisotropo per la forma φ, ossia tale che $V=W\oplus W^{\perp\varphi}$, e consideriamo l'applicazione lineare

$$R_W\colon V\to V \quad \text{tale che} \quad \begin{cases} R_W(v)=v & \text{se } v\in W,\\ R_W(v)=-v & \text{se } v\in W^{\perp\varphi}, \end{cases}$$

chiamata **riflessione** ortogonale rispetto a W. È facile dimostrare che R_W è ortogonale rispetto a φ: dati $x_1,x_2\in V$ possiamo scrivere $x_i=w_i+v_i$ con $w_i\in W$ e $v_i\in W^{\perp\varphi}$ e quindi, siccome $\varphi(w_i,v_j)=0$ per ogni i,j si ha

$$\begin{aligned}\varphi(R_W(x_1),R_W(x_2))&=\varphi(w_1-v_1,w_2-v_2)=\varphi(w_1,w_2)+\varphi(v_1,v_2)\\&=\varphi(w_1+v_1,w_2+v_2)=\varphi(x_1,x_2).\end{aligned}$$

Definizione 2.5.3. Sia $\varphi\colon V\times V\to\mathbb{K}$ bilineare simmetrica non degenere. Per ogni vettore $v\in V$ tale che $\varphi(v,v)\neq 0$ si definisce l'applicazione lineare

$$S_v\colon V\to V,\qquad S_v(x)=x-2\,\frac{\varphi(x,v)}{\varphi(v,v)}\,v.$$

Come per gli spazi euclidei, mostriamo che l'applicazione S_v è φ-ortogonale e, più precisamente, che coincide con la riflessione ortogonale R_W rispetto al sottospazio $W=\operatorname{Span}(v)^{\perp\varphi}$. Dal fatto che $\varphi(v,v)\neq 0$ segue che $\operatorname{Span}(v)\cap W=0$ e quindi $V=$

$\mathrm{Span}(v) \oplus W$; inoltre, $v \in W^{\perp\varphi}$ e siccome $\dim W^{\perp\varphi} = 1$ si ha $W^{\perp\varphi} = \mathrm{Span}(v)$. Dunque, per ogni $w \in W$ ed ogni $a \in \mathbb{K}$ si ha

$$S_v(w) = w - 2\frac{\varphi(w,v)}{\varphi(v,v)}v = w = R_W(w),$$

$$S_v(av) = av - 2\frac{\varphi(av,v)}{\varphi(v,v)}v = -av = R_W(av).$$

Osservazione 2.5.4. Nelle notazioni precedenti, se $x, y \in V$ sono tali che se $\varphi(x,x) = \varphi(y,y)$ e $\varphi(x-y, x-y) \neq 0$, allora $S_{x-y}(x) = y$ e $S_{x-y}(x) = y$. Infatti:

$$\varphi(x, x-y) = \varphi(x,x) - \varphi(x,y) = \varphi(y,y) - \varphi(y,x) = -\varphi(y, x-y),$$
$$\varphi(x-y, x-y) = 2\varphi(x, x-y),$$

e quindi

$$S_{x-y}(x) = x - 2\frac{\varphi(x, x-y)}{\varphi(x-y, x-y)}(x-y) = x - (x-y) = y.$$

Infine, siccome S_{x-y} è uguale al suo inverso, $S_{x-y}(y) = x$.

Possiamo generalizzare parzialmente il risultato del Teorema 1.5.6.

Lemma 2.5.5. *Siano $\mathbb{K}$ un campo di caratteristica diversa da 2, V uno spazio vettoriale di dimensione finita su $\mathbb{K}$, $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare simmetrica non degenere, $W \subseteq V$ un sottospazio vettoriale anisotropo e $g\colon W \to V$ un'applicazione lineare tale che*

$$\varphi(g(x), g(y)) = \varphi(x,y) \qquad \text{per ogni} \quad x, y \in W.$$

Allora esiste un'applicazione $f\colon V \to V$, composizione di al più $2\dim W$ riflessioni del tipo S_v, tale che $f(x) = g(x)$ per ogni $x \in W$. In particolare, ogni $f \in O(V,\varphi)$ è composizione di al più $2\dim V$ riflessioni del tipo S_v.

Dimostrazione. Siano $m = \dim W$ e $e_1, \ldots, e_m \in W$ una base φ-ortogonale. Per ipotesi la restrizione $\varphi\colon W \times W \to \mathbb{K}$ è non degenere e quindi $\Phi(e_i) \neq 0$ per ogni $i = 1, \ldots, m$, dove Φ è la forma quadratica associata a φ.

Denotiamo $v_i = g(e_i)$ e dimostriamo per induzione su $k = 0, \ldots, m$ che esiste un'applicazione $f_k \in O(V,\varphi)$, composizione di al più $2k$ riflessioni S_v e tale che $f(e_i) = v_i$ per ogni $i \le k$; per $k = 0$ il risultato è vero per vacuità di condizioni.

Sia adesso $0 < k \le m$ e supponiamo che esista $h\colon V \to V$, composizione di al più $2(k-1)$ riflessioni di tipo S_v e tale che $h(e_i) = v_i$ per ogni $i < k$. Ponendo $u_k = h(e_k)$, per le ipotesi fatte su g ed h si ha

$$\Phi(u_k) = \Phi(v_k) \neq 0, \quad \varphi(v_i, u_k) = \varphi(v_i, v_k) = \varphi(e_i, e_k) = 0 \quad \forall\, i < k,$$

Dalla formula $\Phi(u_k + v_k) + \Phi(u_k - v_k) = 2(\Phi(u_k) + \Phi(v_k)) = 4\Phi(v_k) \neq 0$ segue che $\Phi(u_k + v_k)$ e $\Phi(u_k - v_k)$ non sono entrambi nulli. Se $\Phi(u_k - v_k) \neq 0$ allora $S_{u_k - v_k}(v_i) = v_i$ per ogni $i < k$, per l'Ossevazione 2.5.4

$$S_{u_k - v_k}(u_k) = u_k - \frac{2\varphi(u_k, u_k - v_k)}{\varphi(u_k - v_k, u_k - v_k)}(u_k - v_k) = v_k,$$

e per dimostrare il passo induttivo basta considerare $f = S_{u_k - v_k}h$. Se invece $\Phi(u_k + v_k) \neq 0$ allora $S_{v_k}S_{u_k+v_k}(v_i) = v_i$ per ogni $i < k$, $S_{v_k}S_{u_k+v_k}(u_k) = v_k$ e per dimostrare il passo induttivo basta quindi considerare l'applicazione $f = S_{v_k}S_{u_k+v_k}h$. $\square$

Teorema 2.5.6 (teorema di Witt). *Siano $\mathbb{K}$ di caratteristica diversa da 2, V uno spazio vettoriale di dimensione finita su $\mathbb{K}$, $\varphi\colon V \times V \to \mathbb{K}$ una forma bilineare simmetrica non degenere, $W \subseteq V$ un sottospazio vettoriale e $g\colon W \to V$ un'applicazione lineare* iniettiva *tale che*

$$\varphi(g(x), g(y)) = \varphi(x, y) \qquad \textit{per ogni} \quad x, y \in W.$$

Allora esiste un'applicazione $f\colon V \to V$, composizione di al più $2\dim V$ riflessioni del tipo S_v, tale che $f(x) = g(x)$ per ogni $x \in W$. In particolare:

(1) *(estensione di Witt) esiste un'applicazione $f\colon V \to V$ lineare bigettiva e φ-ortogonale che estende g;*
(2) *(cancellazione di Witt) esiste un'applicazione $f\colon W^{\perp\varphi} \to g(W)^{\perp\varphi}$ lineare bigettiva e φ-ortogonale.*

Dimostrazione. Per il Lemma 2.5.5 basta dimostrare che esiste un sottospazio anisotropo $U \subseteq V$ ed un'applicazione lineare iniettiva $f\colon U \to V$ tali che $W \subseteq U$, $f_{|W} = g$ e $\varphi(f(x), f(y)) = \varphi(x, y)$ per ogni $x, y \in U$. Ragioniamo per induzione sulla dimensione di $W \cap W^{\perp\varphi}$. Se $W \cap W^{\perp\varphi} = 0$ basta porre $U = W$ e $f = g$. Supponiamo $W \cap W^{\perp\varphi} \neq 0$ e sia $e_1, \ldots, e_m \in W$ una base φ-ortogonale: siccome la restrizione di φ a $W \times W$ è degenere, si ha $W \neq V$ ed esiste almeno un indice i tale che $\varphi(e_i, e_i) = 0$; a meno di permutazioni degli indici possiamo supporre $\varphi(e_m, e_m) = 0$. Per il Lemma 2.3.12, o meglio per quanto visto nella sua dimostrazione, l'applicazione

$$V \to \mathbb{K}^m, \qquad x \mapsto (\varphi(x, e_1), \ldots, \varphi(x, e_m))^T,$$

è surgettiva e quindi possiamo trovare un vettore $e_{m+1} \in V$ tale che $\varphi(e_m, e_{m+1}) = 1$ e $\varphi(e_i, e_{m+1}) = 0$ per ogni $i < m$; siccome $e_m \in W \cap W^{\perp}$ si ha $e_{m+1} \notin W$. Per ogni $s \in \mathbb{K}$ si ha

$$\varphi(e_{m+1} + se_m, e_{m+1} + se_m) = \varphi(e_{m+1}, e_{m+1}) + 2s,$$

e quindi, a meno di aggiungere ad e_{m+1} un opportuno multiplo scalare di e_m, possiamo anche supporre $\varphi(e_{m+1}, e_{m+1}) = 0$. Siccome $g(e_1), \ldots, g(e_m)$ è una base φ-ortogonale di $g(W)$ tale che $\varphi(g(e_m), g(e_m)) = 0$, ripetendo la precedente costruzione possiamo trovare $v_{m+1} \in V - g(W)$ tale che $\varphi(g(e_m), v_{m+1}) = 1$, $\varphi(g(e_i), v_{m+1}) = 0$ per ogni $i < m$ e $\varphi(v_{m+1}, v_{m+1}) = 0$. Se denotiamo con U il sottospazio generato da $e_1, \ldots, e_{m+1}$, allora possiamo estendere g ad un'applicazione iniettiva $f\colon U \to V$ ponendo $f(e_{m+1}) = v_{m+1}$ e $f(e_i) = g(e_i)$ per $i \leq m$.

Per concludere resta da dimostrare che la dimensione di $U \cap U^{\perp\varphi}$ è strettamente minore di quella di $W \cap W^{\perp\varphi}$. Si ha $U \cap U^{\perp\varphi} \subseteq W$: infatti se $x = x_1e_1 + \cdots + x_{m+1}e_{m+1} \in U \cap U^{\perp\varphi}$, allora $\varphi(x, e_m) = x_{m+1} = 0$ e quindi $x \in W$. Siccome $W \subseteq U$ si ha $U^{\perp} \subseteq W^{\perp\varphi}$ e quindi $U \cap U^{\perp\varphi} \subseteq W \cap W^{\perp\varphi}$. D'altra parte $\varphi(e_m, e_{m+1}) = 1 \neq 0$, il vettore e_m non appartiene a $U^{\perp\varphi}$ e quindi $U \cap U^{\perp\varphi} \neq W \cap W^{\perp\varphi}$. □

Corollario 2.5.7. *Siano U, W due sottospazi totalmente isotropi rispetto ad una forma bilineare simmetrica non degenere $\varphi\colon V \times V \to \mathbb{K}$. Se $\dim U \leq \dim W$, allora esiste $f \in O(V, \varphi)$ tale che $f(U) \subseteq W$.*

Dimostrazione. Basta applicare il teorema di Witt ad una qualunque applicazione lineare iniettiva $g\colon U \to W$. □

Esercizi.

87. Determinare tutte le applicazioni ortogonali del piano iperbolico in sé.

88. Provare che l'applicazione indotta da una matrice $A \in M_{n,n}(\mathbb{K})$ è ortogonale rispetto al prodotto interno canonico su $\mathbb{K}^n$ se e solo se $A^T = A^{-1}$.

89. Sia $B = (b_{ij})$ una matrice $n \times n$ simmetrica a coefficienti in un campo $\mathbb{K}$ di caratteristica $\neq 2$. Per ogni vettore riga $a = (a_1, \dots, a_n) \in \mathbb{K}^{(n)}$ sia $H_a \subseteq \mathbb{K}^n$ l'iperpiano di equazione $\sum_{i=1}^n a_i x_i = 0$. Provare che:

(1) H_a è isotropo per la forma $x^T B y$ se e solo se esiste $x \in H_a$ tale che $Bx = a^T$.
(2) H_a è isotropo per la forma $x^T B y$ se e soltanto se

$$\det \begin{pmatrix} b_{11} & \dots & b_{1n} & a_1 \\ \vdots & \ddots & \vdots & \vdots \\ b_{n1} & \dots & b_{nn} & a_n \\ a_1 & \dots & a_n & 0 \end{pmatrix} = 0.$$

Dedurre che $\operatorname{rg}(B) \leq n - 2$ se e solo se ogni iperpiano di $\mathbb{K}^n$ è isotropo.

2.6. Forme quadratiche reali e complesse

In un campo arbitrario, capire quando due forme quadratiche sono congruenti è in generale arduo. Fortunatamente, nei campi dei numeri reali e complessi la situazione è molto più semplice, dove le classi di congruenza sono univocamente determinate da pochi invarianti numerici. Abbiamo già dimostrato che avere lo stesso rango è condizione necessaria affinché due forme quadratiche siano congruenti. Sui numeri complessi tale condizione è anche sufficiente.

Teorema 2.6.1. *Sia V uno spazio vettoriale di dimensione finita su $\mathbb{C}$. Due forme quadratiche $\Phi, \Psi \colon V \to \mathbb{C}$ sono congruenti se e solo se hanno lo stesso rango.*

Dimostrazione. Sia $z_1, \dots, z_n$ un sistema di coordinate su V. Mostriamo che ogni forma quadratica Φ di rango r è congruente alla *forma quadratica standard*

$$I_r(z) = \sum_{i=1}^r z_i^2.$$

Sia φ la forma bilineare associata a Φ e consideriamo una base φ-ortogonale $\epsilon_1, \dots, \epsilon_n$. Abbiamo visto che il rango di φ (=rango di Φ) è uguale al numero di indici i tali che $\Phi(\epsilon_i) \neq 0$; a meno di permutazioni di indici possiamo supporre $\Phi(\epsilon_i) = 0$ se $i > r$ e $\Phi(\epsilon_i) = \lambda_i \neq 0$ se $i \leq r$. Scegliamo per ogni $i \leq r$ una radice quadrata μ_i di λ_i e consideriamo la nuova base

$$v_i = \epsilon_i \text{ se } i > r, \qquad v_i = \frac{\epsilon_i}{\mu_i} \text{ se } i \leq r.$$

Per costruzione vale $\Phi(v_i) = 0$ se $i > r$ e $\Phi(v_i) = 1$ se $i \leq r$ e quindi Φ è congruente alla forma I_r. □

Sui numeri reali due forme quadratiche possono avere lo stesso rango senza tuttavia essere congruenti ed è necessario introdurre ulteriori invarianti.

Definizione 2.6.2. Una forma quadratica $\Phi \colon V \to \mathbb{R}$ su di uno spazio vettoriale reale V si dice:

(1) **definita positiva** (e talvolta scriveremo $\Phi > 0$) se $\Phi(x) > 0$ per ogni $0 \neq x \in V$;
(2) **definita negativa** ($\Phi < 0$) se $\Phi(x) < 0$ per ogni $0 \neq x \in V$, o equivalentemente se $-\Phi$ è definita positiva;
(3) **semidefinita positiva** ($\Phi \geq 0$) se $\Phi(x) \geq 0$ per ogni $x \in V$;
(4) **semidefinita negativa** ($\Phi \leq 0$) se $\Phi(x) \leq 0$ per ogni $x \in V$, o equivalentemente se $-\Phi$ è semidefinita positiva;

(5) **indefinita** in tutti gli altri casi, ossia se esistono $v, w \in V$ tali che $\Phi(v) > 0$ e $\Phi(w) < 0$.

Le stesse denominazioni si applicano alle forme bilineari considerando le forme quadratiche associate.

Sia φ è una forma bilineare simmetrica; in un sistema di coordinate $x_1, \ldots, x_n$ corrispondente ad una base φ-ortogonale, la forma quadratica associata si scrive

$$\Phi(x) = \sum_{i=1}^{n} \lambda_i x_i^2, \qquad \lambda_i \in \mathbb{R}.$$

È immediato osservare che $\Phi > 0$ (risp.: $\Phi \geq 0$, $\Phi < 0$, $\Phi \leq 0$) se e solo se $\lambda_i > 0$ (risp.: $\lambda_i \geq 0$, $\lambda_i < 0$, $\lambda_i \leq 0$) per ogni $i = 1, \ldots, n$.

Definizione 2.6.3. Per ogni forma quadratica reale $\Phi\colon V \to \mathbb{R}$, definiamo:

(1) l'**indice di positività** Φ_+ = massima dimensione di un sottospazio $W \subseteq V$ tale che la restrizione di Φ a W sia definita positiva;
(2) l'**indice di negatività** $\Phi_- = (-\Phi)_+$ = massima dimensione di un sottospazio $W \subseteq V$ tale che la restrizione di Φ a W sia definita negativa;
(3) la **segnatura** come la coppia ordinata $(\Phi_+, \Phi_-) \in \mathbb{N}^2$.

Il prossimo lemma implica che la segnatura di una forma quadratica reale è invariante per congruenza.

Lemma 2.6.4. *Siano $\Phi\colon V \to \mathbb{R}$ una forma quadratica, $f\colon V \to V$ lineare invertibile e $W \subseteq V$ un sottospazio vettoriale. Allora la restrizione di $f^*\Phi$ a W è definita positiva se e solo se lo è anche la restrizione di Φ a $f(W)$.*

Identico risultato si ottiene sostituendo il termine definita con semidefinita e/o positiva con negativa.

Dimostrazione. Per definizione di pull-back, per ogni $w \in W$ vale $f^*\Phi(w) = \Phi(f(w))$. □

Teorema 2.6.5 (Teorema di Sylvester). *Data una forma quadratica reale in forma diagonale (ossia in un sistema di coordinate $x_1, \ldots, x_n$ relativo ad una base ortogonale)*

$$\Phi(x) = \sum_{i=1}^{n} \lambda_i x_i^2, \qquad \lambda_i \in \mathbb{R},$$

si considerino i numeri:

- *p = numero di indici i tali che $\lambda_i > 0$;*
- *q = numero di indici i tali che $\lambda_i < 0$.*

Allora vale $\Phi_+ = p$ e $\Phi_- = q$. In particolare, il rango di φ è uguale alla somma $\Phi_+ + \Phi_-$, gli interi p e q non dipendono dalla particolare base ortogonale scelta e sono invarianti per congruenza.

Dimostrazione. È sufficiente dimostrare che $p = \Phi_+$: infatti considerando la forma quadratica opposta $-\Phi$ i coefficienti λ_i cambiano tutti di segno e $(-\Phi)_+ = \Phi_-$. A meno di permutazione di indici possiamo supporre $\lambda_1, \ldots, \lambda_p > 0$, $\lambda_{p+1}, \ldots, \lambda_{p+q} < 0$ e $\lambda_{p+q+1}, \ldots, \lambda_n = 0$.
Sia $L \subseteq V$ il sottospazio definito dalle equazioni $x_{p+1} = \cdots = x_n = 0$ e denotiamo con $\pi\colon V \to L$ la proiezione sulle prime p coordinate $\pi(x_1, \ldots, x_n) = (x_1, \ldots, x_p, 0, \ldots, 0)$. La restrizione di Φ al sottospazio L è definita positiva e quindi, per definizione di Φ_+, si ha $p \leq \Phi_+$.

Viceversa, sempre per definizione di Φ_+, esiste un sottospazio vettoriale $W \subseteq V$ di dimensione Φ_+ tale che $\Phi(x) > 0$ per ogni $x \in W$, $x \neq 0$. Dimostriamo che l'applicazione lineare $\pi \colon W \to L$ è iniettiva; da questo seguirà che $\Phi_+ = \dim W \leq \dim L = p$.

Sia dunque $x = (x_1, \ldots, x_n) \in W$ tale che $\pi(x) = 0$, ciò significa che $x_1 = \ldots = x_p = 0$ e quindi che $\Phi(x) = \sum_{i>p} \lambda_i x_i^2 \leq 0$. D'altra parte Φ è definita positiva su W e la condizione $\Phi(x) \leq 0$ implica necessariamente $x = 0$. □

Esempio 2.6.6. Calcoliamo rango e segnatura della forma quadratica

$$\Phi \colon \mathbb{R}^4 \to \mathbb{R}, \quad \Phi(x) = x_1^2 + x_2^2 - x_3^2 - x_4^2 + 78x_1x_4 - 36x_2x_3 + 8x_1x_3.$$

La forma è definita positiva sul sottospazio $x_3 = x_4 = 0$ generato dai primi due vettori della base canonica, quindi l'indice di positività è ≥ 2. Similmente, la forma è definita negativa sul sottospazio $x_1 = x_2 = 0$ generato dagli ultimi due vettori della base canonica, quindi l'indice di negatività è ≥ 2. Dunque la forma bilineare è obbligata ad avere rango 4 (il massimo possibile) e segnatura $(2, 2)$.

Corollario 2.6.7. *Due forme quadratiche definite su di uno spazio vettoriale reale V sono congruenti se e solo se hanno la stessa segnatura.*

Dimostrazione. Fissiamo un isomorfismo $V \cong \mathbb{R}^n$ (cioè fissiamo una base); consideriamo poi, per ogni coppia di interi positivi p, q tali che $p + q \leq n$, la forma quadratica

$$I_{p,q}(x) = x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2.$$

e dimostriamo che ogni forma quadratica è congruente ad $I_{p,q}$ per opportuni p, q. Notiamo poi che, per il teorema di Sylvester la coppia (p, q) è uguale alla segnatura.

Sia dunque Φ una forma quadratica; sappiamo che esiste un sistema di coordinate lineari $y_1, \ldots, y_n$ su $\mathbb{R}^n$ rispetto al quale

$$\Phi(y) = \sum_{i=1}^{n} \lambda_i y_i^2$$

A meno di permutazioni di indici possiamo assumere:

(1) $\lambda_i > 0$ se $i \leq p$;
(2) $\lambda_i < 0$ se $p < i \leq p + q$;
(3) $\lambda_i = 0$ se $i > p + q$.

Nel sistema di coordinate

$$\begin{cases} z_i = \sqrt{\lambda_i}\, y_i & \text{se } i \leq p \\ z_i = \sqrt{-\lambda_i}\, y_i & \text{se } p < i \leq p + q \\ z_i = y_i & \text{se } p + q < i \end{cases}$$

la forma quadratica diventa $\Phi(z) = \sum_{i=1}^{p} z_i^2 - \sum_{i=p+1}^{p+q} z_i^2$, che è congruente a $I_{p,q}$. □

Esempio 2.6.8. Calcoliamo, al variare del parametro $\lambda \in \mathbb{R}$, il rango e la segnatura della forma quadratica:

$$\Phi \colon \mathbb{R}^2 \to \mathbb{R}, \qquad \Phi(x) = x_1^2 + 2\lambda x_1 x_2 - x_2^2.$$

Nella base canonica e_1, e_2 la matrice simmetrica corrispondente è

$$B_\lambda = \begin{pmatrix} 1 & \lambda \\ \lambda & -1 \end{pmatrix}$$

Osserviamo che $\Phi(e_1) > 0$, da cui deduciamo $\Phi_+ \geq 1$, e che $\Phi(e_2) < 0$ da cui deduciamo $\Phi_- \geq 1$. Siccome $\Phi_+ + \Phi_- \leq 2$ dovrà necessariamente essere $\Phi_+ = \Phi_- = 1$. Dunque Φ

ha rango 2 e segnatura $(1,1)$, indipendentemente dal valore di λ, e le matrici B_λ sono tutte congruenti tra loro.

Il segno del determinante di una matrice simmetrica è invariante per congruenza. Infatti se B, C sono matrici reali simmetriche congruenti, allora esiste una matrice invertibile A tale che $A^T BA = C$ e quindi

$$\det C = \det(A^T)\det(B)\det(A) = \det(B)\det(A)^2.$$

Siccome ogni matrice simmetrica è congruente ad una matrice diagonale (Teorema 2.4.1), segue immediatamente dal teorema di Sylvester che, per una forma quadratica

$$\Phi\colon \mathbb{R}^n \to \mathbb{R}, \qquad \Phi(x) = x^T Bx,$$

vale:

(1) $\det(B) = 0$ se e solo se $\operatorname{rg}(\Phi) < n$;
(2) $\det(B) > 0$ se e solo se $\operatorname{rg}(\Phi) = n$ e Φ_- pari;
(3) $\det(B) < 0$ se e solo se $\operatorname{rg}(\Phi) = n$ e Φ_- dispari.

Esempio 2.6.9. Segue dal teorema di Sylvester che la segnatura di una matrice simmetrica diagonale a blocchi è la somma (in $\mathbb{N}^2$) delle segnature dei singoli blocchi. Ad esempio, la segnatura della matrice

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)$$

è $(1,1) + (1,1) = (2,2)$.

Esempio 2.6.10. Calcoliamo rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^4 \to \mathbb{R}, \qquad \Phi(x) = x_1x_2 + x_2x_3 + x_3x_4.$$

Mettendo in evidenza x_2 si ha $\Phi(x) = (x_1+x_3)x_2+x_3x_4$ e nelle coordinate $y_1 = x_1+x_3$, $y_i = x_i$ per $i > 1$, la matrice associata è la metà di quella dell'Esempio 2.6.9. Quindi il rango è 4 e la segnatura è $(2,2)$.

Corollario 2.6.11. *Per una matrice simmetrica reale $A \in M_{n,n}(\mathbb{R})$, il numero di autalori positivi (risp.: negativi), contati con molteplicità, è uguale all'indice di positività (risp.: negatività) della forma quadratica $\Phi(x) = x^T Ax$.*

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_n$ gli autovalori di A contati con molteplicità. Per il teorema spettrale esiste una matrice ortogonale E tale che $E^{-1}AE = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$; siccome $E^{-1} = E^T$, la matrice $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ è al tempo stesso simile ad A (e quindi con gli stessi autovalori) e congruente ad A (e quindi con gli stessi indici di positività/negatività). Basta adesso applicare il teorema di Sylvester 2.6.5. □

Riprendendo le notazioni della Sezione 1.4, per ogni matrice $A \in M_{n,n}(\mathbb{R})$ ed ogni $k \leq n$ denotiamo con $A[k]$ il minore principale formata dalle prime k righe e k colonne.

Corollario 2.6.12 (Criterio di Sylvester). *Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica tale che $\det(A[k]) \neq 0$ per ogni $k = 1, \ldots, n$. Allora l'indice di negatività delle forma bilineare $\varphi(x, y) = x^T Ay$, e quindi anche il numero di autovalori negativi di A, è uguale al numero di cambiamenti di segno della successione $1, \det(A[1]), \ldots, \det(A[n])$.*

Dimostrazione. Indichiamo come al solito con $e_1, \ldots, e_n$ la base canonica di $\mathbb{R}^n$ e con $\mathbb{R}^{n-1} \subseteq \mathbb{R}^n$ il sottospazio dei vettori con ultima coordinata nulla, ossia $\mathbb{R}^{n-1} = \mathrm{Span}(e_1, \ldots, e_{n-1})$. Sia $v_1, \ldots, v_{n-1}$ una base φ-ortogonale di $\mathbb{R}^{n-1}$, poniamo $v_n = A^{-1}e_n$ e proviamo che $v_1, \ldots, v_n$ è una base φ-ortogonale di $\mathbb{R}^n$. Per ogni $i < n$ si ha $\varphi(v_i, v_n) = v_i^T A v_n = v_i^T e_n = 0$ dato che $v_i \in \mathbb{R}^{n-1}$. Siccome $A[n-1]$ è invertibile, la restrizione di φ a $\mathbb{R}^{n-1}$ è nondegenere, quindi $\mathbb{R}^{n-1} \cap (\mathbb{R}^{n-1})^{\perp\varphi} = 0$; questo implica $v_n \notin \mathbb{R}^{n-1}$ e quindi $v_1, \ldots, v_n$ è una base di $\mathbb{R}^n$.

Denotiamo $\lambda_i = \varphi(v_i, v_i)$; per il teorema di Sylvester ed induzione su n, il numero dei λ_i negativi con $i < n$ è uguale al numero di cambiamenti di segno della successione $1, \det(A[1]), \ldots, \det(A[n-1])$. Poiché il segno del determinante è invariante per congruenza, si ha che $\det(A[n-1])$ ha lo stesso segno del prodotto $\lambda_1 \cdots \lambda_{n-1}$, e $\det(A[n])$ ha lo stesso segno del prodotto $\lambda_1 \cdots \lambda_n$. Questo equivale a dire che $\lambda_n < 0$ se e solo se $\det(A[n-1])$ e $\det(A[n])$ hanno diverso segno. □

Dunque, il linea di principio, possiamo calcolare gli autovalori di una matrice simmetrica reale per determinare la segnatura di una forma quadratica, anche se questo comporta il calcolo e la scomposizione in fattori di primo grado del polinomio caratteristico. Uno studente masochista potrebbe adottare questa tattica come procedura standard, ignorando le altre più semplici a disposizione (criterio di Sylvester, studio geometrico degli indici di positività/negatività, eliminazione di Gauss simmetrica e, nella peggiore delle ipotesi, regola dei segni di Cartesio). Per non incentivare gli studenti a comportamenti perniciosi per se stessi ed irritanti per i docenti, in questo testo non faremo alcun calcolo di autovalori ai fini dello studio degli invarianti per congruenza.

Esempio 2.6.13. Calcoliamo, in funzione di $\lambda \in \mathbb{R}$, rango e segnatura della matrice simmetrica

$$B_\lambda = \begin{pmatrix} 1 & 1 \\ 1 & \lambda \end{pmatrix},$$

ossia rango e segnatura della forma quadratica associata Φ. Nella base canonica e_1, e_2 sa ha $\Phi(e_1) = 1$ e quindi $1 \leq \Phi_+ \leq 2$ per ogni λ. Il determinante è uguale a $\lambda - 1$ e quindi:

(1) per $\lambda = 1$ il rango è 1 e la segnatura è (1,0);
(2) per $t > 1$ il rango è 2, $|B_\lambda| > 0$ e quindi Φ_- è pari. Necessariamente $\Phi_- = 0$ e la segnatura è $(2, 0)$;
(3) per $\lambda < 1$ il rango è 2, $|B_\lambda| < 0$ e quindi Φ_- è dispari. Necessariamente $\Phi_- = 1$ e la segnatura è $(1, 1)$.

Notiamo anche che per $0 < \lambda < 1$ tutti i coefficienti di B_λ sono positivi, e tuttavia la forma quadratica associata non è definita positiva.

Esempio 2.6.14. Calcoliamo il rango e la segnatura della forma quadratica Φ associata alla matrice simmetrica

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & -3 \end{pmatrix}$$

La diagonale contiene sia valori positivi che negativi e quindi $\Phi_+ \geq 1$, $\Phi_- \geq 1$. Il determinante della matrice è $-5 < 0$, quindi Φ_- è dispari e di conseguenza $\Phi_+ = 2$, $\Phi_- = 1$.

Esempio 2.6.15. Calcoliamo la segnatura della forma quadratica Φ associata alla matrice simmetrica

$$\begin{pmatrix} 1 & 1 & 350! \\ 1 & 2 & \pi \\ 350! & \pi & -3 \end{pmatrix}$$

La diagonale contiene sia valori positivi che negativi e quindi $\Phi_+ \geq 1$, $\Phi_- \geq 1$. La restrizione della forma quadratica al sottospazio U generato dai primi due vettori della base canonica è rappresentata dal minore principale

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

e quindi è definita positiva per quanto visto nell'Esempio 2.6.13. Quindi $\Phi_+ \geq 2$ e dunque $\Phi_+ = 2$ e $\Phi_- = 1$.

Osservazione 2.6.16. Quando la forma quadratica su $\mathbb{K}^n$ è scritta in forma di polinomio omogeneo di secondo grado nelle coordinate, la regola pratica per passare dalla forma quadratica alla matrice associata è mettere sulla diagonale i coefficienti dei quadrati x_i^2 e fuori dalla diagonale le metà dei coefficienti di x_ix_j, con $i < j$.

Esempio 2.6.17. Calcoliamo rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^3 \to \mathbb{R}, \quad \Phi(x) = x_1^2 - 2x_1x_2 - 2x_1x_3,$$

eseguendo l'eliminazione di Gauss simmetrica della matrice associata

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Aggiungiamo la prima rigonna alla seconda

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \xrightarrow{R_2+R_1} \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & -1 \\ -1 & 0 & 0 \end{pmatrix} \xrightarrow{C_2+C_1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & 0 \end{pmatrix},$$

poi la prima alla terza

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & 0 \end{pmatrix} \xrightarrow{R_3+R_1} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \xrightarrow{C_3+C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix},$$

e per finire togliamo la seconda alla terza

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \xrightarrow{R_3-R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_3-C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Quindi, per il teorema di Sylvester, la forma Φ ha rango 2 e segnatura $(1,1)$.

Esempio 2.6.18. Determiniamo rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^4 \to \mathbb{R}, \qquad \Phi(x) = x_1x_2 + x_2^2 + 2x_2x_4 - x_3^2.$$

La matrice associata alla forma quadratica Φ è

$$Q = \begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

e lasciamo al lettore il compito di provare che $\operatorname{rg}(Q) = 3$. Per il criterio di Sylvester, la restrizione di Φ al sottospazio di equazione $x_1 = 0$, ha segnatura $(1,2)$ e quindi $\Phi_+ \geq 1$, $\Phi_- \geq 2$. Siccome $1 + 2 = \operatorname{rg}(\Phi)$ deve necessariamente essere $\Phi_+ = 1$ e $\Phi_- = 2$.

Esercizi.

90 ($\heartsuit$). Ricalcolare rango e segnatura della forma quadratica dell'Esempio 2.6.18, utilizzando la regola dei segni di Cartesio.

91. Calcolare rango e segnatura delle forme quadratiche reali

$$\Phi, \Psi\colon M_{2,2}(\mathbb{R}) \to \mathbb{R}, \qquad \Phi(A) = \det(A), \quad \Psi(A) = \mathrm{Tr}(A^2).$$

92. Sia $(a_{ij}) \in M_{n,n}(\mathbb{R})$ simmetrica e semidefinita positiva. Provare che $\max_{i,j} |a_{ij}| = \max_i a_{ii}$.

93. Si introduce l'*indice* di una forma quadratica reale Φ come la differenza $\Phi_+ - \Phi_-$; è immediato osservare che la segnatura è univocamente determinata da rango e indice. Sia Φ una forma quadratica simmetrica reale di rango r e indice σ. Provare che:

(1) $r - \sigma$ è pari;
(2) Φ è semidefinita positiva se e solo se $r = \sigma$;
(3) Φ è semidefinita negativa se e solo se $r = -\sigma$;
(4) Φ è semidefinita se e solo se ogni vettore isotropo appartiene a $\operatorname{Ker}\Phi$.
(5) Usare i punti 1, 2 e 4 per trovare, senza bisogno di far calcoli, rango e indice della forma quadratica dell'Esempio 2.6.17.

94. Sia $0 < t < 1$ un numero reale e sia $\Phi\colon \mathbb{R}^n \to \mathbb{R}$ la forma quadratica associata alla matrice simmetrica di coefficienti $a_{ij} = t^{(i-1)(j-1)}$, $i, j = 1, \ldots, n$. Mostrare che l'indice $\Phi_+ - \Phi_-$ della forma quadratica dipende solo dalla parità n.

95. Provare che una forma quadratica reale è definita, positiva o negativa, se e solo se non possiede vettori isotropi diversi da 0.

96. Dimostrare che per una matrice reale, simmetrica e ortogonale, l'indice della forma quadratica associata è uguale alla traccia.

97. Sia $H \subseteq M_{3,3}(\mathbb{R})$ il sottospazio vettoriale delle matrici simmetriche a traccia nulla. Calcolare rango e segnatura della forma quadratica $\Phi\colon H \to \mathbb{R}$, $\Phi(A) = \mathrm{Tr}(A^2)$.

98. Per ogni $n \geq 3$ sia $B_n \in M_{n,n}(\mathbb{R})$ la matrice simmetrica di coefficienti

$$b_{ij} = \begin{cases} 0 & \text{se } 2 \leq i \leq n-1 \text{ e } 2 \leq j \leq n-1 \\ 1 & \text{altrimenti} \end{cases}$$

Determinare rango e segnatura di B_3, B_4.

99. Determinare rango e segnatura delle matrici B_n introdotte nell'Esercizio 98, per ogni $n > 4$.

100. Siano $f, g\colon V \to \mathbb{R}$ lineari e linearmente indipendenti come vettori di $\mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$. Mostrare che $\Phi(x) = f(x)g(x)$ è una forma quadratica di rango 2 e segnatura $(1, 1)$. Cosa si può dire se f e g sono linearmente dipendenti?

101. Calcolare rango e segnatura delle forme quadratiche $\Phi, \Psi\colon \mathbb{R}^{350} \to \mathbb{R}$:

$$\Phi(x) = \sum_{i=2}^{350} (2x_1 x_i + x_i^2), \quad \Psi(x) = (x_1 + 2x_2 + 3x_3 + \cdots + 350x_{350})^2 - x_1^2.$$

102. Determinare, al variare di $t \in \mathbb{R}$, rango e segnatura della matrice simmetrica

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & t \end{pmatrix}.$$

103. Si consideri la forma quadratica in $\mathbb{R}^3$ definita da $\Phi(x) = 2x_1x_2 + 6x_2x_3$. Calcolare rango, segnatura e si trovi una base di $\mathbb{R}^3$ in cui Φ è scritta in forma canonica.

104. Nell'Esercizio 83, se $\mathbb{K} = \mathbb{R}$ e $\det B \neq 0$, determinare, in funzione della segnatura di B, la segnatura di φ.

105 (Disuguaglianze di Cauchy–Schwarz). Siano φ una forma bilineare su uno spazio vettoriale reale V di dimensione 2 e si denoti con Φ la forma quadratica associata. Provare che, dati due vettori $u, v \in V$ si ha:

(1) $\varphi(u,v)^2 = \Phi(u)\Phi(v)$ se Φ è degenere oppure se u, v sono linearmente dipendenti;
(2) $\varphi(u,v)^2 < \Phi(u)\Phi(v)$ se Φ è definita e u, v sono linearmente indipendenti;
(3) $\varphi(u,v)^2 > \Phi(u)\Phi(v)$ se Φ è indefinita e u, v sono linearmente indipendenti.

106. Calcolare rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^3 \to \mathbb{R}, \qquad \Phi(x) = x_1^2 + x_2^2 - x_3^2 + 12x_1x_3 + 78x_2x_3.$$

107. Calcolare rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^{700} \to \mathbb{R}, \qquad \Phi(x) = (x_1 + x_2 + x_3 + \cdots + x_{600})^2 - x_1^2.$$

108. Calcolare rango e segnatura delle forme quadratiche $\Phi_i\colon \mathbb{R}^4 \to \mathbb{R}$:

$$\begin{aligned}
&\Phi_1(x) = x_1x_2 + x_2x_3 + x_3x_4,\\
&\Phi_2(x) = x_1^2 - x_3^2 - x_4^2 + 4x_1x_2 - 2x_1x_3 + 4x_2x_3 + 2x_2x_4 - 6x_3x_4,\\
&\Phi_3(x) = x_1^2 - x_2^2 - x_3^2 + 2x_2x_3 + 2x_4x_1,\\
&\Phi_4(x) = x_1^2 - x_3^2 - x_4^2 + 4x_1x_2 - 2x_1x_3 + 4x_2x_3 + 2x_2x_4 - 6x_3x_4,\\
&\Phi_5(x) = x_1^2 - x_3^2 - x_4^2 + 4x_1x_2 + 2x_1x_3 - 4x_2x_3 + 2x_2x_4 - 6x_3x_4,\\
&\Phi_6(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 4(x_1x_2 + x_1x_3 + x_2x_3 + x_2x_4 + x_3x_4),\\
&\Phi_7(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_4 - 2x_1x_2 - 2x_2x_4 - 2x_3x_4.
\end{aligned}$$

109. Calcolare rango e segnatura della forma quadratica

$$\Phi\colon \mathbb{R}^5 \to \mathbb{R}, \qquad \Phi(x) = x_1^2 + 2x_2^2 - x_3^2 - x_5^2 + 2x_1x_2 + 4x_1x_3 + x_2x_3 - 2x_5x_4.$$

110 (♡). Per ogni $n > 0$ sia $B_n \in M_{n,n}(\mathbb{R})$ la matrice simmetrica di coefficienti

$$b_{ij} = i + j - 2, \qquad i, j = 1, \ldots, n.$$

Determinare rango e segnatura di B_1, B_2 e B_3.

111 (☕, ♡). Determinare rango e segnatura della matrice B_n introdotta nell'Esercizio 110 per ogni $n \geq 4$.

112 (☕, ♡). Sia $A = (a_{ij})$ una matrice simmetrica reale di ordine n tale che $2a_{ii} \geq \sum_{j=1}^{n} |a_{ij}|$ per ogni indice i. Provare che A è semidefinita positiva.

2.7. Complementi: prodotti scalari di Minkowski

In relatività ristretta si introduce il cosiddetto prodotto scalare di Minkowski in $\mathbb{R}^4$. Più in generale, un prodotto scalare di Minkowski su uno spazio vettoriale reale V è definito come una forma bilineare simmetrica

$$\langle \cdot, \cdot \rangle\colon V \times V \to \mathbb{R}$$

non degenere di segnatura $(1, \dim V - 1)$.

La prossima proposizione illustra alcune interessanti proprietà matematiche dell'insieme, denotato T, dei vettori 'di tipo tempo'.

PROPOSIZIONE 2.7.1. *Sia $\langle \cdot, \cdot \rangle$ un prodotto scalare di Minkowski in uno spazio vettoriale reale V di dimensione $n+1$. Si considerino i due sottoinsiemi*

$$T = \{v \in V \mid \langle v, v \rangle > 0\}, \qquad \overline{T} = \{v \in V \mid v \neq 0, \ \langle v, v \rangle \geq 0\}.$$

Allora:

(1) *per un sottospazio vettoriale $W \subseteq V$ di dimensione $m+1$, la restrizione di $\langle \cdot, \cdot \rangle$ a W ha segnatura $(1, m)$ se e solo se $W \cap T \neq \emptyset$;*
(2) *per ogni $v \in T$ ed ogni $w \in \overline{T}$ vale $\langle v, w \rangle \neq 0$;*
(3) *per ogni $u \in T$, $v, w \in \overline{T}$ tali che $\langle u, v \rangle > 0$ e $\langle u, w \rangle > 0$ si ha $\langle v, w \rangle \geq 0$.*

In particolare, la relazione $\sim$ sull'insieme T definita come $u \sim v$ se $\langle u, v \rangle > 0$, è di equivalenza.

DIMOSTRAZIONE. Dato che la segnatura è $(1, n)$ si ha $T \neq \emptyset$ ed esiste un iperpiano $N \subseteq V$ in cui la forma bilineare è definita negativa.

(*1*) Dato W di dimensione $m+1$, se la restrizione di $\langle \cdot, \cdot \rangle$ ha segnatura $(1, m)$ allora W contiene almeno un vettore v con $\langle v, v \rangle > 0$ e quindi $T \cap W \neq \emptyset$. Viceversa, se esiste $v \in T \cap W$ allora $\operatorname{Span}(v)$ è un sottospazio di dimensione 1 in cui la forma è definita positiva, mentre $N \cap W$ è un sottospazio di dimensione $\geq m$ in cui la forma è definita negativa.

(*2*) Se fosse per assurdo $\langle v, w \rangle = 0$, con $v \in T$ e $w \in \overline{T}$ allora v, w devono necessariamente essere linearmente indipendenti, e la forma bilineare ristretta a $W = \operatorname{Span}(v, w)$ sarebbe semidefinita positiva, in contraddizione con il fatto che $W \cap N \neq 0$.

(*3*) Se per assurdo esistessero $u \in T$, $v, w \in \overline{T}$ tali che $\langle u, v \rangle > 0$, $\langle u, w \rangle > 0$ e $\langle v, w \rangle < 0$, allora per ogni numero reale non negativo $t \geq 0$ si ha $u + tw \in T$ e tuttavia $\langle v, u + tw \rangle = \langle v, u \rangle + t\langle v, w \rangle$ si annullerebbe per $t = -\langle v, u \rangle / \langle v, w \rangle$ in contraddizione con il punto precedente. □

La relazione $\sim$ introdotta nella precedente proposizione possiede esattamemte due classi di equivalenza. Infatti, sia $v \in T$ un vettore fissato e siano $u, w \in T$ non equivalenti a v, ossia tali che $\langle v, u \rangle \leq 0$ e $\langle v, w \rangle \leq 0$. Per il punto (*2*) si ha $\langle v, u \rangle < 0$ e $\langle v, w \rangle < 0$, quindi $v \sim -u$, $v \sim -w$ e per la proprietà transitiva $-u \sim -w$, che implica immediatamente $u \sim w$.

Si ha dunque una partizione in classi di equivalenza $T = T_+ \cup T_-$, dove

$$T_+ = \{u \in T \mid \langle v, u \rangle > 0\}, \quad T_- = \{u \in T \mid \langle -v, u \rangle > 0\}.$$

Se $u, w \in T_+$ allora $tu + (1-t)w \in T_+$ per ogni $0 \leq t \leq 1$: infatti siccome $\langle v, u \rangle > 0$ e $\langle v, w \rangle > 0$ si ha $\langle v, tu + (1-t)w \rangle = t\langle v, u \rangle + (1-t)\langle v, w \rangle > 0$. Abbiamo quindi dimostrato che ogni classe di equivalenza è un sottoinsieme convesso di V.

Inoltre, dato $v \in T$, il punto (*2*) implica

$$\{w \in \overline{T} \mid \langle v, w \rangle \geq 0\} = \{w \in \overline{T} \mid \langle v, w \rangle > 0\}, \tag{2.7.1}$$

ed il punto (*3*) implica che l'insieme (2.7.1) dipende solo dalla classe di equivalenza di v.

Esercizi.

113. Siano $p \geq 2$, $q \geq 1$ e φ una forma bilineare simmetrica su $\mathbb{R}^{p+q}$ di segnatura (p, q). Provare che la relazione $u \sim v$ se $\varphi(u, v) > 0$, sull'insieme $T = \{u \mid \varphi(u, u) > 0\}$, non è di equivalenza.

114 (☕,♡). Sia $\Phi \colon V \to \mathbb{R}$ una forma quadratica indefinita. Dimostrare che esiste una base di V formata da vettori isotropi per Φ.

Note. La teoria delle forme quadratiche cambia notevolmente a seconda se lo scalare 2 è invertibile o no. Ad esempio, sui campi di caratteristica 2 il teorema di estensione di Witt è falso in generale.

L'aggettivo simplèttico (o simplèctico), dal greco $\sigma\upsilon\mu\pi\lambda\varepsilon\kappa\tau\iota\kappa\acute{o}\varsigma$ "relativo all'intreccio", è stato introdotto in matematica da Hermann Weyl nel 1939.

Per una forma quadratica Φ con indici di positività e negatività Φ_+ e Φ_-, alcuni autori chiamano segnatura la differenza $\Phi_+ - \Phi_-$; altri autori chiamano segnatura la terna (Φ_0, Φ_+, Φ_-) dove Φ_0, detto indice di nullità, è la dimensione del nucleo.

L'interesse dell'autore per la Proposizione 2.7.1 non ha nulla a che fare con la relatività, ma semmai con la teoria di Hodge sulle superfici K3.