Algebra Lineare

Gian Pietro Pirola

Indice

CAPITOLO 1

Dalla geometria euclidea ai vettori

1. Misurare in geometria

Indichiamo con Π il piano della geometria euclidea e ricordiamo che Π è un insieme di punti usualmente denominati con le lettere maiuscole $A, B, C \dots$. Un concetto primitivo nella geometria euclidea è quello di retta. Le rette sono sottoinsiemi del piano e sono denominate con le lettere minuscole $a, b, c \dots$. L'usuale assioma di collineazione dice che per ogni coppia di punti distinti è definita un'unica retta che li contiene. Un punto cruciale della geometria è quello della misura: dei perimetri, delle aree e, nello spazio, dei volumi.
Lo stesso nome Geometria significa: misura ($\mu\epsilon\tau\rho\iota\alpha$) della Terra ($\Gamma\eta\varsigma$). Misurare implica, in linguaggio moderno, dare una relazione di ordine ai nostri oggetti (commisurare). Dobbiamo saper dire quando due grandezze sono eguali o diverse. Seguendo l'intuizione se $A \neq B$ e r è la retta passante per A e B risulta definito (assiomaticamente) un ordinamento totale su r tale che $B \geq A$. Si ha cioè che C è minore di D se partendo da C nella direzione AB si incontra D.
Con l'utilizzo della teoria di Eudosso, il problema della misurazione è risolto negli *Elementi* ad un livello di profondità straordinario. Assodato che il sistema delle frazioni (dei numeri razionali) è insufficiente per misurare per esempio la diagonale del quadrato noto il lato, il problema tecnico fu, in un certo senso, quello di " figurarsi" l'insieme dei numeri reali. Per questo fu necessario rendere rigoroso il concetto di *continuità* o *connessione* della retta.
In termini moderni, fissati due punti distinti O e U di r, si definisce un'applicazione biettiva tra la retta r e il campo dei numeri reali $\mathbb{R}$:

$$f_{OU} : r \to \mathbb{R},$$

in modo tale che $f_{OU}(O) = 0$ e $f_{OU}(U) = 1$. Il valore $f_{OU}(A) = x_A$ è la *coordinata* di A. Questo valore permette di individuare la posizione di A su r. Noi non ci addentreremo nei dettagli della costruzione di $\mathbb{R}$, svolta, di norma, nei corsi di Analisi Matematica. Per onore di cronaca vogliamo ricordare che al tempo di Euclide (330-275 a.C.) si consideravano solo le quantità positive e che i concetti di funzione ed insieme non erano stati inventati.
A partire dalla funzione coordinata f_{OU}, si introduce la funzione differenza $F_{OU} : r \times r \to \mathbb{R}$; se $f_{OU}(A) = x_A$ e $f_{OU}(B) = x_B$, si pone:

$$F_{OU}(A, B) := f_{OU}(B) - f_{OU}(A) = x_B - x_A.$$

La <u>distanza</u> (rispetto a OU) tra C e D con $C, D \in r$ è poi definita come:

$$d_{OU}(A, B) := |F_{OU}(B, A)| = |x_B - x_A|,$$

dove con $|\cdot|$ denotiamo il valore assoluto. Se pensiamo ad r come ad un'asta di misura, il segmento OU è l'unità di misura.

2. Azioni sulla retta euclidea e cambiamenti di unità di misura

Possiamo definire due azioni di $\mathbb{R}$ su r :

- TRASLAZIONE: $T_{OU} : \mathbb{R} \times r \to r$, definita come $T_{OU}(\lambda, P) = (f_{OU})^{(-1)}(\lambda + f(P))$;

- DILATAZIONE: $D_{OU} : \mathbb{R} \times r \to r$, definita come $D_{OU}(\lambda, P) = (f_{OU})^{(-1)}(\lambda f(P))$.

Fissato $\lambda \in \mathbb{R}$, abbiamo rispettivamente le funzioni $T^{\lambda}_{OU}(\cdot) = T_{OU}(\lambda, \cdot)$ (λ-traslazione) e $D^{\lambda}_{OU}(\cdot) = D_{OU}(\lambda, \cdot)$ (λ-dilatazione). Le traslazioni mantengono le distanze, cioè, per ogni $\lambda \in \mathbb{R}$, $dOU(T^{\lambda}_{OU}(P), T^{\lambda}_{OU}(Q)) = d_{OU}(P, Q)$; si noti che il parametro λ che individua la traslazione è determinato dall'immagine di un punto, per esempio dell'origine O. Si ha che $d_{OU}(O, T^{\lambda}_{OU}(O)) = |\lambda|$.
Le dilatazioni mantengono l'origine e il parametro λ che le individua è determinato dall'immagine di U, cioè $U' := D^{\lambda}_{OU}(U) = f_{OU}(-1)(\lambda U)$. Notiamo che se λ è negativo l'orientamento (cioè l'ordinamento) della retta si inverte. Si ha che $d_{OU}(O, U') = |\lambda|$.

Possiamo ora introdurre le funzioni di <u>trasporto</u>. Fissati due punti A e B, definiamo $\tau_{AB} : r \to r$ come la funzione che associa a $P \in r$ l'unico punto Q tale che $x_B - x_A = x_Q - x_P$. Il trasporto mantiene le distanze.

Dobbiamo ora vedere come cambiano le funzioni coordinate f_{OU} al variare del segmento OU. È ragionevole supporre che ci sia compatibilità rispetto alle dilatazioni e alle traslazioni.

(1) **Compatibilità con la traslazione.** Assegnato O', sia $U' = \tau_{OO'}(U)$. Si ha che $f_{OU} = f_{O'U'} + x_{O'}$, e quindi:
$$T_{OU} = T_{O'U'}.$$
(2) **Compatibilità con la dilatazione.** È ragionevole richiedere che nel cambiamento di scala i rapporti tra le grandezze rimangano invariati. Assegnato $U' \neq O$, se $f_{OU} = kf_{OU'}$, allora $f_{OU}(U') = k$. Quindi:
$$D_{OU} = D_{OU'}.$$

ESERCIZIO: Possiamo ricavare la trasformazione generale assumendo: $f_{CD}(P) = af_{AB}(P) + b$ per ogni P. Ora, per ricavare a e b, notiamo che:
$f_{CD}(C) = 0 = af_{AB}(C) + b$
$f_{CD}(D) = 1 = af_{AB}(D) + b$.
Risolvendo il sistema si trova che $a = 1/(f_{AB}(D) - f_{AB}(C))$ e $b = -f_{AB}(C)/(f_{AB}(D) - f_{AB}(C))$; quindi:
$$f_{CD}(P) = \frac{f_{AB}(P) - f_{AB}(C)}{f_{AB}(D) - f_{AB}(C)}.$$

ESERCIZIO: Provare che $d_{CD}(P, Q) \cdot d_{AB}(P, Q) = d_{AB}(C, D)$.

Consideriamo il trasporto in r, cioè la traslazione di punti sulla retta, come definito precedentemente. Osserviamo, lasciandone al lettore la facile verifica, che τ_{AB} non dipende dalla scelta di O nè da quella di U.
$(kf_{OU'})^{(-1)}(x) = (f_{OU'})^{(-1)}(k^{-1}x)$ abbiamo:
Lasciamo come esercizio il completamento della dimostrazione del seguente:

TEOREMA 2.1. Sia $B(r, r)$ l'insieme delle funzioni biettive da r in r. Definiamo l'applicazione $\tau : r \times r \to B(r, r)$ come $\tau(A, B) = \tau_{AB}$; valgono le seguenti proprietà:

(1) $d_{OU}(\tau_{OA}(B), \tau_{OA}(C)) = d_{OU}(B, C)$;
(2) $\tau_{AB}\tau_{OA} = \tau_{OB}$;
(3) $\tau_{OO} =$ identità;
(4) $\tau_{AB} = \tau_{CD} \Leftrightarrow F_{OU}(A, B) = F_{OU}(C, D)$.

Questo conclude il nostro studio della geometria della retta.

3. Trasporto Parallelo nel Piano

Il piano Π della geometria euclidea gode del contestato, ma prezioso assioma delle rette parallele:

V assioma: *Siano r una retta e Q un punto non appartenente ad r; allora esiste ed è unica la retta s passante per Q e parallela ad r, cioè tale che l'intersezione tra s e r sia vuota.*

Utilizzeremo il V assioma per estendere al piano le funzioni di trasporto definite sulle rette. Queste saranno delle funzioni dal piano in sè, i trasporti paralleli, la versione dinamica del V assioma.
Fissati due punti A e B vogliamo definire una applicazione

$$\tau_{AB} : \Pi \to \Pi.$$

Per definizione, τ_{AA} è l'identità. Se $A \neq B$, sia r la retta per A e B. Se $P \in r$ definiamo $\tau_{AB}(P)$ il trasporto parallelo su r (vedi sezione precedente). Se $P \notin r$, sia a la retta passante per A e P; ovviamente $B \notin a$. Vi è un'unica retta a' parallela ad a e passante per B. Sia poi r' la parallela ad r passante per P. Notiamo che a' non è parallela ad r' (questo contraddirebbe il V assioma per B e r'!). Poniamo $\tau_{AB}(P) = Q = r' \cap a'$.

Il trasporto parallelo gode delle seguenti proprietà:

(1) compatibilità con la misura: la lunghezza del segmento che unisce $\tau_{AB}(P)$ e $\tau_{AB}(P)$ è uguale alla lunghezza del segmento PQ;
(2) regola del parallelogramma: dati A, B, C e D in Π; se $\tau_{AB} = \tau_{CD}$, allora $\tau_{AB}(C) = D$ e $\tau_{AC} = \tau_{BD}$.

La proprietà del parallelogramma si rappresenta graficamente:

Domanda: *la regola del parallelogramma è un teorema o un assioma?*

Nota. Il trasporto τ è stato definito nel complementare di r, $\Pi - r$, senza utilizzare la misura sulla retta, cioè la biezione con il campo dei reali, ma soltanto l'assioma delle parallele. Potremmo allora ricorrere ad una specie di trucco utilizzando la regola del parallelogramma. A tal fine, si noti che, posto $Q = \tau_{AB}(P)$, si ha $\tau_{AB}(X) = \tau_{PQ}(X)$ se X non appartiene a $r \cup r'$, dove r è la retta per A e B, e r' è la retta per P e Q. Se allora C è in r, possiamo porre $\tau_{AB}(C) = \tau_{PQ}(C)$, dove il secondo membro dell'equazione è ora definito.

Per dimostrare la regola del parallelogramma $\tau_{AB} = \tau_{PQ}$, è necessario verificare che certi triangoli sono eguali, la quale cosa, implicitamente, equivale al fatto che la misura della distanza è invariante per traslazione sulla retta. È però possibile assumere la regola del parallelogramma come assioma. Per tale via sarebbe possibile fondare una "Geometria non metrica", contraddizione linguistica.

Risposta alla Domanda: *dipende.*

Esercizio: La relazione di parallelismo è una relazione di equivalenza.

Sia $B(\Pi, \Pi)$ l'insieme delle funzioni biettive dal piano in sè e τ il trasporto:

$$\tau : \Pi \times \Pi \to B(\Pi, \Pi).$$

Proposizione 3.1. Sia V l'immagine di τ; V è un sottogruppo abeliano di $B(\Pi, \Pi)$.

L'azione di dilatazione. Fissiamo su una retta r un'origine O e un punto $U \neq O$. Dato un punto B su r e un numero reale λ, sia $B' = D_{OU}(\lambda, B)$. Possiamo allora definire la moltiplicazione $m : \mathbb{R} \times V \to V$ come $m(\lambda, \tau_{OB}) = \lambda\tau_{OB} = \tau_{OB'}$.

Esercizio: La moltiplicazione è ben definita: se $\tau_{OB} = \tau_{AC}$, allora $\lambda\tau_{OB} = \lambda\tau_{AC}$. Valgono le seguenti proprietà della moltiplicazione:

(1) $\lambda(\tau_{OB} + \tau_{AC}) = \lambda\tau_{OB} + \lambda\tau_{AC}$;
(2) $(\lambda + \mu)(\tau_{OB}) = \lambda\tau_{OB} + \mu\tau_{OB}$;
(3) $\lambda\mu(\tau_{OB}) = \lambda(\mu\tau_{OB})$;
(4) $1\tau_{OB} = \tau_{OB}$.

CAPITOLO 2

Introduzione all'Algebra Lineare

1. Spazi vettoriali

Sia K un campo detto degli scalari. Uno spazio vettoriale è un insieme V dotato di due operazioni:

(1) una interna: $f : V \times V \to V$:
(2) una esterna: $m : K \times V \to V$.

Queste devono soddisfare ad alcune proprietà che andiamo a descrivere.

(1) La f è una somma: si pone $f(v, w) := v + w$. Si richiedono le seguenti proprietà:
   (a) associatività: $\forall v, w, h \in V : f(f(v, w), h) = f(v, f(w, h))$, ovvero:
   $$(v + w) + h = v + (w + h);$$
   (b) esistenza dell'elemento neutro: $\exists u$ tale che $\forall v \in V$:
   $$v + u = u + v = v;$$
   (c) esistenza dell'inverso: $\forall v \in V \exists w \in V$ tale che:
   $$v + w = w + v = u;$$
   (d) commutatività: $\forall v, w \in V$:
   $$v + w = w + v.$$

   Le proprietà (a-c) significano che $(V, +)$ è un gruppo, la (d) che tale gruppo è commutativo.
(2) Pensando alla m come ad una moltiplicazione, si pone $m(\lambda, v) = \lambda v$. Si richiedono le seguenti compatibilità: $\forall v, w \in V, \lambda, \mu \in K$
   (a) $\lambda(v + w) = \lambda v + \lambda w$
   (b) $(\lambda + \mu)(v) = \lambda v + \mu v$
   (c) $\lambda\mu(v) = \lambda(\mu v)$
   (d) $1v = v$.

Ricapitolando, un K spazio vettoriale è un gruppo abeliano $(V, +)$ dotato di una moltiplicazione $m : K \times V \to V$ che soddisfa le proprietà (a-d) di 2. Gli elementi di V sono i vettori, gli elementi di K sono gli scalari, la moltiplicazione m è la moltiplicazione scalare per vettore.

Esempi:

(1) lo spazio vettoriale dei trasporti paralleli nel piano;
(2) $(K^n, K) : (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$;
   $\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$;
(3) i polinomi $\mathbb{P}$, le funzioni continue a valori reali;
(4) Lo spazio banale $\{0\}$.

<u>*Discussione sui polinomi:*</u> Un polinomio p di grado n a coefficienti in K si scrive:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in K.$$

Abbiamo per p due interpretazioni:

- come funzione $p : K \to K$;
- come $n+1$-pla dei coefficienti: $p \leftrightarrow (a_0, \ldots, a_n)$.

I polinomi di grado n, con la seconda interpretazione, si identificano con K^{n+1}. Se il campo ha infiniti elementi (e.g. $K = \mathbb{R}$), allora due funzioni polinomiali $q(x)$ e $q'(x)$ sono eguali se i loro coefficienti sono eguali. Per vedere questo si ponga:

$$p = q' - q = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0;$$

sappiamo ora che p deve definire la funzione identicamente nulla. Supponiamo ora per assurdo

$$a_n \neq 0,$$

e ricordiamo che, se x_1 è una radice di p, $p(x_1) = 0$, allora (regola di Ruffini):

$$p(x) = (x - x_1)p'(x).$$

Iterando, se p ha n radici distinte, allora $p(x) = a_n(x - x_1)\ldots(x - x_n)$. In tale caso, se b è una radice di p, $p(b) = 0$, e per la regola di annullamento del prodotto in un campo si ha $b = x_i$ per qualche i. Segue che p ha al massimo n radici distinte. Questo è in contraddizione col fatto che p si annulla su infiniti valori.

2. Primi elementi di algebra lineare

Cominciamo con qualche osservazione elementare. Indicheremo con $\underline{0}$ (vettore nullo) l'elemento neutro di $(V, +)$, mentre 0 sarà lo zero di K (scalare nullo). Inoltre indicheremo con $-v$ l'inverso di v in $(V, +)$, e scriveremo anche $w + (-v) = w - v$. Si ha:

$$\underline{0} + v = v + \underline{0} = v \ ; \ v - v = \underline{0}.$$

Proposizione 2.1. Per ogni $v \in V$ e per ogni $\lambda \in K$:

(1) $0v = 0$;
(2) $\lambda\underline{0} = \underline{0}$;
(3) $(-1)v = -v$;
(4) $\lambda v = \underline{0} \Rightarrow \lambda = 0$ oppure $v = \underline{0}$.

Dimostrazione. (1) $0v = (0+0)v = 0v + 0v$; w è l'inverso di $0v$ nel gruppo $(V, +)$; allora $\underline{0} = 0v + w = (0v + 0v) + w = 0v + (0v + w) = 0v + \underline{0} = 0v$; dunque $\underline{0} = 0v$;
(2) $\lambda\underline{0} = \lambda(\underline{0} + \underline{0}) = \lambda\underline{0} + \lambda\underline{0}$; dunque $\lambda\underline{0} = \underline{0}$;
(3) $\underline{0} = 0v = (-1+1)v = -1v + v$; dunque $-v = -1v$;
(4) se $\lambda v = 0$ e $\lambda \neq 0$, allora $\underline{0} = \lambda^{-1}\underline{0} = \lambda^{-1}(\lambda v) = 1v = v$.

□

Una nozione particolarmente importante è quella di <u>combinazione lineare</u>. Sia V un K spazio vettoriale e siano dati n vettori $v_1, \ldots, v_n$ e n scalari $a_1, \ldots, a_n$, la somma:

$$\sum_{i=1}^{n} a_i v_i = a_1 v_1 + \cdots + a_n v_n$$

è detta combinazione lineare dei vettori v_i per gli scalari a_i. Il vettore $w = \sum_{i=1}^{n} a_i v_i$ è detto allora combinazione lineare dei v_i. Dato un insieme dei vettori $\{v_1, \ldots, v_n\}$, sono problemi tipici:

- quali vettori w di V sono combinazione lineare dei v_i;
- in quanti modi un vettore w è combinazione lineare dei v_i.

Definizione 2.2. (sistema di generatori) Un sottoinsieme S di V si dice un sistema di generatori (s.g.) se ogni vettore di V è combinazione lineare di elementi di S. Lo spazio V si dice finitamente generato se ha un sistema di generatori finito.

Dire che S è un sistema di generatori significa che per ogni v appartenente a V, esistono $v_1, \ldots, v_m$ vettori di S e $a_1, \ldots, a_m$ scalari tali che:

$$v = \sum_{i=1}^{m} a_i v_i.$$

Esempi:

(1) V e $V - \{\underline{0}\}$ (con V non banale) sono insiemi di generatori di V;
(2) sistema di generatori standard in K^n;
(3) caso geometrico.

Esercizio: Sia S un sistema di generatori di V; sia $S' \subset V$ tale che ogni vettore di S sia combinazione di vettori di S'. Dimostrare che S' è un sistema di generatori di V.

La seguente definizione è molto importante:

Definizione 2.3. (indipendenza lineare) I vettori $v_1, \ldots, v_m$ si dicono vettori linearmente indipendenti (l.i.) se e solo se

$$a_1 v_1 + \cdots + a_m v_m = \underline{0} \Rightarrow a_1 = a_2 = \cdots = a_m = 0.$$

Dire che $v_1, \ldots, v_m$ sono linearmente indipendenti significa che il vettore nullo ammette un'unica rappresentazione come combinazione lineare dei v_i, cioè quella banale. Di più abbiamo la seguente:

Proposizione 2.4. (unicità della rappresentazione) Sia $w = a_1 v_1 + \cdots + a_m v_m$; $v_1, \ldots, v_m$ sono linearmente indipendenti se e solo se tale rappresentazione di w come combinazione lineare dei $\{v_i\}$ è unica.

Dimostrazione. Date due combinazioni lineari $w = a_1 v_1 + \cdots + a_m v_m$ e $w = b_1 v_1 + \cdots + b_m v_m$, per differenza troviamo:

$$(a_1 - b_1)v_1 + \cdots + (a_m - b_m)v_m = \underline{0};$$

allora, se i $\{v_i\}$ sono l.i., si ha $a_i = b_i$ per ogni i. Viceversa, se i vettori $\{v_i\}$ non sono l.i., devono esistere degli scalari c_i non tutti nulli tali che $\underline{0} = c_1 v_1 + \cdots + c_m v_m$. Se $w = a_1 v_1 + \cdots + a_m v_m$, allora avremo anche $w = (a_1 + c_1)v_1 + \cdots + (a_m + c_m)v_m$, cioè due rappresentazioni differenti di w. □

Esempi:

(1) $\{\underline{0}\}$;
(2) sistema di generatori standard in K^n;
(3) caso geometrico;
(4) caso di funzioni e polinomi.

Definizione 2.5. (sistema di vettori l.i.) Un insieme S (non necessariamente finito) di vettori di V si dice un sistema di vettori linearmente indipendenti (l.i.) se ogni sottoinsieme finito di S è formato da vettori linearmente indipendenti.

Esempio: $S = \{x^n\}$ nello spazio vettoriale dei polinomi.

Definizione 2.6. (base)Un sottoinsieme S di V è una base di V se è un sistema di generatori linearmente indipendenti.

Esempi: la base standard di K^n e quella dei polinomi.

Proposizione 2.7. (1) Se $\{v_1, \dots, v_n\}$ è un sistema di generatori di V, ma non una base, allora esiste un i tale $\{v_1, \dots, v_n\} - \{v_i\}$ è ancora un sistema di generatori di V.
(2) Se $\{v_1, \dots, v_n\}$ è un sistema di vettori linearmente indipendenti di V, ma non una base, allora esiste v in V tale che $\{v_1, \dots, v_n, v\}$ è un sistema di vettori l.i..

Dimostrazione. (1) Se $a_1v_1 + \cdots + a_iv_i + \cdots + a_nv_n = \underline{0}$, con $a_i \neq 0$, allora $v_i = -(a_i)^{-1}\sum_{j\neq i} a_jv_j$, quindi il sistema $\{v_1, \dots, v_n\} - \{v_i\}$ contiene tutti i vettori $\{v_j\}$ ed è allora un s.g..
(2) Se v non è combinazione lineare di $\{v_1, \dots, v_n\}$, allora $\{v_1, \dots, v_n, v\}$ è un sistema l.i.. Infatti, se $a_1v_1 + \cdots + a_nv_n + av = \underline{0}$, si ha $a = 0$, altrimenti v dipenderebbe dai $\{v_i\}$, ma allora $a_1v_1 + \cdots + a_nv_n = \underline{0}$ e quindi per ogni i, $a_i = 0$.

□

Teorema 2.8. (dello scambio) Sia $\{w_1, \dots, w_m\}$ un insieme di vettori l.i., e $\{v_1, \dots, v_n\}$ un s.g. di V. Esistono m vettori distinti $v_{i_1}, \dots, v_{i_m}$ contenuti in $\{v_1, \dots, v_n\}$ tali che

$$\{w_1, \dots, w_m\} \cup \big(\{v_1, \dots, v_n\} - \{v_{i_1}, \dots, v_{i_m}\}\big)$$

è un s.g..

Il teorema ci dice che possiamo introdurre nel sistema dei generatori $\{v_1, \dots, v_n\}$ i vettori $\{w_1, \dots, w_m\}$, espellendo opportuni m vettori $\{v_{i_j}\}$, in modo che l'insieme trovato dopo lo scambio sia ancora un sistema di generatori.
Prima di dimostrare il teorema illustriamone alcune conseguenze:

Corollario 2.9. $n \geq m$.

Corollario 2.10. Se $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ sono due basi di V, allora $m = n$.

Ci rimane da dimostrare il teorema dello scambio.

Dimostrazione. Procediamo per induzione sul numero m dei vettori l.i..
Passo iniziale: $m = 1$. Dire che $\{w_1\}$ è un insieme di vettori linearmente indipendenti significa $w_1 \neq \underline{0}$. Essendo $\{v_1, \dots, v_n\}$ un sistema di generatori, si ha:

$$w_1 = a_1v_1 + \cdots + a_nv_n.$$

Gli scalari non sono tutti nulli. Se $a_i \neq 0$, $v_i = (a_i)^{-1}(w_1 - \sum_{j\neq i} a_jv_j)$. Le combinazioni lineari di $\{\dots, v_{i-1}, w_1, v_{i+1} \dots\}$ contengono $\{v_1, v_2, \dots, v_n\}$ che è un sistema di generatori. Allora $\{\dots, v_{i-1}, w_1, v_{i+1} \dots\}$ è un sistema di generatori. Notiamo che possiamo scambiare w_1 con ognuno dei v_i per cui $a_i \neq 0$.
Passo induttivo: $m > 1$, e il teorema vale per $m - 1$. Dati $\{w_1, \dots, w_{m-1}, w_m\}$ possiamo scambiare i primi $m - 1$ vettori $\{w_1, \dots, w_{m-1}\}$ con altrettanti vettori del sistema $\{v_1, v_2, \dots, v_n\}$. A meno di riordinare gli indici possiamo supporre che siano i primi $m - 1$ vettori dei $\{v_i\}$, cioè $v_1, \dots, v_{m-1}$. Abbiamo allora che $\{w_1, \dots, w_{m-1}, v_m, \dots\}$ è un s.g.. Allora avremo:

$$w_m = a_1w_1 + \cdots + a_{m-1}w_{m-1} + b_mv_m + \cdots + b_nv_n = \sum_{i<m} a_iw_i + \sum_{j\geq m} b_jv_j.$$

L'unica nostra preoccupazione è quella di scambiare w_m con uno dei v_i (e non con i w_i). Se, per qualche j, $b_j \neq 0$, ripetiamo la procedura del primo passo di induzione. Il problema potrebbe aversi se tutti i b_j fossero nulli. In tale caso però:

$$w_m = a_1w_1 + \cdots + a_{m-1}w_{m-1},$$

quindi $a_1w_1 + \cdots + a_{m-1}w_{m-1} - w_m = \underline{0}$, ma allora tutti i coefficienti sarebbero nulli. D'altra parte sappiamo che i w_j sono l.i.; allora $-1 = 0$, cioè una contraddizione. Questo completa la dimostrazione. □

PROPOSIZIONE 2.11. Se V è finitamente generato allora ammette una base finita. Inoltre si può completare ogni sistema di vettori indipendenti ad una base.

DIMOSTRAZIONE. Si usa la proposizione 2.7. Da un insieme di generatori finito si eliminano uno ad uno i vettori in modo che diventino l.i.. Quindi abbiamo una base di n elementi. Se viceversa abbiamo un insieme S di m vettori indipendenti abbiamo $m \leq n$. Se S non è una base possiamo aggiungere ad S un vettore v in modo che $S \cup \{v\}$ sia l.i.. Questo procedimento non può continuare all'infinito: quindi se un insieme ha più di n elementi non può essere un sistema di vettori l.i.. □

DEFINIZIONE 2.12. (DIMENSIONE) Se V è finitamente generato diremo dimensione di V ($\dim(V)$) il numero di elementi di una qualsiasi sua base. Si pone $\dim(\{\underline{0}\}) = 0$. Uno spazio vettoriale ha dimensione infinita se non è finitamente generato, e porremo $\dim(V) = +\infty$.

Vedremo che la dimensione è l'unico invariante algebrico degli spazi vettoriali finitamente generati. Uno dei problemi principali è allora quello di calcolare le dimensioni e le basi degli spazi vettoriali.

PROPOSIZIONE 2.13. $\dim(K^n) = n$.

CRITERIO. V ha dimensione infinita se e solo se per ogni intero positivo n esistono n vettori linearmente indipendenti in V.

ESEMPIO: $\dim\{Polinomi\} = +\infty$.

Sia V uno spazio vettoriale di dimensione finita n, e sia $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ una base ordinata di V. Sia v un vettore di V; possiamo scriverlo in modo unico come combinazione lineare dei v_i :

$$v = a_1v_1 + \cdots + a_nv_n;$$

ricordiamo che l'esistenza di tale rappresentazione segue dal fatto che i $\{v_i\}$ formano un insieme di generatori, e l'unicità dal fatto che sono linearmente indipendenti. La n-pla (ordinata) $(a_1, \ldots, a_n)$ si dirà n-pla delle coordinate di v rispetto alla base $\mathcal{V}$. In particolare, è definita la funzione delle coordinate $F : V \to K^n$:

$$F(v) := (a_1, \ldots, a_n).$$

DOMANDA: Quali sono le coordinate di v_i?

ESERCIZIO:

(1) Provare che l'applicazione F è biettiva.
(2) Provare che, per ogni $v, w \in V$, e per ogni $\lambda \in K$, si ha: $F(v + w) = F(v) + F(w)$, e $F(\lambda v) = \lambda F(v)$.
(3) Provare che $F(v_i) = e_i := (0, \ldots, 0, 1, 0, \ldots, 0)$.

LA CORRISPONDENZA DI CARTESIO:

La scelta di un' origine O del piano Π ci permette di identificare i vettori con i punti di Π : l'insieme dei $\{\tau_{OP}\}_{P\in\Pi}$ è in corrispondenza biunivoca con $V(\Pi)$, lo spazio dei vettori geometrici sul campo dei reali. Questo definisce una biezione $\Omega : \Pi \to V(\Pi)$. Essendo $\dim(V(\Pi)) = 2$, la scelta di una base, ovvero di due rette distinte per O (gli assi cartesiani), e la scelta di due unità di misura, definiscono una mappa coordinata da $V(\Pi)$ in $\mathbb{R}^2$. Componendo, troviamo la corrispondenza di Cartesio $c : \Pi \to \mathbb{R}^2$.

Abbiamo visto che gli spazi vettoriali di dimensione n sono in corrispondenza biunivoca con K^n. Inoltre la funzione biettiva che determina la corrispondenza, cioè la funzione delle coordinate, preserva le operazioni algebriche degli spazi vettoriali. Tuttavia le coordinate dipendono dalla scelta delle basi.
Dobbiamo capire come cambiano le coordinate quando cambiamo le basi.
Siano $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e $\mathcal{W} = \{w_1, w_2, \dots, w_n\}$ due basi ordinate di V. Sia v un vettore di V. Allora possiamo scrivere:

$$v = \sum_{i=1}^{n} x_i v_i, \quad v = \sum_{j=1}^{n} y_j w_j.$$

Ora, per ogni j, abbiamo:

$$w_j = \sum_{i=1}^{n} a_{ij} v_i.$$

Quindi $v = \sum_{j=1}^{n} y_j (\sum_{i=1}^{n} a_{ij} v_i) = \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} y_j) v_i$, e per l'unicità della rappresentazione dobbiamo avere, per ogni i da 1 a n:

$$x_i = \sum_{j=1}^{n} a_{ij} y_j.$$

Sinteticamente, si costruisce la matrice $n \times n$:

$$\begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix}$$

per cui $Ax = y$, ove:

$$y = \begin{pmatrix} y_1 \\ \dots \\ y_i \\ \dots \\ y_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \dots \\ x_i \\ \dots \\ x_n \end{pmatrix}.$$

La matrice A si chiama matrice di cambiamento dalle coordinate da quelle nella base $\mathcal{W}$ a quelle nella base $\mathcal{V}$.

Osservazione. A è una matrice invertibile. Sia B la matrice di cambiamento di coordinate da $\mathcal{V}$ a $\mathcal{W}$; si ha che $AB = BA = I$, dove I è la matrice identità. Se infatti $Ay = x$ e $Bx = y$ per ogni n-pla x, allora $ABx = Ay = x$, cioè $AB = I$.

Osservazione. A è una matrice invertibile. Sia B la matrice di cambiamento di coordinate da $\mathcal{V}$ a $\mathcal{W}$; si ha che $AB = BA = I$, dove I è la matrice identità. Se infatti $Ay = x$ e $Bx = y$ per ogni n-pla x, allora $ABx = Ay = x$, cioè $AB = I$.

Nota. Le basi ordinate di K^n sono in corrispondenza con le matrici invertibili.

3. Sottospazi vettoriali

Un sottoinsieme W di un K spazio vettoriale V è un sottospazio vettoriale di V se è chiuso rispetto alle operazioni algebriche di V. Questo significa che per ogni coppia di elementi $w_1, w_2 \in W$, e $\lambda_1, \lambda_2 \in K$, anche $\lambda_1 w_1 + \lambda_2 w_2 \in W$. È immediato verificare che W con le operazioni indotte è un K spazio vettoriale. Più formalmente un sottospazio vettoriale $(W, +)$

è un sottogruppo di $(V,+)$ ed è chiuso rispetto alla moltiplicazione scalare-vettore. Il seguente criterio restringe al minimo le verifiche ed è utile nelle applicazioni.

Criterio-Definizione: W è un K sottospazio vettoriale di V se e solo se:

(1) $w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W$;
(2) $w \in W, \lambda \in K \Rightarrow \lambda w \in W$.

Dimostrazione del Criterio. *C.S.*: basta verificare che $(W,+)$ è sottogruppo abeliano; è chiuso rispetto alla somma per 1), contiene l'elemento neutro e l'inverso di ogni elemento: $0w \in W$ e $(-1)w \in W$ per 2).
C.N.: ovvia. □

Gli insiemi $\{\underline{0}\}$ e V sono ovviamente sottospazi vettoriali di V. Un sottospazio vettoriale W si dice proprio se $W \neq \{\underline{0}\}$, $W \neq V$.

Esempi:

(1) K^{n-1} pensato in K^n come l'insieme delle n-ple $(a_1, a_2, \dots, a_{n-1}, 0)$.
(2) Siano $\{v_1, v_2, \dots, v_k\}$ vettori di V; $\mathrm{span}[v_1, v_2, \dots, v_k] := \{v \in V : v = a_1v_1 + \cdots + a_kv_k\}$ è un sottospazio vettoriale di V : lo spazio generato dai v_i. Si noti che $\mathrm{span}[v_1, v_2, \dots, v_k] = V$ se e solo se $\{v_1, v_2, \dots, v_k\}$ è un s.g di V.

Esercizi:

(1) W sottospazio di $V \Rightarrow \dim(W) \leq \dim(V)$.
(2) $\dim(\mathrm{span}[v_1, v_2, \dots, v_k]) \leq k$, e in particolare $\dim(\mathrm{span}[v_1, v_2, \dots, v_k]) = k$ se e solo se i v_i sono l.i..
(3) Dimostrare che ogni spazio di dimensione n contiene sottospazi di dimensione k per ogni intero k t.c. $0 < k < n$.
(4) Sia P lo spazio dei polinomi a coefficienti reali; dire quali dei seguenti sottoinsiemi sono $\mathbb{R}$-sottospazi vettoriali di P :
  (a) i polinomi pari: $X := \{p \in P \mid p(x) = p(-x)\}$;
  (b) i polinomi dispari: $Y := \{p \in P \mid p(x) = -p(-x)\}$;
  (c) $W := \{p \in P \mid p'' - xp' = x^2p\}$;
  (d) $Z := \{p \in P \mid p(1) = 1\}$.

4. Somma e intersezione di sottospazi

Sia $\{V_i\}_{i\in I}$ un insieme di sottospazi vettoriali. Allora $\cap_{i\in I} V_i$ è un sottospazio vettoriale di V. Non è invece vero che l'unione di sottospazi è un sottospazio.

Esercizio: Provare che, se W_1 e W_2 sono sottospazi propri di V, e $W_1 \not\subset W_2$, $W_2 \not\subset W_1$, allora $W_1 \cup W_2$ non è sottospazio vettoriale di V.

Soluzione. Presi $w_1 \in W_1 - W_2$ e $w_2 \in W_2 - W_1$, allora $w_1 + w_2$ non sta in $W_1 \cup W_2$. □

Definiamo il sottospazio somma:

$$W_1 + W_2 := \{w \in V : w = w_1 + w_2,\ w_1 \in W_1, w_2 \in W_2\}.$$

Utilizzando il criterio si verifica che $W_1 + W_2$ è un sottospazio vettoriale. Notiamo che $W_1 + W_2$ è lo spazio vettoriale generato dall'insieme $W_1 \cup W_2$.

Definizione 4.1. (somma diretta) Se $W_1 \cap W_2 = \{\underline{0}\}$ la somma si dice diretta e si scrive $W_1 + W_2 =: W_1 \oplus W_2$.

Se la somma non è diretta allora esiste un $u \neq \underline{0}$ in $W_1 \cap W_2$. Se $w = w_1 + w_2$, allora $w = (w_1 + u) + (w_2 - u)$: la decomposizione di w come somma di un elemento di W_1 e di un elemento di W_2 non è unica. Se la somma è invece diretta, tale rappresentazione è unica. Se infatti $w = w_1 + w_2 = u_1 + u_2$, si ha $u = w_1 - u_1 = u_2 - w_2 \in W_1 \cap W_2$, e quindi nullo; segue che $w_1 = u_1$ e $w_2 = u_2$.

Generalizzando, se $W_1, W_2, \dots, W_k$ sono sottospazi vettoriali di V, definiamo il sottospazio somma:

$$\sum_{i=1}^{k} W_i := \{w \in V : w = \sum_{i=1}^{k} w_i,\ w_i \in W_i\}.$$

Notiamo che $\sum_{i=1}^{k} W_i = \sum_{i=1}^{k-1} W_i + W_k$. Per avere l'unicità della rappresentazione di un vettore dello spazio somma come somma di vettori dei W_i, condizione necessaria e sufficiente è che:

$$\sum_{j=1}^{i-1} W_j \cap W_i = \{\underline{0}\}, \quad \forall i < k.$$

In tale caso diremo che la $\sum_{i=1}^{k} W_i$ è somma diretta, e porremo:

$$\bigoplus_{i=1}^{k} W_i := \sum_{i=1}^{k} W_i.$$

Proposizione 4.2. (formula di Grassmann, o delle dimensioni) Siano W_1, W_2 sottospazi di V, e supponiamo che $\dim(W_1)$ e $\dim(W_2)$ siano entrambe finite. Allora:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \tag{2.1}$$

Dimostrazione. Siano $\dim(W_1 \cap W_2) = a$, $\dim(W_1) = n$ e $\dim(W_2) = m$. Partiamo da una base $\{u_i\}_{i \leq a}$ di $W_1 \cap W_2$ (se $a = 0$ tale base è vuota). I vettori $\{u_i\}_{i \leq a}$ sono linearmente indipendenti e possiamo completare tale base ad una base di W_1 : $\{u_i\}_{i \leq a} \cup \{v_j\}_{a<j\leq n}$. Nello stesso modo costruiamo una base di W_2 : $\{u_i\}_{i \leq a} \cup \{w_k\}_{a<k\leq m}$. Dimostreremo che $\{u_i\}_{i \leq a} \cup \{v_j\}_{a<j\leq n} \cup \{w_k\}_{a<k\leq m}$ è una base di $W_1 + W_2$. Da questo segue subito la (2.1):

$$\dim(W_1 + W_2) = a + (n - a) + (m - a) = m + n - a.$$

Dobbiamo ora verificare:

(1) $\{u_i\}_{i \leq a} \cup \{v_j\}_{a<j\leq n} \cup \{w_k\}_{a<k\leq m}$ è un s.g. di $W_1 + W_2$.

Preso un elemento w di $W_1 + W_2$, abbiamo $w = w_1 + w_2$; ora:

$$w_1 = \sum_{i \leq a} a_i u_i + \sum_{a<j\leq n} b_j v_j,\ w_2 = \sum_{i\leq a} c_i u_i + \sum_{a<k\leq m} d_k w_k;$$

quindi:

$$w = \sum_{i \leq a} (a_i + c_i) u_i + \sum_{a<j\leq n} b_j v_j + \sum_{a<k\leq m} d_k w_k.$$

(2) $\{u_i\}_{i \leq a} \cup \{v_j\}_{a<j\leq n} \cup \{w_k\}_{a<k\leq m}$ sono l.i..

Sia

$$\sum_{i \leq a} (a_i) u_i + \sum_{a<j\leq n} c_j v_j + \sum_{a<k\leq m} d_k w_k = \underline{0}.$$

Poniamo $u := \sum_{a<k\leq m} d_k w_k = -\sum_{i\leq a} a_i u_i - \sum_{a<j\leq n} c_j v_j$; dunque $u \in W_1 \cap W_2$, e $u = \sum_{i\leq a} b_i u_i$ poiché gli u_i sono base di $W_1 \cap W_2$. Quindi, sostituendo:

$$\sum_{i\leq a} (a_i + b_i) u_i + \sum_{a<j\leq n} c_j v_j = \underline{0},$$

e al fatto che $\{u_i\}_{i\le a} \cup \{v_j\}_{a<j\le n}$. è una base di W_1 deduciamo $c_j = 0$ per ogni j.. Ponendo questo nella combinazione lineare iniziale otteniamo:

$$\sum_{i\le a}(a_i)u_i + \sum_{a<k\le m} d_k w_k = \underline{0}.$$

Ora, $\{u_i\}_{i\le a} \cup \{w_k\}_{a<k\le m}$ è una base di W_2, e quindi è un sistema di vettori l.i.: $a_i = 0$ e $d_k = 0$ per ogni i, k.

□

Corollario 4.3. (1) $\dim(W_1+W_2) = \dim(W_1)+\dim(W_2) \iff W_1+W_2 = W_1\oplus W_2$;
(2) $\dim(\bigoplus_{i=1}^k W_i) = \sum_{i=1}^k \dim(W_i)$.

5. Prodotto cartesiano

Siano V e W due spazi vettoriali, il prodotto $V \times W$ è uno spazio vettoriale in maniera naturale. Definiamo le operazioni:
(1) $(v, w) + (v', w') := (v + v', w + w')$;
(2) $a(v, w) := (av, aw)$.

Il prodotto $V \times W$ ha due sottospazi, $V' := V \times \{\underline{0}_W\}$ e $W' := \{\underline{0}_V\} \times W$, che si possono identificare con V e W. Si ha: $V' \cap W' = \{(\underline{0}_V, \underline{0}_W)\} = \{\underline{0}_{V\times W}\}$ e $V' \cup W' = V \times W$. In particolare $\dim(V \times W) = \dim(V)+\dim(W)$.
Osservazione. Si ha $(V \times W) \times Q = V \times (W \times Q)$; in particolare $K^n \times K^m = K^{n+m}$.

6. Spazio quoziente

Veniamo ad una costruzione più delicata, lo spazio vettoriale quoziente. Il problema è che gli oggetti in questione sono classi di equivalenza, oggetti un poco indigesti. Con il dovuto dosaggio, la costruzione è molto interessante.
Sia W un sottospazio vettoriale di V. Definiamo la seguente relazione:

$$v \equiv u \iff v - u \in W.$$

La relazione $\equiv$ è di equivalenza:
- proprietà riflessiva : $v - v = \underline{0} \in W$, quindi $v \equiv v$;
- simmetrica: $v \equiv u \Rightarrow v - u \in W \Rightarrow u - v \in W$, quindi $u \equiv v$;
- transitiva: se $v \equiv u$ e $u \equiv w$, allora $v - u \in W, u - w \in W$, e anche la somma $(v - u) + (u - w) = v - w \in W$, quindi $v \equiv w$.

Indichiamo con V/W l'insieme quoziente. Gli elementi di V/W sono le classi di equivalenza, ovvero gli insiemi contenenti tutti e soli gli elementi tra di loro equivalenti. Nel nostro caso le classi di equivalenza sono del tipo:

$$[v] := \{a \in V | a = v + w, \mathrm{con} w \in W\} = v + W.$$

Si noti che $v_1 \equiv v_2 \Rightarrow [v_1] = [v_2]$. L'insieme quoziente ha una struttura naturale di spazio vettoriale:

$$[v] + [u] := [v + u]; \quad a[v] = [av].$$

Le operazioni sono ben definite: se $v_1 \equiv v_2$ e $w_1 \equiv w_2$, allora $(v_1 - v_2) \in W$ e $(w_1 - w_2) \in W$, quindi $(v_1 - v_2) + (w_1 - w_2) = (v_1 + w_1) - (v_2 + w_2) \in W$, cioè $(v_1 + w_1) \equiv (v_2 + w_2)$. Se $\lambda \in K$ e $v_1 \equiv v_2$, allora $v_1 - v_2 \in W$, quindi $\lambda(v_1 - v_2) = \lambda v_1 - \lambda v_2 \in W$, cioè $\lambda v_1 \equiv \lambda v_2$.
L'applicazione quoziente $G : V \to V/W$ è definita da:

$$G(v) = [v] = v + W.$$

Esercizio: $G(v+u) = G(v) + G(u)$; $G(av) = aG(v)$.

Proposizione 6.1. (dimensione del quoziente) Se $\dim(V) < +\infty$, allora $\dim(V) = \dim(W) + \dim(V/W)$.

Dimostrazione. Sia $\{w_1, \ldots, w_k\}$ una base di W; la si completi quindi ad una base di $V : \{w_1, \ldots, w_k, v_1, \ldots, v_s\}$. Se dimostriamo che $\{[v_1], \ldots, [v_s]\}$ è una base di V/W, avremo: $\dim(W) + \dim(V/W) = k + s = \dim(V)$.

(1) s.g.: sia $[v] \in V/W$; allora $v = \sum_{i=1}^{s} a_i v_i + \sum_{j=1}^{k} b_j w_j = \sum_{i=1}^{s} a_i v_i + w$, con $w \in W$; Quindi: $[v] = [\sum_{i=1}^{s} a_i v_i] = \sum_{i=1}^{s} a_i [v_i]$;
(2) l.i.: sia $\sum_{i=1}^{s} c_i [v_i] = [\underline{0}]$ (chiaramente il vettore nullo di V/W è $[\underline{0}]$); questo significa: $\underline{0} \equiv \sum_{i=1}^{s} c_i v_i$, cioè $\sum_{i=1}^{s} c_i v_i = w \in W$; otteniamo quindi $\sum_{i=1}^{s} c_i v_i + \sum_{j=1}^{k} d_j w_j = \underline{0}$; allora $c_i = 0$ per ogni i.

$\square$

CAPITOLO 3

Applicazioni Lineari

1. Applicazioni Lineari e matrici

La seguente è la più importante nozione dell'algebra lineare.

DEFINIZIONE 1.1. (APPLICAZIONE LINEARE) Siano V e W due K spazi vettoriali; una funzione $F : V \to W$ si dice lineare se per ogni coppia di vettori v_1, v_2 in V e per ogni λ in K si ha:

(1) $F(v_1 + v_2) = F(v_1) + F(v_2)$;
(2) $F(\lambda v_1) = \lambda F(v_1)$.

OSSERVAZIONE. Notiamo che F è lineare se e solo se mantiene le strutture algebriche di spazio vettoriale. Per induzione abbiamo:

$$F(\sum_i a_i v_i) = \sum_i a_i F(v_i).$$

MATRICI. Sia A una matrice $m \times n$; costruiamo $F_A : K^n \times K^m$, definita come:

$$F_A(x) = A^t x,$$

ove $x = (x_1, x_2, \ldots, x_n)$. La F_A è l'<u>applicazione lineare associata</u> ad A.
Ogni applicazione lineare $F : K^n \times K^m$ è associata ad una matrice $m \times n$. Per questo dimostriamo la seguente:

PROPOSIZIONE 1.2. Siano V e W due spazi vettoriali, sia $\{v_i\}_{i \in I}$ una base di V e $\{w_i\}_{i \in I}$ un insieme di vettori di W. Esiste ed è unica una applicazione lineare $F : V \to W$ tale che $F(v_i) = w_i$.

DIMOSTRAZIONE. (1) <u>*Unicità*</u>. Se esiste, F è unica: se v è un vettore di V, si ha $v = \sum_{j=1}^k a_j v_{i_j}$, con $i_j \in I$ per ogni j, quindi deve essere:

$$F(v) = \sum_{j=1}^k a_j F(v_{i_j}) = \sum_{j=1}^k a_j w_{i_j}.$$

(2) <u>*Esistenza*</u>. Per quello che abbiamo visto provando l' unicità, dobbiamo definire, se $v = \sum_{j=1}^k a_j v_{i_j}$:

$$F(v) := \sum_{j=1}^k a_j w_{i_j}.$$

Ora F è ben definita per l'unicità della combinazione lineare. Il fatto F che sia un'applicazione lineare è una semplice verifica.

□

Torniamo al caso $F : K^n \to K^m$; sia $\{e_i\}$ la base canonica di K^n; se

$$F(e_i) = \begin{pmatrix} a_{i1} \\ \dots \\ a_{ij} \\ \dots \\ a_{im} \end{pmatrix},$$

definiamo

$$A = \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mn} \end{pmatrix}.$$

Si ha allora che $F(e_i) = F_A(e_i)$ per ogni i, quindi $F = F_A$. La matrice A così costruita è detta matrice associata a F.

Il legame tra applicazione lineari e matrici è rispiegato ancora. Siano V e W spazi di dimensione finita rispettivamente n ed m; si fissino una base ordinata $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ di V e una $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ di W. Dalla proposizione sappiamo che F è nota se conosciamo $F(v_i)$ e quindi un numero finito di valori. Questo non succede (in generale) per funzioni non lineari. Come prima, se $F(v_i) = \sum_{j=1}^m a_{j,i} w_j$, posto ${}^t(a_{1,i}, a_{2,i}, \dots, a_{m,i}) =: \underline{a}_i^t$, costruiamo:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mi} & \dots & a_{mn} \end{pmatrix}$$

Siano $\underline{x} := (x_1, x_2, \dots, x_n)$ le coordinate di $v = \sum_{i=1}^n x_i v_i$, e $\underline{y} := (y_1, y_2, \dots, y_n)$ le coordinate di $F(v) = \sum_{j=1}^m y_i w_i$. D'altra parte:

$$F(v) = \sum_{i=1}^n x_i F(v_i)$$

è combinazione lineare degli $F(v_i)$, quindi:

$$\underline{y}^t = \sum_{i=1}^n x_i a_i^t = A x^t.$$

Diremo che A è associata (o rappresenta) F tramite le basi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ di V e $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ di W.

Esempi: (applicazioni lineari)

(1) Sia $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ una base di V : la mappa coordinata $F_{\mathcal{V}} : V \to K^n$.
(2) Sia W un sottospazio di V : l'inclusione $i : W \to V$ e il quoziente $q : V \to V/W$.
(3) Le proiezioni $\pi_i : V_1 \times V_2 \to V_i$, definite da $p_i(v_1, v_2) = v_i$, $i = 1, 2$.
(4) Sia $\mathcal{P}$ lo spazio dei polinomi reali: $F : \mathcal{P} \to \mathcal{P}$, definita da $F(p) = p'$.

Proposizione 1.3. Siano V, W, U tre K spazi vettoriali, a e b in K.

(1) Se $F : V \to W$ e $G : V \to W$ sono lineari, $aF + bG : V \to W$, definita da $(aF + bG)(v) := aF(v) + bG(v)$, è lineare.
(2) Se $F : V \to W$ e $H : W \to U$ sono lineari, la composizione $H \circ F : V \to U$ è lineare.
(3) Id: $V \to V$, definita da $\mathrm{Id}(v) := v$, e $0 : V \to W$, definita da $0(v) := \underline{0}_W$, sono lineari.

(4) Se $F: V \to W$ lineare è biettiva, allora $F^{-1}: W \to V$ è lineare.

Osservazione. Indichiamo con $\mathrm{Hom}(V,W)$ lo spazio delle applicazioni lineari da V a W, Hom sta per omomorfismi, sinonimo nel nostro caso di applicazioni lineari. Da 1) della precedente proposizione abbiamo che $\mathrm{Hom}(V,W)$ è dotato in modo naturale della struttura di spazio vettoriale.

Definizione 1.4. (isomorfismo) Un isomorfismo è un' applicazione lineare $F: V \to W$ biettiva.

Notiamo che se V e W sono isomorfi hanno strutture algebriche indistinguibili. In particolare due spazi isomorfi hanno la stessa dimensione. Esempi di isomorfismi sono dati dalle applicazioni coordinate. Sia $v = \{v_1, v_2, \dots, v_n\}$ una base di V $Fv: V \to K^n$ è isomorfismo. La nozione di isomorfismo è chiaramente transitiva simmetrica e riflessiva. Ne segue che due spazi di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione.

Fissiamo ora $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ una base di V e $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ una base di W, con le loro applicazioni coordinate $F_{\mathcal{V}}: V \to K^n$ e $F_{\mathcal{W}}: W \to K^m$. Se $F: V \to W$ è lineare, allora anche:

$$F_{\mathcal{W}} F (F_{\mathcal{V}})^{-1} : K^n \to K^m$$

è lineare, e $F_{\mathcal{W}} F (F_{\mathcal{V}})^{-1} = F_A$, dove la matrice A è la matrice associata alla F. Indicato con $\mathrm{Hom}(V,W)$ l'insieme delle applicazioni lineari da V e W, costruiamo una base di $\mathrm{Hom}(V,W)$: F_{ij} con $i = 1, \dots, n$ e $j = 1, \dots, m$.

$$\begin{aligned} F_{ij}(v_k) &= \underline{0} \quad \text{se} \quad k \neq i \\ F_{ij}(v_i) &= w_j. \end{aligned}$$

La matrice associata ad F_{ij} è:

$$E_{ij} = \begin{pmatrix} 0 & & \dots & & 0 \\ 0 & & \dots & & 0 \\ 0 & \dots & 1 = a_{ij} & \dots & 0 \\ 0 & & \dots & & 0 \\ 0 & & \dots & & 0 \end{pmatrix}.$$

Dato che le $E_{i,j}$ formano una base delle matrici, le $F_{i,j}$ sono una base di $\mathrm{Hom}(V,W)$ (per la proposizione di esistenza e unicità 1.2). Si ha che $\dim(\mathrm{Hom}(V,W)) = nm$. Si sono usate le basi $\mathcal{V}, \mathcal{W}$ per coordinare $\mathrm{Hom}(V,W)$. La corrispondenza $F \to A$ è l'applicazione coordinata relativa alla base $\{F_{i,j}\}_{i,j}$.

Dobbiamo ora vedere come si trasformano le matrici quando cambiano le basi di V e di W. Questo è un cambiamento di coordinate da riscrivere col linguaggio del prodotto di matrici.

Siano $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e $\mathcal{V}' = \{v'_1, v'_2, \dots, v'_n\}$ basi (ordinate) di V e $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ e $\mathcal{W}' = \{w'_1, w'_2, \dots, w'_m\}$ di W. Sia $v \in V$, con coordinate $\underline{x}$ rispetto a $\mathcal{V}$ e $\underline{x}'$ rispetto a $\mathcal{V}'$; sia poi $w = F(v)$, con coordinate $\underline{y}$ rispetto a $\mathcal{W}$ e $\underline{y}'$ rispetto a $\mathcal{W}'$. Sia A la matrice associata a F tramite $\mathcal{V}$ e $\mathcal{W}$, e D tramite $\mathcal{V}'$ e $\mathcal{W}'$. La A e la D sono matrici $m \times n$ per cui vale $\underline{y} = A\underline{x}$ e $\underline{y}' = D\underline{x}'$.

D'altra parte, siano B la matrice di cambiamento di coordinate da $\mathcal{V}$ a $\mathcal{V}'$ e C la matrice di cambiamento da $\mathcal{W}$ a $\mathcal{W}'$: allora $\underline{x}' = B\underline{x}$ e $\underline{y}' = C\underline{y}$. Quindi $C\underline{y} = \underline{y}' = DB\underline{x}$, da cui $\underline{y} = C^{-1}DB\underline{x}$; concludendo, $A = C^{-1}DB$.

Definizione 1.5. (matrici equivalenti) Due matrici $m \times n$ A e B si dicono equivalenti se esistono due matrici invertibili X $n \times n$ e Y $m \times m$ tali che $YAX = B$.

Nota. A e B sono equivalenti se sono associate alla stessa funzione lineare mediante diverse basi.

Definizione 1.6. (operatore) Un' applicazione lineare $F : V \to V$ si dice operatore.

Quando $\dim(V)$ è finita e $\mathcal{V}$ è una base di V, possiamo associare ad un operatore F una matrice A quadrata. Se la base cambia coerentemente da $\mathcal{V}$ a $\mathcal{V}'$ nel dominio e nel codominio, la matrice associata A si trasforma nella matrice $B = C^{-1}AC$, ove C è la matrice di cambiamento di coordinate.

Definizione 1.7. (matrici simili) Due matrici $n \times n$ A e B si dicono simili se esiste una matrice invertibile $n \times n$ Y tale che $Y^{-1}AY = B$.

Due matrici sono simili se e solo sono associate allo stesso operatore $F : V \to V$ mediante cambiamenti coerenti di base.

Nota. Le relazioni indotte dalla similitudine e dalla equivalenza sono relazioni di equivalenza.

Osservazione. Siano $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}, \mathcal{W} = \{w_1, w_2, \ldots, w_m\}$ e $\mathcal{U} = \{u_1, u_2, \ldots, u_l\}$ basi ordinate rispettivamente di V, W e U. Se $F : V \to W$ e $H : W \to U$ sono applicazioni lineari associate rispettivamente alle matrici A e B tramite tali basi, allora $H \circ F$ è associata a BA. Questo rende una nuova ragione della associatività del prodotto di matrici: la composizione di funzioni è infatti associativa. Se F associata ad A è invertibile, allora F^{-1} è associata a A^{-1}. Si noti che la funzione identità Id: $V \to V$ è associata alla matrice I se e solo se la base nel dominio e nel codominio è la stessa.

2. Applicazioni e sottospazi vettoriali

Definizione 2.1. (immagine e nucleo) Sia $F : V \to W$ una applicazione lineare tra due K spazi vettoriali V e W. Definiamo i seguenti sottoinsiemi di W e V:

(1) $\mathrm{Im}(F) := \{w \in W \text{ t.c. } \exists v \in V \text{ per cui } F(v) = w\}$;
(2) $\ker(F) := \{v \in V \text{ t.c. } F(v) = \underline{0}_W\}$.

Nota. ker è una abbreviazione di *kernel=nucleo*, quindi $\ker(F)$ è il nucleo di F, mentre Im è ovviamente abbreviazione di *immagine*.

Proposizione 2.2.

(1) $\mathrm{Im}(F)$ è sottospazio di W e $\ker(F)$ è sottospazio di V;
(2) F è iniettiva $\iff \ker(F) = \{\underline{0}_V\}$;
(3) F è suriettiva $\iff \mathrm{Im}(F) = W$.

Dimostrazione.
(1) Verifica diretta: se $w, w' \in \mathrm{Im}(F)$, allora $F(v) = w, F(v') = w'$ per qualche copia di vettori $v, v' \in V$; allora $F(v + v') = F(v) + F(v') = w + w'$, e per ogni $a \in K$ $F(av) = aF(v) = aw$. Analogamente, se $v, v' \in \ker(F)$, e $a, a' \in K$, allora $F(av + a'v') = aF(v) + a'F(v') = \underline{0}_W + \underline{0}_W = \underline{0}_W$, cioè $av + a'v' \in \ker(F)$.
(2) Se F è iniettiva e $F(v) = F(\underline{0}_V) = \underline{0}_W$, allora $v = \underline{0}_V$, quindi $\ker(F) = \{\underline{0}_V\}$. Se viceversa $\ker(F) = \{\underline{0}_V\}$ e $F(v) = F(w)$, allora $F(v - w) = \underline{0}_W$, dunque $v - w = \underline{0}_V$ e $v = w$.

□

Osservazione. la parte 2) è tipica degli omomorfismi di gruppo (non si usa la struttura del prodotto scalare per vettore).

Se A è una matrice, si pone $\ker(A) = \ker(F_A)$ e $\mathrm{Im}(A) = \mathrm{Im}(F_A)$. Si noti che $\ker(A)$ è esattamente lo spazio delle soluzioni del sistema lineare $Ax = \underline{0}$, mentre $\mathrm{Im}(A)$ è lo spazio generato dalle colonne di A.

TEOREMA 2.3. (DELLE DIMENSIONI) Sia $F : V \to W$ una applicazione lineare tra due K spazi vettoriali V e W. Se $\dim(V) = n < +\infty$, allora:

$$n = \dim(\mathrm{Im}(F)) + \dim(\ker(F)).$$

DIMOSTRAZIONE. Sia $\{u_1, u_2, \ldots, u_k\}$ una base di $\ker(F)$ e $\{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_{n-k}\}$ una base completata di V. Posto $w_i = F(v_i)$, vediamo che $\{w_1, w_2, \ldots, w_{n-k}\}$ è una base di $\mathrm{Im}(F)$.

(1) s.g.: prendiamo w in $\mathrm{Im}(F)$, $w = F(v)$ con $v = \sum_{i=1}^{k} a_i u_i + \sum_{j=1}^{n-k} b_j v_j = u + \sum_{j=1}^{n-k} b_j v_j$, con $u \in \ker(F)$; allora $w = F(v) = \sum_{j=1}^{n-k} b_j F(v_j) = \sum_{j=1}^{n-k} b_j w_j$;
(2) l.i.: supponiamo che $\sum_{j=1}^{n-k} c_j w_j = \underline{0}_W$; allora $F(\sum_{j=1}^{n-k} c_j w_j) = 0$. Da questo segue che $v = \sum_{j=1}^{n-k} c_j w_j \in \ker(F)$; ma allora $v = \sum_{i=1}^{k} d_i u_i$, e quindi $\sum_{j=1}^{n-k} c_j w_j + \sum_{i=1}^{k}(-d_i) u_i = \underline{0}_V$; si ottiene che $c_i = 0$ per ogni i.

□

Avrete certamente notato il parallelismo con la dimostrazione dell'analogo teorema nel caso dello spazio quoziente. La cosa non è casuale: consideriamo il quoziente $V/\ker(F)$ e definiamo $f : V/\ker(F) \to \mathrm{Im}(F)$ nel seguente modo:

$$f([v]) := F(v).$$

La f è ben definita; infatti $v \equiv v'$ (cioè $[v] = [v']$) se e solo se $(v - v') \in \ker(F)$, ma allora $\underline{0}_W = F(v - v') = F(v) - F(v')$, dunque $F(v) = F(v')$.

TEOREMA 2.4. (DI OMOMORFISMO) $f : V/\ker(F) \in \mathrm{Im}(F)$ è un isomorfismo.

DIMOSTRAZIONE. Abbiamo $\mathrm{Im}(f)$ =$\mathrm{Im}(F)$: f è suriettiva; se $f([v]) = F(v) = \underline{0}_W$, allora $v \in \ker(F)$, cioè $[v] = [\underline{0}_V]$: f è iniettiva. □

OSSERVAZIONE. Si noti la struttura di una applicazione lineare $F : V \to W$; definiamo il quoziente (suriettivo) $q : V \to V/\ker(F)$, l'isomorfismo $f : V/\ker(F) \to \mathrm{Im}(F)$, e l'inclusione (iniettiva) i :$\mathrm{Im}(F) \to W$:si ha $F = i \circ f \circ q$.

Veniamo ad alcune importanti conseguenze del teorema delle dimensioni:

COROLLARIO 2.5. Sia $F : V \to W$ una applicazione lineare tra spazi di dimensione finita. Se F è suriettiva, allora $\dim(V) \geq \dim(W)$; se F è iniettiva, allora $\dim(V) \leq dim(W)$. Se $\dim(V)$ =$\dim(W)$, allora F è iniettiva se e solo se è suriettiva.

Se A è una matrice, $\dim(\mathrm{Im}(A))$= $r(A)$. Il teorema diventa:

COROLLARIO 2.6. Sia A una matrice $m \times n$, allora $r(A)$+$\dim(\ker(A) = m$.

COROLLARIO 2.7. Sia A una matrice quadrata $n \times n$, allora sono equivalenti:

(1) A è invertibile;
(2) $\ker(A) = \{\underline{0}\}$;
(3) $r(A) = n$;
(4) le colonne (le righe) di A formano una base di K^n.

ESERCIZI: (vero o falso)

(1) AB invertibile $\Longrightarrow A$ invertibile;
(2) A e B invertibili $\Longrightarrow AB$ invertibile;
(3) se A non è invertibile, AB può essere invertibile;
(4) $AB = I \Longrightarrow BA = I$;

(5) $AB = I$ e $CA = I \Longrightarrow C = B$;
(6) $AB = D$ e $CA = D$, con D invertibile $\Longrightarrow C = B$;
(7) $AB = 0 \Longrightarrow r(A) + r(B) > n$.

Concludiamo questo paragrafo con la seguente:

Proposizione 2.8. (matrici equivalenti e rango) Due matrici $m \times n$ A e B sono equivalenti se e solo se $r(A) = r(B)$.

Dimostrazione. Sia $F_A : K^n \to K^m$; A rappresenta F_A rispetto alle basi standard. Costruiamo delle basi opportune: sia $\{v_1, v_2, \ldots, v_k\}$ una base di $\ker(F_A)$ che completiamo ad una base di K^n con $u_1, \ldots u_{n-k}$; riordinando, abbiamo $\mathcal{V} = \{u_1, \ldots, u_{n-k}, v_1, v_2, \ldots, v_k\}$. Dalla dimostrazione del teorema delle dimensioni $\{w_1 = F_A(u_1), \ldots, w_{n-k} = F_A(u_{n-k})\}$ è una base di $\mathrm{Im}(F)$, completiamola ad una base di K^m $\mathcal{W} = \{w_1, w_2, \ldots, w_m\}$. La matrice associata a F tramite $\mathcal{V}$ e $\mathcal{W}$ è:

$$C(k) = \begin{pmatrix} I_k & 0_{k,m-k} \\ 0_{n-k,k} & 0_{n-k,m-k} \end{pmatrix}.$$

Le matrici A e $C(k)$ sono equivalenti, quindi ogni matrice di rango k è equivalente a $C(k)$. Viceversa, se A e B sono associate ad F, allora $r(A)$ =dim(Im(F))= $r(B)$. □

Nota. Due matrici simili sono equivalenti e hanno lo stesso rango, ma il viceversa non è vero. La matrice identità, per esempio, è simile solo a se stessa.

3. Problemi lineari e sistemi lineari

Un problema si dice problema lineare che richiede di trovare le soluzione dell'equazione, cioè i valori delle x, di

$$F(x) = b, \tag{3.1}$$

dove $F : V \to W$ è una applicazione lineare tra due K spazi vettoriali V e W. Si noti che tale problema ha soluzione se e solo se $b \in \mathrm{Im}(F)$. Le soluzioni dei problemi lineari hanno una struttura particolarmente semplice. Introduciamo il problema lineare omogeneo associato:

$$F(x) = \underline{0}. \tag{3.2}$$

Le soluzioni di (3.2) sono gli elementi del nucleo di F. Se y e z sono due soluzioni di (3.1), allora $m = z - y$ è soluzione del problema omogeneo (infatti $F(m) = F(z) - F(y) = \underline{0}$). Viceversa, se z risolve (3.1) e m risolve (3.2), allora $F(z + m) = F(z) = b$. Quindi:

(1) Il problema (3.1) è risolubile se e solo se $b \in \mathrm{Im}(F)$.
(2) Se (3.1) è risolubile, e z è una soluzione, tutte e sole le soluzioni hanno la forma $x + m$ dove $m \in \ker(F)$, ovvero m risolve (3.2). L' insieme delle soluzioni è la classe di equivalenza $m + \ker(F) = [m]$, $[m] \in V/\ker(F)$.

Se la F è associata alla matrice A, cioè $F = F_A$, abbiamo il sistema lineare $Ax = b$. Si introduce la matrice $A^\star := (A, b)$ aggiungendo alla A la colonna dei termini noti.

Teorema 3.1. (di Rouchè-Capelli)
(1) $Ax = b$ ha soluzioni se e solo se $r(A^\star) = r(A)$.
(2) Se v è una soluzione, allora l'insieme delle soluzioni è della forma $v + \ker(A)$, con $\dim(\ker(A)) = n-$rango(A).

Osservazione. Quando V e W hanno dimensione finita, una scelta delle basi trasforma i problemi lineari in sistemi lineari.

4. Spazio Duale e Coordinate

Lo spazio duale di un K spazio V è per definizione $V^\star$:=Hom(V, K). Supponiamo che $\dim(V) = n < +\infty$; sia $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ una base ordinata di V. Ogni vettore v di V si scrive in maniera unica come combinazione lineare dei vettori della base:

$$v = a_1(v)v_1 + a_2(v)v_2 + \cdots + a_n(v)v_n.$$

La i-esima coordinata di v rispetto alla base $\mathcal{V}$ è $a_i(v)$. Definiamo l'applicazione lineare (i-esima coordinata) $A_i : V \to K$ come:

$$A_i(v) = a_i(v).$$

Nota. Le coordinate sono elementi di $V^\star$. Abbiamo: $A_i(v_j) = 0$ se $i \neq j$, e $A_i(v_i) = 1$.

Proposizione 4.1. (base duale) I vettori di $V^\star$ $A_1, A_2, \ldots, A_n$ formano una base di $V^\star$. $\mathcal{V}^\star = \{A_1, A_2, \ldots, A_n\}$ è detta la base duale di $\mathcal{V}$.

Dimostrazione. (1) l.i.: sia $\sum_{i=1}^n a_iA_i = 0$, dove $0(v) = 0$ per ogni v; allora, per ogni j, $(\sum_{i=1}^n a_iA_i)(v_j) = 0$, ovvero $a_j = 0$;
(2) s.g.: data f in $V^\star$, sia $f(v_i) = b_i$; ora, posto $g = f - (\sum_{i=1}^n b_iA_i)$, si ha $g(v_j) = f(v_j) - (\sum_{i=1}^n b_iA_i)(v_j) = b_j - b_j = 0$, dunque per ogni $v = \sum_{i=1}^n c_iv_i$, $g(v) = \sum_{i=1}^n c_ig(v_i) = 0$; segue $g = 0$, quindi $f = \sum_{i=1}^n b_iA_i$.

□

Sia $F : V \to W$ una applicazione lineare, e $f : W \to K$ un elemento di $W^\star$; allora la composizione $f \circ F : V \to K$ appartiene a $V^\star$. Definiamo $F^\star : W^\star \to V^\star$ come $F^\star(f) = f \circ F$. La dimostrazione che $F^\star$ è lineare è lasciata alla cura del lettore.
Veniamo al caso finito dimensionale. Siano $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ una base di V e $\mathcal{W} = \{w_1, w_2, \ldots, w_m\}$ una base di W. Siano $\mathcal{V}^\star = \{A_1, A_2, \ldots, A_n\}$ e $\mathcal{W}^\star = \{B_1, B_2, \ldots, B_m\}$ le basi di $V^\star$ e $W^\star$ duali di $\mathcal{V}$ e w. Se

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

è la matrice associata ad F : $F(v_i) = a_{1i}w_1 + \cdots + a_{ji}w_j + \cdots + a_{mi}w_m$ per ogni $i = 1, \ldots, n$. Se $X = (x_{ji})$ è la matrice associata ad $F^\star$ tramite le basi duali si ha:

$$F^\star(B_j) = x_{1j}A_1 + \cdots + x_{ij}A_i + \cdots + x_{nj}A_n.$$

Per calcolare le x_{ji} calcoliamo $F^\star(B_j)$ sulle v_i:

$$F^\star(B_j)(v_i) = x_{1j}A_1(v_i) + \cdots + x_{ij}A_i(v_i) + \cdots + x_{nj}A_n(v_i) = x_{ij}.$$

D'altra parte dalla definizione di mappa duale

$$x_{ij} = F^\star(B_j)(v_i) = B_j(F(v_i)) = B_j(a_{1i}w_1 + \cdots + a_{ji}w_j + \cdots + a_{mi}w_m) = a_{ji}.$$

Conclusione: Se A è la matrice associata ad F attraverso le basi $\mathcal{V}$ e $\mathcal{W}$ la matrice associata ad $F^\star$ mediante le basi $\mathcal{V}^\star$ e $\mathcal{W}^\star$ è la trasposta, ${}^tA$, di A.

Esercizio: Siano $F : V \to W, G : W \to U$ lineari e $H = G \circ F : V \circ U$ la composizione, siano $F^\star : W^\star \to V^\star$, $G^\star : U^\star \to W^\star$ e $H^\star : U^\star \to V^\star$. Allora $H^\star = F^\star \circ G^\star$. Se Id:$V \to V$, allora $(\mathrm{Id})^\star$ =Id:$V^\star \to V^\star$.

Osservazione: Passando alle matrici associate abbiamo ${}^t(AB) = {}^tB\ {}^tA$. Regola che conoscevamo. Se $\dim(V) = n$ allora $\dim(V^\star) = n$ e quindi i due spazi sono isomorfi. Se $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ è una base di V, e $\mathcal{V}^\star = \{A_1, A_2, \star, A_n\}$ la base duale, possiamo porre $F(v_i) = A_i$ ed estendere la F per linearità. Tale isomorfismo non è in generale canonico, per esempio non esiste una base naturale nel caso geometrico. Il riflesso di questa non-canonicità viene dal fatto che non vi è un modo naturale di trasformare una applicazione lineare F:V W ad una applicazione G:V* W*. Abbiamo visto infatti che nella dualizzazione naturale F*:W* V* le frecce sono invertite.

Esercizio: *(difficile)* Sia P lo spazio dei polinomi reali; $P^\star$ è isomorfo a P?

La seguente proposizione calcola i sottospazi $\ker(F^\star)$ e $\mathrm{Im}(F^\star)$:

Proposizione 4.2. Sia $F : V \to W$ lineare, e $F^\star : W^\star \to V^\star$ la sua duale. Sia $f : W \to K$ un elemento di $W^\star$, e $g : V \to K$ un elemento di $V^\star$. Allora:

(1) $f \in \ker(F^\star)$ se e solo se $\mathrm{Im}(F) \subset \ker(f)$;
(2) $g \in \mathrm{Im}(F^\star)$ se e solo se $\ker(F) \subset \ker(g)$.

Dimostrazione. (1) Ovvio: $0 = F^\star(f) = f \circ F$ se e solo se $f(F(v)) = \underline{0}$ per ogni v, cioè se e solo se $\mathrm{Im}(F) \subset \ker(f)$.

(2) Sia $g = F^\star(f) = f \circ F$; allora $g(v) = f(F(v))$, quindi se $F(v) = \underline{0}$ si ha $g(v) = 0$. Viceversa, se $\ker(F) \subset \ker(g)$, dobbiamo costruire una f tale che $g = f \circ F$. Cominciamo a definire la f su $\mathrm{Im}(F)$: se $w \in \mathrm{Im}(F)$ e $F(v) = w$ poniamo $f(w) = g(v)$. Se v' è un altro elemento di V tale che $F(v') = w$, allora $F(v - v') = 0$ e quindi $v - v' \in \ker(F) \subset \ker(g)$; segue che $g(v - v') = 0$, cioè $g(v) = g(v')$. Quindi la definizione di f su $\mathrm{Im}(F)$ è ben data.
Ora dobbiamo estendere f a W. Per la nostra dimostrazione è irrilevante sapere come questa estensione si effettui. Infatti nella composizione $F^\star(f) = f \circ F$, la f pesca valori in $\mathrm{Im}(F)$ con totale disinteresse a quello che succede fuori dall'immagine. Però dobbiamo provare che una tale estensione è possibile. Poniamo $U = \mathrm{Im}(F)$. Proviamo ora il seguente:

Lemma 4.3. Sia $f : U \to K$ lineare, con U sottospazio di W. Allora esiste una applicazione lineare $h : W \to K$ tale che $h_{|U} = f$.

Dimostrazione del lemma. Il caso in cui W è finito dimensionale è semplice: si estende una base $\{u_1, \dots, u_k\}$ di U a una di W : $\{u_1, \dots, u_k, w_1, \dots, w_{m-k}\}$. Si pone allora $h(u_i) := f(u_i)$, e (per esempio) $h(w_i) = 0$. Si estende per linearità.
Trattiamo (per scopi didattici) il caso $\dim(W) = +\infty$, per questo si ricorre al Lemma di Zorn. Si considerino tutte le coppie (N, k) dove N è un sottospazio di W che contiene U, e $k : N \to K$ è una applicazione lineare tale che $k_{|U} = h$. Sia S l'insieme di tali coppie. In S abbiamo un ordinamento parziale: $(N, k) \le (L, p)$ se $N \subset L$ e $p_{|L} = k$. Il Lemma di Zorn ci dice che esiste un elemento massimale (M, h). Questo significa che se $(M, h) \le (L, p)$ allora $L = M$ e (ovviamente) $p = h$. Se possiamo provare che $M = W$, la funzione h risolve il nostro problema. Supponiamo (per assurdo) di no, e cioè che esista $v \in W, v \notin M$. Vogliamo provare che M non è massimale. Sia A il sottospazio generato da v, cioè $A = \{av, a \in K\}$. Poniamo $M' = M + A = M \oplus A$. Per costruzione M' contiene M e ogni elemento di w di M' si scrive in modo unico nella forma: $w = m + av$ con $m \in M$. Definiamo $h' : M' \to K$, $h'(w) = h(m)$. Dunque $h'(A) = \{0\}$, e $h'_{|M} = h$; h' è una applicazione lineare. Per costruzione $(M, h) \le (M', h')$, con $M' \ne M$, in contraddizione con la massimalità di (M, h). □

Dal lemma segue la tesi.

□

Corollario 4.4. Nel caso finito dimensionale, $\dim(\mathrm{Im}(F))=\dim(\mathrm{Im}(F^\star))$, quindi $F^\star$ è suriettiva (iniettiva) se e solo se F è iniettiva (suriettiva). Se A è una matrice $r(A) = r({}^tA)$.

Dimostrazione. Supponiamo $F : V \to W$ suriettiva. Dalla proposizione si ha $F^\star$ iniettiva. Quindi dim(Im(F))=dim(W)=dim(ker($F^\star$))+dim(Im($F^\star$))= dim(Im($F^\star$)).

Nel caso generale poniamo U =Im(F) e $F' : V \to U$ l'applicazione indotta dalla F, cioè $F(v) = F'(v)$. Ovviamente $\ker(F) = \ker(F')$ ($F' = F$ con restrizione di codominio). Dalla seconda parte della (?????3.4.2?????) abbiamo Im($F^\star$) =Im($(F')^\star$) e quindi:

$$\dim(\text{ Im } (F^\star)) = \dim(\text{ Im } (F')^\star) = \dim(U) = \dim(\text{ Im } (F)).$$

□

Tentazione: iterare la dualità. Consideriamo il biduale di $V : (V^\star)^\star$. Un elemento di $(V^\star)^\star$ è un funzionale, cioè una funzione sulle funzioni. Se $v \in V$ è fissato, possiamo calcolare tutte le funzioni lineari di $V^\star$ su v. Definiamo:

$$i_v : V^\star \to K, \quad i_v(f) := f(v).$$

Si ha $i_v(f+g) = (f+g)(v) = f(v) + g(v) = i_v(f) + i_v(g)$ iv(lg)=(lg)(v)=lg(v)= liv(g) Quindi $i_v : V^\star \to K$ è lineare: $i_v \in (V^\star)^\star$. Analogamente si verifica che $i_{v+w} = i_v + i_w$ e $i_{\lambda v} = \lambda i_v$ per ogni $v, w \in V$ e $\lambda \in K$. Abbiamo definito una applicazione lineare naturale:

$$i : V \to (V^\star)^\star, \quad i(v) := i_v.$$

Tale funzione è iniettiva: se $i(v) = i_v$ è lo zero di $(V^\star)^\star$, significa che $i_v(f) = f(v) = 0$ per ogni $f : V \to K$ lineare, quindi $v = \underline{0}$. Se v non fosse zero esisterebbe infatti $f : V \to K$ lineare tale che $f(v) = 1 \neq 0$. Se $\dim(V) = n < +\infty$, questo si mostra completando $\{v\}$ ad una base di V; se $\dim(V) = +\infty$ si procede con il Lemma di Zorn come in (????? 3.4.2 ?????). Notiamo che, se $\dim(V) = n < +\infty$, abbiamo $\dim(V)$ =dim($V^\star$) =dim($(V^\star)^\star$) $= n$, quindi $i : V \to (V^\star)^\star$ è biettiva. Abbiamo così dimostrato la seguente:

Proposizione 4.5. L' applicazione naturale $i : V \to (V^\star)^\star$, definita da $i(v) := iv$, dove $iv(f) := f(v)$, è iniettiva; in particolare, se $\dim(V) = n < +\infty$, è un isomorfismo.

Corollario 4.6. Sia $F : V \to W$ una applicazione lineare tra spazi di dimensione finita. Identificando $(V^\star)^\star$ con V e $(W^\star)^\star$ con W, si ha $(F^\star)^\star = F$.

5. Applicazioni multilineari (definizioni)

Concludiamo questo capitolo con una generalizzazione della definizione di applicazione lineare. Tali nozioni saranno riprese nel seguito.
Siano $V_1, \dots, V_s$ e W K spazi vettoriali; considereremo delle applicazioni:

$$f : V_1 \times \cdots \times V_s \to W.$$

Definizione 5.1. (applicazione multilineare) La funzione f si dice multilineare se comunque fissati $v_1 \in V_1, \dots, v_{i-1} \in V_{i-1}, v_{i+1} \in V_{i+1}, \dots, v_s \in V_s$, l'applicazione $f_i : V_i \to W$ definita da:

$$f_i(x) = f(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_s)$$

è lineare, cioè è lineare componente per componente, ovvero se dati comunque $v_i, w_i \in V_i$, e $\lambda_i, \mu_i \in K$, $i = 1, \dots, n$, allora:

$$f(v_1, \dots, \lambda_i v_i + \mu_i w_i, \dots, v_s) = \lambda_i f(v_1, \dots, v_i, \dots, v_s) + \mu_i f(v_1, \dots, w_i, \dots, v_s).$$

Definizione 5.2. (forma multilineare) Quando $W = K$, una applicazione multilineare $f : V_1 \times \cdots \times V_s \to K$ si dice forma multilineare. Nel caso $s = 2$, si dice forma bilineare.

Nota. Quando $s = 1$, una applicazione multilineare è semplicemente una applicazione lineare. Quando $s = 2$, la bilinearità di $f : V \times W \to K$ si esprime come:

$$f(\lambda_1 v_1 + \lambda_2 v_2, \mu_1 w_1 + \mu_2 w_2)$$
$$= \lambda_1\mu_1 f(v_1, w_1) + \lambda_1\mu_2 f(v_1, w_2) + \lambda_2\mu_1 f(v_2, w_1) + \lambda_2\mu_2 f(v_1, w_2).$$

Esempi: (applicazioni bilineari)

(1) Sia A una matrice $m \times n$; $f_A : K^n \times K^m \to K$ definita da $f_A(x, y) :=^t xAy$.
(2) Sia $K = \mathbb{R}$; $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definita da $g((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) := \sum_{i=1}^n x_i y_i$ (prodotto scalare standard.)
(3) Dati V e $V^\star$, $k : V^\star \times V \in K$ definita da $k(f, v) := f(v)$.

Definizione 5.3. Sia $V_1 = \cdots = V_s = V$. Una forma f:V...VK multilineare si dice: simmetrica se per ogni i e j, i¡j f(...,vi ,... ,vj,...)=f(..., vj ,... ,vi,...). alternante (antisimmetrica se $s = 2$) f(...,v i ,... ,vj,...)=-f(...,vj,...,vi,...).

Osservazione. Nell'esempio 1) f_A è simmetrica (antisimmetrica) se e solo se $A =^t A (A = -^t A)$, cioè A è simmetrica (antisimmetrica). Il prodotto scalare dell'esempio 2) è una forma simmetrica ed è eguale a f_I.

Se f è alternante e il campo non è di caratteristica 2, $f(\ldots, v, \ldots, v, \ldots) = 0$. Viceversa se f è multilineare e $f(\ldots, v, \ldots, v, \ldots) = 0$ per ogni v, allora $0 = f(\ldots, v_i + v_j, \ldots, v_i + v_j, \ldots) = f(\ldots, v_i, \ldots, v_i, \ldots) + f(\ldots, v_i, \ldots, v_j, \ldots) + f(\ldots, v_j, \ldots, v_i, \ldots) + f(\ldots, v_j, \ldots, v_j, \ldots) = 0 + f(\ldots, v_i, \ldots, v_j, \ldots) + f(\ldots, v_j, \ldots, v_i, \ldots) + 0$, dunque

$$f(\ldots, v_i, \ldots, v_j, \ldots) = -f(\ldots, v_j, \ldots, v_i, \ldots),$$

ed f è alternante.
Può convenire usare la condizione $f(\ldots, v, \ldots, v, \ldots) = 0$ per definire l'alternanza nel caso multilineare. Notiamo che rispetto alla prima definizione ogni forma in caratteristica due è banalmente alternante, quindi la seconda condizione ha il pregio di funzionare in ogni caratteristica.
Le forme alternanti sono a prima vista bizzarre, vedremo però nel prossimo capitolo come una di esse sia in un certo senso nel cuore di tutta la teoria.

CAPITOLO 4

Determinanti

1. Definizione induttiva

In questo paragrafo fino a diverso avviso tratteremo matrici quadrate di ordine n su un campo K. Lo spazio vettoriale delle matrici quadrate di ordine n su K sarà indicato con $\mathcal{M}(n,k)$, inoltre $\mathcal{M}(1,K)$ si identifica con K, e la matrice $A = (a)$ di ordine 1 si identifica con il suo determinante; poniamo cioè

$$\det(A) = a.$$

Passando alle matrici di ordine 2, se $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, definiamo

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Indicando con $\det(A)$ il determinante di A, abbiamo costruito una funzione

$$\det : \mathcal{M}(2,K) \to K.$$

Per apprezzarne l'importanza, notiamo che $\det(A) \neq 0$ se e solo se il rango di A è uguale a due, cioè A è invertibile. Moltiplicando A per $B = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$, si ottiene $AB = \begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} = \begin{pmatrix} \det(A) & 0 \\ 0 & \det(A) \end{pmatrix}$.
Allora, se $\det(A) \neq 0$, la matrice $\dfrac{1}{\det(A)}B$ è l'inversa di A, e quindi A ha rango 2. Se $\det(A) = 0$, le colonne di B soddisfano il sistema lineare $A(x) = \underline{0}$, ovvero sono nel nucleo di A; se $A \neq 0$, anche $B \neq 0$, e allora $\dim(\ker(A)) > 0$, e $r(A) < 2$.
Il nostro scopo è quello di definire un numero per ogni matrice quadrata che ne determini l'invertibilità. Abbiamo bisogno di premettere la seguente:

Definizione 1.1. Sia $A = (a_{ij})$ una matrice $n \times n$; si denoterà A_{ij} la matrice $(n-1)\times(n-1)$ ottenuta dalla matrice A cancellando la i-esima riga e la j-esima colonna.

Per induzione noi sappiamo calcolare i determinanti delle varie A_{ij}.

Definizione 1.2. (determinante) Se $n = 1$ e $A = (a)$, si pone $\det(A) = a$; se $n > 1$ si pone:

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \cdots + (-1)^{n+1} a_{1n}\det(A_{1n}).$$

$\det(A)$ è per definizione il determinante di A.

Si noti che la definizione è in accordo con la precedente nel caso $n = 2$.

Esempi:

(1) Caso di matrici di ordine 3:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Risulta:

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13})$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}).$$

(2) Esempio numerico: sia $K = \mathbb{C}$, il campo dei numeri complessi ($i^2 = -1$);

$$A = \begin{pmatrix} 1 & 0 & 0 & 2+i \\ 10 & 0 & 0 & i \\ 1 & 0 & 2 & 7 \\ 1 & 2i & 0 & 1 \end{pmatrix}.$$

Allora

$$\det(A) = \det\begin{pmatrix} 0 & 0 & i \\ 0 & 2 & 7 \\ 2i & 0 & 1 \end{pmatrix} - (2+i)\det\begin{pmatrix} 10 & 0 & 0 \\ 1 & 0 & 2 \\ 1 & 2i & 0 \end{pmatrix}$$

$$= i\det\begin{pmatrix} 0 & 2 \\ 2i & 0 \end{pmatrix} - (2+i)10\det\begin{pmatrix} 0 & 2 \\ 2i & 0 \end{pmatrix}$$

$$= i(-4i) - (2+i)10(-4i) = -36 + 80i.$$

(3) Caso in cui A è *triangolare bassa*:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ a_{12} & a_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ a_{1n} & \dots & \dots & \dots & a_{nn} \end{pmatrix}.$$

Si ha:

$$\det(A) = a_{11}a_{22}\dots a_{nn}.$$

In particolare:

$$\det(I) = 1.$$

2. Proprietà del determinante

Il determinante definisce, per ogni intero positivo n, una funzione

$$\det : \mathcal{M}(n, K) \to K.$$

Studieremo le proprietà della funzione determinante che, tra l'altro, aiutano il calcolo esplicito. La prima osservazione è che il determinante è funzione polinomiale dei coefficienti. Inoltre abbiamo già visto la

normalizzazione: $\det(I) = 1$.

Pensando alle matrici A come ad n-ple ordinate di vettori colonna, cioè $A = (v_1, \dots, v_i, \dots, v_n)$, vale la seguente

Proposizione 2.1. ((multi)linearità per colonne) Per ogni i :

$$\det(v_1,\dots,v_i+w_i,\dots,v_n)=\det(v_1,\dots,v_i,\dots,v_n)+\det(v_1,\dots,w_i,\dots,v_n)$$
$$\det(v_1,\dots,av_i,\dots,v_n)=a\det(v_1,\dots,v_i,\dots,v_n).$$

Dimostrazione. Dobbiamo procedere per induzione, essendo il caso $n=1$ evidente. Fissiamo $n-1$ vettori $v_1,\dots,v_{i-1},v_{i+1},\dots,v_n$ e costruiamo l'applicazione $f:K^n\to K$, definita da

$$f(x_{1i},x_{2i},\dots,x_{ni})=\det(v_1,\dots,v_{i-1},x,v_{i+1},\dots,v_n)=\det(A(x))$$
$$=\det\begin{pmatrix}a_{11}&\dots&x_{1i}&\dots&a_{1n}\\ \dots&\dots&\dots&\dots&\dots\\ a_{j1}&\dots&x_{ji}&\dots&a_{jn}\\ \dots&\dots&\dots&\dots&\dots\\ a_{n1}&\dots&x_{ni}&\dots&a_{nn}\end{pmatrix},\quad\text{dove } x=\begin{pmatrix}x_{1i}\\ \dots\\ x_{ji}\\ \dots\\ x_{ni}\end{pmatrix},\ v_k=\begin{pmatrix}a_{1k}\\ \dots\\ a_{jk}\\ \dots\\ a_{nk}\end{pmatrix}.$$

Dobbiamo provare la linearità di f :

$$\det(A(x))=\sum_{k\neq i}a_{1k}\det(A_{1k}(x))+(-1)^{i+1}x_{1i}\det(A_{1i}(x))=\sum_{i=1}^{n}f_k(x).$$

dove:

$$f_k(x)=(-1)^{k+1}a_{1k}\det(A_{1k}(x))\quad\text{se } k\neq i$$
$$f_i(x)=(-1)^{i+1}x_{1i}\det(A_{1i}(x)).$$

Le $f_k(x)$, $k\neq i$, sono lineari per induzione. La matrice $A_{1i}(x)$ invece non dipende da x (la i-esima colonna non appare in $A_{1i}(x)$), allora $\det(A_{1i}(x))=a$ è uno scalare che non dipende da x; $f_i(x)=(-1)^{i+1}ax_{1i}$ è lineare. Quindi la f è somma di funzioni lineari. □

Osservazione. La funzione determinante non è lineare: $\det(A+B)\neq det(A)+det(B)$ in generale, e $\det(aA)=a^n\det(A)$.

Proposizione 2.2. (alternanza (colonne))

i) Se A ha due colonne uguali allora $\det(A)=0$.
ii) Per ogni coppia di indici i e j con $i<j$ si ha:

$$\det(\dots,v_i,\dots,v_j,\dots)=-\det(\dots,v_j,\dots,v_i,\dots).$$

Il determinante, cioè, cambia di segno se scambiamo due colonne.

Dimostrazione. Vediamo come dalla i) segua immediatamente la ii): si ha

$$0=\det(\dots,v_i+v_j,\dots,v_i+v_j,\dots)=\ \text{(per linearità)}\ \det(\dots,v_i,\dots,v_i,\dots)$$

$$+\det(\dots,v_i,\dots,v_j,\dots)+\det(\dots,v_j,\dots,v_i,\dots)+\det(\dots,v_j,v_j,\dots)$$
$$=\det(\dots,v_i,\dots,v_j,\dots)+\det(\dots,v_j,\dots,v_i,\dots).$$

Quindi $\det(\dots,v_i,\dots,v_j,\dots)=-\det(\dots,v_j,\dots,v_i,\dots)$. Dimostriamo la i) nel caso particolare di due colonne vicine: vogliamo provare che $det(\dots,v,v,\dots)=0$, dove il vettore v è piazzato nella k-esima e nella $k+1$-esima colonna. Se l'ordine delle matrici è due la cosa è immediata (per $n=1$ la proposizione è vuota). Procedendo induttivamente, sia $A=(\dots,v,v,\dots)=(a_{ij})$; abbiamo $a_{ik}=a_{i(k+1)}$ per ogni i. Per definizione

$$\det(A)=\sum_{i=1}^{n}(-1)^{j+1}a_{1j}\det(A_{1j}).$$

Si ha $A_{1k} = A_{1(k+1)}$, mentre le A_{1j} con $j \neq k$ e $j \neq k+1$ hanno due colonne uguali. Per induzione i loro determinanti sono zero. Quindi:

$$\det(A) = (-1)^{k+1} a_{1k} \det(A_{1k}) + (-1)^{k+2} a_{1(k+1)} \det(A_{1(k+2)}$$
$$= (-1)^{k+1} \det(A_{1k})(a_{1k} - a_{1(k+1)} = 0 \quad (a_{1k} = a_{1(k+1)}).$$

Avendo provato la i), sappiamo che vale la ii) per colonne vicine. Per dimostrare la i) nel caso generale basta notare che possiamo avvicinare le colonne mediante scambi (cambiando segno):

$$\det(\dots, v, \dots, w, v, \dots) = -\det(\dots, v, \dots, v, w, \dots).$$

Iterando gli scambi: $\det(\dots, v, \dots, v, \dots) = \pm \det(\dots, v, v, \dots) = 0$. □

3. Caratterizzazione del determinante

Le proprietà di linearità e alternanza per colonne e la normalizzazione $\det(\mathrm{I}) = 1$ determinano il determinante.

Definizione 3.1. Una funzione $f : \mathcal{M}(n, K) \to K$ si dice multilineare per colonne se, per ogni i :

(1) $f((v_1, \dots, v_i + w_i, \dots, v_n)) = f((v_1, \dots, v_i, \dots, v_n)) + f(v_1, \dots, w_i, \dots, v_n))$;
(2) $f((v_1, \dots, av_i, \dots, v_n)) = af((v_1, \dots, v_i, \dots, v_n))$.

Si dice alternante se si annulla sulle matrici aventi due colonne eguali, cioè se:

$$f(\dots, v, \dots, v, \dots) = 0.$$

Osservazione. Se f è multilineare e alternante, allora, per ogni coppia di indici i, j con $i < j$, si ha:

$$f(\dots, v_i, \dots, v_j, \dots) = -f(\dots, v_j, \dots, v_i, \dots).$$

Teorema 3.2. (caratterizzazione del determinante) Esiste ed è unica una funzione $f : \mathcal{M}(n, K) \to K$ multilineare per colonne e alternante, tale che $f(\mathrm{I}) = a \in K$. Si ha infatti che $f(A) = a \det(A) = f(I) \det(A)$.

Osservazione. La proposizione dice che det è l'unica funzione multilineare e alternante tale che valga 1 sulla matrice identità.

Dimostrazione del Teorema di caratterizzazione. (1) Esistenza:

$$f(A) = a \det(A).$$

(2) Unicità: Sia:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1i} & \dots a_{1n} & \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{ji} & \dots a_{jn} & \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{ni} & \dots a_{nn} & \end{pmatrix} = (v_1, \dots, v_i, \dots, v_n),$$

dove $v_i =^t (a_{1i}, \dots, a_{ni})$.
Posto $v_1 = a_{11}e_1 + a_{21}e_2 + \dots a_{n1}e_n$, con e_i i-esimo vettore della base canonica di vettori colonna. Si ha, per linearità sulla prima colonna:

$$f(A) = f(a_{11}e_1 + a_{21}e_2 + \dots a_{n1}e_n, v_2, \dots, v_n) = \sum_{j=1}^{n} a_{j1} f(e_j, v_2, \dots, v_n).$$

Procedendo analogamente sulla seconda colonna, dobbiamo ora valutare i vari termini $f(e_j, v_2, \dots, v_n)$; poiché $v_2 = a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n$, si ottiene:

$$f(e_j, v_2, \dots, v_n) = \sum_k a_{k2} f(e_j, e_k, v_3, \dots, v_n).$$

Dunque:

$$f(A) = \sum_{j,k} a_{j1} a_{k2} f(e_j, e_k, v_3, \dots, v_n).$$

Posto $j = j(1)$ e $k = j(2)$, questo diventa:

$$f(A) = \sum a_{j(1),1} a_{j(2),2} f(e_{j(1)}, e_{j(2)}, v_3, \dots, v_n),$$

dove la somma è estesa a tutte le scelte di un valore per $j(1)$ e $j(2)$, cioè a tutte le funzioni $j : \{1, 2\} \to \{1, 2, \dots, n\}$. Data una tale j noi piazziamo il vettore $e_{j(1)}$ sulla prima colonna e il vettore $e_{j(2)}$ sulla seconda. Iterando:

$$f(A) = \sum_j A_j f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)}).$$

Allora la f è nota se sappiamo calcolarla su matrici aventi sulle colonne solo elementi della base canonica. Il coefficiente A_I, irrilevante per la dimostrazione, è facilmente calcolabile come prodotto di n elementi della matrice:

$$f(A) = \sum_j (a_{j(1)1} a_{j(2)2} \dots a_{j(n)n}) f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)}). \tag{4.1}$$

La somma in (4.1) è estesa a tutte le funzioni $j : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. Ad ogni modo $f(A)$ è nota se sappiamo calcolare i vari:

$$f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)}).$$

Ora usiamo l'alternanza: se due colonne sono eguali la f vale zero.
Le espressioni $f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)})$ non nulle sono quelle senza ripetizione di colonne, quindi quelle dove la funzione $j : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ è biettiva. Ripetiamo che, data $j : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$, la ricetta prescrive di piazzare il vettore $e_{j(k)}$ sulla colonna k-esima. Mediante una serie di scambi abbiamo:

$$f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)}) = \pm f(e_1, e_2, \dots, e_n).$$

Ora, il segno più o meno dipende soltanto dalla parità degli scambi che dobbiamo effettuare per ritrovare l'ordinamento naturale, e quindi solo da j :

$$f(e_{j(1)}, e_{j(2)}, e_{j(3)}, \dots, e_{j(n)}) = \sigma(j) f(e_1, e_2, \dots, e_n) = \sigma(j) f(I) = \sigma(j) a, \quad \sigma(j) = \pm 1.$$

<u>*Conclusione*</u>:

$$f(A) = a \sum_j \sigma(j) a_{j(1)1} a_{j(2)2} \dots a_{j(n)n},$$

dove j varia nelle funzioni biettive da $\{1, 2, \dots, n\}$ in sè; la f è univocamente determinata.

$\square$

Osservazione. Una funzione biettiva j sull'insieme dei primi n numeri è detta una permutazione. L'insieme di tali permutazioni forma un gruppo (S_n). Ogni permutazione è prodotto di scambi cioè di quelle permutazioni che scambiano due elementi lasciando fissi tutti gli altri. Il suo segno, $\sigma(j)$, vale 1 se j è prodotto di un numero pari di scambi e -1 se è prodotto di un numero dispari di scambi. Un modo per stabilire che la parità è ben definita è quello di osservare che:

$$\sigma(j) = \det(e_{j(1)}, e_{j(2)}, \ldots, e_{j(n)}),$$

dove il secondo membro dipende solo da j.

Corollario 3.3. **(formula del determinante)** Sia $A = (a_{ij})$; allora:

$$\det(A) = \sum_{j \in S_n} \sigma(j) a_{j(1)1} a_{j(2)2} \ldots a_{j(n)n}. \tag{4.2}$$

4. Teorema di Binet, teorema del rango e loro conseguenze

La formula (4.2) è usata in molti testi come definizione di determinante, e da essa si deducono le varie proprietà. A mio parere essa è meno trasparente della definizione induttiva e meno utile di quando si possa immaginare. Più importante comunque è sicuramente aver provato che la funzione determinante è caratterizzata dalle sue proprietà: la *multilinearità*, l'*alternanza* e la *normalizzazione*. Ora dedurremo da ciò alcune conseguenze di fondamentale importanza.

Teorema 4.1. **(teorema di Binet)** $\det(AB) = \det(A)\det(B) = \det(BA)$.

Dimostrazione. Fissiamo la matrice B e definiamo $f_B : \mathbb{M}(n, K) \to K$

$$f_B(A) = \det(BA).$$

Vediamo subito che f è multilineare e alternante per colonne: se $A = (v_1, \ldots, v_i, \ldots, v_n)$, allora $BA = (Bv_1, \ldots, Bv_i, \ldots, Bv_n)$. Si ha:

$$\det(Bv_1, \ldots, B(v_i + w_i), \ldots, Bv_n) = \det(Bv_1, \ldots, Bv_i, \ldots, Bv_n) + \det(Bv_1, \ldots, Bw_i, \ldots, Bv_n)$$

$$\det(Bv_1, \ldots, Bav_i, \ldots, Bv_n) = a \det(Bv_1, \ldots, Bv_i, \ldots, Bv_n),$$

ovvero:

$$f_B((v_1, \ldots, v_i + w_i, \ldots, v_n)) = f_B((v_1, \ldots, v_i, \ldots, v_n)) + f_B((v_1, \ldots, w_i, \ldots, v_n))$$

$$f_B((v_1, \ldots, av_i, \ldots, v_n)) = a f_B((v_1, \ldots, v_i, \ldots, v_n)).$$

D'altra parte $f_B((\ldots, v, \ldots, v, \ldots)) = \det(\ldots, Bv, \ldots, Bv, \ldots) = 0$. Quindi f_B è multilineare alternante. Allora $f_B(A) = f_B(I)\det(A)$, ma $f_B(I) = \det(IB) = \det(B)$, quindi $\det(BA) = \det(B)\det(A)$. □

Corollario 4.2. **(del teorema di Binet)**

(1) $\det(A^n) = \det(A)^n$;
(2) se A è invertibile, $\det(A^{-1}) = \det(A)^{-1}$;
(3) se A e C sono simili, cioè esiste B tale che $B^{-1}AB = C$, allora $\det(A) = \det(C)$.

Teorema 4.3. **(teorema del rango)** Sia $A \in \mathcal{M}(n, K)$; sono equivalenti:

(1) $\det(A) \neq 0$
(2) $r(A) = n$.

Dimostrazione.
- 1) $\Rightarrow$ 2). Se $r(A) < n$, allora una colonna di A è linearmente dipendente dalle altre. A meno di operare scambi di colonne sulla matrice, possiamo supporre che sia la prima, cioè $A = (v_1, \ldots, v_i, \ldots, v_n) = (a_2v_2 + \cdots + a_nv_n, \ldots, v_i, \ldots, v_n)$. Allora $\det(A) = \det(a_2v_2 + \cdots + a_nv_n, v_2, \ldots, v_i, \ldots, v_n) = \sum_{i>1} a_i \det(v_i, \ldots, v_n)$.

- 2) $\Rightarrow$ 1). Se $r(A) = n$, A è invertibile ed esiste A^{-1}; dal teorema di Binet abbiamo:
$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$
quindi $\det(A) \neq 0$.

$\square$

COROLLARIO 4.4. **(del teorema del rango)** Se A è quadrata, $Ax = b$ ha un'unica soluzione se e solo se $\det(A) \neq 0$. Il sistema omogeneo $Ax = \underline{0}$ ha soluzioni non banali se e solo se $\det(A) = 0$.

Osservazione. I determinanti sono definiti solo per matrici quadrate tuttavia sono utili per calcolare i ranghi delle matrici non quadrate. Sia M una matrice $n \times m$ di rango $r = r(M)$. Una sottomatrice N di M è una matrice ottenuta da M per eliminazione di alcune righe e di alcune colonne. Chiaramente $r(N) \leq r(M)$. Una sottomatrice N di M è un minore di M se N è quadrata. Se $r(M) = r$, allora esiste un minore N di ordine e rango r.
Procedura. Si eliminano $m - r$ colonne da M ottenendo una sottomatrice M' $n \times r$ di M avente rango r; si eliminano $n - r$ righe da M' per trovare N $r \times r$ di rango r (si ricordi che il rango per righe e per colonne è uguale). Per la N trovata allora $\det(N) \neq 0$. Quindi il rango di M è eguale all'ordine massimo dei minori aventi determinanti non nulli. Calcolando i determinanti dei minori si stabilisce il rango di ogni matrice.
ESERCIZIO: Calcolare il rango della matrice reale
$$M(a,b) = \begin{pmatrix} 1 & 2 & a & -a \\ 2 & 4+b & b & 0 \\ a & 2a & 0 & 0 \end{pmatrix},$$
al variare dei parametri reali a e b.

SOLUZIONE. Eliminiamo la seconda colonna: $\det \begin{pmatrix} 1 & a & -a \\ 2 & b & 0 \\ a & 0 & 0 \end{pmatrix} = a^2 b$. Segue che, se a e b sono diversi da zero, $r(M(a,b)) = 3$.
Supponiamo $a = 0$; la matrice diventa: $\begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 4+b & b & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, che ha rango strettamente minore di 3. Poiché $\det \begin{pmatrix} 1 & 0 \\ 2 & b \end{pmatrix} = b$, abbiamo che se $b \neq 0$ il rango è 2. Se $a = b = 0$, $M(0,0)$ ha rango $1:$ infatti tutti i minori di ordine 2 si annullano, e il rango è 1 perché $M(0,0)$ non è nulla.
<u>Risultato:</u> $r = 3$ se $a \neq 0$ e $b \neq 0$, $r = 2$ se $b \neq 0$ e $a = 0$ e $r = 1$ se $a = 0$ e $b = 0$. $\square$

ESERCIZIO: Sia $A = A(a,b) = \begin{pmatrix} 1 & 2 & a \\ 2 & 4+b & b \\ a & 2a & 0 \end{pmatrix}$; per quali valori di a e b reali il sistema $Ax = v$ con $v = \begin{pmatrix} -a \\ 0 \\ 0 \end{pmatrix}$ ammette soluzioni?

Osservazione. Indichiamo con $GL(n,K)$ il gruppo (moltiplicativo) delle matrici invertibili; abbiamo visto che $GL(n,K) = \{A \in \mathcal{M}(n,K) : \det(A) \neq 0\}$. Tale gruppo è comunemente denominato come il *gruppo generale lineare*. Il teorema di Binet dice che la restrizione della funzione determinante
$$\det : GL(n,K)) \to K^\star := K - \{0\}$$

è un omomorfismo di gruppi. Il nucleo di tale applicazione è il sottogruppo $SL(n,K) := \{A : A \in \mathcal{M}(n,K) : \det(A) = 1\}$, detto il *gruppo speciale lineare*.

5. Regole di Laplace

La nostra definizione di determinante è stata di fatto una regola di Laplace. Veniamo ora a vedere che l' aver usato la prima riga non è cosìessenziale:

Teorema 5.1. **(I Regola di Laplace - per righe)** Sia $A = (a_{ij})$; allora $\det(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) = a_{i1}\det(A_{i1}) - a_{i2}\det(A_{i2}) + \cdots + (-1)^{i+n}a_{in}\det(A_{in})$.

Dimostrazione. Si consideri la funzione $\phi : \mathcal{M}(n,K) \to K$, definita come $\phi(A) = \sum_{j=1}^{n} a_{ij}\det(A_{ij})$.
Con la medesima procedura del caso $i = 1$ si ha che ϕ è multilineare e alternante per colonne. Ora $\phi(I_n) = (-1)^{2i}\det(I_{n-1}) = 1$, quindi $\phi(A) = \det(A)$. □

Vogliamo ora dimostrare che procedere per righe o per colonne è irrilevante:

Proposizione 5.2. $\det(A) = \det({}^tA)$.

Dimostrazione. Per gli ordini 1 e 2 la cosa è ovvia. Supponiamo $n > 2$ e definiamo $F : \mathcal{M}(n,K) \to K$:

$$F(A) := \det({}^tA).$$

Chiaramente $F(I) = 1$. La proposizione sarà dimostrata se possiamo provare che F è multilineare e alternante (per colonne). Ora, se $A = (v_1, \dots, v_i, \dots, v_n)$, ${}^tA = \begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tv_i \\ \dots \\ {}^tv_n \end{pmatrix}$. Dobbiamo verificare che:

$$F((v_1, \dots, v_i + w_i, \dots, v_n)) = F((v_1, \dots, v_i, \dots, v_n)) + F((v_1, \dots, w_i, \dots, v_n))$$

$$F((v_1, \dots, av_i, \dots, v_n)) = aF((v_1, \dots, v_i, \dots, v_n)),$$

cioè che:

$$\det\begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tv_i + {}^tw_i \\ \dots \\ {}^tv_n \end{pmatrix} = \det\begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tv_i \\ \dots \\ {}^tv_n \end{pmatrix} + \det\begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tw_i \\ \dots \\ {}^tv_n \end{pmatrix}; \quad \det\begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tav_i \\ \dots \\ {}^tv_n \end{pmatrix} = a\det\begin{pmatrix} {}^tv_1 \\ \dots \\ {}^tv_i \\ \dots \\ {}^tv_n \end{pmatrix},$$

e che:

$$F\big((\dots, v, \dots, v, \dots)\big) = \det\begin{pmatrix} \dots \\ {}^tv \\ \dots \\ {}^tv \\ \dots \end{pmatrix} = 0.$$

Si tratta cioè di verificare che il determinante è multilineare e alternante per righe. Cominciamo con la multilinearità. Dalla formula:

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j})$$

segue subito, per induzione, la multilinearità per tutte le righe diverse dalla prima. Per la prima definiamo:

$$g(x_1, \dots, x_n) = \det \begin{pmatrix} x_1 & \dots & x_i & \dots & x_n \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix} = \sum_{j=1}^{n} (-1)^{j+1} x_j \det(A_{1j}).$$

La g è lineare essendo un polinomio omogeneo di primo grado nelle x_i. Dobbiamo ora verificare l'alternanza per righe: supponiamo che A abbia due righe uguali. Se nessuna di tali righe è la prima riga di A, allora le A_{1j} hanno due righe uguali e quindi i $\det(A_{1j})$ sono tutti 0 per induzione:

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}) = 0.$$

Però $n > 2$, e utilizzando la prima di regola di Laplace, possiamo sviluppare il determinante per una riga distinta dalle due righe in questione. Il precedente ragionamento prova che $\det(A) = 0$. □

Abbiamo la possibilità di sviluppare il determinante per colonne. Vale allora la seguente:

I Regola di Laplace - per colonne:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Veniamo ora alla seconda regola di Laplace; consideriamo la matrice

$$A = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ a_{i1} & \dots & a_{ik} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{jk} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Costruiamo ora B sostituendo alla i-esima riga di A la j-esima, lasciando tutto il resto inalterato.

$$B = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{jk} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \\ a_{j1} & \dots & a_{jk} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Avendo B due righe eguali, il suo determinante si annulla.
Se sviluppiamo il determinante di B secondo la i-esima riga, abbiamo:

$$0 = \sum_{k=1}^{n} (-1)^{i+k} a_{jk} \det(B_{ik}),$$

ma $B_{ik} = A_{ik}$. Quindi vale la seguente:

II Regola di Laplace:

$$\sum_{k=1}^{n}(-1)^{i+k}a_{jk}\det(A_{ik}) = 0 \quad \text{se} \quad i \neq j.$$

Lasciando al lettore il compito di riscrivere le regole di Laplace per colonne, ricapitoliamo:
Regole di Laplace:

I. $\sum_{k=1}^{n}(-1)i + ka_{ik}\det(A_{ik}) = \det(A)$;

II. $\sum_{k=1}^{n}(-1)i + ka_{jk}\det(A_{ik}) = 0$ se $i \neq j$.

Possiamo riorganizzare tali relazioni come segue: si costruisca la matrice $L = (\lambda_{ij})$, ove $\lambda_{ij} = (-1)^{i+j}\det(A_{ij})$.
Tutte le regole di Laplace si riscrivono come:

$$LA = AL = \det(A)\mathrm{I}.$$

Come conseguenza vediamo ancora che se $\det(A) \neq 0$ allora A è invertibile, infatti $A^{-1} = \frac{1}{\det(A)}L$. Ovvero $A^{-1} = (c_{ij})$, dove $c_{ij} = \frac{(-1)^{i+j}\det(A_{ij})}{\det(A)}$.

Osservazione. Tale formula è molto importante teoricamente. Per esempio implica che i coefficienti c_{ij} della matrice inversa sono funzioni razionali (quozienti di polinomi) nei coefficienti della matrice da invertire. È invece meno utile nelle applicazioni numeriche. Per usare la regola di Laplace bisogna calcolare n^2 determinanti. Ogni determinante necessita di una somma di $(n-1)!$ termini, ognuno dei quali con $n-1$ prodotti. Per calcolare l'inversa è più conveniente, dal punto di vista del costo macchina, risolvere gli n sistemi lineari che vengono dall'equazione $AX =$I mediante il metodo di eliminazione. Per convincersi di questo basta invertire manualmente con i due metodi la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

CAPITOLO 5

Operatori

Incominciamo uno studio dettagliato degli operatori. I più semplici da trattare sono sicuramente i multipli dell'identità, gli operatori scalari. La loro azione sullo spazio vettoriale è la moltiplicazione per uno scalare; nel caso geometrico sono le dilatazioni. Per capire il funzionamento di un operatore $f : V \to V$ è fondamentale determinare quelle direzioni (sottospazi) su cui l'azione di f è scalare, cioè è la moltiplicazione per uno scalare.

1. Autovalori e autovettori

Definizione 1.1. (autovalore) Siano V un K spazio vettoriale e $f : V \to V$ un operatore K lineare. Uno scalare $\lambda \in K$ si dice autovalore di f se esiste un vettore $v \in V$, $v \neq \underline{0}$, tale che:

$$f(v) = \lambda v.$$

Osservazione. Il fatto che v sia non nullo è cruciale, altrimenti ogni scalare sarebbe autovalore e la definizione precedente sarebbe inutile.

Definizione 1.2. (autovettore) Siano $f : V \to V$ un operatore e $\lambda \in K$ un autovalore di f. Un vettore $v \in V$ si dice autovettore di f associato a λ (λ-autovettore) se $f(v) = \lambda v$. Il sottospazio $V(\lambda) := \{v \in V : f(v) = \lambda v\} = \ker(f - \lambda \mathrm{id})$ è l'autospazio di f relativo a λ.

Nel caso di matrici quadrate $n \times n$ possiamo ripetere le definizioni:

Definizione 1.3. Uno scalare $\lambda \in K$ si dice autovalore di $A \in \mathcal{M}(n, K)$ se esiste $v \in K^n$, $v \neq \underline{0}$, tale $Av = \lambda v$ (v come colonna). Un elemento v di K^n è un λ-autovettore di A se $Av = \lambda v$, e $V(\lambda) := \{v \in K^n : Av = \lambda v\}$ è il λ-autospazio di A.

Osservazione. Dovrebbe essere chiaro che $v \in K^n$ è un λ-autovettore di $A \in \mathcal{M}(n, K)$ se e solo se v è un λ-autovettore dell'operatore $F_A : K^n \to K^n$ associato ad A. In particolare due matrici simili hanno gli stessi autovalori. La possibilità di passare dal linguaggio delle matrici a quello degli operatori è di grande utilità.

2. Calcolo di autovalori: il polinomio caratteristico

Veniamo al calcolo degli autovalori di una matrice A di ordine n. Dobbiamo stabilire gli scalari x per cui vi sono soluzioni non nulle di:

$$Av - xv = (A - xI)v = \underline{0}. \tag{5.1}$$

Il sistema lineare omogeneo (5.1) dipende dal parametro x ed ha soluzioni non banali se e solo se il rango di $A - xI$ non è n. Questo avviene se e solo se:

$$p(x) = \det(A - xI) = 0. \tag{5.2}$$

La (5.2) è l'*equazione caratteristica*. Analizziamone la struttura; se $A = (a_{ij})$:

$$p(x) = \det \begin{pmatrix} (a_{11}-x) & \dots & a_{1i} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & \dots & (a_{ii}-x) & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1}-x & \dots & a_{ni} & \dots & (a_{1n}-x) \end{pmatrix}.$$

Posto $A(x) = A - xI$, sviluppando il determinante come

$$p(x) = (a_{11}-x)\det(A_{11}(x)) + \sum_{i=2}^{n} (-1)^{i+1} a_{1i} \det(A_{1i}(x)), \tag{5.3}$$

si ottiene induttivamente che $p(x)$ è un polinomio di grado n in x, il polinomio caratteristico di A. Scriviamo:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0.$$

Si ha $a_0 = p(0) = \det(A - 0I) = \det(A)$. Nello sviluppo (5.3) le matrici $A_{1i}(x)$ con $i > 1$ hanno solo $n-2$ colonne contenenti l'incognita x :

$$A_{1i}(x) = \begin{pmatrix} a_{21} & (a_{22}-x) & \dots & a_{2(i-1)} & a_{2(i+1)} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{(i-1)1} & a_{(i-1)2} & \dots & (a_{(i-1)(i-1)}-x) & a_{(i-1)(i+1)} & \dots & a_{(i-1)n} \\ a_{i1} & a_{i2} & \dots & a_{i(i-1)} & a_{i(i+1)} & \dots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & \dots & a_{(i+1)(i-1)} & (a_{(i+1)(i+1)}-x) & \dots & a_{i+1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n(i-1)} & a_{n(i+1)} & \dots & (a_{nn}-x) \end{pmatrix}$$

(nell'esempio $2 < i < n$). Allora i loro determinanti sono polinomi nella x di grado al più $n-2$, quindi:

$$p(x) = (a_{11}-x)\det(A_{11}(x)) + q(x), \quad \deg(q) < n-1.$$

Lavorando nello stesso modo sul determinante di $A_{11}(x) = A_{11} - xI_{n-1}$ abbiamo:

$$p(x) = (a_{11}-x)(a_{22}-x)\det(B_2(x)) + s(x).$$

Il grado di $s(x)$ è minore di $n-1$ e $B_2(x)$ è ottenuta da A eliminando le prime due righe e le prime due colonne. Iterando:

$$p(x) = (a_{11}-x)(a_{22}-x)\dots(a_{nn}-x) + r(x), \quad \deg(r) < n-1.$$

I coefficienti di p di grado n e $n-1$ sono: $a_n = (-1)^n$, $a_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$.

Definizione 2.1. **(traccia)** La somma degli elementi della diagonale principale si chiama traccia di A; si pone: $\mathrm{tr}(A) = a_{11} + a_{22} + \dots + a_{nn}$.

Ricapitolando, abbiamo visto:

Proposizione 2.2. Gli scalari $\lambda \in K$ sono autovalori di A se e solo se sono radici del polinomio caratteristico $p(x) = \det(A - xI)$. Tale polinomio ha grado n, e si ha:

$$\begin{aligned} p(x) &= (-1)^n x^n + (-1)^{n-1}\mathrm{tr}(A)x^{n-1} + \dots + \det(A) \\ &= (-1)^n (x^n - \mathrm{tr}(A)x^{n-1} + \dots + (-1)^n \det(A)). \end{aligned}$$

Osservazione. Se $n = 2$, $p(x) = x^2 - \mathrm{tr}(A)x + \det(A)$.

Il calcolo degli autovalori si riconduce alla determinazione delle radici del polinomio caratteristico, un problema non lineare, anche dal punto di vista computazionale, più insidioso di quanto si possa credere. Il calcolo di autospazi equivale alla risoluzioni di sistemi lineari omogenei. La seguente proposizione mostra la natura operatoriale del polinomio caratteristico:

Proposizione 2.3. Due matrici simili hanno lo stesso polinomio caratteristico.

Dimostrazione. Siano A e C due matrici aventi polinomi caratteristici rispettivamente $p_A(x) = \det(A - xI)$ e $p_C(x) = \det(C - xI)$. Se A e C sono simili, allora esiste B tale che $A = B^{-1}CB$, dunque $p_A(x) = \det(B^{-1}CB - xI) = \det(B^{-1}(C - xI)B) = \det(B^{-1})\det(C - xI)\det(B) = \det(B^{-1})\det(B)\det(C - xI) = \det(C - xI) = p_C(x)$.

(A voler essere pignoli dovremmo aggiungere che si sono calcolati i determinanti in $K(x)$, il campo delle funzioni razionali di K). □

Esercizio: Calcolare autovalori e autovettori di:

$$\begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}; \quad \begin{pmatrix} 2 & 1 & 0 & 4 \\ -1 & 10 & 0 & -2 \\ 0 & 6 & 1 & 0 \\ 2 & 1 & 0 & 4 \end{pmatrix}.$$

Esercizio: Provare che la funzione traccia tr: $\mathcal{M}(n, K) \to K$ è lineare, che tr(A) =tr$({}^tA)$, e che per ogni A e B si ha $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
Esercizi: Indichiamo con $p_A(x)$ il polinomio caratteristico di una matrice A.

(1) Calcolare $p_I(x)$.
(2) Calcolare il polinomio caratteristico e gli autovalori di una matrice triangolare.
(3) Quando 0 è autovalore di A?
(4) Qual è il legame tra $p_A(x)$ e $p_{A+tI}(x)$?

Esercizio: Dimostrare che $p_{AB}(x) = p_{BA}(x)$.

Soluzione. Se B (o A) è invertibile allora $B(AB)B^{-1} = BA$, quindi AB è simile a BA. Se però A e B non sono invertibili non è detto che AB sia simile a BA.
Se per esempio $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ e $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, allora $AB = A$ e BA =0, che non simili perché hanno ranghi diversi. Si noti nell'esempio che i polinomi caratteristici sono eguali.
Si può utilizzare un metodo perturbativo: l'idea intuitiva è quella di deformare B e renderla invertibile. Si considerino le matrici $B(y) = B - yI$, $y \in K$, invertibili per ogni $y \in K$ che non sia autovalore di B. Se assumiamo K infinito (altrimenti immergiamo K in un campo infinito, per esempio il campo delle funzioni razionali), tale condizione è verificata per infiniti valori. Quando $B(y)$ è invertibile, $p_{AB(y)}(x) = p_{B(y)A}(x)$, ovvero $\det(AB - yA - xI) = \det(BA - yA - xI)$. Sia $x_0 \in K$, e si pongano $P(y) = \det(AB - yA - x_0I)$ e $Q(y) = \det(BA - yA - x_0I)$. Ora Q e P sono polinomi, e $Q(y) = P(y)$ per infiniti valori di y : allora $Q = P$, cioè $\det(AB - yA - x_0I) = \det(BA - yA - x_0I)$ per ogni y. Posto $y = 0$, si ha:

$$\det(AB - x_0I) = \det(BA - x_0I).$$

La condizione vale per ogni x_0, quindi $p_{AB}(x) = p_{BA}(x)$.
Notate che, se $K = \mathcal{R}$, dalle identità $\det(AB - yA - xI) = \det(BA - yA - xI)$ per $y \neq 0$, in un intorno dello zero si ottiene:

$$p_{AB}(x) = \lim_{y\to 0} \det(AB - yA - xI) = lim_{y\to 0} \det(BA - yA - xI) = p_{BA}(x)$$

per continuità della funzione determinante. □

Definizione 2.4. **(spettro)** Il sottoinsieme di K

$$\sigma(f) := \{\lambda \in K \ : \ f - \lambda id \text{ non è invertibile}\}$$

si chiama spettro di f.

Osservazione. Si noti che λ è un autovalore se e solo se $\ker(f-\lambda id)\neq\{\underline{0}\}$, cioè se e solo se $f-\lambda id$ non è iniettiva. Nel caso di spazi vettoriali di dimensione finita questo equivale a dire che $f-\lambda id$ non è biettiva, dunque

$$\sigma(f)=\{\text{autovalori di } f\}.$$

Esempi:

(1) Si calcoli lo spettro di matrici triangolari.
(2) Si considerino gli spazi $\mathcal{P}$ dei polinomi reali e $\mathcal{C}^1(\mathbb{R})$. Sia F l'operatore derivata: $F(g)=g'$. Nel caso di $\mathcal{P}$ l'unico autovalore è 0, e $V_0=\{\text{polinomi costanti}\}$. Qual è lo spettro? Nel caso di $\mathcal{C}^1(\mathbb{R})$, ogni λ è autovalore.

Non sempre lo spettro non sempre corrisponde all'insieme degli autovalori. Infatti:

3. Si consideri Rat, lo spazio vettoriale reale delle funzioni razionali reali. Posto ancora $F(g)=g'$, l'unico autovalore è lo zero, ma lo spettro è $\mathbb{R}$. Infatti $F-\lambda id$ non è suriettiva. L'equazione differenziale:
$$g'-\lambda g=\frac{1}{x}$$
non ha soluzione nelle funzioni razionali (perché ?).

Nelle applicazioni matematiche e fisiche gli spazi di dimensione infinita sono molto importanti. È bene, anche per questo, saper lavorare in dimensione finita alla perfezione.

Supponiamo che tutte le radici del polinomio caratteristico di A, $p(x)$, siano nel campo K. Questo è automatico quando K è algebricamente chiuso (e.g. $\mathbb{C}$), cioè se contiene tutte le radici dei polinomi a coefficienti in K. Abbiamo:

$$p(x)=(-1)^n(x-\lambda_1)(x-\lambda_2)\cdot\dots\cdot(x-\lambda_n)=(-1)^n(x^n-\operatorname{tr}(A)x^{n-1}+\dots+(-1)^n\det(A)).$$

Il polinomio $(x-\lambda_1)(x-\lambda_2)\cdot\dots\cdot(x-\lambda_n)$ è monico, cioè il coefficiente di grado massimo è 1. I coefficienti di:

$$(x-\lambda_1)(x-\lambda_2)\dots(x-\lambda_n)=x^n-b_1x^{n-1}+\dots+(-1)^kb_kx^{n-k}+\dots+(-1)^nb_n,$$

dove

$$b_k=\sum_{i_1<\dots<i_k}\lambda_{i_1}\lambda_{i_k}.$$

La traccia e il determinante di A sono rispettivamente la somma e il prodotto dei suoi autovalori: $\mathbf{tr}(A)=b_1=\lambda_1+\dots+\lambda_n$, $\det(A)=b_n=\lambda_1\cdot\dots\cdot\lambda_n$.

3. Diagonalizzazione

Un operatore scalare $f=\lambda id$ commuta con tutti gli operatori, ha un unico autovalore, λ, ogni vettore è autovettore, e, nel caso finito dimensionale, è rappresentato dalla matrice (scalare) λI.
Anche l' azione degli operatori associati alle matrici diagonali è chiara: si moltiplicano (dilatano) le coordinate dei vettori per scalari in generale diversi. È naturale chiedersi quali operatori abbiano questa struttura.

Definizione 3.1. **(matrice diagonalizzabile)** Una matrice A si dice diagonalizzabile se è simile ad una matrice diagonale, cioè se $D=B^{-1}AB$ con D diagonale.

E' immediato verificare la seguente:

Proposizione 3.2. Una matrice A di ordine n è diagonalizzabile se e solo se esiste una base di K^n composta di autovettori di A.

Dimostrazione. Indichiamo con $F_A : K^n \to K^n$ l'operatore associato ad A, cioè $F_A(x) = Ax$, con $x \in K^n$ vettore colonna. Supponiamo che esista una base $\{v_1, v_2, \ldots, v_n\}$ di autovettori di A. Allora $F_A(v_i) = \lambda_i v_i$, e in tale base l'operatore si rappresenta con la matrice:

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ldots & 0 & \ldots & \ldots \\ \ldots & 0 & \lambda_i & 0 & 0 \\ \ldots & \ldots & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

dove D è diagonale. La matrice del cambiamento di base è $B = (v_1, \ldots, v_n)$, e $D = B^{-1}AB$. Sia ora $D = B^{-1}AB$, con D diagonale. Siano v_i, w_i ed e_i rispettivamente l' i-esimo vettore colonna di B, di D e di I. Essendo $BD = AB$ si ha $Av_i = Bw_i$, e $w_i = \lambda_i e_i$ (D è diagonale); dunque $Av_i = B\lambda_i e_i = \lambda_i v_i$. Le colonne di B sono autovettori di A e formano, essendo B invertibile, una base di K^n. □

Definizione 3.3. **(molteplicità geometrica e algebrica)** Siano A una matrice, λ un autovalore di A e $V(\lambda) = \ker(A - \lambda I)$ l'autospazio associato a λ. L'intero:

$$g(\lambda) = \dim(V(l))$$

si dice molteplicità geometrica dell'autovalore λ. La molteplicità algebrica, $a(\lambda)$, di λ è invece la molteplicità di λ come radice di $p(t)$:

$$p(t) = (t - \lambda)^{a(\lambda)} q(t), \quad q(\lambda) \neq 0.$$

Osservazione. Sia K un sottocampo di L, $K \subset L$, (e.g. $\mathbb{R} \subset \mathbb{C}$), e sia $A \in \mathcal{M}(n, K)$. Possiamo pensare ad $A \in \mathcal{M}(n, L)$ definita su L. Gli autovalori di A in L contengono (ma in generale non coincidono con) quelli contenuti in K. Se $\lambda \in K$, la molteplicità algebrica e quella geometrica di λ sono eguali se calcolate in K o in L. Per la molteplicità algebrica la cosa è evidente. Per quella geometrica osserviamo che la dimensione degli autospazi $V(\lambda) \subset K^n$ e $V'(\lambda) \subset L^n$ è $n - r(A - \lambda I)$, indipendente dal campo. Il rango è infatti regolato dall'annularsi di determinanti dei minori, ma il determinante di una matrice è uno scalare che appartiene al campo su cui sono definiti i coefficienti della matrice.

Il seguente lemma sarà utilizzato spesso.

Lemma 3.4. Siano A, B e D matrici quadrate con:

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

dove $\mathbf{0}$ è la matrice nulla; allora:

(1) $\det(A) = \det(B)\det(D)$;
(2) se p_A, p_B e p_D sono i polinomi caratteristici di A, B e D allora $p_A = p_B p_D$.

Dimostrazione. Si applichi il teorema di Binet al prodotto

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \begin{pmatrix} I & C \\ 0 & D \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix}.$$

Ci si riduce ad una elementare applicazione della regola di Laplace. L' affermazione sui polinomi caratteristici segue immediatamente. □

La seguente proposizione permette di stabilire concretamente quando una matrice è diagonalizzabile.

Proposizione 3.5. (1) $1 \leq g(\lambda) \leq a(\lambda) \leq n$.

(2) $A \in \mathcal{M}(n, K)$ è diagonalizzabile se e solo se:
 (a) $p(x) = (-1)^n (x - \lambda_1)^{a(\lambda_1)} \dots (x - \lambda_k)^{a(\lambda_k)}$ (tutti gli autovalori di A sono in K);
 (b) $g(\lambda) = a(\lambda)$ per ogni autovalore λ.

DIMOSTRAZIONE. (1) L'unica disuguaglianza da analizzare è quella $g(\lambda) \leq a(\lambda)$. Posto $g(\lambda) = g$, $a(\lambda) = a$, costruiamo una base di $V : \{v_i, w_j\}$, $0 < i \leq g$, $0 \leq j \leq n - g$, dove $\{v_i\}$ è una base di $V(\lambda)$. La matrice A associata ad f è del tipo $A = \begin{pmatrix} \lambda I_g & B \\ 0 & D \end{pmatrix}$, dove D è quadrata di ordine $n - g$, $A = \lambda I$ se $n = g$. Il polinomio caratteristico è:

$$p(t) = (\lambda - t)^g q(t).$$

Si è usata ricorsivamente la regola di Laplace ovvero il precedente lemma, $q(t)$ è il polinomio caratteristico di D. Ricaviamo allora $g \leq a$.

(2) Noi dimostreremo che **gli autotospazi sono in somma diretta**. Siano $p(t)$ il polinomio caratteristico di A, λ_i, $1 \leq i \leq k$, le sue radici distinte e $V(\lambda_i)$, $1 \leq i \leq k$, gli autospazi di A. Dobbiamo verificare induttivamente che:

$$V(\lambda_i) \cap \{\oplus_{j<i} V(\lambda_j)\} = \{\underline{0}\}.$$

Sia $v \in V(\lambda_i) \cap \{\oplus_{j<i} V(\lambda_j)\}$, $v = \sum_{j<i} v_j$, $v \in V(\lambda_i)$ e $v_j \in V(\lambda_j)$. Quindi $Av = \lambda_i v = \sum_{j<i} \lambda_j v_j$. Si ha $\sum_{j<i} \lambda_i v_j = \sum_{j<i} \lambda_j v_j$ e quindi $\sum_{j<i} \lambda_i v_j = \sum_{j<i} \lambda_j v_j$, ovvero $\sum_{j<i} (\lambda_j - \lambda_j) v_j = 0$. Essendo $\lambda_i \neq \lambda_j$, l'ipotesi di induzione implica $v_j = \underline{0}$: se non fosse così, sia k l'indice più grande per cui $v_k \neq \underline{0}$. Allora

$$v_k = -\sum_{j<k} \frac{\lambda_j - \lambda_i}{\lambda_k - \lambda_i} v_j \in V(\lambda_k) \cap \{\oplus_{j<k} V(\lambda_j)\} = \{\underline{0}\},$$

una contraddizione; allora $v = \underline{0}$. Gli autospazi sono allora in somma diretta; da questo e dalla prima parte della proposizione deduciamo:

$$\dim(\oplus V(\lambda_j)) = \sum_i g_i \leq \sum_i a_i \leq n \ (= \text{ grado di } p(t)).$$

Ma allora $\oplus V(\lambda_j) = K^n$ se e solo se tutte le disuguaglianze precedenti sono eguaglianze. Questo è possibile se e solo se entrambe le condizioni:
 a) $\sum_i a_i = n$, cioè $p(t)$ ha n radici in K,
 b) $g_i = a_i$ per ogni i
sono soddisfatte.
Ora, A ammette una base di autovettori se solo se K^n è somma dei suoi autospazi. Quindi A è diagonalizzabile se e solo se $\oplus V(\lambda_j) = K^n$. Noi abbiamo visto che questo equivale all'insieme delle due condizioni a) e b) .

□

COROLLARIO 3.6. Se il polinomio caratteristico di una matrice A di ordine n ha n radici distinte nel campo K allora A è diagonalizzabile.

ESERCIZIO: Si consideri la seguente matrice reale dipendente da un parametro reale a.

$$X(a) = \begin{pmatrix} 0 & 0 & 3a \\ 1 & 0 & 1 \\ 0 & a & 0 \end{pmatrix}.$$

Dire per quali valori di a la matrice $X(a)$ è diagonalizzabile su $\mathbb{R}$ e per quali su $\mathbb{C}$.

Soluzione. $P_{X(a)}(t) = -(t^3 - at - 3a^2)$. Analizziamo i polinomi del tipo $y(t) = t^3 - bt + c$. Vogliamo sapere per quali valori di b e c si hanno 3 radici reali. Essendo in grado dispari abbiamo sempre una soluzione reale. Per avere tre soluzioni reali $y(t)$ deve avere un punto di minimo ≥ 0 e uno di massimo ≤ 0. Quindi $y' = 3t^2 - b$ deve avere radici reali; allora $b \geq 0$ e le radici sono $\pm w$, dove $w = \sqrt{\frac{b}{3}}$. Dobbiamo avere $y(w) \leq 0 \leq y(-w) : w^3 - bw + c \leq 0 \leq -(w^3 - bw) + c$, da cui $|c| \leq |w^3 + bw| = 2\sqrt{\frac{b^3}{27}}$. Segue che $27c^2 \leq 4b^3$.
Se λ è radice di $y(t)$ con molteplicità 2 deve essere reale, e $y'(\lambda) = 0$. Quindi $\lambda = w$ o $\lambda = -w$, il che avviene se e solo se $27c^2 = 4b^3$.
Spiegazione: due radici complesse coniugate hanno lo stesso modulo, se diventano reali devono diventare eguali. Nel nostro esempio: 3 soluzioni reali se e solo se $3^5a^4 \leq 4a^3$ $0 \leq a \leq (4/3)^5$.
Si analizzi la molteplicità geometrica nel caso delle eguaglianze. □

4. Il teorema di Cayley-Hamilton

Sia V un K spazio vettoriale che supporremo, salvo avviso contrario, di dimensione finita e non nullo: $0 < \dim(V) = n < +\infty$. Una base $\{v_i\}_{1\leq i\leq n}$ di V definisce un isomorfismo da V in K^n e associa ad ogni operatore $f : V \to V$ una matrice A quadrata di ordine n a coefficienti in K, $A \in \mathcal{M}(n, K)$. Il polinomio caratteristico $p(t) = p_A(t)$ di f (e di A) è definito da:

$$p(t) = \det(A - tI),$$

e $p(\lambda) = 0$, $\lambda \in K$, se e solo se λ è un autovalore di f. La definizione è ben data perché due matrici simili hanno lo stesso polinomio caratteristico.
C'è un modo naturale di calcolare i polinomi sugli operatori e sulle matrici quadrate. Per fare questo, sia $q(t) = \sum_i a_i t^i, a_i \in K$, un polinomio a coefficienti in K. Posto $f^0 = id$, indichiamo con $q(f) : V \to V$ l'operatore di V :

$$q(f) = a_0 f^0 + a_1 f^1 + \cdots + a_i f^i + \cdots + a_m f^m.$$

La matrice associata a $q(f)$ rispetto alla base $\{v_i\}_{1\leq i\leq n}$ è allora:

$$q(A) = a_0 I + a_1 A + \cdots + a_i A^i + \cdots + a_m A^m.$$

Esercizi:

(1) Se $p(t) = q(t)s(t)$, allora $p(A) = q(A)s(A) = s(A)q(A)$, e $(p + q)(A) = p(A) + q(A)$.
(2) Se A è simile a B, allora $p(A)$ è simile a $p(B)$.
(3) Se λ è un autovalore di A, allora $p(\lambda)$ è un autovalore di $p(A)$.
(4) Se v è un autovettore di A, allora lo è anche di $p(A)$.
(5) È vero il viceversa di 4? È vero il viceversa di 4 se A è diagonale?

Se $q(A) = \mathbf{0}$ (o $q(f) = 0$) diremo che q annulla A (risp. f). Tale condizione può essere riletta come una combinazione lineare tra le potenze di A (di f):

$$a_0 I + a_1 A + \cdots + a_i A^i + \cdots + a_m A^m = \mathbf{0}.$$

Viceversa, se una tale combinazione sussiste, allora $q(t) = \sum_i a_i t^i$ annulla A. Essendo lo spazio delle matrici di dimensione finita è chiaro che devono esserci molte combinazioni lineari nulle tra le potenze di A. Il teorema di Cayley-Hamilton dice che A è annullata dal suo polinomio caratteristico.

Teorema 4.1. **(di Cayley-Hamilton)** Se $p(t)$ è il polinomio caratteristico di f (risp. di A), allora $p(f) = 0$ (risp. $p(A) = \mathbf{0}$).

Dimostrazione. Sia v un vettore non nullo di V, vogliamo mostrare che $p(f)(v) = 0$, cioè che $p(f)$ è l'operatore nullo. Indichiamo con $w_k = f^{k-1}(v),\ k > 0,\ w_1 = f^0(v) = v$. Essendo $\dim V = n < \infty$, esiste m, $0 < m < n$, tale che i vettori

$$\dots, w_i, \dots, w_m, w_{m+1}$$

sono linearmente dipendenti mentre $\dots w_i, \dots, w_m$ sono ancora linearmente indipendenti. Notiamo che $m = 1$ se e solo se v è un autovalore di f. Questo numero m misura di quanto speciale è v rispetto ad f. In ogni caso esistono degli scalari a_i non tutti nulli tali che:

$$f^m(v) = w_{m+1} = \sum_{i=0}^{m-1} a_i f^i(v) = \sum_{i=1}^{m} a_i w_i.$$

Posto $q(t) = t^m - \ \sum_{i=1}^m a_i t^i$ riscriviamo la precedente relazione:

$$q(f)(v) = 0.$$

Lo scopo sarà quello di dimostrare che $q(t)$ divide il polinomio caratteristico $p(t)$: se $p(t) = s(t)q(t)$ allora

$$p(f)(v) = s(f)(q(f)(v)) = s(f)(0) = 0.$$

Completiamo i $\{w_i\}$ ad una base $\{\dots, w_i, \dots, u_j\}$ $(1 \leq i \leq m,\ 0 \leq j \leq n-m)$, di V. Ora $f(w_i) = w_{i+1}$ e $w_{m+1} = f(w_m) = \sum_{i \leq m} a_i w_i$. La matrice A' associata ha forma:

$$A' = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}, \qquad \text{dove } B = \begin{pmatrix} 0 & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ 0 & 1 & \dots & 0 & a_3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_m \end{pmatrix},$$

con $m > 3$ solo per motivi grafici. Per il solito lemma (5.3.2) $p(t)$ è diviso dal polinomio caratteristico di B. Quest'ultimo è $(-1)^m q(t)$: si sviluppi $\det(B - tI_m)$ rispetto all'ultima colonna. Quindi $q(t)$ divide $p(t)$. $\square$

Il teorema ha conseguenze interessanti.

(1) Permette di dare una formula della matrice inversa con le potenze di A; se $\det(A) \neq 0$:

$$A^{-1} = -\det(A)^{-1}(a_1 I + \dots + a_i A^{i-1} + \dots + (-1)^n A^{n-1});$$

(2) Una matrice A quadrata si dice *nihilpotente* se $A^m = 0$ per qualche $m > 0$. Si ha: A nihilpotente di ordine $n \iff A$ ammette solo l'autovalore $0 \iff A^n = 0$.

Dimostrazione. Se $A^q = 0$ per qualche q e λ è autovalore di A, con v un $\lambda-$autovettore non nullo: $\underline{0} = A^q v = \lambda^q v$, e quindi $\lambda = 0$. Il polinomio caratteristico è allora $(-1)^n t^n$, quindi per il teorema di Cayley-Hamilton $A^n = 0$. $\square$

Esercizi:

(1) Trovare l'errore nella seguente troppo veloce dimostrazione del teorema: $p(t) = \det(A - tI) \Longrightarrow p(A) = \det(A - AI) = \det(A - A) = \det(0) = 0$.
(2) A matrice $2 \times 2 \Longrightarrow A^3 = (t^2 - d)A - tdI$, ove $t =$ traccia di A e $d = \det(A)$.
(3) A ammette un solo autovalore $\lambda \iff A - \lambda I$ è nihilpotente.
(4) Sia $K = \mathbb{R}$; allora A è nihilpotente se e solo se $\mathrm{tr}(A^k) = 0$ per ogni $k = 1, 2, \dots, n$.

Risoluzione di 4. Se A è nihipotente, il suo polinomio caratteristico è $(-1)^n t^n$, e quindi $\mathrm{tr}(A) = 0$. D'altra parte se A è nihipotente anche A^k è nihilpotente, quindi $\mathrm{tr}(A^k) = 0$ per ogni k.

Supponiamo ora $\mathrm{tr}(A^k) = 0$ per $k = 1, 2, \ldots, n$; calcolando la traccia nell'identità del teorema di Cayley-Hamilton:

$$0 = \text{ tr } (\det(A)I + a_1 A + \ldots + a_i A^i + \ldots + (-1)^n A^n) = n \ \det(A).$$

Essendo $K = \mathbb{R}$ di caratteristica zero, $\det(A) = 0$. Esiste allora un vettore non nullo v_1 tale che $Av_1 = \underline{0}$. Si completi v_1 ad una base di $K^n : v_1, \ldots, v_n$. L'operatore nihilpotente F_A rispetto a tale base si rappresenta con una matrice

$$B = \begin{pmatrix} 0 & *** \\ \mathbf{0} & C \end{pmatrix},$$

con $C \in \mathcal{M}(n-1, \mathbb{R})$. Ora, $B^k = \begin{pmatrix} 0 & *** \\ \mathbf{0} & C^k \end{pmatrix}$. Ma $0 = \mathrm{tr}(Ak)$ $=\mathrm{tr}(B^k)$, perché A^k e B^k sono simili. Calcolando: $\mathrm{tr}(C^k) = \mathrm{tr}(B^k) = 0$. Per induzione (il caso $n = 1$ è banale), C^k è nihipotente: $C^{n-1} = 0$. Questo ci dice che $B^{n-1} = \begin{pmatrix} 0 & *** \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, ma allora $B^n = BB^{n-1} = \begin{pmatrix} 0 & *** \\ \mathbf{0} & C \end{pmatrix} \begin{pmatrix} 0 & *** \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = (0)$, e B è nihilpotente. Se la caratteristica del campo fosse 2, allora $\mathrm{tr}(I^k) = \mathrm{tr}(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}) = 2 = 0$, ma I non è nihipotente. □

4.1. Una applicazione del teorema di Cayley-Hamilton. Un polinomio $p(x)$ a coefficienti in K, non ha necessariamente radici in K. Una costruzione algebrica standard prova l'esistenza di un campo L contente K e qualche radice di p. Il teorema di Cayley-Hamilton ne fornisce un modello concreto. Indichiamo con $K[x]$ l'insieme dei polinomi a coefficienti in K, $p(x) \in K[x]$ è riducibile su K se $p(x) = q(x)s(x)$, dove $q(x)$ e $s(x)$ appartengono a $K[x]$ e non sono costanti; $p(x)$ è irriducibile su K se non è riducibile. Indichiamo con $\deg(p)$ il grado di p.

Lemma 4.2. Siano $p(x), s(x) in K[x]$, $\deg(s) = m < n = \deg(p)$. Se p è irriducibile e s non è il polinomio nullo, allora esistono due polinomi $a(x)$ e $b(x)$ tali che $a(x)p(x) + b(x)s(x) = 1$.

Dimostrazione. È un'applicazione diretta dell'algoritmo euclideo di divisione. Sia

$$S := \{f(x) \in K[x] \ : \ f(x) = A(x)p(x) + B(x)s(x) \text{ e } f(x) \neq 0\}.$$

Notiamo che S è non vuoto, infatti $s(x) = 1s(x) + 0p(x)$ appartiene ad S. Sia

$$l(x) = c(x)p(x) + d(x)s(x)$$

un elemento di S avente grado minimo. Dimostriamo che $l(x)$ è una costante (necessariamente) non nulla. Se così non fosse, ponendo $\deg(l(x)) = k$ avremmo

$$1 < k \leq m \quad (s(x) \text{ ha grado} m).$$

Dividendo $p(x)$ per $l(x)$ si ha $p(x) = l(x)q(x) + r(x)$ e $\deg(r(x)) < k$;

$$r(x) = (1 - c(x)q(x))p(x) + (-d(x)q(x))s(x).$$

Se $r(x) \neq 0$, allora $r(x) \in S$, ma questo è in contraddizione con la minimalità di k. D'altra parte, se $r(x) = 0$ allora $p(x) = l(x)q(x)$. Essendo $p(x)$ irriducibile anche questo è impossibile. Allota $l(x) = e \neq 0$. Posto $a(x) = e^{-1}c(x)$ e $b(x) = e^{-1}d(x)$ si ha $a(x)p(x) + b(x)s(x) = 1$. □

Dopo aver dimostrato il lemma la nostra costruzione è molto rapida. Sia $P(x) \in K[x]$ un polinomio irriducibile di grado $m > 1$ e monico, $P(x) = x^n + \dots$. Posto $p(x) = (-1)^n P(x) = (-1)^n x^n + \sum a_i x^{-n-i}$, esiste una matrice A avente $p(x)$ come polinomio caratteristico, per esempio:

$$A = \begin{pmatrix} 0 & 0 & \dots & \dots & a_1 \\ 1 & 0 & \dots & \dots & a_2 \\ 0 & 1 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_n \end{pmatrix}.$$

Sia $L = \{B \in \mathcal{M}(n, K) \; : \; B = q(A), q \in K[x]\}$. Dal teorema di Cayley-Hamilton sappiamo che $L = \langle I, A, A^2, \dots, A^{n-1} \rangle$ è il sottospazio vettoriale di $\mathcal{M}(n, K)$ generato da $I, A, A^2, \dots, A^{n-1}$, e che

$$P(A) = (-1)^n p(A) = 0.$$

Il prodotto di matrici definisce una operazione associativa e commutativa in $L :$ se $X = s(A)$ e $Y = r(A)$ sono in L, allora $XY = YX = (sr)(A) \in L$ per ogni X e Y in L. Identificando gli scalari $\lambda \in K$ con le matrici scalari $\lambda I \in L$, abbiamo un'inclusione $i \colon K \to L$ che rispetta le operazioni, e cioè $i(x + y) = i(x) + i(y)$ e $i(xy) = i(x)i(y)$.

Lemma 4.3.

i) Con la somma e moltiplicazioni di matrici L ùn campo.
ii) L'identificazione di K e $i(K)$ definisce una inclusione di campi $K \subset L$.
iii) $A \in L$ è una radice di $P(x) \in L[x]$.

Dimostrazione.

i) L'unica parte non ovvia è la verifica, che se $X \in L$ e $X \neq 0$ allora X è invertibile e $X^{-1} \in L$. Se $X \neq 0$ allora

$$X = s(A) = b_0 I + b_1 A + b_2 A^2 ... + b_{n-1} A^{n-1}$$

con $\deg s = m < n$ e s non nullo. Allora esistono polinomi $a(x)$ e $b(x)$ tali che

$$1 = a(x)P(x) + b(x)s(x)$$

e quindi

$$I = a(A)P(A) + b(A)s(A) = 0 + b(A)s(A) = b(A)X.$$

Questo prova che$X^{-1} = b(A) \in L$

ii) Identificando gli scalari di K con le matrici scalari $K\dot{I} \subset L$ si identifica K con un sottocampo di L.
iii) Il teorema di Cayley-Hamilton ci dice $P(A) = 0I = 0$ $A \in L$ è radice di $P(x) \in L[x]$

□

Iterando la costruzione precedente possiamo trovare L che contiene tutte le radici di qualsiasi polinomio.Ricapitolando abbiamo:

Proposizione 4.4. Per ogni campo K e polinomio $p(x) \in K[x]$ esiste un'inclusione di campi $K \subset L$ in modo il corrispondente polinomio $p(x) \in L[x]$ ha tutte le radici in L.

Esempio: Sia $K = \mathbb{R}$ e $p(x) = x^2 + 1$ e $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ allora $L = \{X \in M(2, \mathbb{R}) \; : \; X = aI + bA = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\}$. L'applicazione $\tau : \mathbb{C} \to L$, $\tau(a + ib) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, è un isomorfismo di campi. Si ha $\tau(1) = I$ e $\tau(i) = A$. Se $z = a + ib$, $tau(z)$ è la matrice associata all'operatore $f : \mathbb{C} \to \mathbb{C}$, $f(w) = zw$ rispetto alla base 1 e i di $\mathbb{C}$ spazio vettoriale reale di dimensione due.

CAPITOLO 6

Applicazioni bilineari

1. Definizioni e prime proprietà

DEFINIZIONE 1.1. **(applicazione bilineare)** Siano V e W due K spazi vettoriali; un'applicazione:

$$\phi : V \times W \to K$$

si dice bilineare (o forma bilineare) se comunque fissati $v \in V, w \in W$ le funzioni $f_v : W \to K$ e $f_w : V \to K$ definite da:

$$\phi_v(x) := \phi(v, x), \qquad \phi_w(y) = \phi(y, w)$$

sono lineari. Ovvero se, dati comunque v e v' in V, w e w' in W, λ e μ in K :

$$\phi(\lambda v + \mu v', w) = \lambda\phi(v, w) + \mu\phi(v', w), \quad \phi(v, \lambda w + \mu w') = \lambda\phi(v, w) + \mu\phi(v, w').$$

Nota. La bilinearità di $\phi : V \times W \to K$ si esprime:

$$\phi(\lambda_1 v_1 + \lambda_2 v_2, \mu_1 w_1 + \mu_2 w_2) = \lambda_1\mu_1\phi(v_1, w_1) + \lambda_\phi\mu_2\phi(v_1, w_2) + \lambda_2\mu_1\phi v_2, w_1) + \lambda_2\mu_2\phi(v_2, w_2).$$

Quindi ϕ non è lineare, perché $\phi(\lambda(v, w)) = \phi(\lambda v, \lambda w) = \lambda^2\phi(v, w)$, ma è lineare componente per componente. Inoltre induce due applicazione lineari $\phi' : V \to W^*$, definita come $\phi'(v) = \phi(v, \cdot)$, e $\phi'' : W \to V^*$, definita come $\phi''(w) = \phi(\cdot, w)$.
ESEMPI:

(1) Sia A una matrice $m \times n$; è bilineare l'applicazione $\phi_A : K^n \times K^m \to K$ definita da $\phi_A(x, y) :=^t xAy$.
(2) Sia $K = \mathbb{R}$; definiamo $g : \mathbb{R}^n \times \mathbb{R}^n$ come $g\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) := \sum_{i=1}^n x_i y_i$; g è il prodotto scalare standard ($g = \phi_I$).
(3) Dati V e $V^\star$, è bilineare da $V \times V^\star$ in K l'applicazione $k(f, v) := f(v)$.
(4) Sia C^0 lo spazio vettoriale reale delle funzioni continue da $[0, 1]$ in $\mathbb{R}$;

$$I : C^0 \times C^0 \to \mathbb{R}, \quad I(f, g) := \int_0^1 f(x)g(x)dx$$

è bilineare.
(5) Se $\phi : V \times W \to K$ è bilineare, e V', W' sono sottospazi vettoriali rispettivamente di V e W, la restrizione $\phi' : V' \times W' \to K$ è bilineare.

Le applicazioni bilineari sono rappresentate, come nel caso di dimensione finita, da opportune matrici. Come per le applicazioni lineari si devono scegliere delle basi.

PROPOSIZIONE 1.2. Siano $\{v_i\}_{i \in I}$ una base di V e $\{w_j\}_{j \in J}$ una base di W. Per ogni coppia di indici $i \in I$ e $j \in J$, sia fissato $a_{ij} \in K$. Allora esiste ed è unica una applicazione bilineare $\phi : V \times W \to K$ tale che $\phi(v_i, w_j) = a_{ij}$.

DIMOSTRAZIONE. • *Unicità.* Se $v = \sum_i x_i v_i$ e $w = \sum_j y_j w_j$, allora (iterando la bilinearità):

$$\phi(v,w) = \sum_{i,j} x_i y_j \phi(v_i, w_j) = \sum_{i,j} a_{ij} x_i y_j. \tag{6.4}$$

• *Esistenza.* Si usa (6.4) per definire la ϕ.

□

Supponiamo ora $\dim(V) = n < +\infty$ e $\dim(W) = m < +\infty$. Siano $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_m)$ rispettivamente le coordinate di v e w nella basi $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ di V e $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ di W. La proposizione ci mostra che se $A = (a_{ij}) = (\phi(v_i, w_j))$:

$$\phi(v,w) =^t xAy = \sum_{i,j} a_{ij} x_i y_j.$$

DEFINIZIONE 1.3. La matrice $n \times m$ $A = A_\phi = (\phi(v_i, w_j))$ è la matrice associata a ϕ rispetto alle basi $\mathcal{V}$ di V e $\mathcal{W}$ di W.

Le corrispondenze $\phi \mapsto A_\phi$ e $A \mapsto \phi_A$ sono l'una l'inversa dell'altra. Abbiamo:

COROLLARIO 1.4. Tutte e sole le applicazioni bilineari $\phi : K^n \times K^m \to K$ sono associate a matrici $n \times m$:

$$\phi(x,y) = \phi_A(x,y) =^t xAy.$$

ESERCIZIO: Sia ϕ_A un'applicazione bilineare associata ad A matrice $n \times m$. Se $s < n$, $p < m$, $K^s = \{(x_1, \dots, x_s, 0 \dots, 0)\} \subset K^n$ e $K^p = \{(y_1, \dots, y_p, 0, \dots, 0)\} \subset K^m$, qual è la matrice associata alla restrizione di $\phi_A : K^s \times K^p \to K$?

Cambiamento di basi. Dobbiamo vedere come si trasformano le matrici associate quando si cambia base.

Siano $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e $\mathcal{V}' = \{v'_1, \dots, v'_n\}$ basi di V, $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ e $\mathcal{W}' = \{w'_1, \dots, w'_m\}$ di W. Siano $v \in V$ con coordinate x e x' rispetto a $\mathcal{V}$ e $\mathcal{V}'$, $w \in W$ con coordinate y e y' rispetto a $\mathcal{W}$ e $\mathcal{W}'$. Siano B la matrice cambiamento di coordinate da $\mathcal{V}$ a $\mathcal{V}'$ e C quella da $\mathcal{W}$ a $\mathcal{W}'$. Si ha quindi: $x' = Bx$ e $y' = Cy$.

Siano infine A e D le matrici associate all'applicazione bilineare ϕ tramite rispettivamente $\mathcal{V}$, $\mathcal{W}$ e $\mathcal{V}'$, $\mathcal{W}'$.

Si ottiene:

$${}^t xAy = \phi(v,w) =^t x'Dy' =^t (Bx)D(Cy) =^t x({}^t BDC)y.$$

Questo vale per ogni $x \in K^n$ e $y \in K^m$; allora:

$$A =^t BDC.$$

Si noti che $r(A) = r(D)$ perché B e C sono invertibili.

Osservazione. Ad ogni matrice abbiamo associato un'applicazione lineare e una bilineare. Le matrici rappresentano due oggetti diversi. Cambiando basi le leggi di trasformazioni sono diverse. Il caso più importante per le applicazioni bilineari è quello in cui $V = W$.

DEFINIZIONE 1.5. Due matrici quadrate A e D si dicono congruenti (congrue, cogredienti) se esiste una matrice invertibile B tale che $A =^t BDB$.

Due matrici congruenti di ordine n sono associate ad una stessa applicazione bilineare mediante due basi di uno spazio vettoriale di dimensione n.

Nota. Prendendo i determinanti si ha:

$$\det(A) = \det(D)\det(B)^2.$$

Se $K = \mathbb{R}$, il segno del determinante è invariante per congruenza.

Definizione 1.6. Sia $V = W$. Una forma $\phi : V \times V \to K$ bilineare si dice:

- *simmetrica* se per ogni v e w : $\phi(v,w) = \phi(w,v)$;
- *antisimmetrica* se per ogni v e w : $\phi(v,w) = -\phi(w,v)$.

Nota. La simmetria e l'antisimmetria si mantengono per restrizione a sottospazi.
Osservazione. L'applicazione bilineare ϕ_A è simmetrica (rispettivamente antisimmetrica) se e solo $A = {}^tA$, cioè A simmetrica (rispettivamente $A = -{}^tA$, cioè A antisimmetrica). Se $\dim(V) = n < +\infty$, ϕ è simmetrica (antisimm.) se e solo se è associata ad una matrice simmetrica (antisimm.). Il prodotto scalare dell'esempio 2. è un'applicazione simmetrica e eguale a ϕ_I.

Se il campo K è di caratteristica 2 la simmetria equivale all'antisimmetria. Altrimenti (cioè se $2 \neq 0$ in K) per ogni $\phi : V \times V \to K$ definiamo:

$$\phi_s : V \times V \to K \quad : \quad \phi_s(v,w) = \frac{1}{2}(\phi(v,w) + \phi(w,v)) \quad \text{(parte simmetrica)}$$

$$\phi_a : V \times V \to K \quad : \quad \phi_a(v,w) = \frac{1}{2}(\phi(v,w) - \phi(w,v)) \quad \text{(parte antisimmetrica).}$$

Ora ϕ_s e ϕ_a sono bilineari e $\phi = \phi_s + \phi_a$. Nel caso di matrici $A = \dfrac{A + {}^tA}{2} + \dfrac{A - {}^tA}{2}$.

Esercizi:

(1) Dimostrare che la congruenza è una relazione di equivalenza.
(2) Dire quali delle seguenti affermazioni sono vere e quali sono false:
 (a) Se A e B sono congruenti allora sono equivalenti.
 (b) Se A e B sono congruenti allora sono simili.
 (c) Se A e B sono simili allora sono congruenti .
 (d) Se A e B sono equivalenti allora sono congruenti.
(3) ϕ e ψ siano applicazioni bilineari da $V \times V$ in K, λ e μ siano scalari. Dimostrare che $\lambda\phi + \mu\psi$, definita da $(\lambda\phi + \mu\psi)(v,w) = \lambda\phi(v,w) + \mu\psi(v,w)$ è bilineare.
(4) Le operazioni $\phi + \psi$ e $\lambda\phi$ inducono una struttura di spazio vettoriale su $\mathcal{B}(V)$, l'insieme delle applicazioni bilineari su V. Se $\dim(V) = n$, calcolare $\dim(\mathcal{B}(V))$.
(5) Siano $\mathcal{S}(V)$ e $\mathcal{A}(V)$ i sottoinsiemi di $\mathcal{B}(V)$ contenenti rispettivamente le applicazioni simmetriche e quelle antisimmetriche. Provare che $\mathcal{S}(V)$ e $\mathcal{A}(V)$ sono sottospazi vettoriali di $\mathcal{B}(V)$, e $\mathcal{S}(V) \oplus \mathcal{A}(V) = \mathcal{B}(V)$. Calcolare le dimensioni.
(6) Siano A e D congruenti; provare che A è simmetrica (antisimm.)$\Longleftrightarrow$ D lo è.
(7) Se due forme sono congruenti, le loro restrizioni ad un sottospazio sono congruenti?

Definizione 1.7. **(applicazione non degenere)** Una applicazione bilineare $\phi : V \times V \to K$ si dice non degenere se per ogni $v \in V, v \neq \underline{0}$ esiste $w \in V$ t.c. $\phi(v,w) \neq 0$ e $\phi(w,v) \neq 0$.

Osservazione. La definizione $\forall v \in V, v \neq \underline{0}, \exists w, w' \in V : \ \phi(v,w) \neq 0$ e $\phi(w',v) \neq 0$ è solo apparentemente più debole. Se $\phi(v,w') = \phi(w,v) = 0$, allora $\phi(v, w+w') \neq 0$ e $\phi(w+w',v) \neq 0$.

Definizione 1.8. **(spazio nullo e indice di nullità)** Quando ϕ è simmetrica (antisimm.), il sottospazio vettoriale:

$$V(\phi,0) = \{v \in V : \ \phi(v,w) = 0 \ \forall w \in V\} = \{v \in V : \ \phi(w,v) = 0 \ \forall w \in V\}$$

di V è lo spazio nullo di ϕ, e $\dim(V(\phi,0)) = n_0$ è il suo l'indice di nullità. Si ha che $n_0 = 0$ e $V(\phi,0) = \{\underline{0}\}$ se e solo se ϕ è non degenere.

Esercizi:

(1) Dimostrare che ϕ_I è non degenere.
(2) Se A è una matrice simmetrica o antisimm., $V(\phi_A,0) = \ker(A)$.
(3) Se A è simmetrica o antisimm., ϕ_A è non degenere $\Longleftrightarrow$ A è invertibile.

(4) Se $A \in \mathcal{M}(n, K)$ è antisimm. e ϕ_A è non degenere allora n è pari ($\det(A) = 0$ se n è dispari).

2. Forme quadratiche

Le applicazioni bilineari simmetriche rivestono nelle applicazioni geometriche (quadriche) e fisiche (energia cinetica, spazio tempo) una posizione di rilievo.

Definizione 2.1. **(forma quadratica)** Sia $\phi : V \times V \to K$ una applicazione bilineare simmetrica. La funzione $q : V \to K$, definita da $q(v) := \phi(v, v)$, è la forma quadratica associata a ϕ. Quando A è una matrice simmetrica, si indica con q_A la forma associata a f_A, cioè: $q_A(x) =^t xAx$.

Proposizione 2.2. Sia q la forma quadratica associata a ϕ simmetrica. Allora:

(1) per ogni v in V e λ in K : $q(\lambda v) = \lambda^2 q(v)$;
(2) per ogni v e w in V : $q(v + w) - q(v) - q(w) = 2\phi(v, w)$. Se la caratteristica di K, $\mathrm{ch}(K)$, è diversa da 2, $\phi(v, w) = \frac{1}{2}(q(v + w) - q(v) - q(w))$;
(3) se $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ è una base di V e $A = (a_{ij})$, con $a_{ij} = \phi(v_i, v_j)$, allora:

$$q(x_1v_1 + \cdots + x_nv_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j;$$

(4) se $\mathrm{ch}(K) \neq 2$, ogni forma quadratica $q : K^n \to K$ è un polinomio omogeneo di grado 2 (se $\mathrm{ch}(K) = 2$ questa può essere una definizione).

Dimostrazione. Tutto è ovvio; veniamo per esempio alla 2.

$$q(v + w) = \phi(v + w, v + w) = \phi(v, v) + \phi(w, w) + \phi(v, w) + \phi(w, v)$$

$$= q(v) + q(w) + 2\phi(v, w).$$

□

Osservazione. Il passaggio $q \mapsto \phi$ si chiama *polarizzazione della forma quadratica.*
Notiamo che su K^n ($\mathrm{ch}(K) \neq 2$) abbiamo 3 oggetti equivalenti:

a) forme bilineari simmetriche;
b) matrici simmetriche;
c) forme quadratiche.

Il problema ora è quello di cercare delle basi in cui le forme quadratiche siano semplici. Da ora in avanti supporremo $\mathrm{ch}(K) \neq 2$.

Definizione 2.3. Sia $\phi : V \times V \to K$ una applicazione bilineare; un insieme di vettori $\{v_i\}$ di V si dice *ortogonale rispetto a* ϕ se $\phi(v_i, v_j) = 0$ se $i \neq j$.

Notiamo che se $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ è una base di V ortogonale rispetto a ϕ, e $\phi(v_i, v_i) = a_i$, si ha:

$$\phi(\sum_i x_i v_i, \sum_i y_i v_i) = \sum_i a_i x_i y_i;$$

$$q(\sum_i x_i v_i) = \sum_i a_i x_i^2 \quad \text{(forma quadratica associata)}$$

$$D = \begin{pmatrix} a_1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & a_i & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & a_n \end{pmatrix} \quad \text{matrice associata.}$$

Teorema 2.4. (**Lagrange**) Sia ϕ una forma simmetrica e $q : V \to K$ la forma quadratica associata. Esiste una base $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ di K^n ortogonale per ϕ, cioè t.c.:

$$q(y_1 v_1 + y_2 v_2 + \dots + y_n v_n) = \sum_{i=1}^{n} a_i y_i^2.$$

Dimostrazione. Per induzione: se $n = 1$ non vi è nulla da dimostrare. Sia $n > 1$. Se $q = 0$ allora $\phi = 0$, e non vi è nulla da dimostrare; sia v_1 t.c. $q(v_1) = a_1 \neq 0$; completiamo v_1 ad una base $\{v_1, w_2, \dots, w_n\}$ di V e definiamo $L := \{v \in V \ : \ \phi(v_1, v) = 0\}$. Posto $v = x_1 v_1 + x_2 w_2 + \dots + x_n w_n$ e $\phi(v_1, w_j) = b_j$, si ha $\phi(v_1, v) = 0$ se e solo se:

$$a_1 x_1 + b_2 x_2 + \dots + b_n x_n = 0. \tag{6.5}$$

Dunque (6.5) è un problema lineare omogeneo, una equazione (non banale) e n incognite. Quindi L ha dimensione $n - 1$ e $v_1 \notin L$. Restringendo la forma ϕ a L, per induzione costruiamo una base $\{v_2, \dots, v_n\}$ di L ortogonale rispetto a ϕ. Algoritmicamente si procede risolvendo equazioni lineari come la (6.5). Quindi $\phi(v_i, v_j) = 0$ se $i \neq j$ e $i > 1$, $j > 1$. Essendo anche v_i in L per $i > 1$, si ha $\phi(v_i, v_1) = 0$ se $i > 1$. Segue che $\{v_1, v_2, \dots, v_n\}$ è una base di V ortogonale per ϕ. □

Traducendo la proposizione nel linguaggio delle matrici:

Proposizione 2.5. Ogni matrice simmetrica è congruente ad una diagonale.

Dimostrazione. Si tratta di un'applicazione diretta del lemma precedente. □

In pratica la diagonalizzazione delle forme è molto rapida:

Esempio: Sia $q(x, y, z) = 2xy + 2xz + 2yz$, $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

La forma è nulla sulla base canonica; ponendo $X = x + y$, si ha:

$$q(x, y, z) = q'(X, y, z) = 2(X - y)y + 2(X - y)y + 2yz = -2y^2 + 4Xy + 2yz$$

$$= -2(y^2 + 2Xy - yz) = -2(y + X - \frac{1}{2}z)^2 + 2X^2 + \frac{1}{2}z^2;$$

Ponendo poi anche $Y = x + 2y - \frac{1}{2}z$ e $Z = z$, scriviamo:

$$Q(X, Y, Z) = 2X^2 - 2Y^2 + \frac{1}{2}Z^2.$$

Definizione 2.6. Sia q una forma quadratica su V spazio vettoriale su K di dimensione finita. Sia A una matrice che rappresenta q. Il rango di q, $r(q)$, è per definizione quello di A e il suo indice di nullità è $n_0(q) = n - r(q)$. Due forme quadratiche q e q' di K^n si dicono congrue se le loro applicazioni bilineari sono congrue.

Osservazione. Il rango è invariante per congruenza. Quando q è diagonalizzata rispetto ad una base ortogonale, $q = \sum_{i=1}^{n} a_i y_i^2$, $r(q)$ è numero degli a_i non nulli.

Sembrerebbe che la proposizione precedente risolva completamente il problema della congruenza delle forme quadratiche di K^n, cioè delle matrici simmetriche. Non è così ! È ancora da risolvere il problema di quando due diagonali sono congrue. La questione assume tinte drammatiche quando $K = \mathbb{Q}$. Le forme $q = x_1^2 + x_2^2$ e $q' = 2x_1^2 + x_2^2$ (basterebbe considerare x_1^2 e $2x_1^2$) hanno matrici $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, e $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, e non sono congrue. Se lo fossero esisterebbe B t.c. $A =^t BIB =^t BB$; dunque $2 = \det(A) = \det(B)^2$, ma allora $\sqrt{2}$ sarebbe un numero razionale.
Quando $K = \mathbb{C}$, il campo dei numeri complessi, il problema è elementare.

Corollario 2.7. Se $\forall a \in K$ l'equazione $x^2 = a$ ha soluzioni in K, due matrici simmetriche (forme quadratiche su K^n) sono congruenti se e solo se hanno lo stesso rango.

Dimostrazione. Due matrici congruenti hanno lo stesso rango. Viceversa, da 2.5 ogni matrice simmetrica A è congruente ad una diagonale D avente forma quadratica $q(x) = \sum_{i=1}^{r} a_i x_i^2$, dove $r =$ rango di A, $a_i \neq 0$. Sia d_i tale che $d_i^2 = a_i$; allora $q(x) = \sum_{i=1}^{r} (d_i x_i)^2$. Operando il cambiamento di base $v_i' = (d_i)^{-1} v_i$, dove v_i è l'i-esimo vettore della base canonica, se $i \leq r$ e $v_i' = v_i$ se $i > r$. Si ha dunque: $q(\sum_{i=1}^{r} y_i v_i') = \sum_{i=1}^{r} y_i^2$.
La matrice D è congrua a $C(r) = \begin{pmatrix} I_r & 0 \\ 0 & 1 \end{pmatrix}$, □

3. Forme quadratiche reali

In tutto questo paragrafo avremo $K = \mathbb{R}$. La seguente definizione richiede in maniera specifica di poter stabilire quando uno scalare è maggiore di zero.

Definizione 3.1. Sia V uno spazio reale, un'applicazione bilineare simmetrica $\phi : V \times V \to \mathbb{R}$ (una forma quadratica $q : V \to \mathbb{R}$) si dice:

- *semidefinita positiva* ($\phi \geq 0$) se per ogni $v \in V$ si ha: $q(v) = \phi(v, v) \geq 0$;
- *prodotto scalare* o *definita positiva* ($\phi > 0$) se per ogni $v \in V$ si ha: $q(v) = \phi(v, v) \geq 0$ e $q(v) = \phi(v, v) = 0$ se e solo se $v = \underline{0}$;
- *semidefinita negativa* ($\phi \leq 0$) se per ogni $v \in V$ si ha $\phi(v, v) \leq 0$;
- *definita negativa* ($\phi < 0$) se per ogni $v \neq \underline{0}$ si ha $\phi(v, v) < 0$;
- *indefinita* se esistono $v, w \in V$ t.c. $\phi(v, v) < 0 < \phi(w, w)$.

Nota. Un prodotto scalare è una applicazione bilineare simmetrica su uno spazio reale per cui: $q(v) = \phi(v, v) \geq 0$ e $\phi(v, v) = 0 \Leftrightarrow v = \underline{0}$. Chiaramente ϕ è definita negativa se e solo se $-\phi$ è definita positiva.

Definizione 3.2. Una matrice simmetrica A si dice definita positiva se $\phi_A(x, y) =^t xAy$ è definita positiva, cioè se per ogni $x \in \mathbb{R}^n$: ${}^t xAx \geq 0$ e ${}^t xAx = 0 \Leftrightarrow x = \underline{0}$. Si dirà poi semidefinita positiva o negativa se ϕ_A lo è.

Esempi:

(1) Siano $V = \mathbb{R}^n$, $\phi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=n}^{n} a_i x_i y_i$ e $q(x) = \sum_{i=1}^{n} a_i x_i^2$; allora $\phi(q)$ è semidefinita positiva se $a_i \geq 0 \forall i$; è prodotto scalare se $a_i > 0$ $\forall i$.

(2) Su $\mathcal{M}(n,\mathbb{R})$, $g(A,B) :=\mathrm{tr}({}^t\!AB)$ è un prodotto scalare.

DEFINIZIONE 3.3. L'applicazione $\phi((x_1,\dots,x_n),(y_1,\dots,y_n)) = \sum_{i=n}^{n} x_i y_i$ è il prodotto scalare standard di $\mathbb{R}^n$.

(7.3.1)

LEMMA 3.4. Sia V uno spazio vettoriale reale e $\phi\colon V\times V\to\mathbb{R}$ un'applicazione bilineare simmetrica. Siano W e Z due sottospazi vettoriali di V. Siano $\phi|_W$, $\phi|_Z$ le restrizioni di ϕ a W e Z. Se $\phi|_W>0$ e $\phi|_Z\le 0$, allora $W\cap Z=\{\underline{0}\}$, e allora $W+Z=W\oplus Z$. In particolare $\dim(V)\ge\dim(W)+\dim(Z)$.

DIMOSTRAZIONE. Se $v\in W\cap Z$, allora $\phi(v,v)\ge 0$ perché $v\in W$ e $\phi(v,v)\le 0$ perché $v\in Z$, quindi $\phi(v,v)=0$, ma su W ϕ è definita positiva: $v=\underline{0}$. □

Sia V uno spazio vettoriale reale, $\dim(V)=n<+\infty$, e sia $q\colon V\to\mathbb{R}$ una forma quadratica. Consideriamo le seguenti famiglie di sottospazi

$$\mathcal{F}^+=\mathcal{F}^+(q)=\{U\ \text{ sottospazio di } V\mid q_U\colon U\to\mathbb{R} \text{ è definita positiva }\}$$
$$\mathcal{F}^-=\mathcal{F}^+(q)=\{U\ \text{ sottospazio di } V\mid q_U\colon U\to\mathbb{R} \text{ è definita negativa }\}$$

Poniamo $n^+=\max_{U\in\mathcal{F}^+}\dim U$ e $n^+=\max_{U\in\mathcal{F}^-}\dim U$.

DEFINIZIONE 3.5. **(segnatura, rango e indice)** Sia $q:V\to\mathbb{R}$ una forma quadratica, la coppia di interi (non negativi) (n^+,n^-) è la *segnatura*, $r(q)=n^++n^-$ è il *rango*, la differenza $i(q)=n^+-n^-$ è l'*indice*.

Proveremo che la segnatura classifica le congruenze di forme quadratiche reali. Cominciamo a costruire delle forme quadratiche con segnatura assegnata. Dati due numeri interi non negativi s e k con $s+k\le n$, sia $q_{s,k}:\mathbb{R}^n\to\mathbb{R}$ definita da

$$q_{s,k}(x_1,x_2,\dots,x_n)=\sum_{i=1}^{s}x_i^2-\sum_{j=s+1}^{s+k}x_j^2.$$

Le matrici associate a $q_{s,k}$ sono:

$$C(s,k)=\begin{pmatrix} I_s & 0 & 0\\ 0 & -I_k & 0\\ 0 & 0 & 0\end{pmatrix},\quad C(s,0)=\begin{pmatrix} I_s & 0\\ 0 & 0\end{pmatrix},\quad C(0,k)=\begin{pmatrix} -I_k & 0\\ 0 & 0\end{pmatrix}.$$

Nota. Si ha che il rango è $r(q_{s,k})=s+k=r(C(s,k))$, la nullità è $n-(s+k)$, $C(n,0)=I$, $C(0,n)=-I$, $C(s,0)$ è semidefinita positiva, $C(0,k)$ semidefinita negativa e $C(s,k)$, con $s,k>0$, indefinita.

PROPOSIZIONE 3.6. La segnatura di $q_{s,k}$ è (s,k).

DIMOSTRAZIONE. Sia $\mathcal{E}=\{e_1,e_2,\dots,e_n\}$ la base canonica di $\mathbb{R}^n$. Consideriamo gli spazi $U(+):=\mathrm{span}\,[e_1,e_2,\dots,e_s]$, $\dim(U(+))=s$, e $Z:=\mathrm{span}\,[e_{s+1},e_{s+2},\dots,e_n]$, $\dim(Z)=n-s$. Ora q è semidefinita negativa su Z e positiva su $U(+)$, $U(+)\in\mathcal{F}^+$. Quindi:

$$s\le n^+=\max_{U\in\mathcal{F}^+}\dim U.$$

La disuguaglianza inversa segue dal lemma; se $U\in\mathcal{F}^+$, allora $\dim(U)+\dim(Z)\le n$, $\dim(U)\le s$, ma allora $n^+\le s$. Ovvero $n^+=s$.
Passando agli spazi di $\mathcal{F}^-$, si ha che $U(-)=\mathrm{span}\,[e_{s+1},e_{s+2},\dots,e_{s+k}]$; pertanto $n^-=k$. □

COROLLARIO 3.7. La matrice $C(a,b)\in\mathcal{M}(n,\mathbb{R})$ è congruente a $C(k,s)\in\mathcal{M}(n,\mathbb{R})\Longrightarrow a=k$ e $b=s$.

DIMOSTRAZIONE. Le forme quadratiche associate $q_{a,b}$ e $q_{s,k}$ hanno segnature (a, b) e (s, k). La segnatura è invariante per congruenza. $\square$

TEOREMA 3.8. **(teorema di Sylvester)**

i) Ogni matrice simmetrica $A \in \mathcal{M}(n, \mathbb{R})$ è congruente ad una e una sola delle matrici $C(a, b)$.
ii) Siano q e q' due forme quadratiche definite su uno spazio reale V di dimensione finita; q e q' sono congrue se solo se hanno la stessa segnatura.

DIMOSTRAZIONE. Le due affermazioni i) e ii) sono equivalenti: la seconda è la versione intrinseca della prima.

i) Indichiamo con $\phi = \phi_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ l'applicazione definita da $\phi(x, y) =^t xAy$, e con $q(x) =^t xAx$ la forma quadratica associata. Sappiamo che A è congrua ad una diagonale, ovvero che esiste una base $\{v_1, v_2, \ldots, v_n\}$ di $\mathbb{R}^n$ ortogonale rispetto a ϕ, cioè $\phi(v_i, v_j) = 0$ se $i \neq j$. Riordinando, possiamo assumere:

$$\begin{aligned}
&\phi(v_i, v_i) = c_i > 0 \text{ per } i \leq s \\
&\phi(v_i, v_i) = -c_i < 0 \text{ per } s < i \leq k + s \\
&\phi(v_i, v_i) = 0 \text{ per } i > k + s
\end{aligned}$$

Se poniamo allora $w_i := \frac{v_i}{\sqrt{c_i}}$ per $i \leq k + s$ e $w_i = v_i$ per $i > k + s$, avremo:

$$\begin{aligned}
&\phi(w_i, w_i) = 1 \text{ per } i \leq s \\
&\phi(w_i, w_i) = -1 \text{ per } s < i \leq k + s \\
&\phi(w_i, w_i) = 0 \text{ per } i > k + s
\end{aligned}$$

La matrice associata è $C(k, s)$ e la forma quadratica diventa:

$$q'(x_1, \ldots, x_n) = q(x_1w_1 + \cdots + x_nw_n) \sum_{i=1}^{s} x_i^2 - \sum_{j=s+1}^{s+k} x_j^2.$$

ii) Se q e q' sono congruenti allora hanno la stessa segnatura. Viceversa supponiamo abbiano la stessa segnatura (s, k). Sia $\{w_1, w_2, \ldots, w_n\}$ una base di V. Siano A la matrice associata a q e A' quella associata a q'. Si ha:

$$\text{segnatura di } q_A = \text{segnatura di } q = \text{segnatura di } q' = \text{segnatura di } q_{A'} = \ (s, k).$$

Dalla prima parte della proposizione sappiamo che A e A' sono congrue a $C(s, k)$ e quindi sono congrue tra di loro. Questo significa che q e q' sono congrue.

$\square$

Osservazione. Due forme quadratiche sono congruenti se e solo se esistono basi rispetto alle quali la loro espressione analitica è la stessa.

COROLLARIO 3.9. Il rango r di una forma quadratica q di segnatura (s, k) è $s + k$.

DIMOSTRAZIONE. Si passi per congruenza alle matrici $C(s, k)$. $\square$

Nota. L'indice $i(q)$ e il rango $r(q)$ classificano le forme quadratiche reali. Infatti $(s, k) = \left(\frac{1}{2}(r(q) + i(q)), \frac{1}{2}(r(q) - i(q))\right)$ è la segnatura di q. Quando q è non degenere $r(q)$ =dim(V), e l'indice $i(q)$ classifica completamente la forma.

4. Criterio per il calcolo della segnatura

Vogliamo dare un criterio pratico per il calcolo della segnatura della forma quadratica q_A dove A è una matrice simmetrica reale.

Lemma 4.1. Sia $A \in \mathcal{M}(n, \mathbb{R})$ simmetrica. Sia q_A non degenere di segnatura (s, k), $k + s = n$. Allora $\det(A) > 0$ se k è pari e $\det(A) < 0$ se k è dispari.

Dimostrazione. Se q_A è non degenere di segnatura (s, k), allora A è congrua a $C(s, k)$, cioè $A =^t BC(s, k)B$, con B invertibile; quindi $\det(A) = (-1)^k \det(B)^2$. Il segno del determinante è uguale al segno di $(-1)^k$. □

Osservazione. Dal lemma segue che il determinante una matrice associata ad un prodotto scalare è positivo.

Sia W un sottospazio di uno spazio vettoriale reale V di dimensione finita . Supponiamo $0 <\dim(W) = n$ e $\dim(V) = n + 1$. Sia $q : V \to \mathbb{R}$ una forma quadratica su V e $q' : W \to \mathbb{R}$ la sua restrizione. Siano $\mathcal{W} = \{w_1, w_2, \ldots, w_n\}$ una base di W e $\mathcal{V} = \{v_1, \ldots, v_n, v_{n+1}\}$ una base di V. Sia $A \in \mathcal{M}(n+1, \mathbb{R})$ la matrice che rappresenta q rispetto a $\mathcal{V}$ e $B \in \mathcal{M}(n, \mathbb{R})$ la matrice che rappresenta q' rispetto alla base $\mathcal{W}$.

Lemma 4.2. Supponiamo che $q' : W \to \mathbb{R}$ abbia segnatura (s', k') con $s' + k' = n$, q' non degenere. Sia (r, s) la segnatura di q. Abbiamo 3 casi:

(1) $\det(A) = 0 \iff s = s'$ e $k = k'$ (q è degenere);
(2) $\det(A) \neq 0$
   (a) ha lo stesso segno di $\det(B) \iff s = s' + 1$ e $k = k'$;
   (b) ha segno diverso da quello di $\det(B) \iff s = s'$ e $k = k' + 1$.

Dimostrazione. Poiché q' ha segnatura (s', k') con $s' + k' = n$, esistono due sottospazi vettoriali U e Z di W t.c. $\dim(U) = s'$, $\dim(Z) = k'$, $U \oplus Z = W$, e tali che le restrizioni $q' : U \to \mathbb{R}$ e $q' : Z \to \mathbb{R}$ siano rispettivamente definita positiva e definita negativa. Ora $U \in \mathcal{F}^+(q)$ e $Z \in \mathcal{F}^-(q)$, quindi $s \geq s'$ e $k \geq k'$; essendo $s + k \leq n + 1$ rimangono solo tre possibilità:

(1) $s = s'$ e $k = k' \iff r(q) = s + k = n \iff$ rango di q non massimo $\iff$ rango di A non massimo $\iff \det(A) = 0$.
(2) se $\det(A) \neq 0$ si può avere:
   (a) $\text{segno}(\det(A)) = \text{segno}(-1)^k = \text{segno}(\det(B)) \Longrightarrow k = k'$.
   (b) $\text{segno}(\det(A)) = \text{segno}(-1)^k \neq \text{segno}(\det(B)) \iff k = k' + 1$.

□

Sia $A \in \mathcal{M}(n, \mathbb{R})$ una matrice simmetrica $A = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix}$. Consideriamo la successione di matrici simmetriche $A(1), \ldots, A(r), \ldots, A(n) = A$, $A(r) \in \mathcal{M}(r, \mathbb{R})$ la sottomatrice di A di ordine r sulle prime r righe e r colonne. Poniamo $a_0 = 1$ e $a_r = \det(A(r))$ per $r > 0$. Consideriamo la successione :

$$(*) a_0 = 1, a_1, \ldots, a_r, \ldots, a_n. \tag{6.6}$$

Si ha cambiamento di segno nella (6.6) nella posizione i se $a_i a_{i+1} < 0$.

Proposizione 4.3. Assumiamo $a_i = \det(A(i)) \neq 0$, $i = 1, ..., n$. Sia σ il numero dei cambiamenti di segno nella successione (x). Allora la segnatura di q_A è $(n - \sigma, \sigma)$.

Dimostrazione. Sia $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ la base standard; per $r \leq n$, poniamo $V(r) = \text{span}[e_1, \ldots, e_r]$; si ha

$$0 \subset V(1) \subset \cdots \subset V(r) \subset \cdots \subset V(n) = \mathbb{R}^n.$$

Identifichiamo $V(r)$ con $\mathbb{R}^r$ ($\mathbb{R}^r = \{(a_1, \ldots, a_r, 0, \ldots, 0)\} \subset \mathbb{R}^n$). La restrizione di q_A a $\mathbb{R}^r = V(r)$ ha matrice associata $A(r)$ rispetto alla base $\{e_1, \ldots, e_r\}$. Confrontiamo la segnatura delle forme $q_{A(r)}$ e $q_{A(r+1)}$, entrambe non degeneri. Dal Lemma precedente sappiamo che l'intero che regola gli spazi negativi aumenta se e solo se il segno di $\det(A(r))$ è diverso dal segno di $\det(A(r+1))$. Questo avviene se vi è un cambiamento di segno nella successione (6.6). Se $(n-\sigma, \sigma)$ è la segnatura di A (q_A), allora σ è il numero dei cambiamenti di segno in (6.6). □

Proposizione 4.4. *Sia A una matrice simmetrica reale; A è definita positiva se e solo se $det(A(i)) > 0$, $i = 1, \ldots, n$.*

Dimostrazione. Se q_A definita positiva le restrizioni $q_{A(i)}$ a $V(i)$, sono positive, e quindi $det(A(i)) > 0$. Viceversa se $det(A(i)) > 0$ per ogni i allora per la proposizione precedente la segnatura di q_A $(n, 0)$ e quindi la forma definita positiva. □

Osservazione. Il criterio non va usato pedestremente: alcune volte può convenire per esempio considerare i minori partendo dal basso.

Esercizio: Si consideri $X(a) = \begin{pmatrix} 3 & a+3 & 0 & 3 \\ a+3 & -a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix} \in \mathcal{M}(4, \mathbb{R})$, con a parametro reale.

(1) Dire per quali valori di a $X(a)$ è diagonalizzabile su $\mathbb{R}$.
(2) Calcolare, al variare di a, la segnatura della forma quadratica associata ad $X(a)$.
(3) Calcolare autovalori e autovettori di $X(-3)$.

Esercizio: Si consideri la seguente applicazione $\phi(s) : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$:

$\phi(s)\big((x, y, x, t), (x', y', z', t')\big) = 24xx' + syy' + zz' + (5-s)tt' + 2sxy' + 2syx' + zt' + tz' - \cos(s)ty'$.

(1) Dire per quali valori del parametro reale s
  (a) $\phi(s)$ è bilineare;
  (b) $\phi(s)$ è simmetrica;
  (c) $\phi(s)$ è un prodotto scalare.
(2) Nei casi in cui vale (c), costruire una base ortogonale per $\mathbb{R}^4$ rispetto a tale prodotto scalare.

Esercizio: Siano $A \in \mathcal{M}(n, \mathbb{R})$ e $B = {}^tAA$. Dimostrare che B è simmetrica e che q_B è semidefinita positiva, e che è definita positiva se e solo se A è invertibile.

5. Applicazioni bilineari antisimmetriche

Ritorneremo ai prodotti scalari nel prossimo capitolo. Ora vogliamo occuparci delle forme antisimmetriche. Sia V uno spazio vettoriale su K e

$$Q : V \times V \to K$$

una applicazione bilineare antisimmetrica. Noi supporremo la caratteristica di K diversa da 2 ($2 \neq 0$); in questo caso $Q(u, u) = -Q(u, u)$ e allora $Q(u, u) = 0$.

Lemma 5.1. *Siano v e w due vettori di V tali che $Q(v, w) = 1$, e sia $L := \{u \in V : Q(v, u) = Q(w, u) = 0\}$. Allora:*

i) *v e w sono linearmente indipendenti;*

ii) L è un sottospazio vettoriale di V;
iii) $L\oplus$ span$[v,w] = V$.

DIMOSTRAZIONE. i) Se $av+bw = 0$, allora $0 = Q(av+bw, v) = aQ(v, v)+bQ(w, v) = -b$ (quindi $b = 0$), ma anche $0 = Q(av + bw, w) = aQ(v, w) = 0$ (quindi $a = 0$).

ii) È una verifica.

iii) Per provare che $L\oplus$ span$[v,w] = V$, dobbiamo vedere:

(a) $L\cap$ span$[v, w] = \{\underline{0}\}$.
Sia $u \in L\cap$ span$[v, w]$; allora $u = av + bw$; da $Q(u, v) = Q(u, w) = 0$ si trova $a = 0, b = 0$.

(b) $L+$ span$[v, w] = V$.
Sia $u \in V$; siano inoltre $b = Q(v, u)$ e $a = Q(w, u)$. Se $u' = u + (av - bw)$, si ha:

$$Q(v, u') = Q(v, u + av - bw) = Q(v, u) - bQ(v, w);$$
$$Q(w, u') = Q(w, u) + aQ(w, v) = 0.$$

Quindi $u' \in L$ e $u = (av - bw) + u' \in$ span$[v, w] + L$.

□

PROPOSIZIONE 5.2. Supponiamo $\dim V = n < +\infty$, Q bilineare e antisimmetrica su V; allora esiste una base di V $\{v_1, \dots, v_k; w_1, \dots, w_k, u_1, \dots, u_s\}$ tale che $Q(v_i, w_i) = 1$; $Q(w_i, v_i) = -1$; $Q(v_i, w_j) = 0$ se $i \neq j$ $Q(v_i, u_j) = Q(w_i, u_j) = 0$ per ogni i e j. La matrice associata a tale base

$$J(k, s) = \begin{pmatrix} O & I_k & 0 \\ -I_k & 0 & 0 \\ O & O & 0_s \end{pmatrix}$$

DIMOSTRAZIONE. Induzione sulla dimensione.

(1) $n = 1$. Non vi è nulla da dimostrare $Q = 0$.

(2) $n > 1$. Se $Q = 0$ non vi è nulla da dimostrare, altrimenti esistono v e w in V con $Q(v, w) = a \neq 0$, posto $v_1 = v$ e $w_1 = a^{-1}w$ si ha $Q(v_1, w_1) = 1$, sia $L = \{u \in V : Q(v_1, u) = Q(w_1, u) = 0\}$. Dal lemma si ha $L\oplus < v1, w1 >= V$, $dim(L) = n - 2$. Se P è la restrizione di Q a L esiste per induzione una base $\{v_2, \dots, v_k; w_2, \dots, w_k, u_1, \dots, u_s\}$, u1,...,us di L: tale che matrice associata a P è

$$J(k-1, s) = \begin{pmatrix} O & I_{k-1} & 0 \\ -I_{k-1} & 0 & 0 \\ O & O & 0_s \end{pmatrix}$$

La base di V cercata è allora $\{v_1, \dots, v_k; w_1, \dots, w_k, u_1, \dots, u_s\}$

□

Osservazione. L'intero k e quindi il rango classifica la congruenza delle forme. Come corollario immediato si ha:

COROLLARIO 5.3. Ogni matrice antisimmetrica di ordine n è congrua a $J(k, s)$, con $2k+s = n$; in particolare il suo rango è pari.

CAPITOLO 7

Spazi vettoriali Euclidei

In questa sezione analizzeremo le strutture indotte su uno spazio vettoriale reale da un prodotto scalare. In particolare questo permetterà di calcolare le distanze e gli angoli tra due vettori.

Definizione 0.4. La coppia (V, g) con V spazio vettoriale reale e g prodotto scalare su V sarà chiamata *spazio vettoriale euclideo* (s.v.e.).

Nel seguito useremo l'abbreviazione $g(v, w) = (v, w)$. Allora, per ogni v, w in V, $(v, w) \in \mathbb{R}$ e:

(1) $(v, w) = (w, v)$,
(2) $(av + bu, w) = a(v, w) + b(u, w)$,
(3) $(v, v) \geq 0$, e $v \neq \underline{0} \Rightarrow (v, v) > 0$.

La norma di un vettore v sarà per definizione $\|v\| = \sqrt{(v, v)}$.

1. Processo di ortogonalizzazione di Gram-Schmidt

Siano $v_1, \dots, v_k$ vettori linearmente indipendenti di $(V, (\cdot, \cdot))$ (s.v.e.) poniamo induttivamente $w_1 = v_1$; $w_2 = v_2 - \frac{(v_2, v_1)}{(w_1, w_1)} w_1 = v_2 - \frac{(v_2, w_1)}{(w_1, w_1)} v_1$;

$$w_i = v_i - \sum_{j=1}^{i-1} \frac{(v_i, w_j)}{(w_j, w_j)} w_j. \tag{7.1}$$

Proposizione 1.1. Il sistema di vettori $w_1, \dots, w_k$ così costruito è ortogonale; inoltre, per ogni i, si ha: span$[w_1, \dots, w_i]$ =span$[v_1, \dots, v_i]$.

Dimostrazione. Vediamo subito che, per linearità:

$$(w_1, w_2) = (v_1, v_2) - (v_1, \frac{(v_2, w_1)}{(w_1, w_1)} v_1) = (v_1, v_2) - \frac{(v_2, v_1)}{(v_1, v_1)} (v_1, v_1) = 0.$$

Supponendo induttivamente che $(w_j, w_s) = 0$ se $s < i$ e $j < i$, vogliamo vedere che vale $(w_s, w_i) = 0$ se $s < i$.

Supponendo induttivamente che:

$$(w_s, w_i) = (w_s, v_i - \sum j = 1^{i-1} \frac{(v_i, w_j)}{(w_j, w_j)} w_j) = (w_s, v_i) - \sum j = 1^{i-1} \frac{(v_i, w_j)}{(w_j, w_j)} (w_s, w_j).$$

Tutti i termini nella sommatoria sono nulli con l'eccezione del termine $j = s$.

$$(w_s, w_i) = (w_s, v_i) - \frac{(v_i, w_s)}{(w_s, w_s)} (w_s, w_s) = (w_s, v_i) - (w_s, v_i) = 0.$$

Poniamo ora $V(s) = \text{span}[v_1, \dots, v_s]$ e $W(s) = \text{span}[w_1, \dots, w_s]$; chiaramente $V(1) = W(1)$. Supponiamo di aver già dimostrato che $V(j) = W(j)$ per $j < i$ e proviamo che $V(i) = W(i)$.

Intanto sappiamo per induzione che $\mathrm{span}[w_1,\dots,w_{i-1},v_i] = \mathrm{span}[v_1,\dots,v_{i-1},v_i] = V(i)$ e $\mathrm{span}[w_1,\dots,w_{i-1},w_i] = \mathrm{span}[v_1,\dots,v_{i-1},w_i] = W(i)$. Dalla formula (7.1) si ha che w_i è combinazione lineare dei w_j $(j < i)$ e di v_i: dunque $w_i \in \mathrm{span}[w_1,\dots,w_{i-1},v_i] = V(i) \Longrightarrow W(i) \subset V(i)$. D'altra parte:

$$v_i = w_i + \sum_{j=1}^{i-1} \frac{(v_i, w_j)}{(w_j, w_j)} w_j,$$

quindi $v_i \in W(i)$, e $\mathrm{span}[v_1,\dots,v_{i-1},v_i] = V(i) \subset W(i)$; segue che $V(i) = W(i)$. □

Nota. Il processo di ortogonalizzazione lascia fisso ogni sistema ortogonale.

Corollario 1.2.

(1) Se $W \subset V$ un sottospazio e $\dim W = m < +\infty$, esiste una base ortogonale (ortonormale) di W.
(2) Se $\dim V = n < +\infty$, allora si può completare ogni sistema ortogonale (ortonormale) $w_1,\dots,w_k$ ad una base ortogonale (ortonormale) di V.

Dimostrazione.

(1) Si applichi ad una qualsiasi base $\{w_1, w_2,\dots,w_m\}$ di W il processo di ortogonalizzazione.
(2) Si completi $\{w_1, w_2,\dots,w_m\}$ ad una base $\{w_1,\dots,w_m, v_{m+1},\dots,v_n\}$ a cui si applichi il processo di ortogonalizzazione. Per avere basi ortonormali basta dividere ogni vettore per la sua norma.

□

Osservazione. Dal Corollario segue immediatamente che, se $\dim V = n < +\infty$, per ogni prodotto scalare su V esiste una base ortonormale. Ovvero ogni matrice quadrata positiva è congrua all'identità. Questo è già stato dimostrato. Il lettore attento avrà già notato che il processo di ortogonalizzazione non è altro che il procedimento del teorema Lagrange (2.4) semplificato dal fatto che $q(v) \neq 0$ quando $v \neq 0$.

Esercizio: A= $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ è definita positiva. Trovare una base ortonormale del prodotto scalare $(x, y)_A =^t xAy$.

Soluzione. Cominciamo a trovare una base ortogonale: partiamo dalla base standard $\mathcal{E}$. Abbiamo $(e_1, e_1)_A = 1$; $(e_1, e_2)_A = 2$; $(e_1, e_3)_A = 0$; $(e_2, e_2)_A = 5$; $(e_2, e_3)_A = 0$; $(e_3, e_3)_A = 2$. Poniamo $w_1 = e_1, w_2 = e_2 - \frac{(e_2,e_1)_A}{(e_1,e_1)_A} e1 = e_2 - 2e_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$; $w_3 = e_3$ (e_3 è ortogonale a e_1 ed e_2, quindi a w_2). Per normalizzare dobbiamo calcolare le norme: $\|e_1\| = 1; \|e_3\| = \sqrt{2}$ e $(w_2, w_2) =^t w_2 A w_2 = 1$, quindi $\|w_2\| = 1$. La base ortonormale è: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$; $\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$; $\begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$. □

Su $\mathbb{R}^n$ indichiamo con $(x, y)_s =^t xy$ il prodotto scalare standard. La base standard $\mathcal{E} = \{e_1, e_2,\dots,e_n\}$ è ortonormale. La coppia $(\mathbb{R}^n, (\cdot,\cdot)_s)$ è lo spazio vettoriale euclideo standard. Se V ha dimensione n e $\mathcal{V} = \{v_1, v_2,\dots,v_n\}$ è una base ortonormale rispetto a $(\cdot,\cdot)$, l'applicazione coordinata è $F_v : V \to \mathbb{R}^n$; allora $(F_v(v), F_v(w))_s = (v, w)$. Il prodotto scalare scritto nella base ortonormale diventa il prodotto scalare standard, quindi l' isomorfismo $V \leftrightarrow \mathbb{R}^n$ è isomorfismo di spazi vettoriali euclidei.

Definizione 1.3. **(coefficienti di Fourier)** Se $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ è un insieme di vettori non nulli di V (s.v.e). Se $v \in V$, gli scalari $a_i = \frac{(v,v_i)_s}{(v_i,v_i)_s}$ si dicono coefficienti di Fourier di v rispetto a $\mathcal{V}$.

Proposizione 1.4. i) Se $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ è un sistema ortogonale (ortonormale) di V e $v = a_1v_1 + \ldots + a_iv_i + \ldots + a_nv_n$ allora $a_i = \frac{(v,v_i)}{(v_i,v_i}$ (se $\mathcal{V}$ è ortonormale, $a_i = (v, v_i)$).
ii) $\|v\|^2 = (\sum_i \frac{(v_i,v)^2}{(v_i,v_i)})$ (se v è ortonormale, $\|v\|^2 = \sum_i (v_i, v)^2$).

Dimostrazione. i) $(v, v_i) = (a_1v_1 + \ldots + a_iv_i + \ldots + a_nv_n, v_i) = a_i(v_i, v_i)$.
ii) $(v, v) = \sum_i (a_i)^2$.

□

Esempio: Indichiamo con $C^0([0, 2\pi], \mathbb{R})$ l'insieme delle funzioni continue dall'intervallo $[0, 2\pi]$ in $\mathbb{R}$. Sia $(f, g) = \int_0^{2\pi} f(x)g(x)dx$; esso è un prodotto scalare. Infatti $\int_0^{2\pi} f(x)^2 dx \geq 0$, e $\int_0^{2\pi} f(x)^2 dx = 0$ implica $f(x) = 0$. Ora, il sistema (infinito) di vettori

$$\{1; \sin(x), \ldots, \sin(nx) \ldots; \cos(x), \ldots, \cos(nx), \ldots\}$$

è ortogonale. Se $f \in C^0([0, 2\pi], \mathbb{R})$, si possono calcolare i coefficienti di Fourier di f e scrivere la classica *serie di Fourier* di f :

$$a_0 + \sum_n b_n \sin(nx) + \sum_n c_n \cos(nx).$$

Con le opportune definizioni di convergenza, la serie di Fourier converge a f. La terminologia prende appunto il nome da questo importante esempio.

Concludiamo con una applicazione semplice, ma di grande prospettiva:

Teorema 1.5. **(rappresentazione delle forme lineari)** Sia $(V, (\cdot, \cdot)$ uno spazio vettoriale euclideo di dimensione finita. Se $L : V \to \mathbb{R}$ è una applicazione lineare ($L \in V^\star$), allora esiste ed è unico un vettore $v \in V$ tale che:

$$L(w) = (v, w)$$

per ogni $w \in V$.

Dimostrazione. Sia $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ una base ortonormale di V. Se $w = x_1v_1 + x_2v_2 + \cdots + x_nv_n$, si ha:

$$L(w) = x_1L(v_1) + x_2L(v_2) + \cdots + x_nL(v_n).$$

Ponendo $v = L(v_1)v_1 + L(v_2)v_2 + \cdots + L(v_n)v_n$, si trova $L(w) = (v, w)$ per ogni w.
Per l'unicità: se $(v, w) = (v', w)$ per ogni w si ha $(v - v', w) = 0$ per ogni w, ma allora $(v - v', v - v') = 0$ e quindi $v = v'$. □

Osservazione. Sia V uno spazio vettoriale su K e $s : V \times V \to K$ un'applicazione bilineare. Indichiamo con $S : V \to V^\star$ la funzione lineare definita come segue. Per ogni v, $S(v) : V \to K$ è quella funzione lineare tale che per ogni w in V :

$$S(v)(w) = s(v, w).$$

Supponiamo s non degenere, cioè S iniettiva. Nel caso in cui V ha dimensione finita, dall'uguaglianza $\dim(V) = \dim(V^\star)$ si ha che S è biettiva. Allora se $L : V \to K$ è lineare ($L \in V^\star$) esiste ed è unico un vettore v in V tale che $S(v) = L$. Ovvero, per ogni w in V, $s(v, w) = L(w)$. Il teorema di rappresentazione vale.

2. Struttura metrica di uno spazio vettoriale euclideo

Veniamo ora a vedere che gli spazi vettoriali euclidei sono dotati in modo naturale di una metrica e quindi di una topologia. La proposizione chiave è la seguente:

Proposizione 2.1. **(disuguaglianza di Schwarz)** Sia $(V, (\cdot, \cdot))$ uno spazio vettoriale euclideo; per ogni v e w vettori di V si ha:

$$|(v, w)| \leq \|v\| \|w\|,$$

con uguaglianza se e solo se v e w sono linearmente dipendenti.

Prima dimostrazione. Se v o w sono zero non vi è nulla da dimostrare. Altrimenti facciamo partire il breve processo di ortogonalizzazione. Poniamo:

$$u = w - \frac{(v, w)}{(v, v)} v \; ; \quad (u, v) = 0.$$

Allora:

$$0 \leq (u, u) = (u, w - \frac{(v, w)}{(v, v)} v) = (u, w) = (w - \frac{(v, w)}{(v, v)} v, w) = (w, w) - \frac{(v, w)^2}{(v, v)}.$$

Quindi $(v, w)^2 \leq (v, v)(w, w)$. Estraendo la radice quadrata si ottiene la disuguaglianza di Schwarz. L'uguaglianza si ha se e solo se $(u, u) = 0$, ma questo implica che $u = 0$ e quindi $w - \frac{(v,w)}{(v,v)} v = 0$, cioè v e w linearmente dipendenti. $\square$

Seconda dimostrazione. Se v e $w = av$ sono dipendenti, allora:

$$|(v, w)| = |(v, av)| = |a| \|v\|^2 = \|w\| \|v\|.$$

Siano ora v e w indipendenti; sia $W := \text{span}[v, w]$ lo spazio (il piano) generato da v e w ($\dim(W) = 2$). La restrizione del prodotto scalare $(\cdot, \cdot)$ definisce un prodotto scalare su W. Utilizzando la base $\{v, w\}$ per costruire la matrice associata a tale forma bilineare simmetrica troviamo:

$$A = \begin{pmatrix} (v, v) & (v, w) \\ (w, v) & (w, w) \end{pmatrix}.$$

Essendo A una matrice positiva (associata ad una forma positiva), si ha:

$$\det(A) = (v, v)(w, w) - (v, w)^2 > 0.$$

$\square$

Un corollario immediato della disuguaglianza di Schwarz è il seguente:

Corollario 2.2. **(disuguaglianza triangolare)** Per ogni v e w in V : $\|v + w\| \leq \|v\| + \|w\|$.

Dimostrazione. $\|v + w\|^2 = (v + w, v + w) = \|v\|^2 + \|w\|^2 + 2(v, w) \leq \|v\|^2 + \|w\|^2 + 2|(v, w)| \leq \|v\|^2 + \|w\|^2 + 2\|v\| \|w\| = (\|v\| + \|w\|)^2$.
Estraendo la radice quadrata si ottiene: $\|v + w\| \leq \|v\| + \|w\|$. $\square$

Osservazione. Rispetto al prodotto scalare standard di $\mathbb{R}^n$ e in (8.1.4) la disuguaglianza di Schwarz e quella triangolare si scrivono rispettivamente:

(1)

$$|\sum_{i=1}^{n} x_i y_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2};$$

$$\sqrt{\sum_{i=1}^{n}(x_i+y_i)^2} \le \sqrt{\sum_{i=1}^{n} x_i^2} + \sqrt{\sum_{i=1}^{n} y_i^2};$$

(2)

$$|\int_0^{2\pi} f(x)g(x)dx| \le \sqrt{\int_0^{2\pi} f(x)^2dx}\sqrt{\int_0^{2\pi} f(x)^2dx};$$

$$\sqrt{\int_0^{2\pi} (f(x)+g(x))^2dx} \le \sqrt{\int_0^{2\pi} f(x)^2dx} + \sqrt{\int_0^{2\pi} f(x)^2dx}.$$

Dovrebbe essere chiaro che le disuguaglianze (si veda la seconda dimostrazione di quella di Schwarz) sono proprietà del piano vettoriale euclideo. Nel piano, quella triangolare, è quella classica della geometria di Euclide.

Definizione 2.3. Sia $V, (\cdot,\cdot)$ (s.v.e) costruiamo $d : V \times V \times \mathbb{R}$ definita da $d(v,w) := \|v-w\|$; $d(v,w)$ è la *distanza* tra v e w.

(8.2.1)

Proposizione 2.4. La distanza d definisce una metrica su V; per ogni u, v e w in V valgono:

I) $d(v,w) \ge 0$ e $d(v,w) = 0 \iff v = w$;
II) $d(v,w) = d(w,v)$;
III) (disuguaglianza triangolare) $d(v,w) \le d(v,u) + d(u,w)$.

Valgono inoltre

IV) (compatibilità con le traslazioni) $d(v+u, w+u) = d(v,w)$;
V) (compatibilità con le dilatazioni) per ogni $a \in \mathbb{R}$ $d(av, aw) = |a|d(v,w)$.

Dimostrazione. La I, II e la IV, V sono immediate. La III è la disuguaglianza triangolare. Infatti

$$d(v,w) = \|v-w\| = \|(v-u)+(u-w)\| \le \|v-u\| + \|u-w\| = d(v,u) + d(u,w).$$

$\square$

Topologia Euclidea. La disuguaglianza triangolare permette di definire una struttura di spazio topologico su V (s.v.e.). Ricordiamo la costruzione. Sia r un numero reale positivo e v un vettore (punto) di V; sia:

$$D(v,r) = \{w \in V \ : \ d(v,w) < r\}.$$

$D(v,r)$ è il disco aperto di centro v e raggio r. Ora $A \subset V$ si dice aperto se A è unione di dischi aperti oppure A è vuoto. Indichiamo con τ la famiglia di aperti.
Esercizio: Dimostrare che:

a) V e $\emptyset \in \tau$;
b) $\{A_i\}_{i\in I}$ e $A_i \in \tau \Longrightarrow \bigcup_{i\in I} A_i \in \tau$;
c) $A_1 \in \tau$, $A_2 \in \tau \Longrightarrow A_1 \cap A_2 \in \tau$.

Un insieme X dotato di una famiglia τ di sottoinsiemi avente le proprietà a),b) e c) dell'esercizio è uno *spazio topologico*. Una funzione f tra due spazi topologici X e Y si dice *continua* se per ogni aperto A di Y $f^{-1}(A)$ è aperto di X.

Definizione 2.5. Siano $v, w \in V$ e $v \ne \underline{0}, w \ne \underline{0}$. Si pone:

$$\cos(\vartheta(v,w)) = \frac{(v,w)}{\|v\|\|w\|}.$$

Osservazione. La definizione è ben data per la disuguaglianza di Schwarz . Un angolo $\vartheta \in [-\pi, \pi]$ è individuato, a meno del segno, dal coseno ($\cos(\vartheta) = \cos(-\vartheta)$). Conoscere il coseno dell'angolo equivale a determinarne il valore assoluto.

Il piano euclideo. Torniamo al piano Π della geometria euclidea. Attraverso il trasporto parallelo si è associato a Π uno spazio vettoriale reale $V(\Pi)$ di dimensione 2. Quando una unità di misura è fissata su $V(\Pi)$ risulta definito un prodotto scalare. Fissato un punto O (origine) di Π e la biezione $0 : P \to V(P)$. Se $v = O(P) = OP$ poniamo $\|v\| = d(OP)$, la distanza di P dall'origine. Allora se $v = O(P)$ e $w = O(Q)$ e ϑ è l'angolo tra v e w (se v o w sono zero si sceglie un angolo qualsiasi) si pone

$$(v, w) = \|v\|\|w\| \cos(\vartheta).$$

Dimostrare che $(\cdot, \cdot)$ è un prodotto scalare è un esercizio di trigonometria. Procediamo in altro modo, costruiamo due rette perpendicolari passanti per O.

Il teorema di Pitagora dice:

$$d(O, P)^2 = d(O, A)^2 + d(A, P)^2 = d(O, A)^2 + d(O, B)^2.$$

Se $v = OA$ e $w = OB$, si ha

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2;$$

se a e b sono due scalari:

$$\|xv + yw\|^2 = x^2\|v\|^2 + y^2\|w\|^2$$

(Teoremi di Pitagora e di Talete). Ora $xv = O(A')$ e $yw = O(B')$ $xv + yw = O(P')$. Poniamo $a = \|v\|^2$ e $b = \|w\|^2$

$$q(x, y) = ||xv + yw||2 = ax^2 + by^2$$

è una forma quadratica definita positiva. Se $d(OA) = d(OB) = 1$ abbiamo

$$q(x, y) = x^2 + y^2.$$

Utilizzando la corrispondenza di Cartesio con assi ortogonali e di vettori base unitari il piano vettoriale euclideo è isomorfo a quello standard.

3. Matrici ortogonali e isometrie

Riconsideriamo il prodotto di matrici. Sia $(\cdot, \cdot)$ il prodotto scalare standard di $\mathbb{R}^n$ e $\mathcal{E} = \{e_1, e_2, \ldots, e_n\}$ la base standard dei vettori colonna di $\mathbb{R}^n$. Siano $A = (v_1, \ldots, v_n)$ e $B = (w_1, \ldots, w_n)$. A e $B \in \mathcal{M}(n, \mathbb{R})$, dove i v_i e i w_i sono vettori colonna.

$$ {}^tBA = \begin{pmatrix} (w_1, v_1) & \ldots & (w_1, v_j) & \ldots & (w_1, v_n) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (w_i, v_1) & \ldots & (w_i, v_j) & \ldots & (w_n, v_j) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (w_i, v_1) & \ldots & (w_i, v_j) & \ldots & (w_n, v_j) \end{pmatrix}$$

Il prodotto di matrici è una collezione di prodotti scalari. Indichiamo con $O(n)$ il gruppo delle matrici ortogonali reali; A è ortogonale se e solo se:

$$I = {}^tAA = \begin{pmatrix} (v_1, v_1) & \ldots & (v_1, v_j) & \ldots & (v_1, v_n) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (v_i, v_1) & \ldots & (v_i, v_j) & \ldots & (v_n, v_j) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (v_i, v_1) & \ldots & (v_i, v_j) & \ldots & (v_n, v_j) \end{pmatrix}$$

cioè $(v_i, v_i) = 1$ e $(v_i, v_j) = 0$ se $i \neq j$.

Proposizione 3.1. . Sia $A \in M(n, \mathbb{R})$: sono equivalenti

i) $A \in O(n)$;
ii) le colonne (e le righe) di A sono una base ortonormale di $\mathbb{R}^n$
iii) $(Av, Aw) = (v, w)$ per ogni $v, w \in \mathbb{R}^n$.
iv) $\|Av\| = \|v\|$ per ogni $v \in \mathbb{R}^n$.

Dimostrazione.

i) $\Longleftrightarrow$ *ii)* già stata dimostrata.

i) $\Longleftrightarrow$ *iii)* Se A è ortogonale $(Av, Aw) = {}^t v({}^t AA)w =^t vw = (v, w)$. Se $(Av, Aw) = (v, w)$ allora ${}^t v({}^t AA)w =^t v({}^t w$ e quindi

$$ {}^t v({}^t AA - I)w =^t v({}^t B)w = 0 $$

Ove $B =^t AA - I$, ma allora $B = 0$ (la forma bilineare associata a B è nulla).

iii) $\Longleftarrow$ *iv)* Ovvia.

iii) $\Longrightarrow$ *iv)* Da $\|A(v + w)| = \|v + w\|$; $\|Av\| = \|v\|$; $\|A(w)\| = \|w\|$, ricaviamo $(v + w, v + w) = (A(v+w), A(v+w))$: (polarizzazione) $(v, v)+2(v, w)+(w, w) = (Av, Av)+2(Av, Aw)+ (Aw, Aw)$:

$$ (v, w) = (Av, Aw). $$

□

Corollario 3.2. Se $A \in O(n)$ e $\lambda \in \mathbb{R}$ ùn autovalore di A allora $\lambda = +1$ o $\lambda = -1$.

Dimostrazione. Se $Av = \lambda v$ con $v \neq 0$ si ha

$$ \|v\| = \|Av\| = \|\lambda v\| = |\lambda|\|v|. $$

Quindi $|\lambda| - 1)\|v\| = 0$: $|\lambda| = 1$. □

Definizione 3.3. Sia (X, d) uno spazio metrico. Una applicazione suriettiva $f : X \to X$ è una isometria se $d(f(x), f(y)) = d(x, y)$ per ogni coppia x e y di punti di X

Osservazione. Ogni isometria è ovviamente continua; inoltre è iniettiva: se $f(x) = f(y)$, allora $0 = d(f(x), f(y)) = d(x, y)$, quindi $x = y$.

Esercizio: La composizione di due isometrie è una isometria, e l'inversa di una isometria è una isometria.

Se $A \in O(n)$ è $v \in \mathbb{R}^n$ possiamo costruire la funzione $H(A, v) : \mathbb{R}^n \to \mathbb{R}^n$:

$$ H(A, v)(x) \;=\; Ax + v $$

Proposizione 3.4. La funzione $H(A, v)$ è una isometria rispetto alla metrica euclidea.

Dimostrazione. $d(H(A, v)(u), H(A, v)(w)) = \|(Au + v) - (Aw + v)\| = \|A(u - w)\| = \|u - w\| = d(u, w)$. □

Nota. La funzione $H(A, \underline{0})$ è l'operatore associato alla matrice A mentre $H(I, v)$ è la traslazione per il vettore v. Indichiamo con :

$$ G(n) = \{F \in B_n : F = H(A, v), A \in O(n), v \in \mathbb{R}^n\}. $$

Proposizione 3.5. Sia B_n l'insieme delle applicazioni biettive di $\mathbb{R}^n$, $G(n)$ è un sottogruppo di B_n.

Dimostrazione. Componendo: $H(A,v)\circ H(B,w)(x) = H(A,v)(Bx+w) = A(Bx+w)+v = ABx + (Aw+v)$. Quindi

$$H(A,v)\circ H(B,w) = H(AB, Aw+v).$$

Allora $H(A^{-1}, -A^{-1}v)\circ H(A,v) = H(I,\underline{0}) = id$:

$$(H(A,v))^{-1} = H(A^{-1}, -A^{-1}v).$$

□

Il gruppo $G(n)$ è detto il gruppo delle rototraslazioni di $\mathbb{R}^n$ o gruppo galileano. Abbiamo visto che ogni elemento di $G(n)$ è un'isometria di $\mathbb{R}^n$; ora proveremo che $G(n)$ è esattamente il gruppo delle isometrie di $\mathbb{R}^n$.

Esercizio: Dimostrare che $H(A,v)$ è biettiva.

Teorema 3.6. Una funzione $F\colon \mathbb{R}^n \to \mathbb{R}^n$ è una isometria rispetto alla metrica euclidea se e solo se $F \in G(n)$ ovvero $F = H(A,v)$.

Premetteremo alcuni lemmi.

Lemma 3.7. Sia V uno spazio vettoriale euclideo e sia $F : V \to V$ un' isometria tale che $F(0) = 0$ allora

i) $(v,w) = (F(v), F(w))$;
ii) F è lineare.

Dimostrazione. i) Notiamo che $d(0,F(v)) = d(0,v)$ $d(0,F(w)) = d(0,w)$, allora $\|F(v)\|^2 = \|v\|^2$ e $\|F(w)\|^2 = \|w\|^2$, inoltre

$$\|F(w)-F(v)\|^2 = d(F(w),F(v))^2 = d(w,v)^2 = \|w-v\|^2$$

$$\|F(w)-F(v)\|^2 = \|F(w)\|^2 + \|F(v)\|^2 - 2(F(v),F(w)) = \|w\|^2 + \|v\|^2 - 2(v,w) = \|w-v\|^2$$

Quindi

$$2(v,w) = 2(F(v),F(w)).$$

ii) $F(v+w) = F(v) + F(w) \Longleftrightarrow \|F(v+w) - F(v) - F(w)\|^2 = 0.$

Ora

$$\begin{aligned}\|F(v+w)-F(v)-F(w)\|^2 &= (F(v+w),F(v+w)) + (F(v),F(v)) \\ &\quad +(F(w),F(w)) - 2(F(v+w),F(v)) \\ &\quad -2(F(v+w),F(w)) + 2(F(v),F(w)),\end{aligned}$$

e, usando la prima parte del lemma, risulta:

$$\begin{aligned}\|F(v+w)-F(v)-F(w)\|^2 &= (v+w,v+w) + (v,v) \\ &\quad +(w,w) - 2(v+w,v) - 2(v+w,w) + 2(v,w) \\ &= 0.\end{aligned}$$

Si ha:

$$F(\lambda v) = \lambda F(v) \Longleftrightarrow \|F(\lambda v) - \lambda F(v)\|^2 = 0.$$

Ora $\|F(\lambda v) - \lambda F(v)\|^2 = (F(\lambda v) - \lambda F(v), F(\lambda v) - \lambda F(v)) = (F(\lambda v), F(\lambda v)) + \lambda^2(F(v),F(v)) - 2\lambda(F(\lambda v), F(v)) =$ (usando la prima parte del lemma) $(\lambda v, \lambda v) + \lambda^2(v,v) - 2\lambda(\lambda v, v) = 0.$

□

Conclusione della dimostrazione del teorema. Se F è una isometria di $\mathbb{R}^n$, sia $K : RnRn, K(x) = F(x) - F(0)$. Allora K una isometria e K(0)=0. Ne segue che esiste una matrice ortogonale A tale che

$$F(x) - F(0) = FA(x) = Ax$$

Quindi F(x)=Ax+F(0), poniamo F(0)=v si ha F=HA,v. □

4. Proiezioni ortogonali

Sia $(V, (\cdot, \cdot))$ uno spazio vettoriale euclideo. Sia W un sottospazio di V sia:

$$W^{\perp} := \{v \in V \ : \ (v, w) = 0 \ \forall \ w \in W\}.$$

Si Noti che W è un sottospazio di V, se v e u sono in W , a e b sono scalari e $w \in W$ allora

$$(av + bu, w) = a(v, w) + b(v, u) = 0 + 0 = 0.$$

Notiamo che $W^{\perp}$ contiene tutti i vettori ortogonali a tutti i vettori di W.

Definizione 4.1. Lo spazio $W^{\perp}$ detto lo spazio ortogonale a W.

(8.4.1)

Proposizione 4.2.
- i) $W \cap W^{\perp} = \underline{0}$.
- ii) $(A + B)^{\perp} = A^{\perp} \cap B^{\perp}$.
- iii) $W \subset (W^{\perp})^{\perp}$
- iv) Se $\dim W) = n < +\infty$ allora $V = W \oplus W^{\perp}$.
- v) Se $\dim W) = n < +\infty$ allora $W = (W^{\perp})^{\perp}$.

Dimostrazione.
- i) Se $v \in W \cap W^{\perp}$ allora $(v, v) = 0$ $(v \in W$ e $v \in W^{\perp}$. Quindi $v = 0$.
- ii) Se $v \in (A + B)^{\perp}$ allora $(v, a + b) = 0$ per ogni $a \in A$ e $binB$, posto $a = 0$ abbiamo $v \in B^{\perp}$, prendendo $b = 0$ si ha $v \in A^{\perp}$. Se viceversa $v \in A$∪B allora $(v, a + b) = (v, a) + (v, b) = 0$ e $v \in (A + B)^{\perp}$.
- iii) ovvio.
- iv) Sia $w_1, \ldots, w_1 n$ una base ortonormale di W si ha allora
$$v \in W\perp \Longleftrightarrow (v, w_i) = 0, \ i = 1, 2, \ldots n.$$
Ora se $v \in V$ poniamo
$$w = (v, w_1)w_1 + ... + (v, w_i)w_i + ... + (v, w_n)w_n,$$
$w \in W$. Sia $u = v - w$ ora $(u, w_i) = (v - w, w_i) = (v, w_i) - (w, w_i) = 0$. Quindi $u \in W^{\perp}$ e $v = w + u$ ne segue che $V = W \oplus W^{\perp}$ la somma diretta per i).
- v) $W \oplus W^{\perp} = V$ e $W \subset (W^{\perp})^{\perp}$. Se $v \in ((W^{\perp})^{\perp}$ scriviamo $v = w + u$ con $w \in W$ e $u \in W^{\perp}$, $u = v - w \in W^{\perp} \cap (W^{\perp})^{\perp} = \underline{0}$.

□

Corollario 4.3. Se $\dim(V) = n < +\infty$ e $\dim(W) = k$, allora $\dim(W^{\perp}) = n - k$.

Dimostrazione. Ovvia. □

Definizione 4.4. **(decomposizione ortogonale)** Una decomposizione $V = A_1 \oplus A_2 \oplus \cdots \oplus A_k$ si dice ortogonale se per ogni $v_i \in A_i$ e $v_j \in A_j$, con $i \neq j$, si ha $(v_i, v_j) = 0$, e gli A_i non sono banali.

Osservazione. La decomposizione $V = A_1 \oplus A_2 \oplus \cdots \oplus A_k$ è ortogonale se e solo se $(A_i)^\perp = A_1 \oplus \cdots \oplus A_{i-1} \oplus A_{i+1} \oplus \cdots \oplus A_k$. In particolare, nel caso di dimensione finita, se W è non banale allora $W \oplus W^\perp$ è una decomposizione ortogonale di V.

Proiezione. Se $V = W \oplus W^\perp$, ogni vettore $v \in V$ si decompone in modo unico

$$v = w + w' \ w \in W \ e \ w' \in W^\perp.$$

Poniamo allora $p = p_W : V \to V$,

$$p(w + w') = w,$$

p ̀la proiezione ortogonale su W.

Proposizione 4.5.

i) La funzione p è lineare.
ii) $Im(p) = W$.
iii) $ker(p) = W^\perp$,
iv) $p = p^2$;
v) $d(v, p(v))dim(v, w)$ per ogni $w \in W$ e se vale $d(v, p(v)) = dim(v, w)$ allora $p(v) = w$.

Dimostrazione.

i), ii) e *iii)* ovvie.
iv) $p^2(v) = p(p(v) + 0) = p(v)$.
v) Se $u \in W$ e $u' \in W^\perp$ $\|u + u'\|^2 = \|u||2 + ||u'|^2$ (Pitagora). Se $w \in W$, $d(v, w)^2 = \|v - w\|^2 = \|p(v) - w + w')\|^2 = \|p(v) - w\|^2 + \|w'\|^2 \ \|w'\|^2 = \|v - p(v)\|^2 = d(v, p(v))$.

□

Osservazione. La proiezione ortogonale $p(v)$ è il punto di minima distanza da v ad un sottospazio vettoriale.
Se $V = A_1 \oplus \cdots \oplus A_k$ è una decomposizione ortogonale, e $p_i = p_{A_i} : V \to V$ è la proiezione su A_i, allora $v = p_1(v) + \cdots + p_k(v)$.

5. Il gruppo ortogonale

In questa sezione vogliamo studiare il gruppo ortogonale $O(n)$. Ci è difficile sopravalutarne la sua importanza geometrica, abbiamo già visto come esso sia in qualche modo il cuore del gruppo di Galileo delle isometrie degli spazi euclidei.
Considerazioni topologiche elementari. Se identifichiamo $\mathcal{M} = \mathcal{M}(n, \mathbb{R})$ con $\mathbb{R}^{n^2}$, possiamo definire su $\mathcal{M}$ la topologia euclidea. Questa topologia è indotta dal prodotto scalare

$$g(A, B) := \text{ tr } ({}^t\!AB);$$

sia $\|A\|_\mathcal{M}$ la norma indotta. Questa ha alcune importanti proprietà: se $A =^t (w_1, \ldots, w_n)$,

$$\|Av\| = \|\big((v, w_1), \ldots, (v, w_n)\big)\| = \sqrt{\sum_{i=1}^n (v, w_i)^2}$$

$$\overset{\text{Schwartz}}{\leq} \sqrt{\sum_{i=1}^n \|w_i\|^2 \|v\|^2} = \|v\| \sqrt{\sum_{i=1}^n \|w_i\|^2} = \|A\|_\mathcal{M} \|v\|.$$

Inoltre, se $B = (v_1, \dots, v_n)$:

$$\|AB\|_{\mathcal{M}} = \sqrt{\sum_{i,j=1}^{n} (v_i, w_j)} \le \sqrt{\sum_{i,j=1}^{n} \|v_i\|^2 \|w_j\|^2} = \sqrt{(\sum_{i=1}^{n} \|v_i\|^2)(\sum_{j=1}^{n} \|w_j\|^2)} = \|A\|_{\mathcal{M}} \|B\|_{\mathcal{M}}.$$

Ricapitolando:

$$\|Av\| \le \|A\|_{\mathcal{M}} \|v\| \quad \text{e} \quad \|AB\|_{\mathcal{M}} \le \|A\|_{\mathcal{M}} \|B\|_{\mathcal{M}}. \tag{7.2}$$

PROPOSIZIONE 5.1. Il sottoinsieme $O(n)$ di $\mathcal{M}$ è:

i) compatto;
ii) non connesso.

DIMOSTRAZIONE. i) Dobbiamo provare che $O(n)$ è chiuso e limitato.

(a) *Chiusura.* Si consideri l'applicazione $F : \mathcal{M} \to \mathcal{M}$ definita da $F(A) = {}^t AA$; F è continua (verifica: F si ottiene per composizione di $A \mapsto (A, {}^tA)$ e della moltiplicazione $(A, B) \mapsto AB$) e $O(n) = F^{-1}(I)$: la controimmagine continua di un chiuso è chiusa.

(b) *Limitatezza.* Se $A \in O(n)$, $\|A\|_{\mathcal{M}} = \sqrt{\text{tr } {}^tAA} = \sqrt{n}$.

ii) La funzione determinate è continua (polinomiale) e $\det(O(n)) = \{1, -1\}$.
Se $J = \begin{pmatrix} -1 & 0 \\ 0 & I_{n-1} \end{pmatrix}$, si ha che $\det(I) = 1$ e $\det(J) = -1$. Se $O(n)$ fosse connesso esisterebbe una matrice A in $O(n)$ tale che $\det(A) = 0$, una contraddizione.

$\square$

Osservazione. La moltiplicazione $m : O(n) \times O(n) \to O(n)$, definita da $m(A, B) = AB$, e l'inversa $c : O(n) \to O(n)$, definita da $c(A) = A^{-1} = {}^t A$, sono continue; questo rende $O(n)$ (e $GL(n, \mathbb{R})$) un *gruppo topologico* cioè un gruppo dotato di una topologia le cui operazioni algebriche sono continue.

DEFINIZIONE 5.2. **(gruppo ortogonale speciale)** Sia $SO(n) := \{A \in O(n) \ : \ \det(A) = 1\}$; $SO(n)$ è un sottogruppo di $O(n)$ detto gruppo delle matrici ortogonali speciali.

$SO(2)$. Sappiamo che ogni matrice ortogonale con determinante 1 è una matrice di rotazione $A(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$; si ha che $A(\vartheta)A(\omega) = A(\vartheta + \omega)$. L'operatore associato F_ϑ è una rotazione di angolo ϑ nel piano e ha come autovalori $e^{i\vartheta}$. Possiamo allora identificare $SO(2)$ con il cerchio unitario di centro l'origine nel piano, ovvero con gli angoli, mediante la corrispondenza $F_\vartheta \mapsto (\cos(\vartheta), \sin(\vartheta))$.

Orientazione di uno spazio vettoriale reale. Vogliamo affrontare il problema della misurazione degli angoli. Premettiamo la seguente definizione valida per spazi reali.

DEFINIZIONE 5.3. Sia V uno spazio vettoriale reale (euclideo) t.c. $0 < \dim(V) = n < +\infty$.

- Due basi ordinate di V, $\mathcal{V}$ e $\mathcal{W}$, definiscono la stessa orientazione se il determinate della matrice del cambiamento di basi è positivo; in tale caso poniamo $\mathcal{V} \approx \mathcal{W}$.
- La $\approx$ definisce una relazione di equivalenza sulle basi ordinate di V.
- Una classe di equivalenza di $\approx$ è una orientazione di V.
- Uno spazio vettoriale è orientato quando una orientazione di V è stata scelta.

Osservazione. Su V vi sono sempre due sole orientazioni, precisamente $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ e $\mathcal{V}' = \{-v_1, \dots, v_n\}$. L'orientazione standard di $\mathbb{R}^n$ è quella definita dalla base standard, su $\mathbb{R}$ le due orientazioni sono date da $+1$ e -1.

Ritorniamo (un tormentone) al piano della geometria Euclidea. Indichiamo con $V(\Pi)$ lo spazio vettoriale della geometria Euclidea e con $(\cdot,\cdot)$ il prodotto scalare indotto dalla distanza euclidea. Fissiamo una base ortonormale $\{E,F\}$ di $V(\Pi)$ e quindi una sua orientazione.

LEMMA 5.4. Se v è un vettore unitario di $V(\Pi)$ esiste ed è unico un vettore w unitario t.c. $\{v,w\}$ sia ortonormale e definisca la stessa orientazione di $\{E,F\}$.

DIMOSTRAZIONE. Essendo v unitario, $v=\cos(\vartheta)E+\sin(\vartheta)F$.
Esistono solo due vettori w e w' di norma 1 tali che $(v,w)=0$: $w=-\sin(\vartheta)E+\cos(\vartheta)F$ e $w'=-w$. Solo $\{v,w\}$ definisce la stessa orientazione di $\{E,F\}$. □

Dati due vettori non nulli σ e ζ vogliamo ora definire l'angolo $\vartheta(\sigma,\zeta)\in[-\pi,\pi[$ tra σ e ζ. Siano $v=\frac{\sigma}{\|\sigma\|}$ e $u=\frac{\zeta}{\|\zeta\|}$. Preso w t.c.$\{v,w\}$ sia la base ortonormale costruita nel Lemma precedente, si ha:

$$u=\cos(\vartheta)v+\sin(\vartheta)w. \tag{7.3}$$

Quindi possiamo determinare $\vartheta(v,u)=\vartheta(\sigma,\zeta)$ in $[-\pi,\pi[$ per cui vale la (7.3).
Conclusione. Nel piano euclideo il prodotto scalare permette la definizione di $|\vartheta(\sigma,\zeta)|$, l'orientazione permette la scelta del segno di $\vartheta(\sigma,\zeta)$.

SO(3). La struttura di $SO(3)$ è più complessa di quella di $SO(2)$ (si noti che $SO(2)$ è abeliano). Tuttavia la sua descrizione è agevole.

LEMMA 5.5. Ogni matrice A di $SO(3)$ ammette l'autovalore $+1$.

DIMOSTRAZIONE. Il determinante di A è il prodotto dei suoi autovalori, e $\det(A)=1$. La matrice A può avere 3 autovalori reali x,y e t, oppure un autovalore reale x e due complessi coniugati z e $\overline{z}$. Quindi avremo o $xyt=1$ o $xz\overline{z}=x\|z\|^2=1$. In entrambi i casi un autovalore reale è positivo e quindi è $+1$. □

Quindi se $A\in SO(3)$ esiste v con $\|v\|=1$ tale che $Av=v$. Siano $L=\text{span }[v]$ lo spazio generato da v e $L^\perp$ lo spazio ortogonale a L.

LEMMA 5.6. (1) Se $w\in L^\perp$, allora $Aw\in L^\perp$, cioè $L^\perp$ è invariante rispetto all'operatore associato ad A.
(2) L'operatore indotto $F':L^\perp\to L^\perp$ è ortogonale e (ogni sua matrice associata) ha determinante 1.
(3) Se $A\neq I$, allora la molteplicità algebrica dell'autovalore $+1$ di Aè uno.

DIMOSTRAZIONE. (1) Sia $w\in L^\perp$: allora $0=(v,w)=(Av,Aw)=(v,Aw)$.
(2) Se scegliamo una base ortonormale di $\mathbb{R}^3$, $\{v,w,u\}$, la matrice associata a F_A mediante tale base è $B=\begin{pmatrix}1&0&0\\0&a&b\\0&c&d\end{pmatrix}$. La restrizione $F':L^\perp\to L^\perp$ è una isometria e, nella base ortonormale $\{w,u\}$, si rappresenta con la matrice $B'=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ che è ortogonale. Ora $\det(B)=\det(B')=\det(A)=1$. Quindi $B'\in SO(2)$.
(3) La matrice $B=\begin{pmatrix}1&0&0\\0&\cos(\vartheta)&\sin(\vartheta)\\0&-\sin(\vartheta)&\cos(\vartheta)\end{pmatrix}$ è simile ad A; tutti i suoi autovalori sono reali se e solo se $\sin(\vartheta)=0$, e quindi $\cos(\vartheta)=\pm 1$, quindi, se la molteplicità algebrica di $+1$ è maggiore di uno, $B=I$, ma allora $A=I$.

□

Abbiamo visto che la decomposizione ortogonale $L \oplus L^{\perp}$ è A-invariante. La retta L è l'asse di rotazione e $\vartheta \in]-\pi, \pi]$ è l'angolo di rotazione. Si ha allora:

Proposizione 5.7. Ogni matrice $A \in SO(3)$, $A \neq I$, è univocamente determinata dal suo asse e dal suo angolo di rotazione.

Dimostrazione. L'asse di rotazione L è unico quando $A \neq I$. La restrizione $F' : L^{\perp} \to L^{\perp}$ è una rotazione. □

Abbiamo visto che è possibile rappresentare ogni matrice di $SO(3) - \{I\}$ con un asse e un angolo. Questo permette di costruire un modello topologico. Sia:

$$D = \{v \in \mathbb{R}^3 \ : \ \|v\| \leq \pi\}$$

il disco chiuso di $\mathbb{R}^3$ di centro $\underline{0} = (0, 0, 0)$ e raggio π. Se $v \neq \underline{0}$, sia $L(v) = \text{span}[v]$ lo spazio generato da v e $\vartheta(v) = \|v\| \in]0, \pi]$. Definiamo $H : D \to SO(3)$:

$$H(v) = \begin{cases} \text{rotazione di angolo } \vartheta(v) \text{ e asse } L(v) & \text{se } v \neq \underline{0} \\ I \quad \text{se } v = \underline{0}. \end{cases}$$

Proposizione 5.8. (1) La funzione H è suriettiva;
(2) $H(v) = H(v')$ se e solo se $v = v'$ oppure $\|v\| = \|v'\| = \pi$ e $v' = -v$.

Dimostrazione. Riscrittura della proposizione precedente. □

Osservazione. È un po' laborioso dimostrare che H è una funzione continua. Invece di far questo, che del resto è intuitivo, vediamo come possiamo immaginarci $SO(3)$. Sia $D' := \{v \in \mathbb{R}^3 : \|v\| < \pi\}$ il disco aperto; $H : D' \to H(D')$ è biettiva. Sul bordo $\partial D := \{v \in \mathbb{R}^3 : \|v\| = \pi\}$, cioè la sfera di centro l'origine e raggio π, si ha invece $H(v) = H(-v)$. Possiamo pensare $SO(3)$ ottenuto da D incollando i punti antipodali del bordo. È istruttivo riflettere sulla superficie che si ottiene identificando i punti antipodali di una sfera. Se eliminate due calotte (circoli polari) e incollate i punti antipodali rimanenti, troverete un nastro di Moebious. Questa non banalità topologica di $SO(3)$ è legata alla teoria spinori.

CAPITOLO 8

Il teorema spettrale

In questa sezione daremo una dimostrazione del teorema spettrale. Studieremo l'interazione tra forme bilineari e gli autovaolri. Per questo il campo fissato è il campo dei numeri complessi $\mathbb{C}$. Dovremo anche modificare la nozione di prodotto scalare in modo da conservare i vantaggi della positività che avevamo mel caso euclideo.

1. Prodotti Hermitiani

Sia V uno spazio vettoriale definito su $\mathbb{C}$.

DEFINIZIONE 1.1. **(prodotto hermitiano)** Un prodotto hermitiano su V è una funzione $H : V \times V \to \mathbb{C}$ tale che:

a. $H(v + w, u) = H(v, u) + H(w, u)$;
b. $H(v, w + u) = H(v, u) + H(v, u)$;
c. $H(v, w) = \overline{H(w, v)}$ (la barra denota la coniugazione);
d. $H(v, \lambda w) = \lambda H(v, w)$ e $H(\lambda v, w) = \overline{\lambda} H(v, w)$;
e. $H(v, v) \geq 0$, e $H(v, v) = 0$ se e solo se $v = \underline{0}$.

DEFINIZIONE 1.2. Uno spazio vettoriale hermitiano è una coppia (V, H) dove V è uno spazio vettoriale complesso e H un prodotto hermitiano su V. Useremo nel seguito la notazione $H(v, w) = \langle v, w \rangle$.

Osservazione. Dalla c. risulta che $H(v, v) = \langle v, v \rangle$ è reale e quindi che la e ha senso. La norma di un vettore v è allora definita da: $\|v\| = \sqrt{\langle v, v \rangle}$. Come nel caso del prodotto scalare si ha: $\|\lambda v\| = |\lambda| \|v\|$, ove $|\cdot|$ è la norma in $\mathbb{C}$. Valgono analogamente la disuguaglianza triangolare, $\|v + w\| \leq \|v\| + \|w\|$, e quella di Schwarz, $|\langle v, w \rangle| \leq \|v\| \|w\|$.
ESEMPIO: Il prodotto hermitiano standard su $\mathbb{C}^n$ è definito da:

$$\langle x, y \rangle_s :=^t \overline{x} \cdot y = \sum_{i=1}^{n} \overline{x}_i y_i,$$

ove $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$. Con il processo di ortonormalizzazione di Gram-Schmidt (con la dovuta attenzione), se $\dim(V) = n < +\infty$ si può trovare una base unitaria $\{v_i\}$ di V, cioè tale che $\langle v_i, v_j \rangle = 0$ se $i \neq j$, $\langle v_i, v_i \rangle = 1$.

(Nel contesto hermitiano si preferisce il termine unitario ad ortonormale). Se allora $v = \sum_i x_i v_i$, $w = \sum_i y_i v_i$, $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ sono i vettori coordinate in $\mathbb{C}^n$, si ha:

$$\langle v, w \rangle = \langle \sum_i x_i v_i, \sum_i y_i v_i \rangle = \sum_{i=1}^{n} \overline{x}_i y_i =^t \overline{x} \cdot y = \langle x, y \rangle_s.$$

In altre parole ogni prodotto hermitiano su V, via l'isomorfismo $V \longleftrightarrow \mathbb{C}^n$ definito dalla mappa delle coordinate in una base unitaria, diventa il prodotto hermitiano standard.

La seguente proposizione introduce un concetto fondamentale.

Proposizione 1.3. Sia f un operatore di V, se $\dim(V) = n < +\infty$ esiste ed è unico un operatore f^* tale che per ogni coppia di vettori v e w :

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

Dimostrazione. (1) *Unicità.* (Non si usa $\dim(V) = n < +\infty$.) Se f^* e g^* soddisfano la proprietà: $\langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, g^{ast}(w)\rangle$, allora:

$$\langle v, f^*(w) - g^*(w)\rangle = 0$$

per ogni v e w; posto $v = f^*(w) - g^*(w)$, troviamo:

$$\langle f^*(w) - g^*(w), f^*(w) - g^*(w)\rangle = 0.$$

Quindi, per la e. di 1.1, $f^*(w) - g^*(w) = 0$, e $f^*(w) = g^*(w)$. Poiché questo vale per ogni w, segue $f^* = g^*$.

(2) *Esistenza.* Fissiamo una base unitaria $\{v_i\}_i$ di V e la matrice A associata ad f via tale base. Si consideri allora:

$$A^* :=^t \overline{A};$$

risulta:

$$\langle Ax, y\rangle =^t (\overline{A}\overline{x}) \cdot y =^t \overline{x} A^* y = \langle x, A^* y\rangle.$$

Se f^* è l'operatore associato a A^*, ne segue allora:

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

□

Definizione 1.4. **(aggiunto)** L'operatore f^* definito dalla precedente proposizione si dice l'aggiunto di f rispetto al prodotto hermitiano $\langle \cdot, \cdot\rangle$.

Definizione 1.5. Un operatore si dice *autoaggiunto* se $f^* = f$, *antiautoggiunto* se $f^* = -f$ e *unitario* se $f^* = f^{-1}$; f si dice **normale** se $ff^* = f^* f$ cioè f commuta con la sua aggiunta.

Sia $V = \mathbb{C}^n$ e $\langle \cdot, \cdot\rangle$ il prodotto hermitiano standard. Se f è un operatore di V associato alla matrice A, allora l'operatore aggiunto f^* è associato ad $A^* =^t \overline{A}$, detta matrice aggiunta. Una matrice si dice hermitiana se $A^* = A$, antihermitiana se $A^* = -A$, unitaria se $A^* = A^{-1}$, normale se $AA^* = A^* A$.

Nota. Vi è un'ovvia corrispondenza: se f è autoaggiunta (antiautoaggiunta, unitaria o normale) la matrice associata è hermitiana (antihermitiana, unitaria o normale) rispetto ad una base unitaria.

Esercizi: Scriviamo ogni matrice complessa A nella forma $A = C + iD$, ove C e D sono matrici reali, dette rispettivamente la parte reale e la parte immaginaria di A; A è reale se $D = \mathbf{0}$.

(1) Se A è hermitiana allora C è simmetrica e D antisimmetrica.
(2) Se A è unitaria e reale allora A è ortogonale.
(3) Le matrici hermitiane, antihermitiane e unitarie sono normali.
(4) Se ${}^t\overline{A} = p(A)$ con p polinomio allora A è normale.
(5) Vale il viceversa di 4?
(6) Verificare che l'insieme delle matrici unitarie $U(n) := \{U \in GL(n) : UU^* = I\}$ e l'insieme delle matrici unitarie speciali $SU(n) := \{U \in U(n) : \det(U) = 1\}$ formano gruppo rispetto alla moltiplicazione.
(7) Dire se le matrici normali invertibili formano gruppo moltiplicativo.
(8) A è unitaria se e solo se le sue colonne (righe) sono una base unitaria.
(9) Verificare: $(f^*)^* = f$, $(f+g)^* = f^* + g^*$, $(\lambda f)^* = \overline{\lambda} f^*$, $id^* = id$, $(f - \lambda id)^* = f^* - \overline{\lambda} id$, $(fg)^* = g^* f^*$.

2. Il teorema

Enunciamo ora il teorema spettrale.

Teorema 2.1. **(Teorema Spettrale I)** L'operatore f è normale rispetto a $\langle \cdot, \cdot \rangle$ se e solo se esiste una base unitaria composta di autovettori di f.

Il suo corrispettivo ed equivalente enunciato matriciale è il seguente:

Teorema 2.2. **(Teorema Spettrale II)** Una matrice A è diagonalizzabile attraverso matrici unitarie se e solo se è normale.

Metà della dimostrazione è semplice: se A è diagonalizzabile mediante matrici unitarie allora $U^{-1}AU = D$, ove D è diagonale e $U^{-1} = {}^t\overline{U}$; allora $A = UDU^{-1}$ e $A^* = {}^t\overline{A} = U\overline{D}U^{-1}$. Dunque, dato che due matrici diagonali commutano, $AA^* = UD\overline{D}U^{-1} = U\overline{D}DU^{-1} = A^*A$.

La parte più interessante è il viceversa. Lo dimostreremo usando la notazione operatoriale, che è più flessibile di quella matriciale. Noi dimostreremo una serie di lemmi elementari nei quali supporremo sempre f normale ($f^*f = ff^*$).

Lemma 2.3. Per ogni v e w, $\langle f(v), f(w)\rangle = \langle f^*(v), f^*(w)\rangle$.

Dimostrazione. $\langle f(v), f(w)\rangle = \langle v, f^*f(w)\rangle = \langle v, ff^*(w)\rangle = \langle f^*(v), f^*(w)\rangle$, dal momento che $(f^*)^* = f$. $\square$

Lemma 2.4. $\ker(f) = \ker(f^*)$.

Dimostrazione. Per il Lemma 2.3, $\langle f(v), f(v)\rangle = \langle f^*(v), f^*(v)\rangle$, e quindi $\|f(v)\| = \|f^*(v)\|$. Ora $f(v) = \underline{0} \Leftrightarrow 0 = \|f(v)\| = \|f^*(v)\| \Leftrightarrow f^*(v) = \underline{0}$. $\square$

Esercizio: Se $\|f(v)\| = \|f^*(v)\|$ per ogni v, allora f è normale.

Lemma 2.5. Se v è autovettore di f relativo all'autovalore λ, allora v è autovettore di f^* relativo all'autovalore $\overline{\lambda}$.

Dimostrazione. Sia $g = f - \lambda id$; allora $g^* = f^* - \overline{\lambda} id$; g è normale perché l'identità commuta con ogni operatore. Applicando il Lemma 2.4 si ha $\ker(g) = \ker(g^*)$, ma $\ker(g)$ e $\ker(g^*)$ sono rispettivamente il λ-autospazio di f e il $\overline{\lambda}$-autospazio di f^*. $\square$

Arriviamo al punto chiave:

Lemma 2.6. Esiste una base di autovettori di f. (f è s.s. o diagonalizzabile).

Dimostrazione. Se A non fosse diagonalizzabile, per il criterio ???(2.20) esisterebbero λ, v e w con $f(w) = \lambda w + v$, $f(v) = \lambda v$ e v non nullo. Ma allora, per il Lemma 2.5, $f^*(v) = \overline{\lambda} v$, e dall'eguaglianza $\langle f(w), v\rangle = \langle v, f^*(v)\rangle$ otterremmo: $\langle \lambda w + v, v\rangle = \langle w, \overline{\lambda} v\rangle$. Poiché $\langle \lambda w, v\rangle = \langle w, \overline{\lambda} v\rangle$, abbiamo $\langle v, v\rangle = 0$, che contraddice $v \neq \underline{0}$. $\square$

Lemma 2.7. Se λ e μ sono due autovalori distinti di f, allora i relativi autospazi, $V(\lambda)$ e $V(\mu)$, sono ortogonali, cioè se $v \in V(\lambda)$ e $w \in V(\mu)$, allora $\langle v, w\rangle = 0$.

Dimostrazione. Presi v e w nei due autospazi, abbiamo $f(v) = \lambda v$ e $f(w) = \mu w$; inoltre $f^*(v) = \overline{\lambda} v$. Dunque:

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

Quindi $(\mu - \lambda)\langle v, w\rangle = 0$, ed essendo $(\mu - \lambda) \neq 0$ si ha allora $\langle v, w\rangle = 0$. $\square$

Conclusione della dimostrazione del teorema spettrale. Scegliamo in ogni autospazio di f una base unitaria; l'unione di tali basi, per il Lemma 2.6, è una base di V, unitaria per il Lemma 2.7. □

Veniamo a qualche risultato speciale:

Teorema 2.8. (1) f è autoaggiunta (A hermitiana) $\iff$ $f(A)$ è normale e i suoi autovalori sono tutti reali.
(2) f è antiautoaggiunta (A antihermitiana) $\iff$ $f(A)$ è normale e i suoi autovalori sono tutti immaginari puri.
(3) f è unitaria (A unitaria) $\iff$ $f(A)$ è normale e i suoi autovalori hanno modulo 1.

Dimostrazione. $\implies$) 1. Sia $f(v) = \lambda v$, con $v \neq \underline{0}$. Essendo f autoaggiunta, si ha: $\langle v, f(v)\rangle = \langle v, f^*(v)\rangle$; allora $\langle v, \lambda v\rangle = \langle v, \overline{\lambda} v\rangle$ (per il Lemma 2.5) e $(\overline{\lambda} - \lambda)\langle v, v\rangle = 0$; dunque $\overline{\lambda} = \lambda$, cioè λ reale.
2. Se $f^* = -f$, allora (if) è autoaggiunta.
3. Se f è unitaria, $\langle v, f^*(v)\rangle = \langle v, f^{-1}(v)\rangle$. Sia $f(v) = \lambda v$, con $v \neq \underline{0}$. Da $\langle v, \overline{\lambda} v\rangle = \langle v, \lambda^{-1} v\rangle$ otteniamo $(\overline{\lambda} - \lambda^{-1})\langle v, v\rangle = 0$, e quindi $\overline{\lambda} = \lambda^{-1}$, ovvero $|\lambda| = 1$.
$\impliedby$) Semplici verifiche.

□

Proposizione 2.9. Siano $(V, \langle\cdot,\cdot\rangle)$ uno spazio hermitiano di dimensione finita, $f : V \to V$ un operatore normale e W un sottospazio invariante. Se la dimensione di V è finita, allora $W^\perp = \{v \in V : \langle v, w\rangle = 0 \ \forall w \in W\}$ è f-invariante.

Dimostrazione. L'operatore f è semisemplice, quindi la sua restrizione $f : W \to W$ ammette una base di autovettori, $w_1, \ldots, w_s$. Allora $f(w_i) = \lambda_i w_i$, e $W^\perp = \{v \in V : \langle v, w_i\rangle = 0, \ i = 1, \ldots, s\}$.
Sia ora $v \in W^\perp$; allora $\langle f(v), w_i\rangle = \langle v, f^*(w_i)\rangle = \langle v, \overline{\lambda}_i w_i\rangle = \overline{\lambda}_i \langle v, w_i\rangle = 0$, quindi $f(v) \in W^\perp$. □

Esercizio: Con le notazioni della proposizione precedente, dimostrare che $f : W \to W$ è normale.
Esercizio: Si considerino i vettori reali $v = (2, a+1, 0)$ e $w = (a, 1, 0)$ intesi come vettori colonna, al variare di $a \in \mathbb{R}$. Determinare le matrici ortogonali reali A di ordine 3 tali che v e w sono autovettori di A.

3. Teorema Spettrale Reale

Teorema 3.1. Sia $A \in \mathcal{M}(n, \mathbb{R})$; allora A è diagonalizzabile mediante matrici ortogonali se e solo se A è simmetrica. In particolare ogni matrice simmetrica reale è diagonalizzabile su $\mathbb{R}$.

Dimostrazione. Sia A tale che $O^{-1}AO = D$, ove D è diagonale (necessariamente reale) e O è ortogonale. Allora $O^{-1} =^t O$, $A = ODO^{-1}$, quindi ${}^tA =^t (ODO^{-1}) =^t O^{-1}\ {}^tD^tO = ODO^{-1} = A$.
Supponiamo ora ${}^tA = A$; sul campo complesso A è hermitiana, dunque è diagonalizzabile su $\mathbb{C}$ (perché è normale), e i suoi autovalori sono reali (perché è hermitiana). Ma allora tutti gli autovalori di A sono in $\mathbb{R}$. La dimensione degli autospazi è la stessa calcolata su $\mathbb{R}$ e su $\mathbb{C}$. Infatti le soluzioni del sistema lineare $Ax - \lambda x = \underline{0}$ dipendono solo dal rango di $A - \lambda I$, che non dipende dal campo. Quindi le molteplicità algebriche e geometriche degli autovalori di A

sono eguali. Ne segue che A è diagonalizzabile su $\mathbb{R}$. Notiamo infine che il prodotto hermitiano standard ristretto ad $\mathbb{R}^n$ dà il prodotto ortogonale standard:

$$< x, y >_s =^t x \cdot y = (x, y)_s.$$

Gli autospazi reali distinti sono ortogonali nel senso usuale, quindi possiamo trovare una base ortogonale di vettori reali che diagonalizza A. $\square$

Diamo ora la versione operatoriale del teorema spettrale.

Definizione 3.2. Siano V uno spazio vettoriale reale, $(\cdot, \cdot)$ un prodotto scalare su V, $f : V \to V$ un operatore; un operatore f^* di V si dice *aggiunto* di f rispetto a $(\cdot, \cdot)$ se per ogni v e w in V si ha $(f(v), w) = (v, f^*(w))$; f si dice *autoaggiunto* se $f^* = f$.

Lasciamo per esercizio la verifica delle varie proprietà delle aggiunte e la loro esistenza nel caso finito dimensionale. Nel caso del prodotto euclideo standard, se $f(x) = Ax$, con A matrice di ordine n, allora $f^*(x) =^t A$:

$$(x, {}^tAy)_s =^t x({}^tAy) = ({}^tx{}^tA)y =^t (Ax)y = (Ax, y)_s.$$

Teorema 3.3. **(teorema spettrale operatoriale)** Se V ha dimensione finita su $\mathbb{R}$, f è autoaggiunto rispetto a $(\cdot, \cdot) \iff$ esiste una base ortonormale di autovettori di f.

Dimostrazione. Si costruisca una base ortonormale di V; allora f è autoaggiunto se e solo se la sua matrice associata è simmetrica. $\square$

Nota. Il teorema spettrale reale è in qualche modo sorprendente. Eccone un'altra sua versione. Sia $\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ un' applicazione bilineare simmetrica; allora $\phi(x, y) =^t xAy$, con A simmetrica. Ora, l'esistenza di una matrice ortogonale che la diagonalizza ci dice che $O^{-1}AO = D$, ma anche che ${}^tOAO = D$, quindi A è cogrediente e simultaneamente simile ad una diagonale. Le colonne di O sono allora una base ortonormale rispetto al prodotto scalare standard $(\cdot, \cdot)_s$ e ortogonale rispetto a ϕ.
Esercizio: Sia Φ simmetrica; la segnatura (k, s) della forma associata ϕ è data dal numero k di autovalori positivi e dal numero s di autovalori negativi.

4. Matrici Positive

Se A è simmetrica e definita positiva scriveremo $A > 0$.
Esercizi:

(1) Se A è invertibile allora ${}^tAA > 0$.
(2) Se $A > 0$ allora $A^2 > 0$ e $A^{-1} > 0$.

Proposizione 4.1. (1) Data $A > 0$, esiste unica una matrice $B > 0$ tale che $B^2 = A$; usualmente si pone $B = \sqrt{A}$.
(2) Data $G \in GL(n, \mathbb{R})$, esistono uniche due matrici $O \in GL(n, \mathbb{R})$ ortogonale e $B > 0$ tali che $G = OB$.

Dimostrazione. (1) Utilizzando il teorema spettrale scriviamo $D = O^{-1}AO$, con

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ t0 & \dots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

Si pone

$$\sqrt{D} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{pmatrix},$$

e $B := O\sqrt{D}O^{-1}$; dunque $B^2 = A$ ($B =: \sqrt{A}$).
Veniamo all'unicità: se $C^2 = A$ con $C > 0$, allora la seguente affermazione vale:

$$W \text{è il } \lambda\text{-autospazio di } C \iff W \text{è il } \lambda^2\text{-autospazio di } A.$$

Quindi se $C^2 = A = B^2$, con $B > 0$ e $C > 0$, gli operatori associati a B e C coincidono sugli autospazi; essendo le matrici diagonalizzabili, ne segue $B = C$.

(2) Se $G = OB$, con $B > 0$ e O ortogonale, allora, ponendo $A =^t GG$, si ha $A = B^2$ (dunque $B = \sqrt{A}$) e $O = GB^{-1}$; questo dimostra l'unicità. Per l'esistenza verifichiamo che la O cosìdefinita è ortogonale:

$$^tOO =^t (GB^{-1})GB^{-1} =^t B^{-1}(^tGG)B^{-1} = B^{-1}B^2B^{-1} = I.$$

□